STI 2014 Leiden

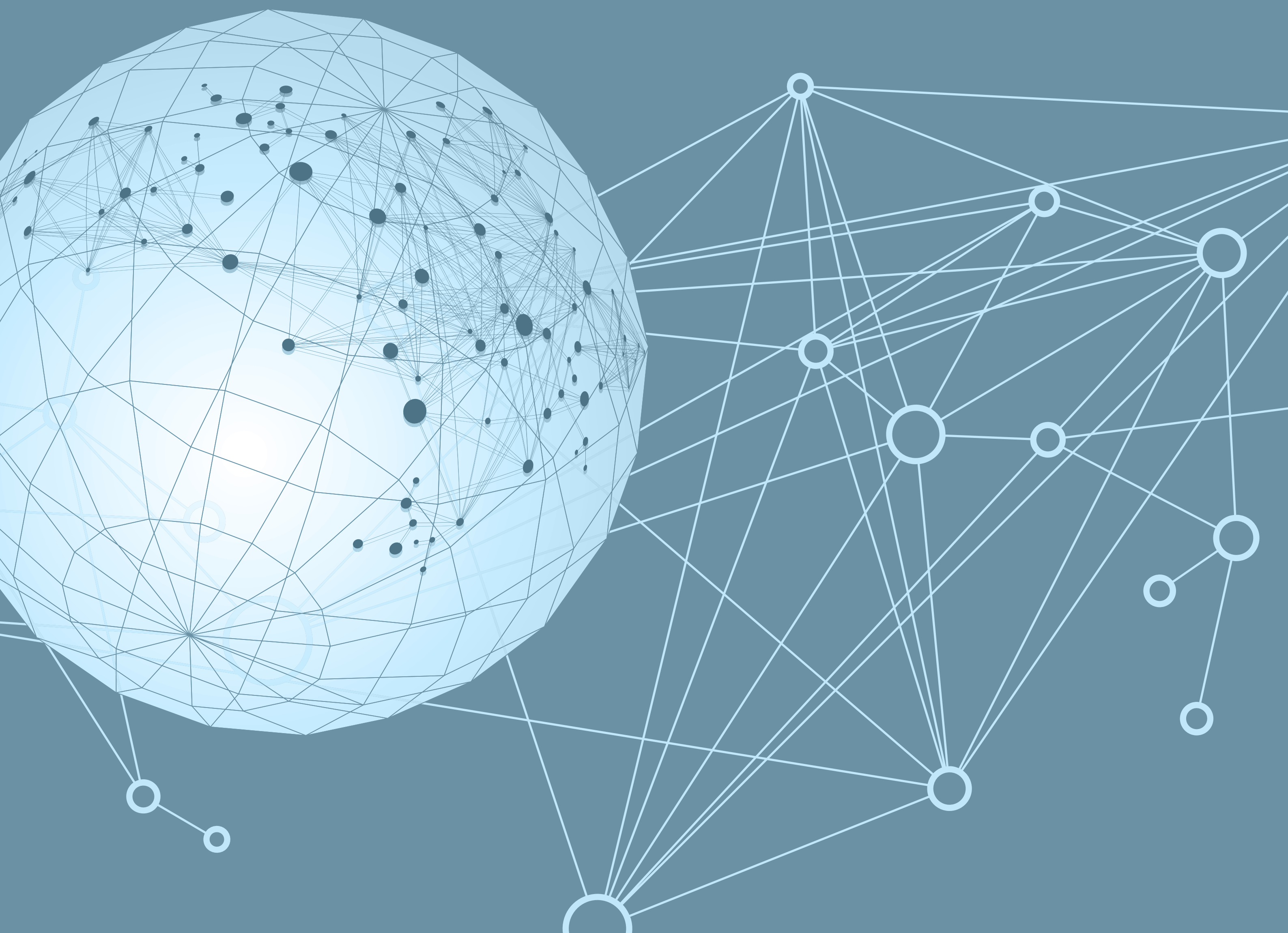
Proceedings of the science and technology indicators conference 2014 Leiden

"Context Counts: Pathways to Master Big and Little Data"

3 - 5 September 2014 in Leiden, the Netherlands

Electronic version - Edited by Ed Noyons

Published by Universiteit Leiden - CWTS 2014



Preface

This year, the Science and Technology Indicators (STI) conference is held in Leiden, the Netherlands, in collaboration with the European Network of Indicators Developers (ENID). The conference takes place in a period of historic transformations to the scientific and scholarly system. The conference motto “Context Counts – Pathways to Master Big and Little Data” aptly captures some of the most important changes.

First, we are witnessing the rise of new paradigms with respect to the economic and societal role of research. This is for example visible in the emphasis on societal relevance, the policy speak about Grand Challenges in Europe and the US, and the practices of new (and older) generations of researchers who try to combine breakthrough fundamental work with contributions to the solution of urgent problems. Although blue-sky research will remain crucial for scientific and scholarly progress, the new generations of researchers will work in a very different context from the generation that came out of World War II.

Second, the cumulative creation of data-generating machines and scientific instruments has led to a flood of data -- all challenging, not all meaningful. This data flood also has ramifications for our own field. With the shift towards web-based and computer-supported work in virtually all disciplines, the traces researchers leave in their daily work can increasingly be turned into data and indicators. In addition, social media are creating more (pressure on) the communicative activities of researchers, as exemplified by the rising sub-field of altmetrics.

Combined, the changing economic and societal role of research and the increasing availability of digital information lead to a rising demand for scientometric expertise. The present hunger for data and for indicators also lays bare a need for a meaningful interpretation. Scientometricians can no longer merely be data providers or indicator builders. They need to be able to put the data in the right context. And increasingly, they will also need to self-critically examine the use of their own products by the scientific and scholarly communities at large.

Indeed, context counts – in more than one way.

For the STI-ENID 2014 conference 125 papers were submitted. We accepted 70 oral presentations and 30 posters. Along with the regular indicators topics, the two trends discussed above are well represented in various sessions and in the 5 special events we scheduled on top of the regular program.

We are grateful to all authors for submitting their papers, posters and special events as well as to all members of the scientific committee for reviewing them. We also wish to thank Suze van der Luijt for producing and editing this book of proceedings.

Paul Wouters (Conference chair)
Ed Noyons (Editor)

Context counts: proceedings of the STI 2014

STI 2014 Organization and Program Committee

Conference chair
Paul Wouters

Local Organizing Committee
Maria Klijn
Ton van Raan
Mirella Imthorn
Inge van der Weijden
Suze van der Luijt
Ed Noyons

Program board
Judit Bar-Ilan
Vincent Lariviere
Jordi Molas
Ed Noyons (program chair)
Emanuela Reale
Inge van der Weijden (poster chair)

Scientific Committee
Jonathan Adams
Per Ahlgren
Fredrik Åström
Judit Bar-Ilan
Aparna Basu
María Bordons
Kevin W. Boyack
Clara Calero
Rodrigo Costas
Rickard Danell
Hans-Dieter Daniel
Koenraad Debackere
Gemma Derrick
Ying Ding
Tim Engels
Rainer Frietsch
Stephan Gauch
Yves Gingras
Wolfgang Glänzel
Jochen Gläser
Isabel Gomez
Juan Gorraiz
Stefanie Haustein
Diana Hicks
Sybille Hinze
Kim Holmberg
Marianne Hörlesberger
Stefan Hornbostel
Masatsura Igami
Margriet Jansz
Evaristo Jiménez-Contreras
Ben Jongbloed
Koen Jonkers
Sylvan Katz
Richard Klavans
Kayvan Kousha
Philippe Laredo
Vincent Larivière
Birger Larsen
Grit Laudel
Benedetto Lepori
Jacqueline Leta
Grant Lewison
Terttu Luukkonen
Marc Luwel
Valentina Markusova
Göran Melin
Martin Meyer
Stasa Milojevic
Henk Moed
Johann Mouton
Ulle Must
Anton Nederhof
Maria Nedeva
Ed Noyons
Bluma C. Peritz
Isabella Peters
Ismael Rafols
Emanuela Reale
Ed Rinia
Ronald Rousseau
Jane Russell
Elias Sanz-Casado
Andrea Scharnhorst
Edgar Schiebel
Christian Schloegl
Ulrich Schmoch
Jesper Schneider
András Schubert
Anne Sigogneau
Arnauld Sillet
Giorgio Sirilli
Daniel Sirtes
Gunnar Sivertsen
Stig Slipersæter
Cassidy Sugimoto
Mike Thelwall
Bart Thijs
Robert Tijssen
Peter Van Den Besselaar
Inge van der Weijden
Nees Jan Van Eck
Thed Van Leeuwen
Anthony Van Raan
Karen Vandevelde
Liwen Vaughan
Peter Vinkler
Martijn Visser
Ludo Waltman
Matthias Weber
Jos Winnink
Matthias Winterhager
Paul Wouters
K. Brad Wray
Michel Zitt
Alesia Zuccala

Table of Contents

Presentations at conference

All presentations (both oral and posters) in alphabetical order .

Enrichment of Bibliometric Databases by Assigning Region Information by means of the Web

Mehmet Ali Abdulhayoglu* and Bart Thijs*

[*]*Mehmetali.abdulhayoglu@kuleuven.be*
[*]*Bart.thijs@kuleuven.be*
Centre for R&D Monitoring (ECOOM) and Dept. MSI, KU Leuven, Waaistraat 6, Leuven, 3000 (Belgium)

Introduction

Promising are the recent experiments to use the Web as a tool for cleaning and correcting of address information. It is now possible to use various geographical open data sources such as GeoNames, GooglePlaces and Wikipedia to construct geographic information systems. Van Canneyt et al. (2013) states that the databases mentioned above have become increasingly popular to identify given user-specified places.

In this study, we aim at retrieving regional information through web services for a given place by using city name, postal code and country name indexed in Web of Science. Boulos, M.N.K (2012) studied such a similar work by enhancing the PubMed by means of GeoNames. We applied GeoNames and GooglePlaces both providing data that can easily be processed. Besides these services, we use Wikipedia when they fail to assign given address component to a region.

Data

Data in this study stems from a project where we provide indicators on sixteen regions in nine countries. All addresses from publications indexed in WoS for these countries in the period 1991-2011 were processed and about 10% (1.7 million addresses) could be assigned to one of the selected regions through a manual cleaning procedure. Our aim is to automate this process by applying unique combinations of country name, city name and postal code that occur at least in ten different addresses. This results in 28.488 combinations which represents 97.6% of all addresses.

Sources

GeoNames

The GeoNames offers some Web services to access regional data such as *postalCodeSearch* or *Search.* More information can be accessed via http://www.geonames.org. Figure2 gives the result for the query *'3000, Leuven, Belgium'*.

Figure2: Result of the GeoNames *postalCodeSearch*


```
▼<geonames>
   <totalResultsCount>1</totalResultsCount>
  ▼<code>
     <postalcode>3000</postalcode>
     <name>Leuven</name>
     <countryCode>BE</countryCode>
     <lat>50.87959</lat>
     <lng>4.70093</lng>
     <adminCode1>VLG</adminCode1>
     <adminName1>Vlaanderen</adminName1>
     <adminCode2/>
     <adminName2>Vlaams Brabant</adminName2>
     <adminCode3/>
     <adminName3>Leuven</adminName3>
   </code>
 </geonames>
```

To retrieve the correct match from the XML-formatted results, first postal code and country must exactly match with the ones in the query. Second, we control the city name since the result might not always give identical city name or the city name indexed in the database might be erroneous. To grab the similarity, we apply the Jaro-Winkler string similarity which is effective especially for short strings (Bilenko et al., 2003). We observe that matches with a Jaro-Winkler score of at least 0.80 are reliable. On the other hand, GeoNames might return a town of the city given in the query depending on the postal code or might return the city name in its own language which might result in a very small score.

GooglePlaces
It is applied as a complementary application to GeoNames for those records could not be matched by GeoNames *postalCodeSearch*. The website of this service is https://developers.google.com/places/. Figure3 shows an XML-formatted result.

Figure3: Result of the GooglePlaces


```
▼<address_component>
   <long_name>3000</long_name>
   <short_name>3000</short_name>
   <type>postal_code</type>
 </address_component>
▼<address_component>
   <long_name>Leuven</long_name>
   <short_name>Leuven</short_name>
   <type>locality</type>
   <type>political</type>
 </address_component>
▼<address_component>
   <long_name>Vlaams-Brabant</long_name>
   <short_name>VB</short_name>
   <type>administrative_area_level_2</type>
   <type>political</type>
 </address_component>
▼<address_component>
   <long_name>Vlaams Gewest</long_name>
   <short_name>Vlaams Gewest</short_name>
   <type>administrative_area_level_1</type>
   <type>political</type>
 </address_component>
▼<address_component>
   <long_name>België</long_name>
```


As in the previous application, the postal code with city and country names is checked. Only the region information for exact matches having identical postal codes and country names are retained. These region data are also matched with the GeoNames *Search* service for the same cities and countries. Among those only the records having the identical region names are taken as correct assignments.

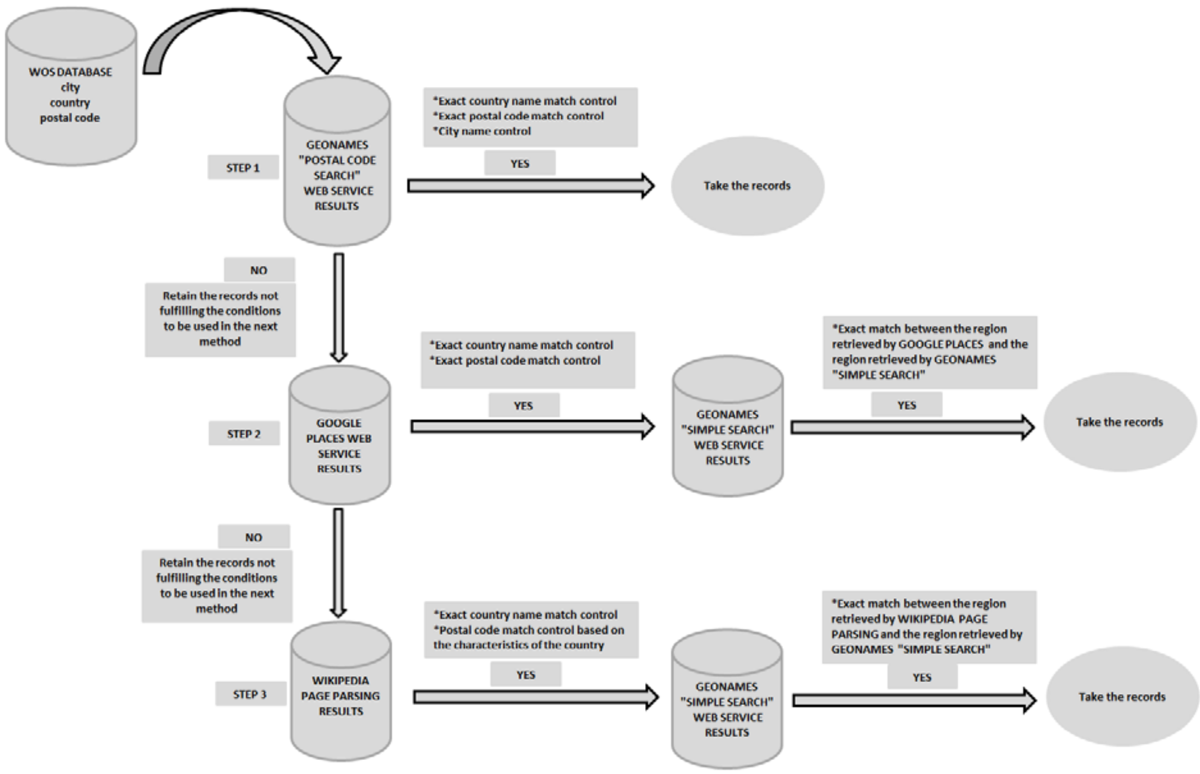
Wikipedia
To obtain region and postal code information for South Korea and Finland, we use the related links in the Wikipedia pages containing list of postal codes (http://en.wikipedia.org/wiki/List_of_postal_codes_in_South_Korea and http://en.wikipedia.org/wiki/List_of_postal_codes_in_Finland). In these pages the corresponding areas for each postal codes are given with their related Wikipedia link. By parsing these pages first, we get the corresponding Wikipedia page link. After accessing the Wikipedia page of the related places, we try to retrieve the region information by parsing the XML format of the page and confirm it by GeoNames *Search* service. Only those results matching with our database records and having exact region matches are accepted as a correct assignment. Figure4 gives an example of the XML-document containing the country and region name.

Figure4: A sample XML part of a Wikipedia page

```
<tr class="mergedtoprow">
<th scope="row" style="text-align:left;">Country</th>
<td>Belgium</td>
</tr>
<tr class="mergedrow">
<th scope="row" style="text-align:left;"><a href="/wiki/Communities,_regions_and_language_areas_of_Belgium" title="Communities, regions and language areas of Belgium">Community</a></th>
<td><a href="/wiki/Flemish_Community" title="Flemish Community">Flemish Community</a></td>
</tr>
<tr class="mergedrow">
<th scope="row" style="text-align:left;"><a href="/wiki/Communities,_regions_and_language_areas_of_Belgium" title="Communities, regions and language areas of Belgium">Region</a></th>
<td><a href="/wiki/Flemish_Region" title="Flemish Region">Flemish Region</a></td>
</tr>
<tr class="mergedrow">
<th scope="row" style="text-align:left;"><a href="/wiki/Provinces_of_Belgium" title="Provinces of Belgium">Province</a></th>
<td><a href="/wiki/Flemish_Brabant" title="Flemish Brabant">Flemish Brabant</a></td>
</tr>
```

Figure6 gives an overview of our methodology.

Figure6: Summary of the Process of Retrieving Regional Information



Results

We obtain promising results by applying GeoNames, GooglePlaces and Wikipedia parsing methods on retrieving regional information. Table1 and Table2 give the total number of combinations in its *"Total"* column.

Table1. Retrieved number of addresses and their percentages by GeoNames and GooglePlaces.

Countries STEP1&STEP2	#Components by GeoNames (STEP1)	#Components by Google Web Service (STEP2)	#STEP1+STEP2	Total	STEP1%	STEP1+STEP2%
Australia	1094	366	1460	2395	52.40	82.06
Austria	165	58	223	399	44.07	96.85
Belgium	322	93	415	599	67.17	92.82
Germany	2571	197	2768	5674	54.17	60.43
Netherlands	1785	-	1785	2275	82.06	82.06
Poland	1033	-	1033	1326	83.41	83.41
Spain	2178	274	2452	2939	83.37	83.60

The second, the third and the fourth columns of the Table1 give the number of combinations for which correct regional information could be retrieved by using GeoNames, Google and their sum, respectively. The percentages of addresses related to the matched postal code-city-country combinations are also given in the last two columns. Table1 shows that the application of GeoNames and GooglePlaces are powerful to retrieve correct region information for those countries except for Germany with more than 82% accuracy.

Table2. Retrieved number of addresses and their percentages through Wikipedia parsing (Step3).

Countries STEP3	#Components by Wikipedia Parsing (STEP3)	Total	STEP3%
South Korea	2120	3025	84.25
Finland	923	1092	99.73

Table2 gives the results for South Korea and Finland whose regions are assigned by parsing Wikipedia with a confirmatory service, Geonames *Search*. Table2 shows that the results are promising as 84.25% of the addresses from South Korea can be correctly assigned to a region while that is 99.73% for Finland.

Conclusion

We developed a promising method applying *GeoNames*, *GooglePlaces* and *Wikipedia parsing* to assign precise region information to the addresses from a set of countries on publications indexed in our WoS database. The percentages of the correct assignments for each country are high. Finally, the results based on our suggestions for retrieving region information are highly consistent with our previous study.

References

Bilenko, M., Mooney, R., Cohen, W., Ravikumar, P., & Fienberg, S. (2003). Adaptive name matching in information integration. *IEEE Intelligent Systems*,18(5), 16-23.

Boulos, M. N. K. (2012). Geosemantically-enhanced PubMed queries using the GeoNames ontology and Web Services. *In Geoscience and Remote Sensing Symposium (IGARSS), 2012 IEEE International*, 2891-2893.

Van Canneyt, S., Schockaert, S., & Dhoedt, B. (2013). Discovering and Characterizing Places of Interest Using Flickr and Twitter. *International Journal on Semantic Web and Information Systems (IJSWIS)*, 9(3), 77-104.

The SNIP indicator in relation to the Norwegian model[1]

Per Ahlgren

per.ahlgren@sub.su.se
University Library, Stockholm University, SE-106 91 Stockholm, Sweden

Introduction

Several citation-based indicators of journal impact exist. Perhaps the most well-known of these is the *Journal Impact Factor* (JIF). A major drawback of JIF, however, is its lack of field (subject) normalization. Differences in citation volumes between different fields are not taken into account. Recently, the Centre for Science and Technology Studies (CWTS), Leiden University, presented a field normalized citation-based indicator of journal/series impact, *source normalized impact per paper* (SNIP) (Moed, 2010; Waltman, van Eck, van Leeuwen, & Visser, 2013). SNIP belongs to a set of indicators that are based on the idea that citations to publications should be normalized against the length of the reference lists of the citing publications (e.g., Glänzel, Schubert, Thijs, & Debackere, 2011; Leydesdorff & Bornmann, 2011; Zitt, 2010). These *source normalized* indicators utilize the fact that the typical reference list length vary across fields. Clearly, this citing side normalization contrast with the traditional approach to field normalization, where a classification scheme is used (e.g., Braun & Glänzel, 1990; Moed, De Bruin, & van Leeuwen, 1995; van Raan, 1996). In this approach, each publication is assigned to one or more of the fields of the scheme. An example of an indicator that relies on a classification scheme is the *mean normalized citation score* (MNCS) (Waltman, van Eck, van Leeuwen, Visser, & van Raan, 2011a, 2011b). For this indicator, citation scores of the target publications (the publications under evaluation) are compared to expected citation scores for publications in the fields to which the publications belong. The fields used are the Thomson Reuters subject categories of journals. Source normalized indicators do not require a field classification scheme, which might be their main advantage. Instead, the field of a source is determined by publications that cite the source.

In the remaining part of this work, we let the term "source" stand for journals and series. SNIP is a quotient of which the numerator, the *raw impact per paper* (RIP), gives the average number of citations to the publications of a given source, where the publications are published in one of the years n, $n + 1$ and $n + 2$, and where the citing publications are published year $n + 3$ (the year of analysis) in sources covered by the database under consideration.

The denominator of SNIP is the *database citation potential* (DCP) of the source. DCP is based on the idea that sources such that their citing publications tend to have long reference lists have a higher potential to be cited compared to sources such that their citing publications tend to have short reference lists. In the definition of DCP, only active references are taken into account. An *active reference* in a publication, published in the year of analysis in a source covered by the database, is defined as a reference to a publication, published in a source covered by the database, during the three preceding years, relative to the year of analysis. For both RIP and DCP, cited and citing publications are included only if they are of the document types *article*, *conference paper* or *review*. Moreover, citations originating from certain sources are not counted in the calculation of SNIP values. Examples of such sources, *non-*

[1] The author would like thank Ludo Waltman for valuable remarks.

citing sources, are trade journals and sources with a small amount of references to other sources. For details on DCP and on the source exclusion rules, we refer the reader to Waltman et al. (2013).

The Norwegian model for evaluation of publications is applied yearly in Norway. The subjects of the evaluation are the Norwegian universities and university colleges. Research resources are distributed to these entities according to the result of the evaluation.

The Norwegian model can be said to combine production and impact. For the latter, though, citations are not used. Instead, the model considers the extent to which publications are published in channels, like journals, with large scientific prestige. A large number of channels have been assessed in Norway and assigned to exactly one of three levels:

0: Non-scientific publishing channel.
1: Scientific publishing channel.
2: Publishing channel with extra large scientific prestige.

For more information on the model, see Schneider (2009) and Sivertsen (2010).

One of the criteria used in Norway when journals are manually assigned to levels is the extent to which journals are cited. In earlier research, it has been shown that field normalized journal citation impact, measured on the basis on the subject categories of Thomson Reuters, correlate rather well with the manual assignments of journals to levels that are performed in Norway (Ahlgren, Colliander, & Persson, 2012). However, taken into consideration that the approach to field normalization underlying SNIP is considerably different from the approach that uses the subject categories of Thomson Reuters, it is reasonable to ask whether the SNIP values of sources tend to correspond to the levels of the sources, where we in this study, in addition to the three levels, take into consideration sources that have not been assessed in Norway. These sources form a separate category in the study. The purpose of the study, which involves more than 15,000 sources, is to investigate the relation between SNIP and the levels of the Norwegian model/the category of non-assessed sources within different subject area categories and across such categories.

Data and methods

Three lists of sources were utilized in the study:

- *CWTS Journal Indicators list*, September 2013 (CWTSList)
- *Scopus title list*, September 2013 (ScopusList)
- *The Norwegian list*, March 2013 (NoList)

CWTSList reports, among other things, SNIP values per source and year. All sources in CWTSList are indexed in Scopus. The list does not give subject information for the sources. However, such information is present in ScopusList. Each source in this list has been assigned to one or more of Scopus' subject area categories, 27 in number. NoList is a list, updated one time per year, over sources that have been assessed in Norway. Each entry in the list is associated with a level (0, 1 or 2).

From CWTSList, each source (a) with a SNIP value for year 2012, (b) with a print ISSN, and (c) classified as a citing source was extracted. The extracted sources were matched against ScopusList in order to get the subject area categories for each source. 15,177 sources are included in the dataset. Of these are 14,972 journals and 205 series (139 book series and 66 conference proceedings). These sources were matched against NoList in order to get levels for the sources. If a source was not found in NoList, the source was assigned the value -1. In this study, we regard -1 ("not in NoList") and the levels 0, 1 and 2 as categories on a nominal variable. The extraction and matching operations resulted in a list in which each source has a SNIP value, is associated with exactly one of the categories -1, 0, 1 or 2, and is associated with one or more subject area categories.

In the first part of the study, the SNIP value for a given source was weighted on the basis of the number of publications, of the three types referred to in the preceding section and published in the period 2009-2011, belonging to the source.[2] Let S be a subject category and C one of the four categories. We define the *weighted SNIP mean for S with respect to C*, SNIP(S, C), as

$$\mathrm{SNIP}(S,C) = \frac{\sum_{s \in S \cap C} (1/m_s) n_s SNIP(s)}{\sum_{s \in S \cap C} (1/m_s) n_s} \qquad (1)$$

where s is a source, m_s the number of subject area categories for s, n_s the number of publications belonging to s, and SNIP(s) the SNIP value for s.

Weighted SNIP means were obtained also for the four categories without regard to subject area categories. Such a mean is given by Equation (1), if "$S \cap C$" is replaced by "C" and the leftmost factors of the numerator and denominator are deleted.

In the second part of the study, a multinominal logistic regression analysis was performed in order to investigate the ability of SNIP to predict Norwegian model level/category of non-assessed sources.

Results

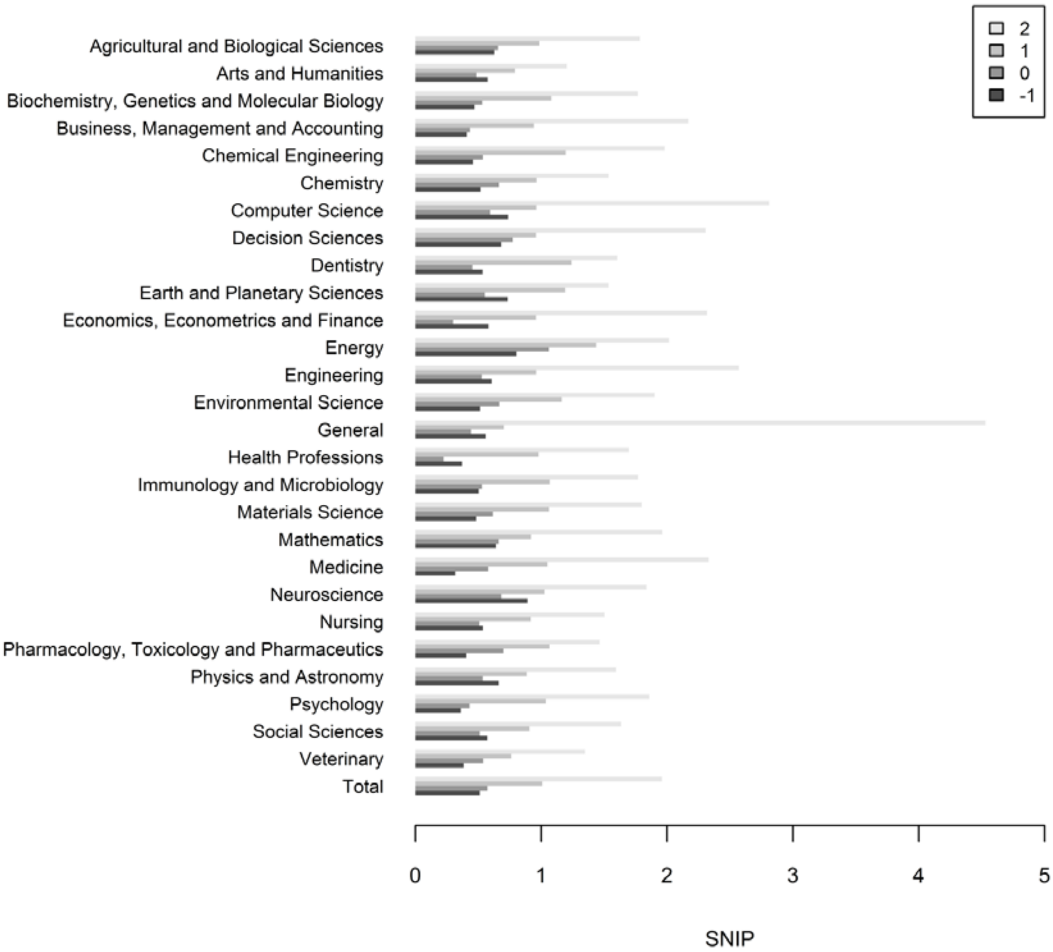
Figure 1 visualizes, for each subject area category, weighted SNIP means for the categories -1 ("not in NoList"), 0 (non-scientific publishing source), 1 (scientific publishing source) and 2 (publishing source with extra large scientific prestige). Also weighted SNIP means for the four categories without regard to subject area categories are indicated ("Total" on the vertical axis). For the latter, it is evident that the SNIP means increase consistently when we move from category -1[3], via the categories 0 and 1, to category 2.

For the 27 subject area categories it holds that the SNIP means for category 2 are consistently higher than the corresponding values for category 1, whereas the category 1 SNIP means are consistently higher than the corresponding values for category 0. The subject area category Neuroscience stands out: its SNIP mean for category -1 is about 30% higher than its mean for category 0 and not much lower than the corresponding value for category 1. The two prestige journals *Nature* (SNIP = 8.58) and *Science* (SNIP = 8.06) give rise to the high SNIP mean for the subject area category General and category 2, a combination with only three sources.

[2] Mathematically, the ability of SNIP to properly correct for field differences assumes weighted SNIP means (Waltman et al., 2013).

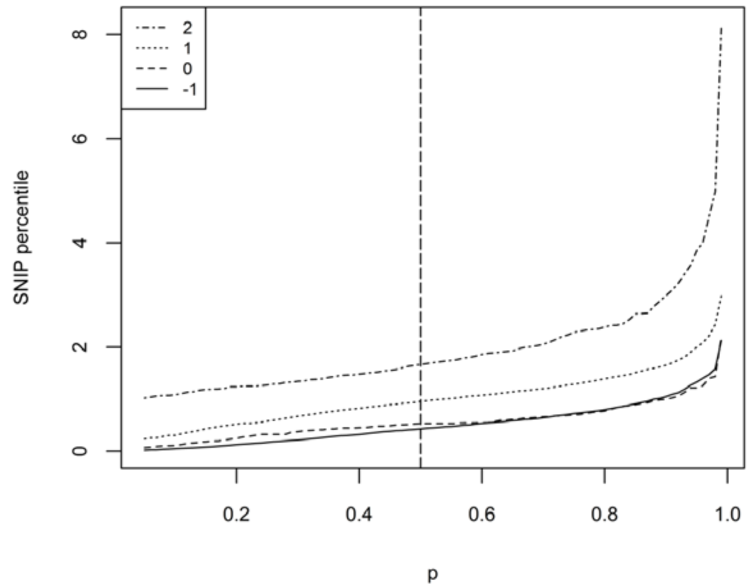
[3] 3,451 of the 15,177 sources (22.7%) belong to category -1 and are thus absent from NoList.

Figure 1: Weighted SNIP means for 27 subject area categories for the four categories -1, 0, 1 and 2. ”Total” concerns the SNIP means for the four categories without regard to subject area categories.



Percentiles at $p = 0.05, 0.06, \ldots, 0.99$ were calculated for the four distributions of SNIP values corresponding to the four categories -1, 0, 1 and 2, in order to complement the picture given by the means. Note that the SNIP value for a source *s* occurs n_s (the number of publications belonging to *s*) times in the distribution for the category of *s*. Figure 2 shows the outcome of the calculations. The curve for category 2 lies consistently above the curve for category 1, and the latter curve lies consistently above the curves for categories 0 and -1. The median, i.e. the percentile at $p = 0.5$, for category 2 is 1.67, whereas the medians for the categories 1, 0, and -1 are 0.96, 0.53 and 0.43, respectively. The same pattern is thus obtained as in the case of means.

Figure 2: SNIP percentiles at p = 0.05, 0.06, …, 0.99 for all four categories.



Several interesting observations regarding deviating cases can be done. The journal *Foundations and Trends in Machine Learning* belongs to category -1, and thereby does not occur in NoList. The category has the weighted SNIP mean 0.51, and the journal has the SNIP value 12.40, which corresponds to rank 9 when the 15,177 sources are ranked after SNIP values. Further, the two humanistic journals *Antiquite Tardive* and *Revue Romane*, both in category 2, have the SNIP value 0.

Multinominal logistic regression analysis

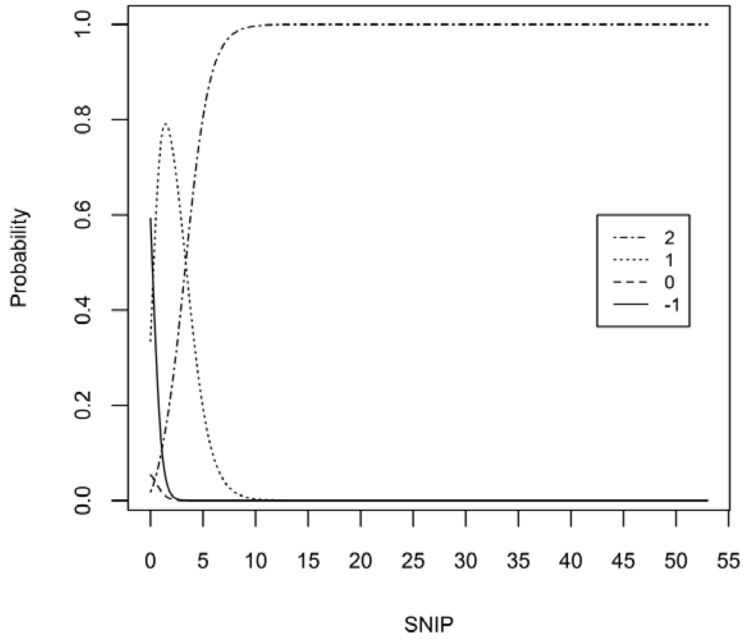
7,589 sources were randomly selected from the set of 15,177 sources. The sources in the resulting set were used as training data for the generation of the regression model, whereas the remaining 7,588 sources were used to test to which extent the model correctly classifies them.

Since we have four categories on the dependent variable, the regression model contains three regression coefficients for SNIP. The category -1 (”not in NoList”) was used as reference category, i.e. the category to which the other categories are compared. Let P(category = x) be the probability that a source belongs to category x, where x = 0, 1, 2, and let P(category = -1) be the probability that a source belongs to category -1. A given regression coefficient for SNIP, β, estimates how much ln(P(category = x)/P(category = -1)) changes when SNIP increases by 1.

ln(P(category = 2)/P(category = -1)) increases with β = 3.21 when SNIP increases by 1 (95% CI [3.03, 3.39]), and the odds becomes about 25 times greater with such an increase of SNIP. β for ln(P(category = 1)/P(category = -1)) is equal to 2.34 (95% CI [2.18, 2.50]). Interestingly, also β for ln(P(category = 0)/P(category = -1)) is positive, 0.52 (95% CI [0.14, 0.90]).

Figure 3 shows how the regression model predicts probabilities for category membership for sources at different SNIP values (0, 0.1, 0.2, 0.3, …, 53). For a given SNIP value, the model predicts four probabilities, one for each category, and the sum of these probabilities is equal to 1. At SNIP = 3.4 are the model predicted probabilities for membership in categories 1 and 2 approximately 0.5, and thereby are the predicted probabilities for membership in the other two categories close to 0. When SNIP increases from 3.4, the predicted probabilities for membership in category 1 decrease, while the corresponding probabilities for category 2 increase.

Figure 3: Predicted probabilities for category membership over SNIP values (0, 0.1, 0.2, 0.3, …, 53).



In Table 1, the classification accuracy of the regression model, with respect to the 7,588 sources in the set that was not used for model generation, is reported. For a given source with its SNIP value, the model predicts four probabilities, one for each category, and the source is assigned to the category with the highest predicted probability. 69.2% of the sources are correctly classified (last row/last column). For comparison with chance, the proportional chance criterion, where the sources are randomly, and proportionally, distributed over the categories, is 48.5%, whereas the maximum chance criterion, i.e. the share of sources in the largest category (category 1), is higher, 65.2% (Huberty, 1984; Morrison, 1969).

Table 1. Classification table for sources, which were not used for model generation. Number of sources = 7,588.

	Predicted				
Observed	-1	0	1	2	% correct
-1	606	0	1072	8	35.9
0	55	0	145	1	0.0
1	319	0	4579	49	92.6
2	4	0	686	64	8.5
Total, %	13.0	0.0	85.4	1.6	*69.2*

Discussion and conclusions

In this study, we have dealt with the relation between the SNIP indicator and the levels of the Norwegian model/the category of non-assessed sources. The result shows that there is a correlation between SNIP values and the four categories. This is perhaps not unexpected, since one of the criteria used in Norway when journals are manually assigned to levels is the extent to which journals are cited. Nevertheless, the correlation supports the standpoint that the manual assignments of sources to levels that are performed in Norway are reasonable, given that SNIP is considered as an indicator with a high degree of validity. Inversely, if the manual assignments are considered to be in the main reasonable, one can assert that the study supports that SNIP is an indicator with a high degree of validity. One should be aware of, although, that the results of the study are consistent with the possibility that both SNIP and the Norwegian assignments have defects that covary.

The Norwegian model has been criticized for under-coverage of sources. Under the assumption that SNIP is an indicator with a high degree of validity, this criticism is to some extent weakened by the study: sources that do not occur in NoList are associated with smaller SNIP values compared to sources that have been assessed as non-scientific.

As is reported in the results section, there are cases that deviate from the general pattern. For instance, journals with high SNIP values that have not been assessed in Norway. It might be a good idea for those involved in the assessments to take a closer look on deviating cases.

For future research, a comparison of SNIP and the revised SCImago Journal Rank indicator, SJR2 (Guerrero-Bote & Moya-Anegon, 2012), with respect to the ability to predict membership at level 2 of the Norwegian model is planned.

References

Ahlgren, P., Colliander, C., & Persson, O. (2012). Field normalized citation rates, field normalized journal impact and Norwegian weights for allocation of university research funds. *Scientometrics, 92*(3), 767-780.

Braun, T., & Glänzel, W. (1990). United Germany - the new scientific superpower. *Scientometrics, 19*(5-6), 513-521.

Glänzel, W., Schubert, A., Thijs, B., & Debackere, K. (2011). A priori vs. a posteriori normalisation of citation indicators. The case of journal ranking. *Scientometrics, 87*(2), 415-424.

Guerrero-Bote, V. P., & Moya-Anegon, F. (2012). A further step forward in measuring journals' scientific prestige: The SJR2 indicator. *Journal of Informetrics, 6*(4), 674-688.

Huberty, C. J. (1984). Issues in the use and interpretation of discriminant-analysis. *Psychological* Bulletin, *95*(1), 156-171.

Leydesdorff, L., & Bornmann, L. (2011). How Fractional Counting of Citations Affects the Impact Factor: Normalization in Terms of Differences in Citation Potentials Among Fields of Science. *Journal of the American Society for Information Science and Technology, 62*(2), 217-229.

Moed, H. F. (2010). Measuring contextual citation impact of scientific journals. *Journal of Informetrics, 4*(3), 265-277.

Moed, H. F., De Bruin, R. E., & van Leeuwen, T. N. (1995). New bibliometric tools for the assessment ofnational research performance: Database description, overview of indicators and first applications. *Scientometrics, 33*(3), 381-422.

Morrison, D. G. (1969). Interpretation of discriminant analysis. *Journal of Marketing Research, 6*(2), 156-163.

Schneider, J. W. (2009). an outline of the bibliometric indicator used for performance-based funding of research institutions in norway. *European Political Science, 8*(3), 364-378.

Sivertsen, G. (2010). A performance indicator based on complete data for the scientific publication output at research institutions. *ISSI Newsletter*, *6*(1), 22-28.

van Raan, A. F. J. (1996). Advanced bibliometric methods as quantitative core of peer review based evaluation and foresight exercises. *Scientometrics, 36*(3), 397-420.

Waltman, L., van Eck, N. J., van Leeuwen, T. N., & Visser, M. S. (2013). Some modifications to the SNIP journal impact indicator. *Journal of Informetrics*, *7*(2), 272-285.

Waltman, L., van Eck, N. J., van Leeuwen, T. N., Visser, M. S., & van Raan, A. F. J. (2011a). Towards a new crown indicator: an empirical analysis. *Scientometrics, 87*(3), 467-481.

Waltman, L., van Eck, N. J., van Leeuwen, T. N., Visser, M. S., & van Raan, A. F. J. (2011b). Towards a new crown indicator: Some theoretical considerations. *Journal of Informetrics*, *5*(1), 37-47.

Zitt, M. (2010). Citing-side normalization of journal impact: A robust variant of the Audience Factor. *Journal of Informetrics, 4*(3), 392-406.

Theoretical foundations and aplications: a study of normalized indicators Salton's Cosine and Jaccard Index in Author Co-citation Analysis

Bruno Henrique Alves* and Ely Francina Tannuri de Oliveira**

**bruninkmkt@hotmail.com*
UNESP – Univ. Estadual Paulista, 737 Hygino Muzzi Filho Avenue, 17525-900 Marília (Brazil)

*** etannuri@gmail.com*
UNESP – Univ. Estadual Paulista, 737 Hygino Muzzi Filho Avenue, 17525-900 Marília (Brazil)

Introduction

Studies on Author Co-citation Analysis (ACA) aim to identify influential authors and show their interrelations from citations (White; Griffith, 1981; White; McCain, 1998). ACA analyzes the intellectual and social structure of an area, scientific field or set of researchers.

When comparative studies are intended, given the specificities of each area, the importance of normalized indicators, which standardize the units of measure and reveal aspects not explained in absolutes, are emphasized.

According to the studies of Luukkonen et al. (1993), absolute and normalized measures carry different types of information: the first shows the central "actors" of the networks, while the latter shows the intensity of relations and reveal aspects that are not identifiable in the absolute frequencies. Among relative indices, Pearson's Correlation Coefficiengt Salton's Cosine, and Jaccard Index are cited (Leydersdoff; Vaughan, 2006).

Pearson's r was the standard measure before the studies of Ahlgren, Jarneving & Rousseau (2003), who criticized its use, showing that it does not satisfy as similarity and proximity measures.

This research aims to deepen the study on normalized indicators of ACA. Specifically, it presents and analyzes normalized indicators such as Ss and JI, and compares the similarities between them via identification of normalized relations applied to Information Science.

Salton's Cosine (Ss) and Jaccard Index (JI) are stressed. These two normalized indices are calculated from the co-occurrence matrix of absolute data, according to Luukkonen et al. (1993).

In the studies by Hamers et. al. (1989), co-occurrences represent co-citations, Ss is then expressed (Equation 1):

$$Ss_{(a,b)} = \frac{\text{Cocit}_{(a,b)}}{(\text{cit}_{(a)} \,.\, \text{cit}_{(b)})\ ^{1/2}}$$

Where:
$coc_{(a,b)}$= total of co-occurences of authors a and b
$cit_{(a)}$ = total of citations received by author a
$cit_{(b)=}$ total of citations received by author b

Luukkonen et al (1993), express JI by (Equation 2):

$$\mathrm{IJ} = \frac{\mathrm{Cocit}\ (a, b)}{\mathrm{Cit}\ (a) + \mathrm{Cit}\ (b) - \mathrm{Cit}(\ a \cap b)}$$

Both Ss as IJ vary between zero and one: the closer to one, more similar are the two authors (with theoretical-methodological proximity, similarity, complementarity, overlap, or opposed ideas or even co-authorship); the closer to zero, the farther is the association between the two authors.

Methodological procedures
Firstly, a theoretical study on ACA and its indicators initiated the research. Data was extracted from 110 articles published in the 2007-2011 period, from ENANCIBs[1] proceedings, in Brazil. We identified 1242 cited researchers, 2003 references, composing a target group of 20 researchers cited at least 12 times.

A 20x20 square matrix was built, from the most cited authors, with absolute co-citation frequency. Ss and JI was applied. We used Microsoft Excel macros built in "Visual Basic for Applications" (VBA).

We comparatively analyzed the results of the two normalized matrices using Ss and JI, evidencing the proximities, similarities and differences between the present values and the intensities of connections in the networks.

Presentation and analysis of data
The two normalized matrices using Ss (Equation 1) and JI (Equation 2) are presented in Tables 1 and 2.

[1] ENANCIB- National Meeting on Research in Information Science in Brazil.

Table 1. Ss Normalized Matrix

	BUFREM	LETA	M-CHAPULA	MEADOWS	MENEGHINI	MUELLER	MUGNAINI	NORONHA	OLIVEIRA	PACKER	PINHEIRO	POBLACIÓN	PRICE	ROUSSEAU	SANTOS	SPINAK	STUMPF	TARGINO	VANZ	VELHO
BUFREM	1,00	0,00	0,25	0,08	0,06	0,28	0,19	0,25	0,31	0,07	0,19	0,19	0,11	0,13	0,42	0,19	0,12	0,13	0,17	0,06
LETA		1,00	0,00	0,21	0,40	0,18	0,08	0,00	0,00	0,26	0,12	0,00	0,35	0,00	0,00	0,08	0,15	0,00	0,15	0,16
M-CHAPULA			1,00	0,15	0,15	0,13	0,23	0,08	0,15	0,08	0,17	0,00	0,07	0,16	0,30	0,40	0,14	0,08	0,07	0,15
MEADOWS				1,00	0,35	0,54	0,15	0,15	0,25	0,27	0,41	0,21	0,39	0,16	0,05	0,10	0,32	0,31	0,19	0,30
MENEGHINI					1,00	0,31	0,31	0,23	0,08	0,84	0,12	0,16	0,20	0,16	0,00	0,08	0,29	0,16	0,07	0,31
MUELLER						1,00	0,35	0,26	0,31	0,29	0,43	0,27	0,35	0,09	0,25	0,00	0,29	0,32	0,08	0,31
MUGNAINI							1,00	0,15	0,31	0,33	0,12	0,24	0,07	0,08	0,44	0,08	0,21	0,32	0,14	0,08
NORONHA								1,00	0,23	0,25	0,12	0,48	0,07	0,00	0,30	0,08	0,29	0,32	0,21	0,15
OLIVEIRA									1,00	0,08	0,52	0,56	0,13	0,00	0,30	0,08	0,07	0,24	0,21	0,15
PACKER										1,00	0,06	0,17	0,07	0,09	0,00	0,00	0,31	0,17	0,08	0,25
PINHEIRO											1,00	0,30	0,30	0,00	0,17	0,06	0,05	0,24	0,16	0,06
POBLACIÓN												1,00	0,07	0,00	0,23	0,00	0,22	0,25	0,22	0,24
PRICE													1,00	0,14	0,06	0,21	0,19	0,07	0,19	0,13
ROUSSEAU														1,00	0,08	0,50	0,15	0,08	0,22	0,16
SANTOS															1,00	0,23	0,21	0,23	0,14	0,15
SPINAK																1,00	0,00	0,08	0,15	0,16
STUMPF																	1,00	0,37	0,47	0,29
TARGINO																		1,00	0,15	0,16
VANZ																			1,00	0,07
VELHO																				1

Table 2. JI Normalized Matrix

	BUFREM	LETA	M-CHAPULA	MEADOWS	MENEGHINI	MUELLER	MUGNAINI	NORONHA	OLIVEIRA	PACKER	PINHEIRO	POBLACIÓN	PRICE	ROUSSEAU	SANTOS	SPINAK	STUMPF	TARGINO	VANZ	VELHO
BUFREM	1,00	0,00	0,14	0,04	0,03	0,15	0,10	0,14	0,18	0,03	0,10	0,10	0,06	0,07	0,26	0,10	0,06	0,07	0,09	0,03
LETA		1,00	0,00	0,10	0,25	0,08	0,04	0,00	0,00	0,15	0,06	0,00	0,21	0,00	0,00	0,04	0,08	0,00	0,08	0,09
M-CHAPULA			1,00	0,07	0,08	0,06	0,13	0,04	0,08	0,04	0,09	0,00	0,03	0,09	0,17	0,25	0,08	0,04	0,04	0,08
MEADOWS				1,00	0,19	0,37	0,07	0,07	0,13	0,14	0,26	0,10	0,23	0,08	0,02	0,05	0,18	0,16	0,10	0,16
MENEGHINI					1,00	0,15	0,18	0,13	0,04	0,71	0,06	0,09	0,11	0,09	0,00	0,04	0,17	0,09	0,04	0,18
MUELLER						1,00	0,18	0,13	0,15	0,13	0,26	0,13	0,19	0,04	0,13	0,00	0,15	0,16	0,04	0,15
MUGNAINI							1,00	0,08	0,18	0,20	0,06	0,14	0,03	0,04	0,29	0,04	0,12	0,19	0,08	0,04
NORONHA								1,00	0,13	0,14	0,06	0,32	0,03	0,00	0,17	0,04	0,17	0,19	0,12	0,08
OLIVEIRA									1,00	0,04	0,33	0,39	0,07	0,00	0,17	0,04	0,04	0,14	0,12	0,08
PACKER										1,00	0,03	0,10	0,04	0,05	0,00	0,00	0,18	0,10	0,04	0,14
PINHEIRO											1,00	0,17	0,18	0,00	0,09	0,03	0,03	0,13	0,09	0,03
POBLACIÓN												1,00	0,04	0,00	0,13	0,00	0,13	0,14	0,13	0,14
PRICE													1,00	0,07	0,03	0,12	0,10	0,04	0,10	0,07
ROUSSEAU														1,00	0,04	0,33	0,08	0,04	0,13	0,09
SANTOS															1,00	0,13	0,12	0,13	0,07	0,08
SPINAK																1,00	0,00	0,04	0,08	0,09
STUMPF																	1,00	0,23	0,30	0,17
TARGINO																		1,00	0,08	0,09
VANZ																			1,00	0,04
VELHO																				1

In the analysis of Tables 1 and 2, we initially highlighted Meneghini and Packer with the highest value for Ss equal to 0.84 and JI equal to 0.71, observing that in the absolute matrix (not presented here) the co-citation between these two authors is 10, with 13 citations made to Meneghini and 11 to Packer. The number of co-citations is relativized by citations made to the two authors. In Figures 1 and 2, the links between these authors are strongly highlighted.

Meadows and Mueller present Ss equal to 0.54, JI equal to 0.37 and the absolute number of co-citations equal to 19 with 31 citations made to Meadows and 40 citations to Mueller, which justifies the relativized median value for Ss, and lower to JI. In Figure 1, the link between these two researchers is much more highlighted than in Figure 2.

Researchers Leta and Spinak present 0.08 for Ss and 0.04 for JI with absolute co-citation value equal to 1, with 12 citations to Leta and 12 citations to Spinak, which explains the low relativized values. In Figures 1 and 2 the connections for both Ss and JI present their links slightly differentiated.

Figure 1: Ss Normalized Network

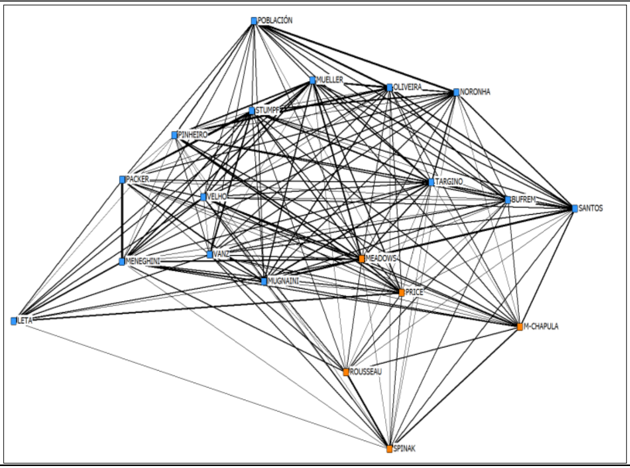
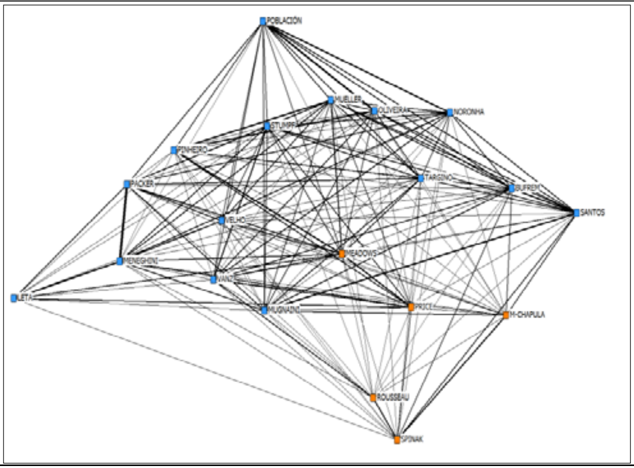


Figure 2: JI Normalized Network



The highlights ratify Hamers et.al. (1989), when they claim that Ss formula often produces a relative similarity measure which is twice the number obtained by JI. Extending this analysis to other values, it is observed that the higher the Ss, the closer JI will be to it, and above half of Ss (the example of Meneghini and Packer; Meadows and Mueller), and the lower the Ss, the JI will be closer or will be the very half (the example of Leta and Spinak).

Final considerations

This study has validated the analyzes already made by other scholars and advanced on existing analyzes between Ss and JI, showing when there is a tendency of proximity. They exhibit similar behavior and the choice of using either index does not present a conclusive position on the pointed question, and consequently, the appropriate methodology to establish ACA is not fully consolidated.

References

Algren, P., Jarneving, B. & Rousseau, R. (2003). Requirements for a Cocitation Similarity Measure, with Special Reference to Pearson’s Correlation Coefficient. *Journal of the American Society for Information Science and Technology,* 54, 6, 550-560.

Hamers, L. et al. (1989). Similarity measures in scientometric research: the Jaccard index versus Salton’s cosine formula. *Information Processing & Management*, 25, 3, 315-318.

Leydesdorff, L. & Vaughan, L. (2006). Co-occurrence Matrices and their applications in Information Science: Extending ACA to the Web environment. *Journal of the American Society for Information Science and Technology*, 57, 12, 1616-1628.

Luukkonen, T. et al. (1993). The measurement of international scientific collaboration. *Scientometrics*, Amsterdam, 28, 1, 15-36.

White, H.D., & Griffith, B. (1981). Author cocitation: A literature measure of intellectual structure. *Journal of the American Society for Information Science*, 32, 163–171.

White, H.D. & Mccain, K.W. (1998). Visualizing a discipline: an author co-citation analysis of Information Science, 1972-1995. *Journal of the American Society for Information Science,* 49, 4, 327-355.

Differences between examiner and applicant citations in the European Patent Office: a first approach

Joaquín M. Azagra-Caro and Elena M. Tur*

* *jazagra@ingenio.upv.es, elmatu@ingenio.upv.es*
INGENIO (CSIC-UPV), Ciudad Politécnica de la Innovación Edif. 8E 4º, Camino de Vera s/n, 46022 Valencia

Abstract
In the US Patent Office, examiners add extra shares of citations to foreign applicants. We explore a similar country club effect in the European Patent Office (EPO). Using EPO data of over 3,500,000 citations in years 1997-2007, we find national variation in the probability of an applicant originating a citation rather than the examiner. Symmetrically to the US case, EPO examiners add extra citations to non-signatory member states. Moreover, if examiners are likely to come from the same country of the applicants, applicant-citation shares increase, pointing to the existence of national bias in EPO patent examiners. These results hold after controlling for sub-national characteristics of the patenting process.

Keywords
Citations, knowledge flows, national biases

Introduction
The geography of innovation makes extensive use of backward citations in patents to measure knowledge flows (Jaffe et al., 1993). Several works emphasize the importance of distinguishing the origins of citations because, in theory, citations inserted by patent examiners are likely to be less localized than applicant citations. United States Patent and Trademark Office (USPTO) data mostly tend to confirm this for the US case (Thompson, 2006) although there are some differences for some specific measures of distance (Alcácer & Gittelman, 2006). European data confirm it for some European countries (Criscuolo & Verspagen, 2008), but not some regions with low absorptive capacity (Azagra-Caro et al., 2009). These studies focus on the match or distance between citing and cited country. However there is another geographic concern that has been largely unexplored, i.e. what are the characteristics of the citing country? Do patent examiners add more citations to patent applications from specific countries?
This is a relevant question because the answer might reveal underlying economic forces that are subject to policy influence, or uncover individual questionable examiner practices. There is some evidence suggesting that, for the USPTO, geographic origin of the applicant matters, e.g. US examiners add more citations to foreign applications (Alcácer et al., 2009). However, there is a lack of research on a similar 'club effect' in the case of the European Patent Office (EPO). This is unfortunate because the EPO is frequently used as a benchmark against the USPTO, and is considered one of the highest quality patent systems due to its rigorous granting process and flexibility applied to later stages in a patent's life (Saint-George & van Pottelsberghe, 2013). This paper focuses on the EPO. By comparing with the USPTO, we should be able to identify whether there is a symmetrical geographical effect, namely whether EPO patent examiners are more likely to add citations to foreign applications: Do EPO examiners add extra citations to applications from countries outside the European Patent

Organization (EPOrg)? And do EPO examiners add extra citations to applications from countries other than their own?

Model, data and variables

We estimate the following model:

$$\Pr(appcit_{ijklt}) = f(\alpha X_{it}, \beta X_{jt}, \gamma X_{kt}, \delta X_{lt}) + \varepsilon_{ijklt} \quad (2)$$

where *appcit* is equal to 1 if the citation is inserted by the applicant and 0 if inserted by the examiner. The probability varies according to the characteristics of the citation *i*, the patent *j*, the applicant *k* and the applicant country *l*. The year of the patent application *t*, is lagged two periods for national economic and research and development (R&D) characteristics to prevent endogeneity.

Data on patents and citations come from Patstat (October 2012 edition). We selected patents where the publication authority was the EPO –almost 2.5 million. After removing those with missing or unreliable information for application year and technology class (represented by the International Patent Classification IPC), and those without citations, we were left with 2 million patents.

Those patents contained over 12 million citations. Patstat classifies them into origin types, i.e. the moment in the examination process when the citation was inserted. There are ten types of origins (coded 0-9), but only some are relevant for this study, i.e. those indicating that either patent applicant or examiner could have inserted the citation (see section **Error! Reference source not found.** for further details): origins coded 0 (citations introduced during search), 2 (citations introduced during examination) and 5 (citations from the International Search Report). They represent most (82%) of the citations.

Patstat differentiates who inserted the citation by classifying citations with origins 0, 2 and 5 into several categories. Categories (coded with single letters, A, X, Y, etc.), refer to the relevance of prior art to invalidate claims of novelty. Criscuolo & Verspagen (2008) call category D ‘applicant citations’ and sum the other categories as ‘examiner citations’. We follow this method.

In the estimations, the number of observations is not the number of citations for two reasons. First, duplicates are created if the patent has more than one applicant. We deal with this econometrically by weighting the observations by the inverse number of applicants. Second, we match Patstat to other databases on national characteristics that do not have full information for all countries and years. The sample includes over 3.6 million observations. The proportion of D-citations in the total is our dependent variable, computable for over 7 million citations.

Table 1 provides information on the econometric model variables.

The dependent variable is a dummy that takes the value 1 if the citation comes from the examiner. A logit model is appropriate for this kind of data.

Table 1. Variable definitions and descriptive statistics (n=3,663,276)

Vector	Name	Source	Variables	Description	Mean	Std. Dev.	Min	Max
$appcit_{ijklt}$	Applicant citation	Patstat	Citation category D	1 if citation category is D, 0 if other category	0.07	0.26	0.00	1.00
X_{it}	Citation characteristics	Patstat	Non-patent literature	1 if non-patent literature, 0 if patent literature	0.36	0.48	0.00	1.00
			European search report	1 if origin in search report	0.85	0.36	0.00	1.00
			Examination report	1 if origin in examination	0.00	0.06	0.00	1.00
X_{jt}	Patent characteristics	Patstat	Euro-PCT	1 if EPO-PCT, 0 if direct EPO	0.46	0.50	0.00	1.00
			Grant	1 if granted, 0 otherwise	0.18	0.39	0.00	1.00
			Filing year	Application year	2001.94	3.03	1997.00	2007.00
			A Human Necessities	1 if IPC code is A Human Necessities	0.21	0.41	0.00	1.00
			B Performing Operations; Transporting	B Performing Operations; Transporting	0.26	0.44	0.00	1.00
			C Chemistry; Metallurgy	C Chemistry; Metallurgy	0.22	0.41	0.00	1.00
			D Textiles; Paper	D Textiles; Paper	0.02	0.14	0.00	1.00
			E Fixed Constructions	E Fixed Constructions	0.04	0.19	0.00	1.00
			F Mechanical Engineering; Lighting; Heating; Weapons; Blasting	F Mechanical Engineering; Lighting; Heating; Weapons; Blasting	0.14	0.34	0.00	1.00
			G Physics	G Physics	0.26	0.44	0.00	1.00
			H Electricity	H Electricity	0.25	0.43	0.00	1.00
X_{kt}	Applicant characteristics	ECOOM*	Individual	1 if institutional sector is… Individual only	0.08	0.26	0.00	1.00
			Government	Government only	0.03	0.16	0.00	1.00
			University	University only	0.02	0.15	0.00	1.00
			Hospital	Hospital only	0.00	0.04	0.00	1.00
			Company-government	Company and government	0.00	0.04	0.00	1.00
			Company-university	Company and university	0.00	0.00	0.00	1.00
			Company-hospital	Company and hospital	0.00	0.01	0.00	1.00
			Government-university	Government and university	0.00	0.00	0.00	1.00
			# applications	Number of applications (millions)	0.00	0.00	0.00	0.15
X_{lt}	Country of applicant characteristics – economic and R&D	OECD R&D Statistics	GDP	Real Gross Domestic Product (GDP): billion Euro	0.04	0.04	0.00	0.13
			GDP per capita	GDP: Euro per inhabitant (millions)	0.03	0.01	0.00	0.07
			GERD intensity	Total intramural Gross R&D expenditure (GERD): Millions of Purchasing Power Standards (PPS) at 2000 prices	2.51	0.47	0.28	4.58
			% business funding of R&D	Business R&D funding: Share of GERD	0.64	0.09	0.17	0.91
	Country of applicant characteristics – related to EPO	EPO Annual Reports	Prob EPO exam same country	Probability of examiner from same nationality	0.10	0.10	0.00	0.26
			EPOrg member	EPO member (yes/no)	0.44	0.50	0.00	1.00

* Methodology for construction of ECOOM data explained in DuPlessis et al. (2009), Magerman et al. (2009) and Peeters et al. (2009).

Results

Table 2 presents the estimations. Column 1 includes the specification of Equation 1 with citation and patent characteristics only; the remaining columns include the variables progressively.

Citation and patent characteristics

The results for the sub-national variables are consistent across estimations. Citations are coded to indicate whether the origin is a Euro-PCT (not a direct EPO) application, and whether it is the European search report or the examiner report (rather than the international search report). The coefficient of "Euro-PCT" is negative and significant, indicating that this longer procedure leads to higher numbers of examiner citations. The coefficient of "European search report" is negative and significant, implying that citations in this second phase are more likely to be associated with examiners than if there was an international search report in the first phase. The coefficient of "Examiner report" is also negative and significant and higher than the coefficient of "European search report", meaning that citations in this third phase are most likely to come from examiners.

The sample includes applications and grants. This is controlled for in the models by the dummy variable "Grant". The estimated coefficient is positive and significant. Hence, we can confirm a link between receiving relatively fewer examiner citations and having the patent granted. In part, this is intuitive. It becomes more interesting if we consider that, in the USPTO, this does not necessarily apply. In the USPTO, more experienced examiners, and examiners that systematically cite less prior art, are more likely to award patent grants (Lemley & Sampat, 2012). Moreover, USPTO examiners rarely use applicant citations to reject a grant (Cotropia et al., 2013). Hence, examiner citation shares are not associated with denial of a grant in the USPTO but they are in the EPO. This and other signs may indicate the superiority of the EPO patent system (Saint-George & van Pottelsberghe, 2013).

We test whether applicants are more likely than examiners to cite non-patent literature, extrapolating from US evidence that examiners rarely cite non-patent literature (Sampat, 2004). The positive and significant sign of "Non-patent literature" shows that this is the case. Applicants are probably more familiar with the fundamental knowledge base underpinning their inventions, while examiners are often engineers whose expertise is related more to parcels of applied knowledge.

Applicant characteristics

Dummies for organizational type of the applicant (models 2-3) can be used to validate empirically which one matters more. "Company only" is the benchmark. The positive, significant coefficients of "Government only" and "University only" indicate that these institutions generate more reliability than corporate patents. The coefficients of "Individuals only" and "Hospital only" are negative and significant, which means that citations are less likely to originate in applicants than in the case of firms. Individuals may show lower citation shares because institutions facilitate settings where citing is more common practice, i.e. through sharing of references and codified knowledge. Examiner citation shares may be larger for hospitals because they do not have a tradition of patenting, and on patents related to clinical practice which are less related to science.

Table 2. Logistic regression of the probability of an applicant originating a citation rather than the examiner

	1 Citation and patent characteristics	2 + Applicant characteristics	3 + Country characteristics
Euro-PCT	-0.68*** (0.01)	-0.68*** (0.01)	-0.48*** (0.01)
European search report	-0.93*** (0.01)	-0.94*** (0.01)	-0.57*** (0.01)
Examination report	-2.73*** (0.09)	-2.74*** (0.09)	-2.43*** (0.09)
Grant	0.30*** (0.00)	0.29*** (0.00)	0.29*** (0.00)
Non-patent literature	0.06*** (0.01)	0.05*** (0.01)	0.10*** (0.01)
Individual		-0.15*** (0.01)	-0.21*** (0.01)
University		0.04*** (0.01)	0.08*** (0.01)
Government		0.13*** (0.01)	0.05*** (0.01)
Hospital		-0.39*** (0.07)	-0.31*** (0.07)
Company-government		-0.09* (0.05)	-0.10* (0.05)
Company-university		1.16*** (0.29)	1.14*** (0.30)
Company-hospital		0.48* (0.27)	0.31 (0.27)
Government-university		-0.17 (0.52)	-0.40 (0.53)
# applications		-0.91 (0.56)	-10.64*** (0.57)
GDP			0.84*** (0.14)
Per capita GDP			18.77*** (0.86)
GERD intensity			0.27*** (0.01)
% business funding of R&D			-0.88*** (0.04)
Prob EPO exam same country			0.61*** (0.04)
EPOrg member			0.64*** (0.01)
Constant	37.53*** (1.44)	38.74*** (1.45)	67.19*** (1.64)
Observations	3,663,276	3,663,276	3,663,276
Log likelihood	-848,023	-847,774	-838,745
χ^2	54,181	54,658	75,414
Prob>χ^2	0.000	0.000	0.000

* p<0.1; ** p<0.05; *** p<0.01. Robust standard errors in parenthesis. No collinearity according to Variance Inflation Factors. All models include a trend and eight IPC section dummies. Weight: share of number of applicant countries.

Models 2-3 include dummies for types of organizational interactions (taking “Company only” as benchmark). University-company co-applications for patents are strongly associated with a higher probability of an applicant rather than the examiner including a citation. Somewhat surprisingly, government-company co-application for patents is negatively related to that probability. A possible reason might be that organizations in the category government have heterogeneous missions. Government labs with an industry orientation are more likely to engage in partnerships with firms that lead to patents, than labs with an academic orientation, and the government-company dummy captures this type of partnership. This double industry orientation receives a higher share of examiner citations. For other interactions (“Company-hospital”, “University-government”) the dependent variable does not change significantly.
The number of applicant citations decreases with the increase in the number of applications. Alcácer et al. (2009) found the same in the USPTO case. Their explanation is that large applicants prefer “broad patent portfolios, with relatively low value placed on any single invention” (p. 426). Alternatively, it might be that applicants include unrelated cites after the invention or omit relevant cites for strategic reasons (Breschi & Lissoni, 2005). Perhaps experienced applicants learn how to “cheat”, and hide a higher number of relevant references.

National characteristics
The variables GDP, per capita GDP and GERD intensity test the assumption that larger, wealthier and scientifically stronger countries are more likely to create conditions favorable to the appearance of novelty. Their positive, significant coefficients provide evidence to support it. Hence, we observe that countries with these favorable endowments benefit from lower examiner citation shares.
The coefficient of the share of business funding variable in model 3 is negative and significant, supporting this expectation. Examiner citation shares are higher in patents from national contexts where the research orientation is towards more applied research.
Country block effects may also play a role in the model. Specifically, we are interested in whether there is a club effect similar to the one shown by Alcácer et al. (2009) in the USPTO case: US applicants receive fewer examiner citation shares than non-US ones. In our EPO sample, this club effect would not be strictly national since the EPO is international. Instead, we propose that such an effect might be visible for countries belonging to the EPOrg. In the model, the dummy is equal to 1 if the applicant country belongs to EPOrg, to capture this phenomenon. The estimation (positive and significant) verifies that there is a lower propensity for EPOrg member states to receive cites from the examiner. Hence, the EPO is similar to the USPTO: outsiders are less warmly received.
Having isolated a club effect, the nationality of examiners might be influential. Collins & Wyatt (1988) detected national chauvinism in citations to non-patent literature in US genetics patents: “it appears that every country is its own best citer” (p.73). However, Meyer (2000) finds no signs of national chauvinism in nanotechnology patent applications to the USPTO from Swedish applicants. In our estimation, the positive, significant coefficient of the probability of an application being examined by an examiner from the same country as the patent applicant provides support for the national bias assumption.

Conclusions
The literature on the geography of knowledge flows has shown that the probability of an applicant rather than the examiner originating a citation depends on differences between citing and cited countries. Our contribution to this stream of literature is that the conditions of the citing country also matter to predict that probability. Our findings show that better national economic and scientific endowments increase applicant citation shares, whereas higher proportions of business funding of R&D foster examiner citation shares. Future

research could test which group of determinants (citing country characteristics or citing-cited country differences) matter more.

Previous analyses of the characteristics of applicant versus examiner citation shares found differences across patent and applicant. We show the presence of additional disparities across citation characteristics, namely procedural aspects of the patenting process and knowledge base of the patent. Our results for procedural aspects increase our understanding of the generation of citations in the various phases of the life of an EPO application. Our results for knowledge base suggest the importance of science to provide credibility to applications.

The use of a sample based on EPO applications allowed comparison with earlier works exploiting USPTO evidence. It suggests that large applicant citation shares are more clearly associated with being awarded a patent by the EPO than the USPTO. It also signals that there are similar club effects, which favor EPOrg members at the EPO and US residents at the USPTO. Since the methods used by Alcácer et al. (2009) and those applied in this study differ, interpretation of this comparison should be cautious. A possible avenue of further inquiry could be designing an experiment to enable direct comparison between both data sources.

References

Alcácer, J., and M. Gittelman, "Patent citations as a measurement of knowledge flows: The influence of examiner citations," Review of Economics and Statistics 88:4 (2006), 774-779.

Alcácer, J., M. Gittelman, and B.N. Sampat, "Applicant and examiner citations in U.S. patents: An overview and analysis," Research Policy 38 (2009), 415-427.

Azagra-Caro, J.M., I. Fernández de Lucio, F. Perruchas, and P. Mattsson, "What do patent examiner inserted citations indicate for a region with low absorptive capacity?" Scientometrics 80:2 (2009), 443–457.

Azagra-Caro, J.M., P. Mattsson, and F. Perruchas, "Smoothing the lies: The distinctive effects of patent characteristics on examiner and applicant citations," Journal of the American Society for Information Science and Technology 62:9 (2011), 1727-1740.

Breschi, S., and F. Lissoni, "Knowledge networks from patent data," In H. F. Moed, W. Glänzel and U. Schmoch (Eds.), Handbook of quantitative science and technology research (Netherlands: Kluwer Acad. Publ.. 2005).

Collins, P., and S. Wyatt, "Citations in patents to the basic research literature," Research Policy 17 (1988), 65-74.

Cotropia, C.A., M.A. Lemley, and B.N. Sampat, "Do applicant patent citations matter?" Research Policy 42:4 (2013), 844-854.

Criscuolo, P., and B. Verspagen, "Does it matter where patent citations come from? Inventor vs. examiner citations in European patents," Research Policy 37 (2008), 1892-1908.

Du Plessis, M., B. van Looy, X. Song, and T. Magerman, "Data Production Methods for Harmonized Patent Indicators: Assignee sector allocation," EUROSTAT Working Paper and Studies (2009), Luxembourg.

Jaffe, A. B., M. Trajtenberg, and R. Henderson, "Geographic localization of knowledge spillovers as evidenced by patent citations," Quarterly Journal of Economics 108:3 (1993), 577-598.

Lemley, M., and B.N. Sampat, "Examiner Characteristics and Patent Office Outcomes," Review of Economics and Statistics 94:3 (2012), 817-827.

Magerman T, J. Grouwels, X. Song, and B. van Looy, “Data Production Methods for Harmonized Patent Indicators: Patentee Name Harmonization,” EUROSTAT Working Paper and Studies (2009), Luxembourg.

Meyer, M., “Does science push technology? Patents citing scientific literature,” Research Policy 29 (2000), 409-434.

Peeters B., X. Song, J. Callaert, J. Grouwels, and B. van Looy, “Harmonizing harmonized patentee names: an exploratory assessment of top patentees,” EUROSTAT working paper and Studies (2009), Luxembourg.

Saint-George, M., and B. van Pottelsberghe de la Potterie, “A quality index for patent systems,” Research Policy 42 (2013), 704– 719.

Sampat, B.. “Examining Patent Examination: An Analysis of Examiner and Applicant Generated Prior Art, Working Paper (2004), School of Public Policy, Georgia Institute of Technology.

Thompson, P. “Patent citations and the geography of knowledge spillovers: Evidence from inventor- and examiner-added citations,” Review of Economics and Statistics 88:2 (2006), 383-388.

Altmetric gender bias? – Preliminary results[1]

Judit Bar-Ilan* and Inge van der Weijden**

*Judit.Bar-Ilan@biu.ac.il
Department of Information Science, Bar-Ilan University, Ramat Gan, 5290002, Israel

**i.c.m.van.der.weijden@cwts.leidenuniv.nl
CWTS, Leiden University, P.O. Box 905, Leiden, 2300AX, The Netherlands

Introduction

Gender bias in science has been studied extensively. Several studies pointed to a gender gap between men and women in terms of number of publications and citations (e.g., Aksnes, Rorstad, Piro & Sivertsen, 2011; West, Jacquet, King, Corell & Bergstrom, 2013; Larivière, Ni, Gingras, Cronin & Sugimoto, 2013; Van der Weijden & Calero Medina, 2014) regardless their academic position and research field. Interestingly, literature showed no gender differences regarding research impact (van der Weijden & Calero Medina, 2014). In this study we set out to examine whether there are gender specific differences, when instead of citations we consider an altmetric, more specifically Mendeley readership counts.

Online dissemination of knowledge

The web has provided new opportunities for academics to disseminate their research results. Online CV's, homepages or publication lists for the scholarly related activities of academics are examples. They can include wider publication types (e.g. journal of conference papers, books, and reports) and pre-prints, which would not be indexed by major scientific databases. In this way, Academic Web CVs or online lists of publications (institutional or personal) can be a significant method to facilitate knowledge transfer (Kousha & Thelwall, 2013). Furthermore, online CVs or resumes can be updated frequently and share bibliographic information, abstract or even the full-text of published or in press research through personal or institutional self-archiving practices. In an earlier study (Van der Weijden & Calero Medina, 2014) it was shown that gender has impact on the Web presence (having an online CV or an individual webpage for publication lists) of academics across fields: males are more active compared to females.

Reference managers such as Mendeley are online tools that can help researchers to disseminate and organize their research. Mendeley (mendeley.com) is a free and widely used online reference manager that provides aggregated counts of the number of users who bookmarked an item. Mendeley calls this the number of "readers", although we cannot be certain that users who save items to their libraries actually read them. Zahedi, Costas and Wouters (2013) provided a characterization of Mendeley users. Previous studies showed that Mendeley readership counts is one of the most promising altmetric (Zahedi, Costas & Wouters, in press), both because of Mendeley's large coverage and because several studies showed significant, medium strength correlations between readership and citation counts (e.g. Li, Thelwall & Giustini, 2012; Haustein et al., in press). In this study we investigate gender-

[1] This work was partially supported by the EU FP7 ACUMEN project (Grant agreement: 266632). We would like to thank Mike Thelwall for developing and updating Webometric Analyst 2.0. We are greatly indebted to Clara Calero Medina and Rodrigo Costas Comesana for collecting parts of the data.

specific differences in Mendeley readership counts. To the best of our knowledge this is the first study to investigate gender in altmetrics.

Research setup

Our dataset is based on the common dataset of the EU funded ACUMEN project. The data are comprised of a set of 494 astronomers and astrophysicists from 14 EU countries and Israel. The gender of all researchers was verified. Publications of these researchers were retrieved from Thomson Reuters' Web of Science (WOS) using the "Large scale author name disambiguation using rule-based scoring and clustering" algorithm developed at CWTS to detect publications per researcher. This step resulted in a list of 27,645 publications. Some of the publications are repeated in the set, because it lists for every researcher his full list of publications, and a number of publications were co-authored by several authors in the dataset.

For 60% of the publications in the list WOS provided DOIs. When submitting the DOI of a publication, the Mendeley API retrieves the number of readers of this publication. Thus this method covers 60% (16,791) publications. To cover the publications for which WOS did not provide a DOI, and also to retrieve readership counts of items covered by Mendeley, but for which no DOI is provided in Mendeley, we conducted title searches using Webometric Analyst 2.0 developed by Mike Thelwall (http://lexiurl.wlv.ac.uk/) on a subset of 12,000 publications. It should be noted that title searches are not straightforward, because special characters are not always recognized by Mendeley, and titles are not always written identically on WOS and Mendeley. Mendeley is built by its users, and users do not always provide accurately the metadata describing the item. Thus partial titles were searched and these were matched with the original list of titles, checking source, year, author and DOI when available.

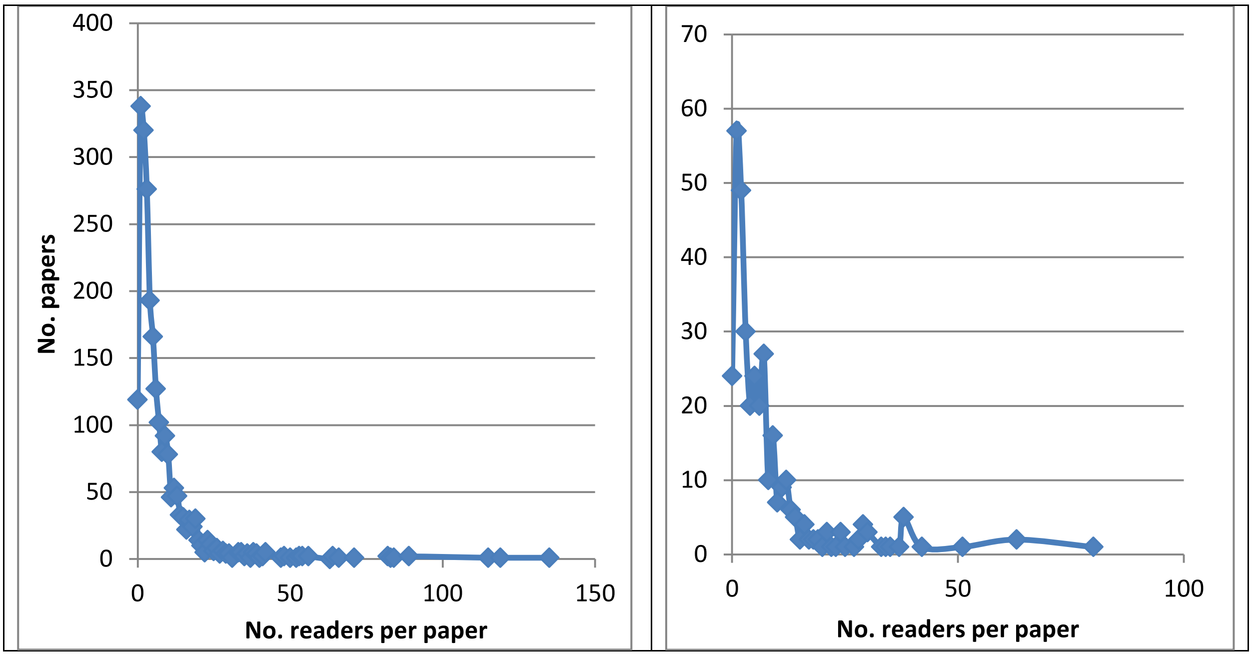
Preliminary results

We present here results based on 12,000 WOS records out of the 27,645 records for which title searches were conducted. Out of the 293 researchers in this subset, 60 were women (20%) and 233 men. Women authored 1778 publications (15%). The percentage of women and their publication share in the subset are almost identical to the respective percentages in the whole set.

We located 2,711 publications in Mendeley (23%). Out of the female authored publications, 360 were found in Mendeley (20%), compared with 2351 male authored publications (out of 10,222, 23%). Thus there seems to be a slight "advantage" of male authored publications to be found on Mendeley, at least in this subset. Male author's publications indexed by Mendeley, were "read" by 7.1 readers on average, while for female authored papers indexed by Mendeley the average number of readers was 7.7.

Thus the preliminary findings show that in terms of the percentage of publications found on Mendeley men have a slight advantage, but in terms of the average number of readers, the women are doing better, the median number of readers is 4 for men and 4.5 for women. On the other hand the most-highly read publications in the dataset were authored by men, as can be seen from Figure 1. The distributions are similar in shape, but the most-read item by a male author was read 135, while the most "read" item by a female author was only "read" 80 times.

Figure 1: Reader distribution by gender of the author: Men on the left, women on the right



Limitations

In this study we searched for publications of specific authors, thus the dataset is not necessarily representative. In addition most papers in astrophysics are multi-authored, thus a paper assigned to a male author might have had a female co-author and vice versa.

Next research steps

In this research-in-progress paper we only considered publications by astrophysicists. We have similar sized datasets, in terms of the number of researchers, in three additional fields: philosophy, public health and environmental engineering. We plan to conduct similar studies in these three fields.

References

Aksnes, D. W., Rorstad, K., Piro, F., & Sivertsen, G. (2011). Are female researchers less cited? A large-scale study of Norwegian scientists. *Journal of the American Society for Information Science and Technology*, 62(4): 628-636.

Haustein, S., Peters, I., Bar-Ilan, J., Priem, J., Shema, H., & Terliesner, J. (in press). Coverage and adoption of altmetric sources in the bibliometric community. *Scientometrics*, doi:10.1007/s11192-013-1221-3.

Kousha, K. & Thelwall, M. (2013). Evaluating the web research dissemination of EU academics: A multi-discipline outlink analysis of online CVs. In *Proceedings of the 14th International Conference Society of Scientometrics and Informetrics Conference* 2013. Vienna, Austria, pp. 705-719.

Larivière, V., Ni, C., Gingras, Y., Cronin, B., & Sugimoto, C. (2013). Bibliometrics: Global gender disparities in science. Nature, 204, 211-213.

Li, X., Thelwall, M., & Giustini, D. (2012). Validating online reference managers for scholarly impact measurement. *Scientometrics*, *91*(2), 461-471.

Van der Weijden, I. & Calero Medina, C. (2014). Gender effects on evaluation indicators. *ACUMEN Deliverable*. February 2014.

West, J. D., Jacquet, J., King, N. M., Corell, S. J., & Bergstrom, C. T. (2013). The role of gender in scholarly authorship. PLOS *ONE*, 8(7), e66212.

Zahedi, Z., Costas, R., & Wouters, P. (2013). What is the impact of the publications read by the different Mendeley users? Could they help to identify alternative types of impact? In *PLoS ALM Workshop*, 7-9 October, 2013, San Francisco. Slides retrieved from: http://lanyrd.com/profile/zohrehzahedi/slides/

Zahedi, Z., Costas, R., & Wouters, P. (in press). How well developed are altmetrics? A cross disciplinary analysis of the presence of 'alternative metrics' in scientific publications. *Scientometrics*, DOI:10.1007/s11192-014-1264-0

Paving the way for a new composite indicator on business model innovations[1]

Franz Barjak*, Marc Bill**

**franz.barjak@fhnw.ch*
***marc.bill@fhnw.ch*
School of Business, University of Applied Sciences and Arts Northwestern Switzerland,
Riggenbachstrasse 16, CH-4600 Olten, Switzerland

Introduction

In the USA, 40% of the 27 companies founded in the last 25 years, that grew their way into the Fortune 500 in the past 10 years did so through business model innovation (Johnson, Christensen, & Kagermann, 2008). David Teece (2010) suggested that the more radical a technological innovation, the greater the need for business model innovation (BMI) in order to capture (part of) the value created by the new technology. Henry Chesbrough (2007, p. 12) seconds: "Today, innovation must include business models, rather than just technology and R&D."

Overall, there is a growing focus on business models and business model innovations (BMI) (Zott, Amit, & Massa, 2011). However, academic research seems to lag behind business practice (ibid.) and we currently know rather little on business model innovations. A big part of the growing literature on BMI is conceptual (see the reviews in Morris, Schindehutte, & Allen, 2005; Osterwalder, Pigneur, & Tucci, 2005; Zott, et al., 2011). Others have developed instruments for using the concept in business practice and consulting (Osterwalder & Pigneur, 2009). Empirical evidence on BMI results mainly from case studies and very few ad-hoc and mostly non-scientific surveys.

- Case studies can capture a broad set of influences within the innovating companies as well as in their environment and are important for developing theory. Usually the case studies are limited to a small number of cases (see e.g. Bucherer, Eisert, & Gassmann, 2012; Casadesus-Masanell & Tarzijan, 2012; Chesbrough & Rosenbloom, 2002; Rohrbeck, Günzel, & Uliyanova, 2012). It is impossible to gather from this line of work how important BMI are in different economies, whether there are specific barriers against it in national research and innovation systems, or what macro-economic consequences BMI have.
- Drawing on a unique, manually collected dataset Zott and Amit (2008) find that novelty-centered business models – coupled with product market strategies that emphasize differentiation, cost leadership, or early market entry – can enhance firm performance. A recent study on Australian pension funds collated a study on 64 companies (pension funds) and measured the degree of BMI as the total of up to seven innovations which should impact the business model (Hartmann, Oriani, & Bateman, 2013). The analysis found appositive impact of BMI on operational pension fund performance. Non-scientific surveys implemented by consultancies have suggested that business model innovators are more successful than other types of innovators, see for instance the BCG innovation survey (Lindgardt, Reeves, Stalk, & Deimler, 2009)

[1] This work was supported by the European Commission under contract SC-RTD/C2/2013/SI2.655784.

and the IBM CEO survey (IBM Institute for Business Value, 2012). However, at least with regard to the IBM survey, the conceptualisation and the underlying sample introduce considerable uncertainty about the validity of this result.

Methodologically stronger innovation surveys, such as the harmonized European Community Innovation Survey (CIS) 2010, the Japanese National Innovation Survey 2012 or the US Business R&D and Innovation Survey (BRDIS) 2010 do not know the concept of BMI (see Barjak, Niedermann, & Perrett, 2013). The same applies for the Oslo Manual, the OECD guidelines for collecting innovation data, which defines and describes four types of innovation but excludes BMI in its most recent edition (OECD, 2005).

CIS experts have complained about the low use and impact of the CIS dataset, the most comprehensive multi-country data set on corporate innovation (Arundel, 2007; Bloch & Lopez-Bassols, 2009). The development and analysis of complex indicators can be a remedy to this, raising the policy relevance of CIS survey questions (Arundel, 2007). A number of such indicators have been suggested to identify different innovation modes or types (Frenz & Lambert, 2012), however, the construct of BMI is also omitted in this line of work.

The present paper aims to close this gap by

- linking the BMI construct conceptually and empirically to established innovation surveys and their definitions,
- identifying gaps in the survey coverage with regard to the BMI construct,
- developing suggestions on how to close these gaps.

We first introduce our understanding of business models and business model innovations in the next section. In section 3 we implement this definition, develop a composite indicator for BMI and measure it with data from CIS 2008 and CIS 2010. The last section summarizes and concludes the paper.

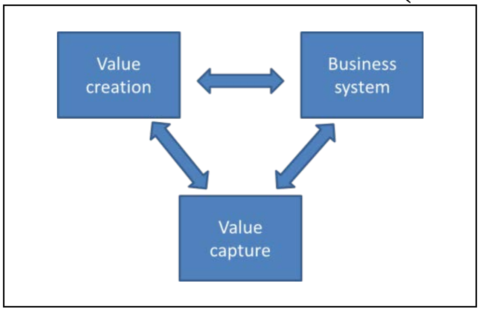
Business models and business model innovation

Business models

In science, as Baden-Fuller and Morgan (2010) point out, models are organisms for investigation. For instance, the laboratory mouse is a model that is representative for its class of mammals and experimenting with lab-mice generates insights that are relevant for mammals. In analogy, business models can be considered as representatives of certain genres of firms that can be studied. A number of scholars have suggested using three aspects of value to define the business model construct and distinguish different genres of firms (see Figure 1):[2]

[2] See in particular Osterwalder & Pigneur (2009), Teece (2010), Yunus et al. (2010) or Zott, Amit, & Massa (2011).

Figure 1: Conceptualization of business models (Source: Authors).



- *Value creation* refers to how and for whom a company (or other organisation) creates value (Morris, Schindehutte, & Allen, 2005). Without value and benefits, users or customers are unlikely and a compelling value proposition is one of the elements of a good business model (Teece, 2010). Brandenburger and Stuart (1996) define the value created in an organisation as the buyers' willingness-to-pay for the products of this organisation minus the organisation's suppliers' opportunity costs. Hence, an organisation can create more value by raising downstream willingness-to-pay or reducing upstream opportunity costs in the value chain. The total maximum value that is created in a value chain depends on the end consumers' willingness-to-pay.
- *Business system.* For a sustainable business model the value needs to be delivered to customers and the costs of doing so need to be lower than the generated revenues. Part of the business model is the entire business system which has been defined as "the 'system of works' (the production/delivery system) that a firm designs - within and beyond its boundaries - to produce and deliver its goods or services to its target customers" (Itami & Nishino, 2010, p. 364). The business system reflects the business architecture and how the organisation internally mobilises its capabilities and organises its activities. It also includes the division of labour between the organisation and its external trading partners and how this is controlled.
- *Value capture.* The third crucial element of a sustainable business model was mentioned already: it is the logic of how to capture value from whatever group of users or customers who benefit from the value created (Chesbrough & Rosenbloom, 2002). The value appropriation has been depicted as the outcome of bargaining between the clients, the firm and the firm's suppliers (Brandenburger & Stuart, 1996). This bargaining results in a distribution of shares of value. However, important is not only who appropriates how much, but also what influences the bargaining position and what contributions justify value claims.

Business model innovation

Experimenting with the business model is common management practice. Managers conduct thought experiments, simulations or real experiments in order to find out whether changes to the business model would raise overall success (Baden-Fuller & Morgan, 2010). Following our definition of business models, we consider business model innovations (BMI) as changes of all three components of business models, 1) value creation, 2) business systems, and 3) revenue generation. This includes innovations in the form of newly introduced goods or services (Mitchell & Coles, 2003) or changes to processes of producing and delivering products, but it requires also that these technological innovations are complemented by "organizational and business model changes as well as alterations in the business network" and how these are linked (Rohrbeck, Günzel, & Uliyanova, 2012, pp. 9-10). BMI is then a composite type of innovation combining more basic types of innovation (Björkdahl & Holmén, 2013).

In addition to combining changes in different areas of the business and its partner network, and creating and appropriating value in a different way, the literature generally agrees on the fundamental character of the changes (Bock, Opsahl, George, & Gann, 2012; Cavalcante, Kesting, & Ulhoi, 2011; Markides, 2006; Yunus, et al., 2010).[3] BMI go beyond "only" doing business in a new way in the company.

Examples of business model innovations have been discussed frequently in the literature. BMI that mainly employ a new approach to *creating value* to the customers are, for instance, shifts from products to services. Chesbrough (2007) points to GE Aircraft, where the engines unit switched the value proposition from selling jet engines to its clients to selling flight hours with the engines rented from GE Aircraft and serviced by the company, shifting the risk of downtime from the airline customer to GE. Similar approaches have taken hold in other industries. An example for a very successful business model innovation that applied a different *business system* than the one dominating at that time in the industry is Dell Computer's direct-to-user (consumers and businesses) business model (Teece, 2010). Good examples of innovative approaches to *generating revenues* and capturing some of the value in the company are again the sponsor-based business model of Google (main revenue from advertisers, see Casadesus-Masanell & Zhu, 2013, on this type of business models) or "freemium" business model of Skype (cheap premium services on top of a free service which helped to scale up the user base in a short time period).

We now try to operationalize this understanding of BMI with existing data on innovation in firms.

Mapping business model innovation in Europe and beyond by means of innovation survey data

Mapping the BMI construct on the existing types of innovations as defined by the OECD and others and implemented in national innovation surveys has clear advantages, as it makes use of existing strong datasets on innovation, contributes to raising the relevance of such data and is more cost-effective than doing an ad hoc survey

Methodology

In order to measure BMI we need to obtain data on innovations that change the value proposition, how the value is created and delivered to users and clients, and how some of this value leads to revenues which are captured by the firm. National innovation surveys do not use the value concept, but they distinguish up to four other types of innovations as suggested by the OECD (2005). In order to map the three components of our business model definition on the four innovation types distinguished by the OECD, we developed three propositions (see Table 1).

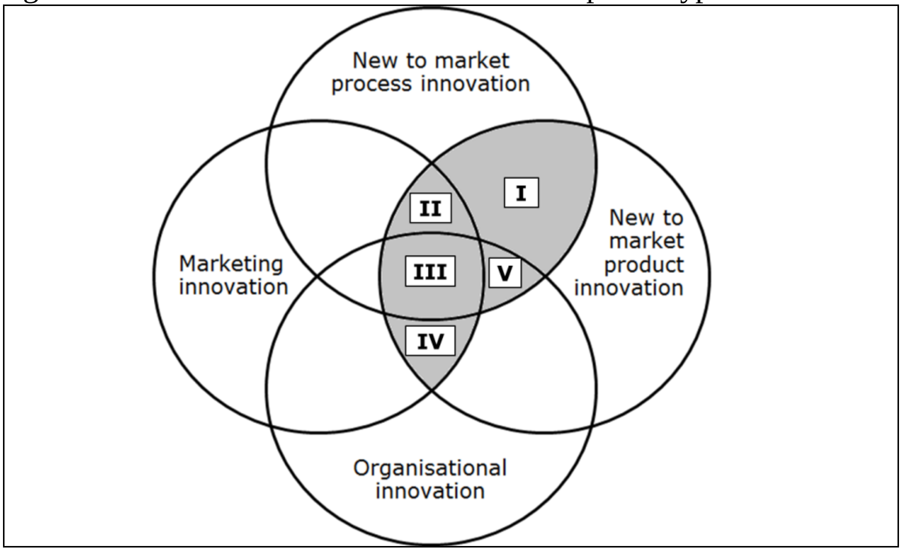
[3] This is challenged by Bucherer, Eisert & Gassmann (2012) who, however, employ a rather narrow definition of radical innovations as characterised by a "discontinuity along the two most important dimensions on a macro-level perspective" (ibid., p. 192) which are industry and market. Using a softer definition and setting radical innovations equal to new to the market/industry, the innovations which they described as incremental would also qualify as radical.

Table 1. Mapping of the business model construct on innovation types

Business model component	**Innovation types**	**Proposition**
Value creation	Product innovation	1. New value propositions will in many, if not in most cases, coincide with product innovations.
Business system	Process innovation, organisational innovation	2. Changes of business systems can be in the form of changes in the production processes as well as internal and external organisation and division of labour along the value chain.
Value capture	Process innovation, marketing innovation	3. A new approach for capturing value will coincide with a process and/or marketing innovation.

This results in a delimitation of business model innovations as a composite type of innovation at the intersection of the four types of innovation defined by the OECD as shown in the figure below (grey area covered by segments I-V).

Figure 2. Business model innovation as a composite type of innovation



This operationalization encounters two challenges:

- A BMI requires in our understanding that the different types of innovations are not implemented independently of each other, but they need to be connected. In order to reduce the risk of including companies with disconnected innovations we limit the analysis to SMEs (firms with <250 employees). This lowers the number of false positives that is companies which introduced different, unconnected innovation types. This also reduces the impact of firm-size differences on country-level indicators, which has been found to influence how the innovation questions in innovation surveys are interpreted by respondents (Arundel, O'Brien, & Torugsa, 2013).
- As we have argued above, business model innovations should be perceived as fundamentally novel and radical changes of how innovating companies do business (and not just as an incremental adjustment). The OECD (2005) suggests three increasing degrees of novelty: new to the firm, new to the market, and new to the world. The available surveys, however, uses the full range of novelty measures only for product innovations; for process innovations they only asks for new to the firm and new to the market. For organisational and marketing innovations it is limited to new to

the firm (Eurostat, 2010). Relying on the existing measures, we restrict the understanding of radical innovations to products and processes introduced as market firsts.

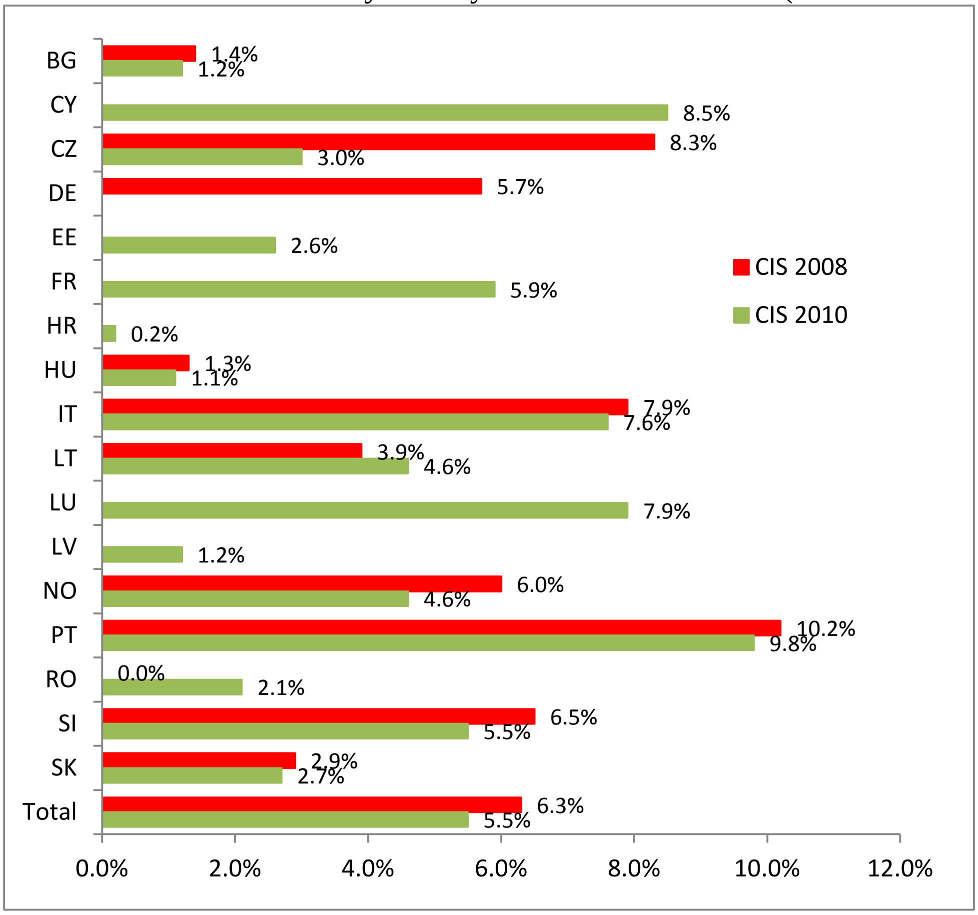
The only survey that included sufficient questions to measure BMI according to this operationalization is the European Community Innovation Survey CIS. We used CIS 2008 and 2010 microdata for the available European countries (CIS 2008: 11 and CIS 2010: 16).

Results of the mapping
Overall 6.3% of the small and medium sized enterprises from 11 countries and different sectors were classified as business model innovators according to CIS 2008 (see figure 3). For a slightly different selection of 16 countries the share of business model innovators went down to 5.5% in CIS 2010. Across countries we find Portugal having the highest share of business model innovators with approximately 10% of all SMEs and it is notable that Portugal has high shares for all industries. Taking CIS 2010 Cyprus, Italy and Luxembourg have rather large shares of BMI as well. In Romania, Hungary, Latvia and Bulgaria the share of business model innovators is lowest with less than 2% of all SMEs. Whereas in most countries for which data in both data sets is available the share of BMI has gone down, it rose in Latvia. Drastic changes, like the drop in the BMI rate in the Czech Republic from 8.3% (the second highest) in 2008 to 3.0% in 2010 require further analyses.

Arundel (2007) explains the implausibly high innovation indicators of some countries, like Portugal and Spain, with the markets which they take as reference points: firms serving less developed domestic markets will more often state that they introduced new products than firms serving more sophisticated international markets. Arundel suggests including only firms which are active on comparable markets, e.g. international markets. Implementing this with CIS 2008 and calculating the indicator for exporting firms only, we get an overall ratio of business model innovators of 9.1% of all exporting SMEs, or +2.8 percentage points compared to all SMEs. Though Portugal still has the highest ratio of BMI (12.1%), other countries in the sample are closer by, in particular the Czech Republic (11.8%), Norway (11%), and Italy (11.1%). It seems that the varying sophistication of the companies' target markets can explain some of the cross-country variation but not all.

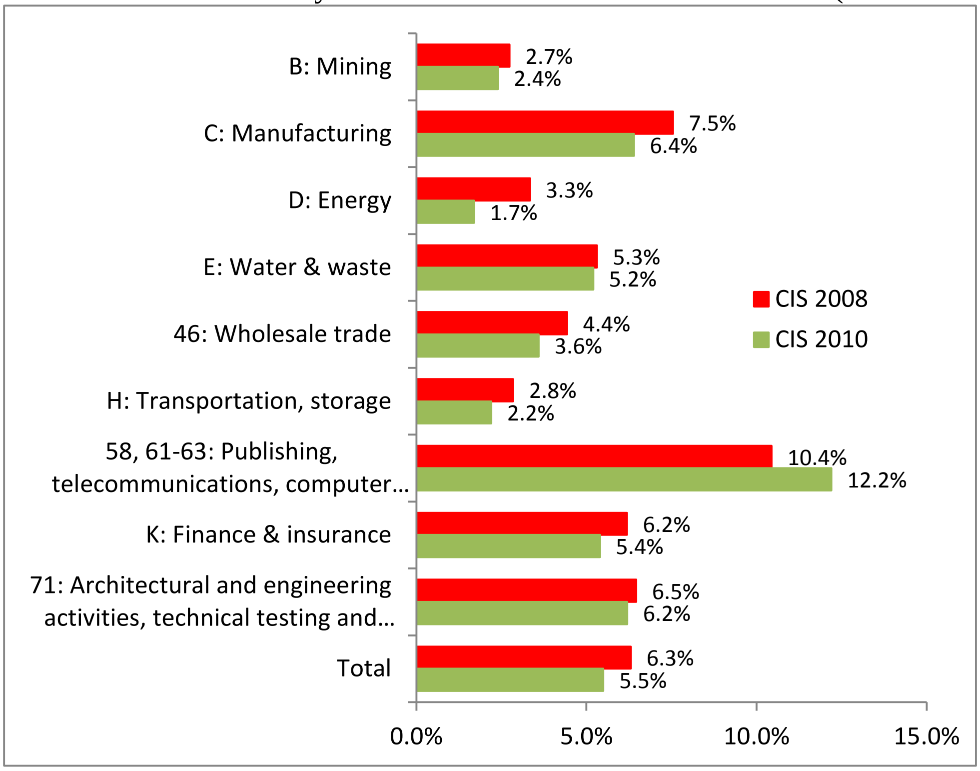
The share of BMI varies between NACE divisions from 1.7% in energy to 12.2% in publishing, telecommunications, computer programming & consultancy and information services. This industry is also the only one in the dataset showing a rise of the share of BMI between 2008 and 2010 (up from 10.4% in CIS 2008).

Figure 3. Percentage of companies with less than 250 employees and a business model innovation by country in CIS 2010 and 2008 (Source: Authors)



We lack good sources for comparing this data in order to evaluate its reliability. One possible source is the above mentioned IBM survey which for different reasons should be used with care (see above). Bock et al. (2012, p. 286) had access to the 2006 survey and based on their data we get a share of business model innovators of 19.2% (=107/556) across all survey respondents. The share varies between 16.5% and 25% according to firm size classes and industries without any consistent pattern. It is highest in Japan with 30%, followed by 22.6% in the Americas. In Europe and China the share is lowest, with less than 15% of all surveyed companies having been identified as business model innovators. However, we do not know whether the IBM data set is reliable and whether the shown magnitude of BMI among large firms is plausible. In order to generate a better basis for comparison, we also measured the share of business model innovators according to our operationalization among all CIS 2010 respondents with at least 250 employees (results not shown). The BMI share among large innovative companies in Europe is 3.4 times higher than among SMEs (without knowing whether the innovations were really introduced in connection to each other) and it is 3.8 percentage points higher than in the IBM surveys; in both surveys finance companies are most often business model innovators. The patterns point into the right direction and raise our trust in the CIS results.

Figure 4. Percentage of companies with less than 250 employees and a business model innovation by NACE division in CIS 2010 and 2008 (Source: Authors)



Conclusions

First, the paper deduced from the literature on business models and business model innovations (BMI) a composite indicator to identify business model innovators and measure BMI across countries. The composite indicator uses the definitions and data on innovations resulting from the work of OECD and Eurostat working groups. It operationalizes BMI as a combination of new to the market product innovations and new to the market process innovations, or new to the market product innovations, organisational innovations, and marketing innovations.

The implementation of the indicator with data from the Community Innovation Surveys CIS 2008 and 2010 shows that approximately one out of 20 SMEs has introduced a business model innovation within the previous three years before the survey. The share of business model innovators decreased slightly from 2008 to 2010 and it varies considerably across countries and industries. At industry level, we see the highest share as well as a rise of BMI in the publishing, telecommunications, computer programming & consultancy business. The most conservative sector is the energy sector, where the rate of BMI even went down from 3.3% in 2008 to 1.7% in 2010. The differences of the incidence of BMI across countries are generally in line with the differences found for other types of innovation with data from the Community Innovation Survey. However, the external validity of the CIS-based indicators still requires further analyses.

If policy makers want to improve the conditions for BMI and lower the barriers against it, they are well advised to first improve the information basis by providing the resources for a better measurement and quantification of business model innovators in both SMEs and large companies.

References

Arundel, A. (2007). Innovation survey indicators: What impact on innovation policy? In OECD (Ed.), *Science, Technology and Innovation Indicators in a Changing World* (pp. 49-64): OECD Publishing.

Arundel, A., O'Brien, K., & Torugsa, A. (2013). How firm managers understand innovation: implications for the design of innovation surveys. In F. Gault (Ed.), *Handbook of Innovation Indicators and Measurement* (pp. 88-108). Cheltenham, UK: Edward Elgar.

Baden-Fuller, C., & Morgan, M. S. (2010). Business Models as Models. *Long Range Planning, 43*(2–3), 156-171.

Barjak, F., Niedermann, A., & Perrett, P. (2013). *The Need for Innovations in Business Models. Final Policy Brief (Deliverable 5).* Retrieved from http://ec.europa.eu/research/innovation-union/pdf/expert-groups/ERIAB-BMI_PB_new_business_models.pdf

Bloch, C., & Lopez-Bassols, V. (2009). Innovation Indicators. In OECD (Ed.), *Innovation in Firms*: OECD Publishing.

Bock, A. J., Opsahl, T., George, G., & Gann, D. M. (2012). The Effects of Culture and Structure on Strategic Flexibility during Business Model Innovation. *Journal of Management Studies, 49*(2), 279-305.

Brandenburger, A. M., & Stuart, H. W. (1996). Value-based Business Strategy. Journal of *Economics & Management Strategy, 5*(1), 5-24.

Bucherer, E., Eisert, U., & Gassmann, O. (2012). Towards Systematic Business Model Innovation: Lessons from Product Innovation Management. *Creativity and Innovation Management, 21*(2), 183-198.

Casadesus-Masanell, R., & Tarzijan, J. (2012). When one business model isn't enough. *Harvard Business Review*.

Casadesus-Masanell, R., & Zhu, F. (2013). Business model innovation and competitive imitation: The case of sponsor-based business models. *Strategic management journal, 34*(4), 464-482.

Cavalcante, S., Kesting, P., & Ulhoi, J. (2011). Business model dynamics and innovation: (re)establishing the missing linkages. *Management Decision, 49*(7-8), 1327-1342.

Chesbrough, H. (2007). Business model innovation: it's not just about technology anymore. *Strategy & Leadership, 35*(6), 12-17.

Chesbrough, H., & Rosenbloom, R. S. (2002). The role of the business model in capturing value from innovation: evidence from Xerox Corporation's technology spin-off companies. *Industrial and Corporate Change, 11*(3), 529-555.

Eurostat. (2010). *The Community Innovation Survey 2010 (CIS 2010). The harmonised survey questionnaire.* Retrieved 25.02.2014, from http://epp.eurostat.ec.europa.eu/portal/page/portal/microdata/documents/CIS_Survey_form_2010.pdf

Frenz, M., & Lambert, R. (2012). *Mixed modes of innovation: an empiric approach to capturing firms'* innovation *behavior.* STI Working Paper 2012/6: OECD Publishing.

IBM Institute for Business Value. (2012). *Leading Through Connections Insights from the Global Chief Executive Officer Study.* Retrieved from http://public.dhe.ibm.com/common/ssi/ecm/en/gbe03485gben/GBE03485GBEN.PDF

Itami, H., & Nishino, K. (2010). Killing Two Birds with One Stone: Profit for Now and Learning for the Future. *Long Range Planning*, *43*(2–3), 364-369.

Johnson, M. W., Christensen, C. M., & Kagermann, H. (2008). Reinventing Your Business Model. *Harvard* Business *Review*.

Lindgardt, Z., Reeves, M., Stalk, G., & Deimler, M. (December, 2009). *White Paper: Business Model* Innovation*: When the Game Gets Tough, Change the Game.* New York: Boston Consulting Group.

Markides, C. (2006). Disruptive Innovation: In Need of Better Theory. *The Journal of Product Innovation Management*, *23*, 19-25.

Mitchell, D., & Coles, C. (2003). The ultimate competitive advantage of continuing business model innovation. *Journal of* Business *Strategy, 24*(5), 15-21.

Morris, M., Schindehutte, M., & Allen, J. (2005). The entrepreneur's business model: toward a unified perspective. *Journal of Business Research, 58*(6), 726-735.

OECD. (2005). *Proposed Guidelines for Collecting and Interpreting Technological Innovation Data: Oslo* Manual. (3 ed.). Paris: OECD.

Osterwalder, A., & Pigneur, Y. (2009). *Business Model Generation. A Handbook for Visionaries, Game* Changers*, and Challengers.* Wiley: Hoboken.

Osterwalder, A., Pigneur, Y., & Tucci, C. L. (2005). Clarifying business models: Origins, present, and future of the concept. *Communications of the Association for Information Systems, 16*(1), 1-25.

Rohrbeck, R., Günzel, F., & Uliyanova, A. (2012). Business Model Innovation: the Role of Experimentation. Paper *presented at the R&D Management Conference*, Grenoble, France. http://ssrn.com/abstract=2074315.

Teece, D. J. (2010). Business models, business strategy and innovation. *Long range planning, 43*(2), 172-194.

Yunus, M., Moingeon, B., & Lehmann-Ortega, L. (2010). Building Social Business Models: Lessons from the Grameen Experience. *Long Range Planning*, *43*(2–3), 308-325.

Zott, C., Amit, R., & Massa, L. (2011). The business model: Recent developments and future research. Journal *of Management*, *37*(4), 1019-1042.

A typology of countries based on efficiency of publication and patenting with respect to manpower and expenditure[1]

Aparna Basu*

aparnabasu.dr@gmail.com
CSIR-National Institute of Science Technology and Development Studies
Dr. K.S. Krishnan Marg, New Delhi 110075 (INDIA)

Abstract
Developed countries have successfully used their scientific knowledge to fuel economic growth. We ask if they are also more efficient in terms of measurable outputs of the scientific system taken as a function of inputs. A model by Albuquerque in 2005 showed that developed and developing countries had distinct behaviors with respect to the publication of scientific papers and patents, suggesting that developed countries are more efficient in converting their accumulated scientific knowledge for economic benefit. Vinkler (2005) on the other hand, suggested that poorer countries make more efficient use of their resources. Using multiple definitions of efficiency corresponding to multiple resource inputs and outputs in the science and innovation ecosystem, we create a typology of countries along multiple dimensions of efficiency. The typology suggests that the simple categories of developing and developed used by Albuquerque may no longer be sufficient in this context as some countries are moving away from publishing towards patenting, apprently fuelled by high expenditures in the business sector.

Keywords: Patents, Publications, Development, Albuquerque model, Efficiency

Introduction
At a time when recession has forced many countries to control R&D expenditure, the question of efficient use of resources becomes important. While European countries have curtailed expenditure, Asian countries such as China and Korea are investing more in research. China has also substantially increased its output of scientific papers and manpower so that now it is second only to the USA. Under the present circumstances, one may like to ask which countries are more efficient in the use of resources used towards meeting their scientific goals?

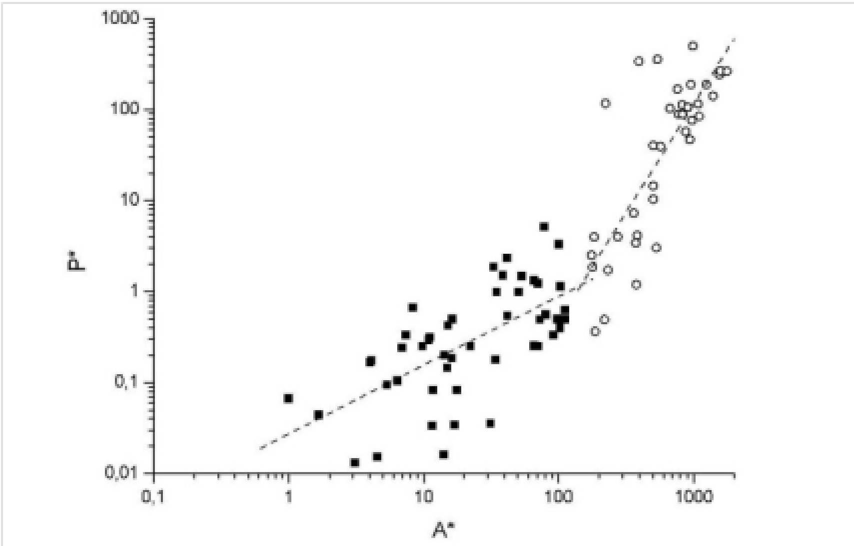
Efficiency estimation requires the use of both inputs and outputs, but they have rarely been combined. As pointed out by Wendt (2012), traditional reports on S&T such as the Main Science and Technology Indicators (MSTI), the European Report on Science and Technology Indicators, or the report from the US National Science Board prefer to report data on input resources and outputs separately rather than combine values to obtain efficiency. This is likely due to the fact that measurement of inputs like research expenditure arose in the context of OECD's efforts, while ouputs like publications arose in the context of information science (Leydesdorff & Wagner, 2009), but may additionally be due to problems of comparability of data across nations. Regardless, there have been attempts at examining the question of national research efficiency. Two articles in prominent journals are the studies by Robert M. May (1997): "Scientific Wealth of Nations", in *Science* and David A. King (2004): "The Scientific Impact of Nations", in *Nature.* Several other studies by Rousseau (1998); Vinkler

[1] This work was supported by a grant from CSIR Extra Mural Division under the Emeritus Scientist scheme.

(2005, 2008); Shelton (2008); Leydesdorff & Wagner (2009); Shelton & Leydesdorff (2011) examined different aspects of this question. Rousseau (1998) considered both publications and patents as outputs and R&D expenditure and manpower as inputs using Data Envelopment Analysis to obtain effectiveness of European countries. Vinkler (2005) concludes from his analysis that poorer countries make more efficient use of their resources. Leydesdorff & Wagner (2009) note that countries differ considerably in terms of efficiency of turning financial inputs into bibliometrically measurable outputs and compute the cost of a research publication in different countries. Shelton (2008) compares input shares and output shares using regression models. Shelton & Leydesdorff (2011) show that there is a trade-off between publications and patenting.

Most of these studies have focused on European countries, Japan, USA and China. Developing countries were included by Basu (2013, 2014a). It was suggested by Vinkler (2005) that poorer countries make more efficient use of their resources, but this has not been verified. At about the same time, in an empirical study across 151 countries, Albuquerque (2005) linked the state of development of a country to its scientific output, i.e. papers and patents. He found that when patents were plotted against scientific papers (both normalized by the population) the developed countries lay beyond a certain threshold in terms of papers published, and were distributed around a line of higher slope as compared to the developing countries (Fig. 1). Developed countries published more papers, and at this stage they were able to convert the knowledge to economically useful products and processes that required them to obtain patents. Albuquerque termed the ratio of papers to patents as *efficiency*, a measure that decreased with development. Basu (2014a) took a direct definition of efficiency as the ratio of outputs to inputs to the science system, obtaining a multidimensional entity for efficiency that took into account multiple inputs, such as expenditure on R&D or manpower, and outputs such as publications or patents. She showed that some countries deviate from an Albuquerque type distribution, indicating new trends in the evolution of research priorities of nations. In this paper, we have created a typology of countries that simultaneously accounts for a multidimensional efficiency parameter in the form of a colour coded ‘heat map’. Based on the typology we group similar countries to answer the question of whether developed countries are more efficient in the production of science.

Fig.1: The Albuquerque model of the distribution of developed (open circles) and developing (dark squares) countries on a plot of papers (A) vs. patents (P) both normalized by population.



Data and Methodology

Data for this study has been taken from OECD and UNESCO publications (OECD 2011; UNESCO, 2010, Hollanders and Soete, 2010). For scientific papers and patents, data are from SCI-Expanded and the USPTO for a set of selected countries from across all continents for the years 2007 and 2008. Restricting to the USPTO gives a ‘home advantage’ to USA in terms of patents (Criscuolo, 2006), which may be expected to give relatively higher patent values for the USA. To factor this out other patent databases such as the PCT or Triadic patents may also be considered (Shelton & Leydesdorff, 2011). However for this study we have only taken data from the USPTO. For papers, articles, reviews and letters are taken into account, and fractional counting has been used for the allocation of collaborative papers to nation states.

The values for the Gross expenditure on R&D (GERD) are adjusted to Purchasing Power Parity (PPP$) for the year 2007. This adjustment accounts for differences in the cost of living index in different countries and facilitates comparison of expenditures across countries. Manpower values have been taken in terms of Full Time Equivalents (FTEs) of persons employed in R&D. Data are shown in Table 1.

Table 1: Research expenditure GERD, Manpower, Papers and Patents for selected countries

	GERD $bnPPP	Manpower (FTE's)	Papers SCI-E	Patents USPTO
Australia	15.36	87140	28313	1516
Brazil	20.20	133266	26482	124
Canada	23.96	139011	43539	3806
China	102.40	1423380	104968	7362
France	42.89	215755	57133	3631
Germany	72.24	290853	76368	9713
India	24.79	154827	36261	741
Italy	22.12	96303	45273	1836
Japan	147.90	709974	74618	33572
Korea	41.30	221928	32781	6424
Mexico	55.90	37930	8262	81
Russia	23.40	451213	27083	286
Spain	19.34	130896	35739	363
UK	41.04	261406	71302	4007
USA	398.00	1425550	272879	81811

Efficiency has been simply defined, following Basu (2013) as the ratio of the outputs of the science and innovation ecosystems, viz., papers and patents, to the inputs, expenditure in R&D and manpower. With two outputs and two inputs there are four dimensions of efficiency corresponding to the four equations below.
The efficiency for paper production has two *values EE(Pap)* and *ME(Pap),*
Expenditure Efficiency EE(Pap) = Papers/GERD (1)
Manpower Efficiency ME(Pap) = Papers/Manpower (2)

The efficiency for patent production also has two values *EE(Pat)* and *ME(Pat),*
Patent Expenditure Efficiency EE(Pat) = Patents/GERD (3)
Patent Manpower Efficiency ME(Pat) = Patents/Manpower (4)

2-dimensional plots of the efficiency of paper production and patenting respectively are used to locate countries in an efficiency space. To create a typology, the countries are categorized in terms of where they lie with respect to the average efficiency of the set in each of four dimensions, each category differentiated by half the standard deviation.
No time lags have been taken in this analysis, although it is expected that money allocated in R&D will result in publications only after some time. In the case of patents with a more protracted system of applications and verification, results are expected after even longer time intervals. The justification for this is that the country wise data do not change very rapidly with time, with the exception of China (Shelton, 2008).

Analysis

The efficiencies of scientific publication with respect to manpower and expenditure, ME(Pap) and EE(Pap) obtained from Eqns.1 and 2, and plotted in Fig.2.

Fig. 2 Plot of country wise efficiency of publication with respect to manpower and expenditure

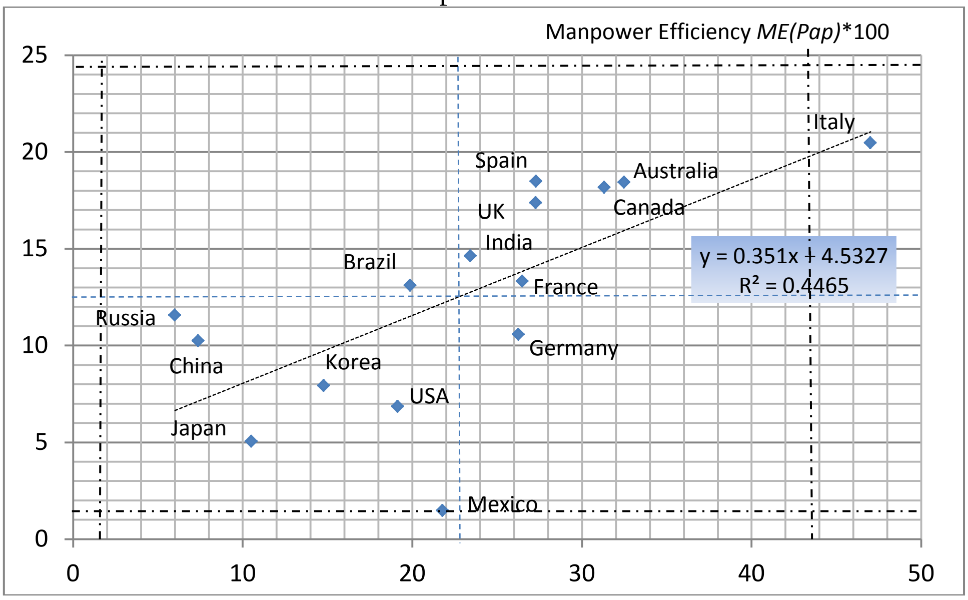
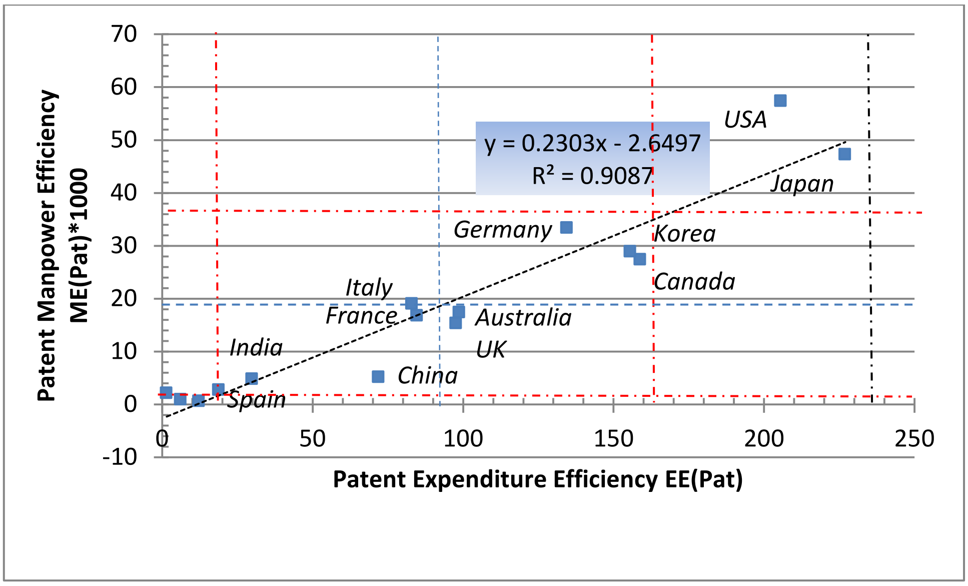


Fig.2 shows that Italy has the highest efficiency in publication (Basu 2014b). In terms of both expenditure efficiency and manpower efficiency, measured in terms of papers per dollar and papers per person, Spain, UK, Canada and Australia are well above average, (indicated by the lines in the figure), while France and India are just above average. Brazil is just below average on manpower efficiency but above average in expenditure efficiency, while Germany is above average on manpower efficiency and below average on expenditure efficiency. Mexico is well below average on expenditure efficiency and just below average on manpower efficiency. Russia is well below average in terms of manpower efficiency and just below average on expenditure efficiency. USA is just below average in manpower efficiency and below average on expenditure efficiency. China Korea and Japan are all below average on both the dimensions, expenditure efficiency and manpower efficiency. The fit of the line to the data shows that expenditure efficiency and manpower efficiency of publication are not highly correlated. Some correlation is expected as a proportion of allocated funds go toward salaries of persons engaged in R&D.

Fig. 3 Plot of country wise efficiency of patenting with respect to manpower and expenditure



*Russia, Mexico and Brazil are unlabelled and seen at the left cormer of the plot.

In a similar visual analysis for patenting efficiencies EE(Pat) and ME(Pat) obtained using Eqns. 3 and 4, (Fig.3) we find that the USA and Japan are among the highest in both expenditure and manpower efficiency of patenting. Germany, Korea and Japan are all above average along both dimensions of patenting. Italy, France, UK and Australia are all just below average on manpower efficiency, while only UK and Australia are above average on expenditure efficiency. China is below average on expenditure efficiency, but well below average on manpower efficiency. India and Spain are low on both the dimensions, while Russia, Mexico and Brazil (shown together in the lower left corner of the figure) are also well below average along both the dimensions. The high degree of collinearity in the data as shown by the fitted line indicates that there is a degree of dependence in the expenditure and manpower in patenting. The implication is that the dimensions of efficiency are not entirely independent – nevertheless they can be used to create a typology of the countries.

Results

Based on the position of each country along the four dimensions of efficiency in scientific production as a ratio of inputs, manpower and expenditure, we place them in categories to create a typology (Fig. 4). The colour-coded categorization shows a steady progression of the countries from being very highly efficient in publication and above average in patenting to countries that are low in publication efficiency but highly efficient in patenting. In between there are countries which are average or low in both publication and patenting.

Fig. 4 Typology of countries along a 4-dimensional efficiency parameter, efficiency of publication with respect to manpower and expenditure ME(Pap) and EE(Pap), and efficiency of patenting ME(Pat) and EE(Pat) w.r.t. manpower and expenditure.

	ME(Pap)	*EE(Pap)*	*EE(Pat)*	*ME(Pat)*
Italy	5+	3+	1-	1+
Canada	2+	3+	2+	2+
Australia	2+	3+	1+	1-
UK	1+	2+	1+	1-
France	1+	1+	1-	1-
Spain	1+	3+	3-	2-
India	1+	1+	2-	2-
Brazil	1-	1+	3-	3-
Mexico	1-	4-	3-	2-
Russia	4-	1-	3-	3-
China	3-	1-	1-	2-
Germany	1+	1-	2+	2+
Korea	2-	2-	2+	2+
USA	1-	3-	4+	5+
Japan	3-	3-	4+	4+

Scale	
5+	Very high
4+	High
3+	Medium high
2+	Medium ++
1+	Medium ++
1-	Medium -
2-	Medim --
3-	Medium low
4-	Low

Canada lies above average on all four dimensions of publication and patenting efficiency. Italy is the highest on manpower efficiency of publication and also expenditure efficiency. All four countries Italy, Canada, Australia and UK are high on publication efficiency and moderate on patenting efficiency. Together they may be called Group 1 or the ‘classical developed’ countries.

France is around average in both patenting and publication, being slightly more efficient in publication. It is labeled Group 2.

In Group 3, Spain is high on expenditure efficiency but moderate on manpower efficiency with respect to publication but low in patenting efficiency. India is moderate (above average) in publication efficiency but low on patenting efficiency. Brazil is about average in expenditure efficiency but low on manpower efficiency in publication and very low in patenting. Together they act as a single group while Spain can be regarded as a transitional country between Group 3 and the ‘classical developed’ countries, Group 1.

Mexico and Russia are placed together in Group 4, as they are below average on almost all efficiency dimensions, though their characteristics are different. Both are low on patenting efficiency, but while Mexico is very low on expenditure efficiency, Russia is very low on manpower efficiency. This is due to the very high levels of manpower in Russia.

China has been placed alone in Group 5. It is very low on manpower efficiency of publication and also below average on the other counts. This may be because of the very high manpower values in China, which could be a data discrepancy that might arise if the OECD manpower

categories are not properly harmonized across countries. In general Chinese efficiency appears low because of massive resource allocations.
Group 6 has four countries, Germany, Korea, USA and Japan which are graded with respect to publication and patenting efficiency. In general they have lower levels of publication efficiency in comparison to patenting efficiency which is high to very high. This last group consists of countries which have very high expenditure on R&D in the business sector, or the business component of GERD (BERD). All the countries spend over 70% of their R&D allocation in the business sector. This leads to high levels of patenting. The trend of resource allocations favouring patenting over publications is a relatively new trend, and this group may be called the Ínnovators'.
In summary, the typology groups countries which are more efficient in publication to those that are more efficient in patenting. There are differences in output efficiency with respect to manpower and with respect to expenditure. France can be considered as a nation that lies between Groups 1 and 6, with lower levels of efficiency in publication than Group 1 but higher than Group 6, and correspondingly lower levels of patenting efficiency than Group 6. China is close to Russia in some respects but is also tending toward Group 6.

Discussion

The empirical distribution defined by Albuquerque (2005) has been used here as an underlying model that distinguishes between publication and patenting characteristics of developing and developed countries. According to Albuquerque, developed countries have higher levels of publications per capita, together with much higher levels of patenting. We have modified the approach to include inputs, R&D expenditure and manpower in R&D to obtain efficiency of S&T for different countries. Combining input and output indicators across countries can present difficulties since collection of input statistics and their categorization into statistical units may vary from country to country (Luwel, 2004, Wendt, 2012). However, results obtained here appear to be consistent with perceived performance of different countries. It has been observed by Basu (2014a) using data of 2007-2008, that Japan, USA, South Korea and Germany have moved away from an Albuquerque type distribution. The typology of publication and patenting efficiencies with respect to manpower and expenditure shows that this group, called the Innovative countries, has a very high efficiency in patenting but with a correspondingly low efficiency in publication. In contrast, the Classical Developed countries including those in Europe, UK, Australia and Canada have high publication efficiencies and medium to low level patenting efficiencies. This trade-off between publications and patents has been noted earlier by Shelton and Leydesdorff (2011), who have also noted that expenditure in different sectors of the economy such as the Higher Education sector (HERD) is correlated to publications and expenditure in the business sector (BERD) is correlated to patents. (The funding for R&D or GERD is usually broken down into four sectors – the business sector, the government sector, the higher education sector and the non-profit sector.) If we look at the composition of expenditure on R&D, we find that 78.2% of total R&D expenditure in Japan comes from the business sector. It is 67.3% in the US, 67.6% in Germany, only 45.1% in the UK (2008 figures) (OECD, 2011; Eurostat, 2011) and 76.8% for Korea (2005 figures) (Adams, et al, 2013). R&D activities in the business sector are expected to be more closely related to the production of new products and processes, and therefore patenting. Japan and Korea are among the countries that have the highest percentage shares from the business sector (BERD) with the exception of a few countries like Israel, not included in our study. It would appear that high efficiency in patenting is fuelled in these four countries by high research expenditures in the business sector. One area that has received considerable attention in recent years is the structural difference in R & D funding between Europe and its main competitors. Policymakers in

Europe have tried to increase R & D business expenditure so that it is more in line with relative contributions observed in Japan or the United States (Eurostat, 2012). From the typology it appears that Germany has already become part of this group.
Another policy target of the EU is the Lisbon strategy, to take research intensity or GERD as a ratio of GDP to 3% by 2020. It is referred to as the headline indicator by the European Commission (Eurostat, 2012). Rising spends without a concomitant rise in outputs will dilute efficiency. For example, in our study Italy has the highest efficiency in publications. Its publications are on par with other European countries, but its research expenditure as a proportion of GDP, GERD/GDP is low, more in line with developing countries (Basu, 2014b).
Russia and the emerging economies, India, Brazil and Mexico have low levels of patenting efficiency, but medium levels of publication efficiency. Russia has low values of manpower efficiency due to a large manpower base. China's efficiency is also diluted by a very large workforce and high expenditure. China is currently close to Russia, but is on the way to transition to Group 6, that is Germany USA, Korea and Japan, as also France.
With the emergence of these groups identified by the typology, it is clear that developing and developed are not sufficient as categories in this context, and new categories are required that distinguish between countries such as the Classical Developed countries and the Innovator countries. Developing countries are moderate to low in efficiency. We also conclude that countries with high R&D expenditure in the business sector are also highly efficient in patenting. Developed countries that show high efficiency in patenting show a lessened efficiency in publications output pointing to emerging research priorities in these countries.
Some limitations of our study relate to the difficulties related to comparability of data across nations, the use of the USPTO database that gives the USA an advantage in patent applications and the fact that no time lags have been included in our definition of efficiency.

References

Adams, J., D. Pendlebury, B. Stembridge (2013) *Building Bricks*, Thomson Reuters Report

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* (pp759-778) Kluwer Academic Publishers.

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, forthcoming*

Basu, A. (2014b) Is Italy highly efficient in science? (Submitted for publication)

Criscuolo, P. (2006). The 'home advantage' effect and patent families. A comparison of OECD triadic patents, the USPTO and EPO. *Scientometrics, 66*(1), 23-41.

Eurostat (2012) *Statistics Explained*, European Commission, http://epp.eurostat.ec.europa.eu/statistics_explained/index.php/R_%26_D_expenditure: Retrieved July, 2012

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

King D.A. (2004) The scientific impact of nations, *Nature*, 430, 311-316 (15 July 2004)

Leydesdorff L., C. Wagner (2009) Macro-level indicators of the relations between research funding and research output, *Journal of Informetrics*, Vol 3, Issue 4, Oct 2009, pp 353–362

Luwel, M. (2004) The use of input data in the performance analysis of R&D systems: Potentialities and pitfalls, in The Handbook of Quantitative Science and Technology Research, H. Moed, W. Glanzel, U. Smoch (Eds.), pp 315-338

May, R.M. (1997) The Scientific Wealth of Nations, *Science* 7 February 1997: 793-796

Hollanders, H., L. Soete (2010) The growing role of knowledge in the global economy. In *UNESCO Science Report 2010*, UNESCO Publishing.

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: Retrieved July 2012

Rousseau, S, R. Rousseau (1998) The Scientific wealth of European nations: Taking effectiveness into account, *Scientometrics*, Volume 42, Number 1, pp 75-87

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., L. Leydesdorff (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., D. W. Aksnes, G.Sivertsen 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. http://2012.sticonference.org/index.php?page=proc

Disciplinary Profiles and Performance of Research Systems: a World Comparison at the Country level[1]

Irene Bongioanni*, Cinzia Daraio**, H. F. Moed*** and Giancarlo Ruocco*

[*] *irene.bongioanni@gmail.com, Giancarlo.Ruocco@roma1.infn.it*
*Sapienza University of Rome, Department of Physics, Rome, (Italy)

[**] *daraio@dis.uniroma1.it (Corresponding Author)*
*Sapienza University of Rome, Department of Computer Control and Management Engineering, Via Ariosto, 25, Rome, 00185 (Italy)

[**] *H.Moed@elsevier.com*
Informetric Research Group, Elsevier, Amsterdam (The Netherlands)

Introduction

The disciplinary structure of the scientific production of countries has been much studied in the literature (see e.g. Almeida, 2009, Tian et al. 2008 and Glanzel, 2008). Several studies have analysed national publication profiles. Such profiles indeed show interesting features of a country's research system and its national scientific policy. A commonly used approach is based on the study of publication profiles by discipline. Within this framework, the world's scientific output is divided into major scientific fields, and the relative contribution of each country with respect to each field is illustrated on a radar chart (see e.g. Glanzel, 2000 and King, 2004). The publication profile of a national research system is then measured by the Relative Specialization Index (RSI) which indicates whether a country has a relatively low or high share in world publications in a given discipline compared to the overall share of world total publications.

Zhou et al. (2012) proposed to use the classical Gini index as a measure of diversity within systems and to measure similarities between systems with the popular Salton's cosine measure. More recently, Bongioanni, Daraio and Ruocco (2013, 2014) have investigated the quantitative evaluation of disciplinary profiles of European countries and their evolution over time in a general framework in which the scientific production is modelled as a complex system. They proposed a more general measure of similarity of disciplinary profiles between systems, borrowed from the physics of complex systems (from spin-glasses systems), where it is named "overlap". Spin glass models which are conceived as the prototype of a complex system, are increasingly applied in a wide range of empirical contexts in other fields, such as biology, computer science, and the economics of financial markets.

Furthermore, their use offers the opportunity to investigate the dynamics of the overall system over time, that is whether the system converges towards a unique disciplinary profile or it diverges to a differentiated configuration.

Research questions and policy relevance of the analysis

In this paper we extend the analysis carried out in Bongioanni et al. (2013, 2014) to the world and include also the consideration of productivity at disciplinary level in the analysis.

The main research questions addressed in the paper are:

[1] This work was supported by Elsevier that provided the data within the Elsevier Bibliometric Research 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".

1) Is the global research system converging towards a unique disciplinary profile or towards a differentiated disciplinary specialization?
2) What is the disciplinary profile of countries whose profiles are similar to the world standard, what is the profile of countries far away from this standard?
3) What is the degree of similarity among European countries, and how do their profiles compare to those of large non-European countries, and to the world standard?
4) What happens to the global research system if we consider the closeness of disciplinary profiles with respect to the best performers (top 25%) as measured by the scientific production (number of publications, citations and so on) per author at disciplinary level? Does the global research system converge towards a unique disciplinary profile or a differentiated disciplinary specialization emerge?

These questions are relevant for policy makers in charge of planning investments in R&D at country level, to have a quantitative evaluation and an empirical understanding on the actual state and the tendency of the global research system over time.

Method

Bongioanni, Daraio and Ruocco (2013, 2014) propose 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 Pa(i) i.e. the shares of articles published in a subject category i for a given country 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 1 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, Pa(i) and Pb(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

$$\sigma_a(i) = \frac{P_a(i) - \langle P_a(i)\rangle}{\sqrt{\langle P_a(i)^2\rangle - \langle P_a(i)\rangle^2}}$$

in which <A> stands for average of A, and $\sigma_a(i)$ and $\sigma_b(i)$ represent the normalised shares of the indicator considered, for country a and country b, respectively; and D is the number of subjects or disciplines analysed, which in this study amounts to 27. A full list of these disciplines, derived from Scopus, is given in the Appendix.

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.

In this paper we computed the overlaps:

- Of each main country in the world[2] against all other countries;

[2] The main countries analysed are the 42 countries grouped in Table 3 for descriptive purpose.

- Of each country against the world reference;
- Of each country with respect to the top 25% performers in terms of productivity, defined as the number of articles (or citations and so on), divided by average number of authors by field.

Data

Data was extracted from the Scopus database and refer to the scientific production of world countries and 27 Scopus subject categories (disciplines) from 1996 to 2012. The available indicators are reported in Table 1.

Table 1. Presentation of the indicators analysed in the paper

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.

Descriptive analysis

Before applying our approach to the full set of data (i.e. for all the countries and all the subject categories), we present as a first illustration of the data a descriptive analysis based on 4 groups of disciplines, namely Medicine, Science, Social Science and Engineering (that were built by aggregating the 27 Scopus subject categories according to Table 2) and 8 groups of countries defined in terms of their geographical location and volume of scientific production, labelled as EU1, EU2, EU3, NA, SA, OC, FE, E. Table 3 provides a definition of these sets. See the Appendix for the full list of 42 countries analysed and the full list of subject categories' acronym.

Table 2 Groups of disciplines

Group of disciplines	*Scopus subject categories included*
Med	BIOC-IMMU-MEDI-NEUR-NURS-PHAR-VETE-DENT-HEAL
Sci	AGRI-CHEM-EART-ENVI-MATE-MATH-PHYS
SocSci	ARTS-BUSI-DECI-ECON-PSYC-SOCI
Eng	CENG-COMP-ENER-ENGI

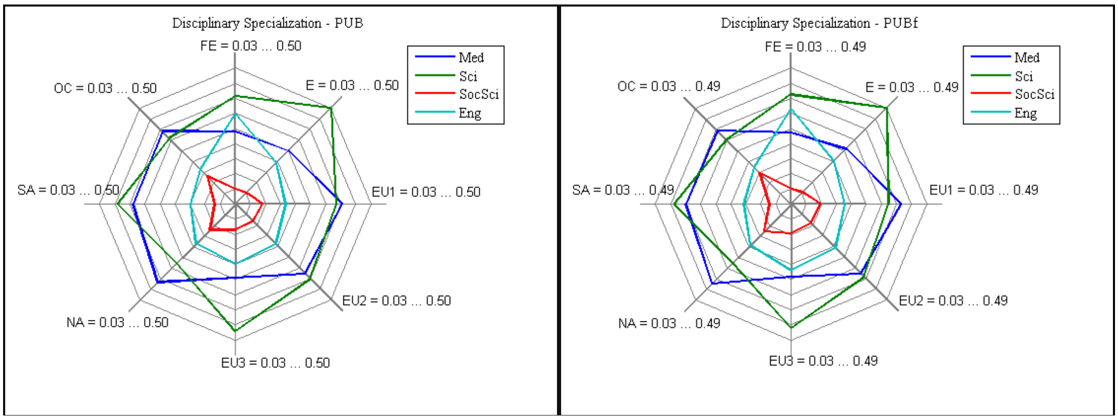
Table 3 Groups of countries analysed in the descriptive analysis

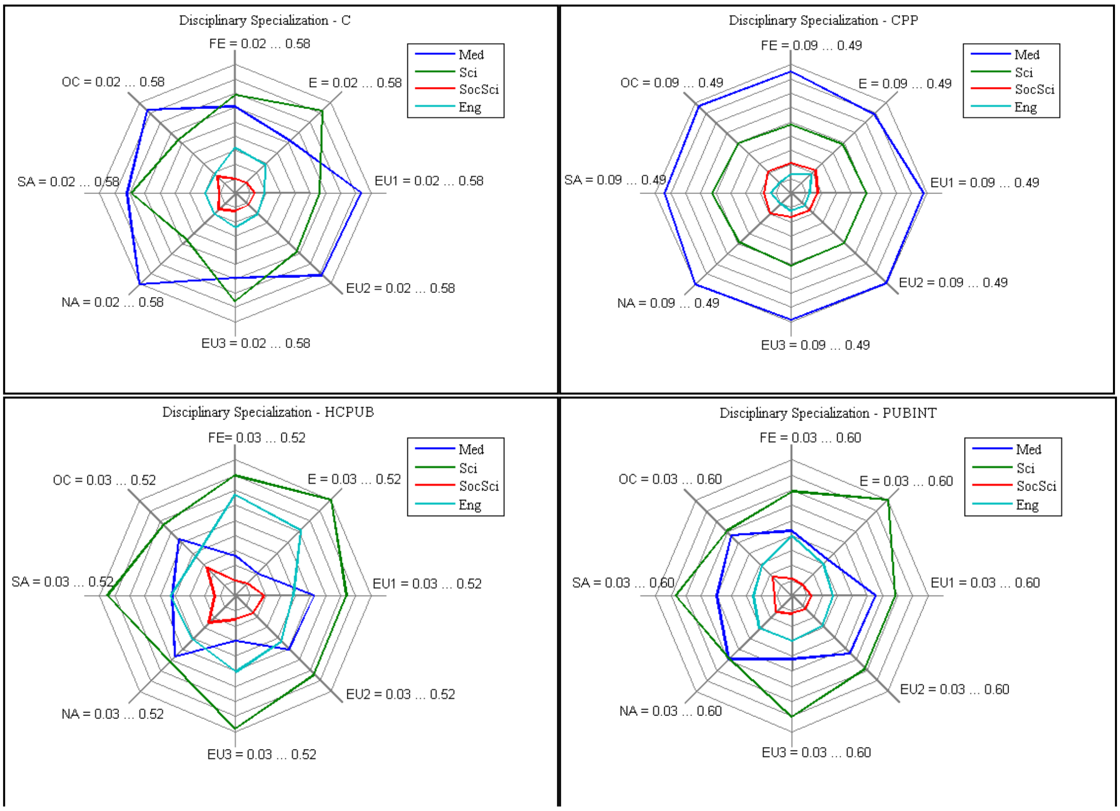
Acronym	*Countries included*
EU1	GBR, DEU, FRA, ITA, ESP, NLD, SWE, POL, BEL
EU2	AUT, DNK, FIN, GRC, CZE, PRT, HUN, IRL, ROU
EU3	SVK, SVN, BGR, LTU, EST, LVA, CYP, LUX, MLT
EU27	EU1, EU2, EU3
NA	North America: Canada, USA
SA	South America: Brazil, Argentina, Chile, Mexico
OC	Oceania: Australia, Indonesia
FE	Far East: China, Japan, Taiwan, Korea
E	East: Russia, Turkey, India

Note: European countries are grouped on the base of their volume of publications.

A first investigation of the disciplinary specialization of countries can be provided by radar plots. Figure 1 illustrates the radar plots calculated over the disciplines defined in Table 2, country sets from table 3, for a selection of indicators listed in Table 1.

Figure 1 Radar plots of selected indicators PUB, PUBf, C, CPP, HCPUB and PUBINT (see Table 1).





An inspection of Figure 1 shows that large differences exist in the profiles of the various country sets. Results for the indicators PUB and PUBf are almost identical although in the latter the values tend to be lower. The figure shows for instance that country sets EU3 (smaller European countries) and E (Russia, India, Turkey), have a strong activity in science fields, and NA (North America) in medical sciences. For all countries the activity in social sciences is relatively low. Interestingly, as far as the CPP indicator is concerned, we observe a strong regularity in the disciplinary profiles of countries because the radar diagram of CPP has almost all concentric circles.

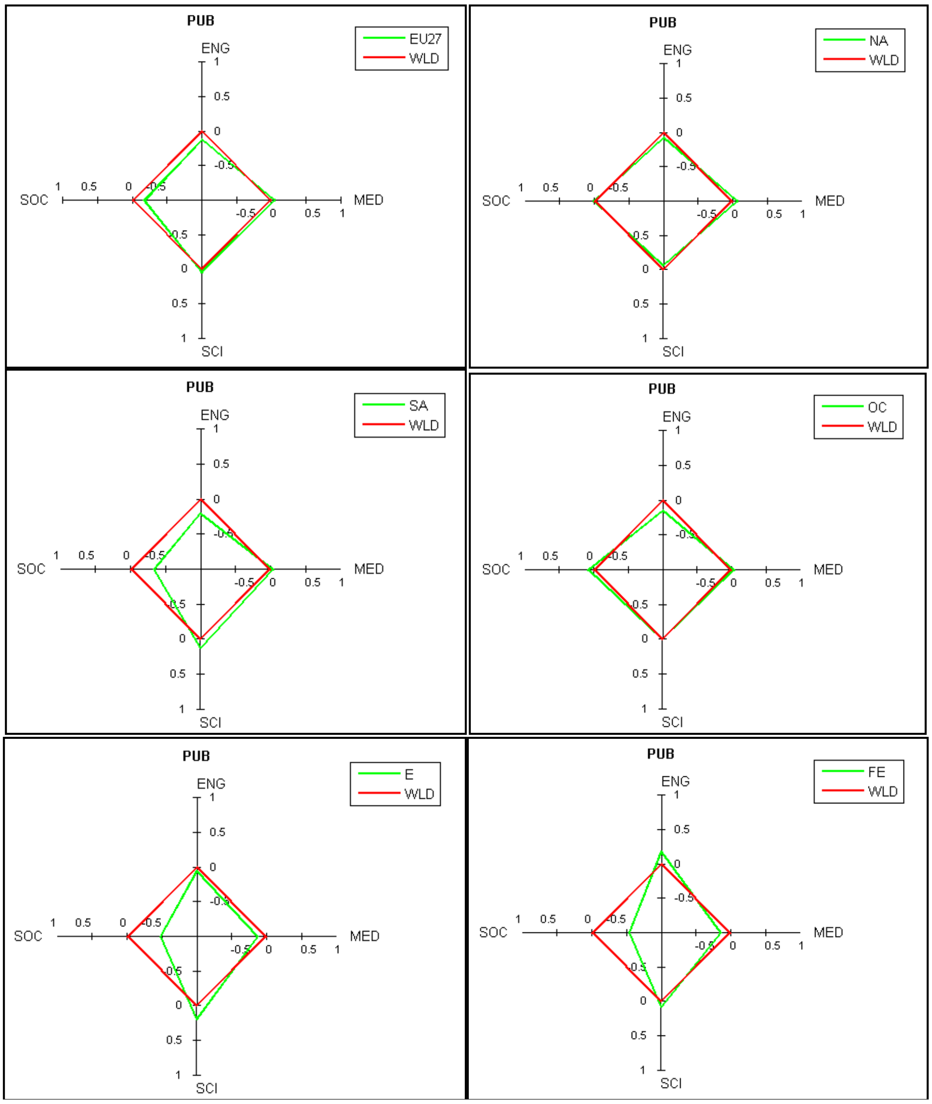
Differences among countries emerge also from the inspection of the following Table 4.
Table 4 reports the RSI calculated according to Glanzel (2000) which ranges from [-1,1].

Table 4. Relative Specialization Indices (RSI) of the total scientific production over 1996-2012 (PUB) by groups of countries.

	EU27	NA	SA	E	FE	OC
GENE	-0,168	0,086	-0,162	0,096	-0,001	-0,164
AGRI	0,044	0,018	0,424	0,085	-0,158	0,265
ARTS	-0,352	-0,255	-0,502	-0,754	-0,861	-0,234
BIOC	0,085	0,118	-0,007	-0,050	-0,019	0,014
BUSI	-0,189	-0,022	-0,493	-0,283	-0,242	0,146
CENG	-0,098	-0,159	-0,027	0,124	0,123	-0,193
CHEM	0,036	-0,150	0,018	0,298	0,141	-0,200
COMP	-0,036	-0,047	-0,173	-0,154	0,180	-0,007
DECI	0,008	0,051	-0,074	-0,141	0,058	0,098
EART	0,099	0,062	0,143	0,157	-0,035	0,231
ECON	0,033	0,124	-0,271	-0,423	-0,491	0,191
ENER	-0,166	-0,136	-0,096	0,113	0,159	-0,191
ENGI	-0,158	-0,064	-0,271	-0,081	0,198	-0,192
ENVI	0,025	0,033	0,110	0,060	-0,151	0,212
IMMU	0,125	0,087	0,224	-0,078	-0,129	0,093
MATE	-0,041	-0,194	-0,095	0,203	0,234	-0,220
MATH	0,091	-0,042	0,061	0,098	0,050	-0,040
MEDI	0,045	0,047	0,001	-0,198	-0,186	0,043
NEUR	0,122	0,190	0,025	-0,364	-0,211	0,098
NURS	-0,099	0,212	-0,018	-0,648	-0,609	0,264
PHAR	-0,002	-0,003	0,049	0,170	0,001	-0,131
PHYS	0,110	-0,049	0,079	0,276	0,156	-0,181
PSYC	-0,083	0,188	-0,331	-0,616	-0,731	0,137
SOCI	-0,175	0,046	-0,267	-0,456	-0,573	0,137
VETE	0,038	-0,030	0,483	0,231	-0,436	0,073
DENT	-0,025	-0,038	0,525	0,070	-0,211	-0,074
HEAL	-0,008	0,188	-0,260	-0,405	-0,402	0,200

The following Figure 2 reports the radar plots of RSI calculated by groups of countries and by groups of disciplines.

Figure 2. Radar plots of the RSI of total scientific publications (PUB) over 1996-2012 by groups of countries and disciplines.



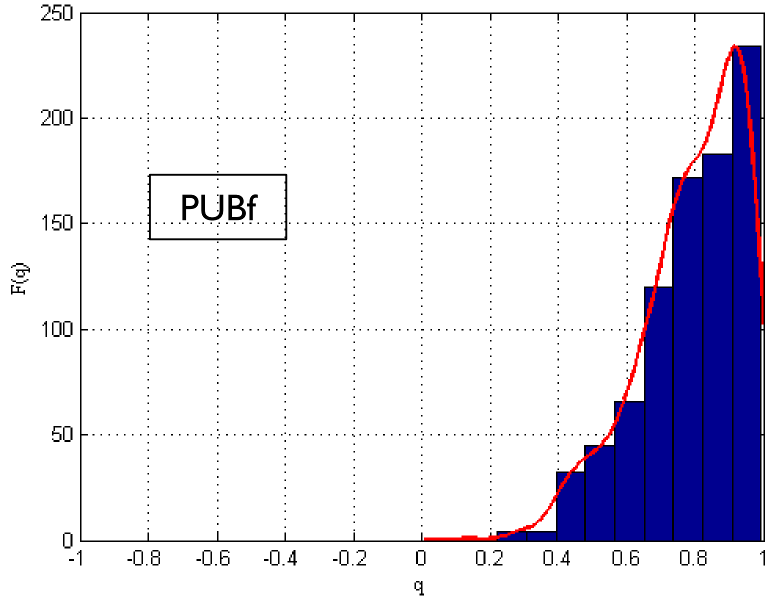
Results

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' profiles relates to their distribution. The distribution of the overlap reveals whether there is a *converge*nce 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 (2013, 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.

Figure 3 illustrates the distribution of the overlap values calculated among all main countries of the world analysed. The overlaps are calculated over the volume of publications in fractional count (PUBf). The distribution of the overlaps clearly shows a pick on one, reflecting the existence of a *convergence towards a unique disciplinary profile*. We observe however that there are countries which are a bit far away (those with overlap values around zero) and a certain heterogeneity among countries exist.

Figure 3. Distribution of the overlaps calculated on each country against all other countries in the world for the indicator PUBf.



The detailed results of the overlap of the disciplinary profile of each country against all other countries in the world are reported in the Appendix. Please note that the values reported in the Appendix refer to the volume of publications and compare the disciplinary specialization of countries among them, without making reference to the world standard or to productivity (these analyses are performed later in this section).
For the interpretation of the results, the overlap measure as recalled above, can varies between -1 and 1 indicating opposite profile (-1) or exactly the same disciplinary profile (1), with zero corresponding to independent profile and intermediate values indicating in-between levels of similarity or dissimilarities.
Overall, we observe a trend towards a *globalization of science* in the world as all overlap values are positive and a clear peak on one emerged.
Moreover, main European countries such as ITA, DEU, GBR, FRA, are very close to the main scientific world producers USA, CAN, AUS. But BRIC countries Brazil, China, India and especially Russia have profiles that are rather different from that of the USA and other main countries.
We compare also the overlap of each of the 42 countries (listed in the Appendix) against the world reference (total sum over all the countries in the world) and found that the top 10% of countries, i.e. those that have the highest values of overlap, - in other words, those that are more similar to the world standard, are the ones listed below:

Country	**Overlap**
USA	0.976
FRA	0.975
IRL	0.975
CAN	0.972

On the contrary, the bottom 10% of the countries, i.e. those with lowest value of the overlaps and so more far away from the world standard are reported below:

Country	**Overlap**
LVA	0.691
CYP	0.689
RUS	0.546
TUV	0.171

Table 5 shows the disciplinary profiles of countries with the highest values of the overlap (TO, Top Overlap), that are the countries closer to the world reference compared with the disciplinary profile of countries with the lowest values of the overlap (BO, Bottom Overlap).

Table 5 Overlap calculation with respect to the World reference. Disciplinary profiles of countries with top overlap (TO) values (first 10% of highest values of overlap) and bottom overlap (BO) values (lowest 10% values of overlap)

	PUB		**PUBf**		**C**		**PUBINT**	
Discipline	**TO**	**BO**	**TO**	**BO**	**TO**	**BO**	**TO**	**BO**
AGRI	5.60	0.00	5.53	0.00	4.87	0.00	6.37	0.00
ARTS	1.01	0.00	1.58	0.00	0.21	0.00	0.43	0.00
BIOC	9.33	12.79	9.52	11.28	17.64	15.71	11.17	14.38
BUSI	1.19	0.00	0.99	0.00	0.43	0.00	0.80	0.00
CENG	2.19	0.00	2.24	0.00	1.57	0.00	1.93	0.00
CHEM	4.29	0.00	4.63	0.00	5.48	0.00	5.73	0.00
COMP	6.76	11.83	6.03	14.74	2.36	6.33	5.49	9.69
DECI	0.60	0.00	0.54	0.00	0.25	0.00	0.53	0.00
EART	3.25	7.06	3.02	6.43	3.23	12.23	4.91	11.73
ECON	0.95	0.00	0.95	0.00	0.49	0.00	0.84	0.00
ENER	1.08	0.00	1.08	0.00	0.55	0.00	0.97	0.00
ENGI	10.36	25.58	10.52	27.18	3.83	12.14	8.09	19.65
ENVI	3.39	10.05	2.95	10.43	2.56	12.50	3.07	10.07
IMMU	2.64	4.41	2.63	2.73	5.59	5.17	3.42	5.85
MATE	4.29	0.00	4.42	0.00	3.31	0.00	5.24	0.00
MATH	3.86	0.00	3.90	0.00	1.78	0.00	4.62	0.00
MEDI	18.33	17.02	17.75	12.45	23.70	25.22	14.31	19.42
NEUR	2.02	0.00	2.38	0.00	3.86	0.00	2.49	0.00
NURS	0.98	0.00	1.16	0.00	0.66	0.00	0.54	0.00
PHAR	2.14	3.89	2.16	1.97	2.60	5.24	2.00	4.64
PHYS	7.86	0.00	7.22	0.00	8.01	0.00	11.56	0.00
PSYC	1.52	0.00	1.94	0.00	1.34	0.00	1.21	0.00
SOCI	3.87	7.37	4.49	12.78	1.26	5.46	1.92	4.56
VETE	0.60	0.00	0.70	0.00	0.51	0.00	0.70	0.00
DENT	0.44	0.00	0.23	0.00	0.24	0.00	0.27	0.00
HEAL	1.12	0.00	1.14	0.00	1.00	0.00	0.78	0.00
GENE	0.31	0.00	0.31	0.00	2.68	0.00	0.62	0.00
TOTAL	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00

Table 5 shows that the largest discrepancy between the profiles of countries showing the largest overlap with the world standard and those revealing the lowest overlap, is that the latter group has a disproportionally large activity in the field of engineering, and, to a lesser extent, in earth sciences, environmental sciences and computer science.

We proceed with the analysis by investigating the productivity of disciplinary profiles of countries by using the Number of Publishing Authors (NA variable) at disciplinary level, calculated in Scopus.
In order to investigate the productivity of the disciplinary profiles of countries, for each country and discipline, we calculated the average Number of Authors over the time span and divided the indicators such as PUB, PUBf, C, HCPUB and PUBINT by this average. We did not consider very small countries that account altogether for less than 0.5% of the overall scientific production. In a next step, we computed the values of the 75th percentile of the productivity distribution in each discipline, which corresponds to the top 25% in terms of productivity, and finally calculated the overlap between each country and this value

As already recalled above, the distribution of the overlaps reveals whether there is a *convergence* of the overall system towards a unique disciplinary profile or there is a divergence of the system towards different disciplinary configurations: in particular, one pick on one shows a convergence towards the *same* disciplinary profile for all countries, two picks point to two *different* configurations of disciplinary profiles.

Some preliminary results seem to show that, when comparing the disciplinary profiles of countries on a productivity base dividing the scientific production by the number of publishing authors in a discipline, a certain *polarization of science* seems to emerge on the global scale. Further analyses however are needed to confirm this preliminary findings and reveal whether the great divide that appeared is that between scientifically developing and scientifically developed countries, or whether there are other explaining factors.

Conclusions

The research questions analysed in the paper are relevant for governments in charge of policy for research and innovation to have an empirical understanding about the specialization and disciplinary profile of their country, the relationships between their disciplinary specialization and the rest of the world specialization, to decide on which discipline the R&D policy of their country can best be concentrated, taking into account the comparison of their disciplinary specialization with respect to the other European countries and the main other competitors in the world.
We provided a first illustration of how the newly and advanced indicators on the comparison of disciplinary profiles, proposed by Bongioanni, Ruocco and Daraio (2013, 2014) and based on a physics of complex system approach, could be relevant for science policy.
The next developments of the analysis will include a systematic investigation of the dynamics of disciplinary profiles of best performing countries and a deeper understanding of how productive the disciplinary configuration of each country is with respect to the world reference and to major European and non-European countries.

References

Almeida J.A.S., Pais A.A.C.C. & Formosinho, S.J. (2009). Science indicators and science patterns in Europe. *Journal of Informetrics*, 3, 134-142.

Bongioanni I., Daraio C. & Ruocco G. (2013), Comparing national disciplinary structures: a quantitative approach. Proceedings of ISSI 2013 (eds. J. Gorraiz, E. Schiebel, C. Gumpenberger, M. Horlesberger and H. Moed), Vienna, Vol. 2, 1900-1904. (ISBN: 978-3-200-03135-7).

Bongioanni I., Daraio C. & Ruocco G. (2014), A Quantitative Measure to Compare the Disciplinary Profiles of Research Systems and Their Evolution Over Time, Technical Report n. 10, 2013, DIAG Sapienza university of Rome, *forthcoming in Journal of Informetrics*.

Tian Y., Wen C. and Hong S. (2008). Global scientific production on GIS research by bibliometric analysis from 1997 to 2006. *Journal of Informetrics*, 2, 65-74.

Glanzel W., Debackere K. & Meyer M. (2008). ‘Triad’ or ‘tetrad’? Onglobal changes in a dynamic world. *Scientometrics*, 74, 71-88.

Glanzel W. (2000). Science in Scandinavia: A Bibliometric Approach. *Scientometrics* 48, 121-150.

King D.A. (2004). The scientific impact of nations. *Nature* 430, 311-316.

Zhou Q., Rousseau R., Yang L., Yue T. & Yang G. (2012). A general framework for describing diversity within systems and similarity between systems with applications in informetrics. *Scientometrics*, 93, 787-812.

Appendix.

List of subject categories and countries' Acronyms.

Subject Category	Description
AGRI	Agricultural and Biological Sciences
ARTS	Arts and Humanities
BIOC	Biochemistry, Genetics and Molecular Biology
BUSI	Business, Management and Accounting
CENG	Chemical Engineering
CHEM	Chemistry
COMP	Computer Science
DECI	Decision Sciences
EART	Earth and Planetary Sciences
ECON	Economics, Econometrics and Finance
ENER	Energy
ENGI	Engineering
ENVI	Environmental Science
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
DENT	Dentistry
HEAL	Health Professions
GENE	General

Code	Country
ARG	Argentina
AUS	Australia
AUT	Austria
BEL	Belgium
BGR	Bulgaria
BRA	Brazil
CAN	Canada
CHL	Chile
CHN	China
CYP	Cyprus
CZE	Czech Republic
DEU	Deutschland
DNK	Denmark
ESP	Spain
EST	Estonia
FIN	Finland
FRA	France
GBR	United Kingdom
GRC	Greece
HUN	Hungary
IDN	Indonesia
IND	India
IRL	Ireland
ITA	Italy
JPN	Japan
KOR	Korea
LTU	Lithuania
LUX	Luxembourg
LVA	Latvia
MEX	Mexico
MLT	Malta
NLD	The Netherlands
POL	Poland
PRT	Portugal
ROU	Romania
RUS	Russia
SVK	Slovakia
SVN	Slovenia
SWE	Sweden
TUR	Turkey
USA	United States

Appendix: Overlap values between main countries (42 obs). Indicator PUBf.

	ARG	AUS	AUT	BEL	BGR	BRA	CAN	CHL	CHN	CYP	CZE	DEU	DNK	ESP	EST	FIN	FRA	GBR	GRC	HUN	IDN	IND	IRL	ITA	JPN	KOR	LTU	LUX	LVA	MEX	MLT	NLD	POL	PRT	ROU	RUS	SVK	SVN	SWE	TUR	TWN
AUS	0,808																																								
AUT	0,774	0,934																																							
BEL	0,814	0,949	0,991																																						
BGR	0,789	0,701	0,836	0,844																																					
BRA	0,927	0,912	0,902	0,928	0,799																																				
CAN	0,785	0,973	0,961	0,970	0,773	0,893																																			
CHL	0,936	0,928	0,867	0,894	0,772	0,944	0,891																																		
CHN	0,390	0,518	0,626	0,610	0,723	0,494	0,659	0,468																																	
CYP	0,273	0,599	0,599	0,582	0,512	0,416	0,665	0,489	0,758																																
CZE	0,876	0,873	0,939	0,946	0,939	0,926	0,910	0,912	0,701	0,578																															
DEU	0,793	0,877	0,974	0,973	0,923	0,882	0,927	0,849	0,689	0,576	0,963																														
DNK	0,854	0,954	0,955	0,968	0,752	0,935	0,942	0,904	0,439	0,441	0,889	0,910																													
ESP	0,858	0,941	0,976	0,983	0,857	0,944	0,952	0,928	0,589	0,565	0,968	0,956	0,961																												
EST	0,806	0,765	0,711	0,738	0,763	0,748	0,794	0,843	0,694	0,612	0,831	0,754	0,690	0,761																											
FIN	0,786	0,963	0,971	0,972	0,796	0,904	0,989	0,883	0,680	0,669	0,926	0,939	0,939	0,959	0,796																										
FRA	0,801	0,899	0,982	0,981	0,910	0,892	0,945	0,871	0,693	0,615	0,967	0,995	0,920	0,967	0,768	0,953																									
GBR	0,751	0,971	0,958	0,972	0,742	0,876	0,966	0,875	0,529	0,594	0,870	0,921	0,958	0,944	0,695	0,950	0,935																								
GRC	0,718	0,931	0,979	0,964	0,794	0,872	0,965	0,845	0,708	0,696	0,920	0,938	0,912	0,955	0,733	0,980	0,954	0,935																							
HUN	0,875	0,847	0,918	0,931	0,927	0,900	0,895	0,890	0,636	0,561	0,972	0,944	0,888	0,955	0,791	0,897	0,952	0,861	0,882																						
IDN	0,637	0,764	0,690	0,699	0,614	0,719	0,792	0,734	0,782	0,761	0,748	0,663	0,635	0,709	0,846	0,814	0,697	0,668	0,785	0,676																					
IND	0,796	0,721	0,806	0,823	0,940	0,814	0,792	0,777	0,782	0,527	0,931	0,872	0,739	0,854	0,819	0,816	0,864	0,721	0,808	0,906	0,736																				
IRL	0,781	0,965	0,974	0,982	0,805	0,911	0,983	0,891	0,656	0,691	0,930	0,941	0,941	0,968	0,762	0,985	0,960	0,968	0,975	0,910	0,788	0,804																			
ITA	0,792	0,918	0,992	0,988	0,845	0,903	0,951	0,868	0,608	0,547	0,936	0,979	0,957	0,971	0,704	0,955	0,984	0,953	0,961	0,927	0,654	0,809	0,958																		
JPN	0,755	0,805	0,926	0,930	0,940	0,841	0,888	0,774	0,770	0,556	0,942	0,977	0,848	0,906	0,740	0,900	0,965	0,858	0,898	0,923	0,670	0,915	0,897	0,935																	
KOR	0,562	0,631	0,774	0,767	0,881	0,654	0,761	0,592	0,940	0,694	0,838	0,848	0,619	0,743	0,719	0,788	0,841	0,672	0,799	0,798	0,729	0,887	0,777	0,763	0,917																
LTU	0,508	0,547	0,642	0,645	0,812	0,561	0,643	0,581	0,898	0,722	0,762	0,739	0,475	0,625	0,776	0,668	0,737	0,565	0,676	0,688	0,726	0,797	0,665	0,631	0,782	0,894															
LUX	0,497	0,727	0,805	0,764	0,697	0,636	0,792	0,635	0,766	0,858	0,780	0,764	0,658	0,766	0,662	0,829	0,794	0,710	0,857	0,749	0,779	0,707	0,829	0,749	0,743	0,803	0,705														
LVA	0,419	0,437	0,559	0,554	0,770	0,477	0,573	0,455	0,958	0,719	0,695	0,660	0,370	0,543	0,715	0,605	0,657	0,448	0,618	0,634	0,736	0,789	0,593	0,542	0,745	0,928	0,946	0,718													
MEX	0,892	0,861	0,886	0,903	0,896	0,927	0,887	0,919	0,708	0,569	0,962	0,917	0,847	0,912	0,895	0,908	0,924	0,829	0,879	0,915	0,817	0,901	0,896	0,886	0,899	0,814	0,796	0,725	0,722												
MLT	0,510	0,900	0,858	0,853	0,533	0,725	0,884	0,732	0,527	0,713	0,709	0,766	0,816	0,815	0,563	0,873	0,798	0,916	0,895	0,673	0,692	0,550	0,898	0,826	0,694	0,581	0,487	0,751	0,397	0,676											
NLD	0,759	0,961	0,971	0,976	0,730	0,891	0,957	0,863	0,490	0,524	0,869	0,923	0,977	0,952	0,647	0,950	0,932	0,986	0,942	0,857	0,630	0,707	0,957	0,968	0,855	0,647	0,506	0,699	0,397	0,820	0,891										
POL	0,809	0,796	0,910	0,913	0,975	0,855	0,859	0,838	0,744	0,566	0,975	0,969	0,826	0,920	0,806	0,880	0,962	0,827	0,883	0,950	0,677	0,940	0,878	0,918	0,965	0,883	0,820	0,745	0,752	0,940	0,649	0,821									
PRT	0,731	0,771	0,841	0,839	0,894	0,784	0,857	0,779	0,877	0,749	0,928	0,876	0,734	0,858	0,863	0,882	0,886	0,751	0,876	0,895	0,852	0,932	0,866	0,820	0,897	0,934	0,861	0,875	0,867	0,909	0,654	0,730	0,919								
ROU	0,401	0,484	0,625	0,612	0,802	0,489	0,609	0,484	0,941	0,752	0,734	0,715	0,418	0,606	0,677	0,639	0,711	0,519	0,674	0,672	0,672	0,801	0,639	0,602	0,774	0,933	0,949	0,768	0,963	0,716	0,482	0,471	0,798	0,876							
RUS	0,492	0,291	0,457	0,463	0,817	0,405	0,405	0,438	0,703	0,401	0,646	0,631	0,313	0,468	0,655	0,426	0,605	0,349	0,424	0,624	0,408	0,732	0,424	0,485	0,689	0,765	0,843	0,434	0,822	0,658	0,136	0,302	0,760	0,693	0,804						
SVK	0,870	0,796	0,867	0,886	0,956	0,871	0,854	0,859	0,731	0,543	0,975	0,923	0,820	0,908	0,866	0,871	0,921	0,792	0,844	0,952	0,735	0,965	0,865	0,867	0,933	0,867	0,797	0,742	0,751	0,945	0,594	0,780	0,966	0,941	0,769	0,728					
SVN	0,631	0,741	0,801	0,808	0,864	0,716	0,829	0,718	0,913	0,800	0,878	0,847	0,673	0,804	0,826	0,839	0,857	0,748	0,842	0,838	0,830	0,905	0,843	0,778	0,883	0,947	0,912	0,849	0,899	0,859	0,691	0,697	0,885	0,964	0,924	0,701	0,894				
SWE	0,807	0,952	0,982	0,990	0,806	0,917	0,968	0,876	0,564	0,523	0,916	0,958	0,982	0,968	0,712	0,967	0,962	0,976	0,950	0,903	0,668	0,788	0,966	0,981	0,915	0,730	0,585	0,718	0,493	0,873	0,846	0,985	0,880	0,798	0,548	0,408	0,855	0,761			
TUR	0,747	0,932	0,946	0,949	0,722	0,909	0,911	0,862	0,485	0,465	0,868	0,894	0,942	0,941	0,614	0,915	0,899	0,944	0,929	0,823	0,622	0,725	0,922	0,938	0,830	0,626	0,509	0,654	0,395	0,824	0,873	0,966	0,818	0,712	0,485	0,283	0,765	0,688	0,950		
TWN	0,459	0,658	0,761	0,744	0,762	0,614	0,778	0,567	0,971	0,807	0,782	0,793	0,589	0,714	0,699	0,804	0,800	0,677	0,833	0,719	0,816	0,804	0,788	0,738	0,846	0,956	0,879	0,851	0,913	0,773	0,683	0,649	0,807	0,911	0,918	0,632	0,777	0,937	0,704	0,637	
USA	0,754	0,957	0,968	0,979	0,784	0,876	0,986	0,857	0,631	0,630	0,894	0,945	0,950	0,944	0,733	0,972	0,956	0,984	0,955	0,886	0,721	0,772	0,978	0,967	0,908	0,759	0,631	0,752	0,547	0,860	0,888	0,973	0,862	0,816	0,593	0,418	0,835	0,805	0,982	0,919	0,761

Atypical combinations are confounded by disciplinary effects

Kevin W. Boyack* and Richard Klavans**

[*] *kboyack@mapofscience.com*
SciTech Strategies, Inc., Albuquerque, NM, 87122 (USA)

[**] *rklavans@mapofscience.com*
SciTech Strategies, Inc., Berwyn, PA, 19312 (USA)

Abstract
Uzzi et al. (2013) recently argued that the highest impact articles are likely to reference novel combinations of existing knowledge while still building upon typical combinations. In this study we replicate this intriguing finding using slightly different methods. We also show, however, that the findings are not free from disciplinary effects. For example, physics builds primarily on typical combinations, while multidisciplinary journals participate much more often in atypical combinations. We strongly suspect that atypical co-cited journal combinations, and thus citation rates, are highly dependent on discipline and journal effects.

Introduction

Two new indicators for innovative high impact papers were recently introduced in a *Science* article by Uzzi et al. (2013), hereafter referred to as UMSJ. The authors used co-cited journal-journal relationships to determine whether any pair of cited references is typical or atypical. Using cited references from nearly 18 million articles, they calculated actual and expected counts for each co-cited journal pair, and converted those counts into Z-scores. Negative Z-scores indicate that actual counts are less than expected, and reflect atypical knowledge relationships. Positive Z-scores indicate the opposite – typical knowledge relationships. The authors show that articles that have higher than average typical relationships (using the median Z-score) combined with a high level of atypical relationships (using the left 10th percentile Z-scores) are twice as likely to be highly cited as the average article.

The UMSJ study was designed to test the premise that innovation is often based on original or novel combinations of existing knowledge (Chen et al., 2009; Guimera, Uzzi, Spiro, & Amaral, 2005), while at the same time being strongly based in an existing and well-established paradigm that is robust enough to incorporate new knowledge.

The purpose of this study is to replicate the UMSJ study using a slightly different technique, and to further explore the relationship between novelty (building on atypical knowledge relationships), convention (building on typical knowledge relationships) and citation rates. This paper proceeds as follows. First, we provide detail about the differences between the UMSJ method and our method, and show our replication of their primary results and findings. This is followed a preliminary analysis of disciplinary effects. The paper concludes with a discussion of possible effects from journal impact which may negate their central findings.

Replication

UMSJ calculated Z-scores as $Z = (N_{actual} - N_{expected}) / N_{variance}$ for pairs of co-cited journals where N are journal co-citation counts. Their calculations were based on 17.9 million research articles (1950-2000) from the Web of Science (WoS), and the 302 million references (edges) from these articles to 15,613 cited journals. This formulation gives a negative Z-score to any journal pair where the actual counts are less than the expected counts. Ten Monte Carlo simulations were run that reassigned edges in a random way, while preserving temporal and distributional characteristics of the original citation network at the paper level. Expected co-citation count values and variances for co-cited journal pairs were calculated from the results of these Monte Carlo simulations.

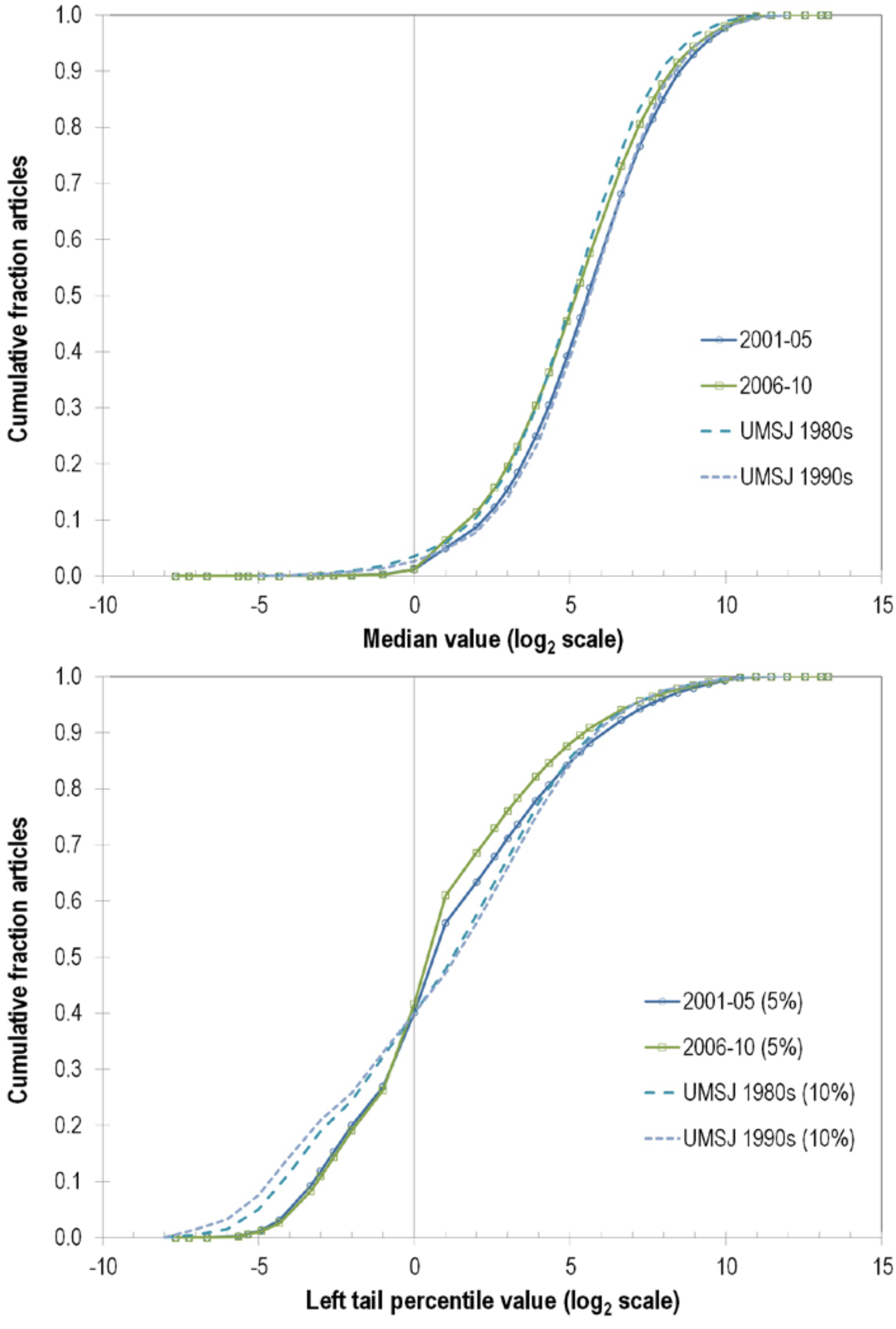
Using these Z-scores, UMSJ then calculated 10^{th} percentile (left tail) and median Z-scores for each article after ordering the Z scores corresponding to their co-cited journal pairs from lowest to highest. The resulting cumulative probability distributions showed that half of WoS articles had a median Z-score greater than 64, while 41% of those articles had a 10^{th} percentile Z-score that was negative. These two statistics were used as the basis for two indicators. The median Z-score for an article was used to signal conventionality; articles with a median score of greater than the overall median were designated as "high convention". The 10^{th} percentile Z-score was used to signal novelty; articles with a negative 10^{th} percentile Z-score were designated as "high novelty". Upon testing the top 5% highly cited articles (by year), UMSJ found that articles with high convention and high novelty are twice as likely to be highly cited as the average article. Although UMSJ also tested different definitions of novelty (e.g., 1%, 10%) and explored the effect of authorship structure on their results, those additional experiments did not change the overall results. Thus, our study focuses on replicating the primary typical vs. atypical distributions and indicators of convention and novelty that are the basis for the findings of the UMSJ study.

The methodology we used to replicate the UMSJ study differs from theirs in several respects. First, we used Scopus data rather than WoS data. Second, we used a more current ten year dataset (2001-2010) rather than the historical 50 year dataset (1950-2000) used by UMSJ. Our dataset is thus smaller than the one used by UMSJ (12.0M articles + 226M references vs. 17.9M articles + 302M references), but certainly still large enough to provide for valid results. The difference in time window is not expected to be an issue since UMSJ showed results that were comparable for multiple time periods. Third, while UMSJ used articles only, we used articles and conference papers. Scopus indexes much more conference material than does WoS, and since articles and conference papers are both aimed primarily at reporting original research we felt justified in including both document types. Finally, we used a different formulation to calculate typical and atypical relationships. Rather than using Z-scores and Monte Carlo simulations, we calculated K50 statistics for co-cited journal pairs (Klavans & Boyack, 2006). K50 has the same general formulation as the UMSJ Z-scores, $(N_{actual} - N_{expected}) / normalization$. The difference is that the expected and normalization values for K50 are calculated using the row and column sums from the square co-citation count matrix rather than using a Monte Carlo technique. This difference leads to a savings in computation – calculating row and column sums is much less expensive computationally than using multiple Monte Carlo runs. Our K50 distributions are very similar to Uzzi's Z-score distributions, thus suggesting that the additional computation required by multiple Monte Carlo calculations may be unnecessary.

Distributions

Figure 1 compares the distributional characteristics of median and left tail percentile statistics from our study with those of UMSJ. Z-score curves were obtained by transcribing data from Figures 1B, C of Uzzi et al. (2013). Our K50 values have been scaled (multiplied by 10^4) to fall within the same range as the UMSJ Z-scores. Figure 1a shows that while the fraction of papers with negative median K50 values is lower than the UMSJ values, the K50 curves fall between the two UMSJ curves over most of the range. Thus, use of median statistics to designate articles as "high convention" should work similarly with K50 values as it does for the UMSJ Z-scores. For the left tail values, we found that only 30% of articles had a 10th percentile K50 value that was negative, while 40% of articles had a 5th percentile K50 value that was negative. Figure 1b compares K50 values at the 5th percentile with UMSJ 10th percentile values, and shows that the K50 curves are very similar to the UMSJ Z-score curves. Thus, our use of 5th percentile K50 statistics to designate articles as "high novelty" should perform similarly to the UMSJ 10th percentile Z-scores.

Figure 1. Comparison of median and left tail distributions from K50 statistics with the same distributions based on UMSJ Z-scores.



The K50 distributions are remarkably similar to the UMSJ distributions given that we used a different database, a different metric, and included conference papers along with articles in our calculations. Based on this similarity between distributions, both in the principles behind their calculation and in practice, replication of additional results from UMSJ using K50 statistics is justified.

Indicators

UMSJ proposed a method for identifying "hit" papers using the principles of novelty and conventionality based on Z-scores and their distributions. To test this method, a 2x2 categorization based on median and 10th percentile Z-scores was used to classify the top 5% highly cited papers (citation counts as of 8 years after publication). We followed the same procedure with some differences. We computed all citation counts to papers as of 2011; thus papers published in 2001 had a ten year citation window while papers published in 2005 had only a 6 year window in which to accrue citations. Also, we used the 5th percentile (rather than the 10th percentile) K50 score as the basis for distinguishing between high novelty and low novelty.

As with UMSJ, our analysis was limited to the top 5% highly cited articles by year. Despite the differences in our test samples, we get 2x2 matrix probabilities that are similar to UMSJ (see Table 1). The differentiation between our high/high (N+C+) and low/low (N-C-) pairs is even higher than that obtained by UMSJ. In addition, the fraction of articles that end up in the N+C+ bin is slightly higher using our method (9.5% vs. 6.7%), suggesting that our calculations can identify even more highly cited papers than can the UMSJ method. Note that the N+C- bin also has a probability of greater than 5% (0.0659), which suggests that novelty plays a greater role than conventionality in the formulation of a "hit" or highly cited article.

Table 1. Probabilities of "hit" papers (top 5% highly cited).

	UMSJ (1990-2000)		**This study (2001-2005)**	
	% sample	**Prob**	**% sample**	**Prob**
High Novelty, High Convention (N+C+)	6.7%	0.0911	9.5%	0.0959
High Novelty, Low Convention (N+C-)	26%	0.0533	30.6%	0.0659
Low Novelty, High Convention (N-C+)	44%	0.0582	40.5%	0.0433
Low Novelty, Low Convention (N-C-)	23%	0.0205	19.4%	0.0205

In summary, we have replicated the distributions and hit paper probabilities introduced in Uzzi et al. (2013) to a high degree, despite differences in methodology. This replication suggests that our process is sufficiently accurate to be used to more deeply explore the relationships between novelty, convention, and citation rates.

Disciplinary effects

As mentioned above, UMSJ tested multiple definitions of novelty and explored the effect of authorship structure on their results. They also explored the effect of disciplines on their results by examining central tendencies for median and 10th percentile statistics by WoS subject category. They looked at the relationships between novelty, convention, and hit papers for each category, and found that the overall relationships generally held true. However, their detailed results showed that the N+C+ bin in the 2x2 matrix had the highest probability of containing a hit paper for only 64.4% of 243 WoS subject categories. Although this is consistent with the main result on the whole, the fact that this number is not close to 100%

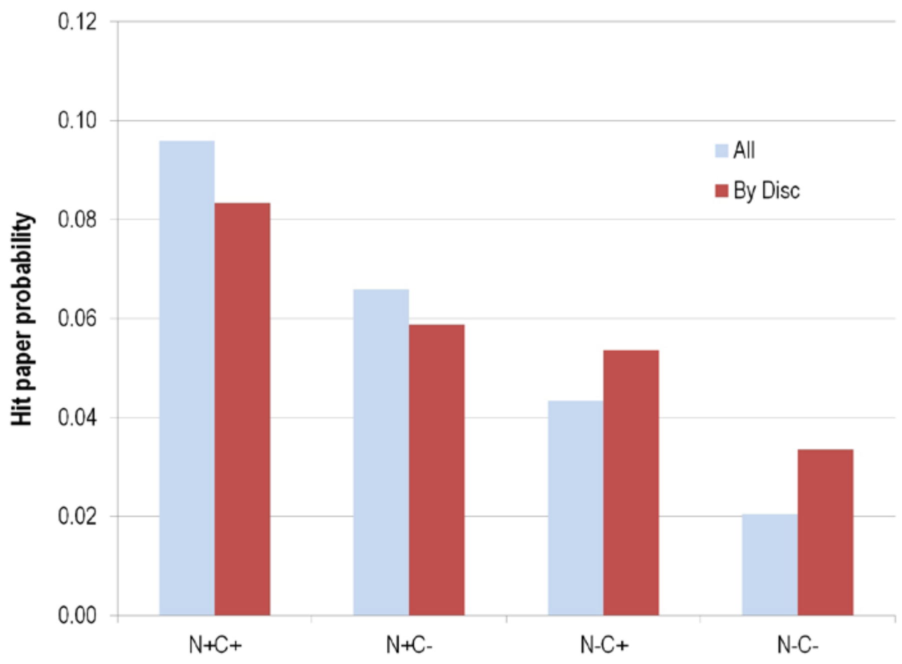
suggests that their method is not free from disciplinary effects. It is also well known that impact by discipline is nonlinearly related to size (Katz, 1999, 2000). Thus, we felt it prudent to more deeply explore potential disciplinary effects on the indicators proposed by UMSJ.

Discipline-based sampling

The first, and simplest, test was to calculate 2x2 matrix probabilities using the top 5% highly cited articles where the top 5% was sampled by discipline rather than over the entire sample. We expected different results because the top 5% sample over all disciplines used by UMSJ is naturally enriched in papers from disciplines with high citation rates (e.g., biochemistry, physics) and depleted in papers from disciplines with lower citation rates (e.g., social sciences, engineering). Sampling by discipline will introduce papers with smaller numbers of citations from these lower cited disciplines into our sample at the expense of more highly cited papers from highly cited disciplines.

We took the top 5% of highly cited papers by discipline using the article-based (as opposed to journal-based) discipline-level structure introduced in Boyack and Klavans (2014) and calculated 2x2 matrix probabilities. Figure 2 shows that while discipline-based sampling preserves the probability ordering of bins (i.e., N+C+ highest, N-C- lowest), the separation between the highest and lowest probabilities is much less than for the non-discipline based case. This degradation suggests that the higher probability associated with the non-discipline based case is due to the enrichment of that sample with articles from highly cited disciplines, and is evidence of a larger disciplinary effect than is acknowledged by Uzzi et al. (2013). This does not detract from the fact that, even when disciplines are considered, the combination of typical and atypical combinations associated with these indicators leads to a higher than average incidence of highly cited papers. However, when disciplines are considered the effect is less prominent.

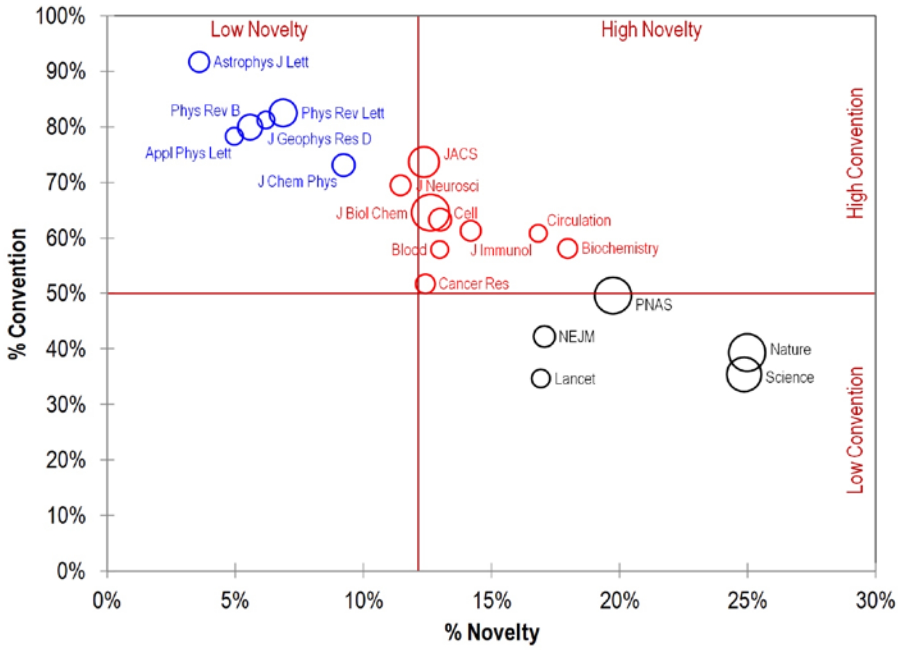
Figure 2. Effect of sampling the top 5% highly cited papers by discipline on probabilities of hit papers based on novelty and conventionality indicators.



Top 20 knowledge areas
Since UMSJ used journals as surrogates for knowledge areas, we also used journals as the base unit of analysis in our replications of their results. As with disciplines, journals vary widely in size and influence. Thus, we decided to take a closer look at those journals that contributed most to the system of knowledge interactions.

A total of 58,020 separate Scopus journal identifiers were cited by the 12 million articles in our dataset. Although this number seems much larger than the 15,613 journals analyzed by UMSJ, the signal is highly concentrated in a much smaller number of journals. The top 300 journals account for half of the total number of co-citations in the system, while the top 15,600 journals account for 99.6% of the total number of co-citations. Thus, the existence of a long tail in our data has almost no effect on the overall system. We limited our analysis to the top 20 journals, which participated in 15.9% of the co-citations in the system. Four of these journals (*J Biol Chem*, *Nature*, *Science*, and *PNAS*) each participated in more than 1.5% of the total co-citations.

Figure 3. Top 20 co-cited journals plotted as a function of novelty and convention. Circle sizes reflect numbers of co-citations.



Percentages of novel and conventional K50s were calculated for each journal, where %Novel is the fraction of negative K50s, and %Convention is the fraction of K50s above the median (0.00421226) for the entire system. Figure 3 shows these 20 journals, each plotted as a function of their %Novel and %Convention values. The figure has been divided into four quadrants that correspond to the four groupings in the 2x2 matrix mentioned earlier. The dividing line for novelty is at 12.16%, which is the fraction of all co-citations across the system with negative K50 values. Among these 20 journals, three groups can be easily distinguished. Six journals, all of which are highly related to physics, are closely grouped in the low novelty, high convention quadrant (upper left). Nine journals, all of which are related to biochemistry or medicine, are grouped in or very near the high novelty, high convention quadrant (upper right). The remaining five journals are all in the high novelty, low convention quadrant. Three of these journals are clearly multidisciplinary while the other two (*NEJM*, *Lancet*) are broad medical journals, and thus more multidisciplinary than other medical journals. The type of knowledge relationships associated with these prominent journals clearly varies by discipline. Physics is highly associated with typical relationships. Biochemistry and medicine are associated with the pair of relationships promoted by UMSJ – a combination of typical and atypical relationships. Multidisciplinary journals are more highly associated with atypical knowledge relationships.

We note that this analysis accounts for only 15.9% of the co-citations in the system, and only applies to a few of the top cited disciplines in science. A detailed investigation of the rest of the system may show different effects. Nevertheless, the fact that a large disciplinary effect is seen in the top few journals (which comprise a significant fraction of the overall signal representing typical and atypical relationships) suggests that discipline may be a significant confounding effect as regards these relationships.

Summary

We have replicated the distributions and hit paper probabilities from UMSJ using a slightly different methodology. This replication allows us to proceed to more deeply explore how the notions of novelty and convention might be measured using citation data and our metrics.

The analysis of disciplinary effects above is preliminary; a much more detailed analysis is needed. In addition, the fact that three high impact multidisciplinary journals (*Nature*, *Science*, *PNAS*) account for 9.4% of all of atypical combinations (negative K50 values) suggests that there may be significant journal-level effects as well.

The idea that measurement of novelty might lead to a paper-level indicator of impact type has been intriguing to us for some time (Klavans & Boyack, 2013). While we point out some potential problems with specifics of the UMSJ study, we believe that their underlying logic – that of creating an indicator based on the notion of novelty and distribution tails – is sound. What remains is to identify and test other potential measurements of novelty that are relatively independent of discipline and journal effects.

References

Boyack, K. W., & Klavans, R. (2014). Creation of a highly detailed, dynamic, global model and map of science. *Journal of the Association for Information Science and Technology, DOI: 10.1002/asi.22990*.

Chen, C., Chen, Y., Horowitz, M., Hou, H., Liu, Z., & Pelligrino, D. (2009). Towards an explanatory and computational theory of scientific discovery. *Journal of Informetrics, 3*, 191-209.

Guimera, R., Uzzi, B., Spiro, J., & Amaral, L. A. N. (2005). Team assembly mechanisms determine collaboration network structure and team performance. *Science, 308*, 697-702.

Katz, J. S. (1999). The self-similar science system. *Research Policy, 28*, 501-517.

Katz, J. S. (2000). Scale-independent indicators and research evaluation. *Science and Public Policy, 27*(1), 23-36.

Klavans, R., & Boyack, K. W. (2006). Identifying a better measure of relatedness for mapping science. *Journal of the American Society for Information Science and Technology, 57*(2), 251-263.

Klavans, R., & Boyack, K. W. (2013). *Towards the development of an article-level indicator of conformity, innovation and deviation*. Paper presented at the 18th International Conference on Science and Technology Indicators.

Uzzi, B., Mukherjee, S., Stringer, M., & Jones, B. (2013). Atypical combinations and scientific impact. Science, *342*, 468-472.

Normalized Citation Indexes: a theoretical methodological study applied to science

Maria Cláudia Cabrini Grácio* and Ely Francina Tannuri de Oliveira**

*cabrini@marilia.unesp.br
** etannuri@gmail.com
UNESP – Univ Estadual Paulista, 737 Hygino Muzzi Filho Avenue, 17525-900 Marília (Brazil)

Introduction

Absolute citation indicators have limitations when used to compare the impact across areas due to their specificities as the disciplines differ in the practice of citation.

Persson, Glanzel & Danell (2004) suggest the use of normalized indicators which make it possible to eliminate the dependence on the context of the area, since they standardize the measure units (Li et al. 2013). In this context, Vinkler (2012) stresses that it appears to be acceptable to apply relative scientometric indicators to comparative evaluations and that the normalisation processes of the impact indicators have been widely applied in scientometrics for a long time. This author points out the existence of various type of relative indices depending on the standard. Among them, the RCR-type indices (Schubert & Braun, 1986), which use the impact data of the publishing journals and the "crown" index (Van Raan, 2004) and RW-index (Vinkler, 1986) which use the impact data of the corresponding field.

Among the procedures, we highlight the normalization by mean area (*Ma*) and median (*Md*) (Moed, 2009; Li et al, 2013.). Another procedure may be obtained from the average of the 10% most productive ($Ma_{10\%}$), an adaptation of Moed (2010), in which the author refers to the 10% most cited.

A normalized indicator is calculated by:

$$IN_j = \frac{I_j}{PN_g}$$

where: IN_j = normalized index for the individual j;
I_j = absolute indicator value for the individual j;
PN_g = normalization parameter - *Ma*, *Md* or $Ma_{10\%}$.

Values below 1 mean that the individual is below the overall trend in the field and above 1 suggest that the performance is above the reference behavior (*Ma*, *Md ou* $Ma_{10\%}$).

This investigation aims to perform a theoretical methodological study of the contribution of normalized citation indexes to visualize the impact of science, from the Brazilian presence perspective in 27 areas of knowledge, presented by SCImago Journal & Country Rank for published documents in 1996-2007.

More specifically, we analyze and correlate the results of applying the three presented procedures for the normalization of the citation index per document and determine the linear regression model of the indexes IN_{md} and $IN_{ma_{10\%}}$ expressed in function of IN_{ma} in order to predict the behavior of the first two.

Methodological procedures

SCImago JR allowed data retrieval, for each area, regarding the total number of documents published during 1996-2007 and the average citations received by these documents until 2012 by producing countries.

For each area, we calculated *Ma*, *Md e* $Ma_{10\%}$ for the number of citations per document. Then, we calculated the normalized index of Brazil by IN_{ma}, IN_{md} and $IN_{ma_{10\%}}$. Next, we calculated Pearson correlations between the normalized indexes by the three procedures.

Finally, we determined the regression equation of IN_{md} and $IN_{ma_{10\%}}$ in function of IN_{ma}.

Presentation and analysis of data

Table 1 shows the normalized citation indexes in order by IN_{ma}.

Table 1. Normalized Citation Indexes per document.

Area	IN_{ma}	IN_{md}	$IN_{ma10\%}$
Nursing	1.8	2.5	0.8
Econ._Econometr&Fin	1.8	1.9	1.0
Dentistry	1.4	1.6	1.0
Business._Manag&Acc	1.4	1.9	0.8
Environmental_Sci	1.3	1.4	1.0
Psychology	1.3	1.6	0.8
Decision_Sciences	1.3	1.4	0.9
Mathematics	1.3	1.4	0.8
Pharmacology_Toxicology&Pharmaceutics	1.2	1.2	0.8
Chemistry	1.1	1.3	0.8
Health_Professions	1.1	1.3	0.7
Physics_&Astronomy	1.1	1.3	0.7
Chemical_Engineering	1.1	1.3	0.9
Materials_Science	1.1	1.1	0.8
Social_Sciences	1.1	1.2	0.7
Immun.&Microbiol	1.0	1.0	0.7
Biochem._Gen&Mol_Biol	1.0	1.1	0.6
Earth&Planet_Sci	1.0	1.2	0.8
Engineering	1.0	1.3	0.8
Arts_&Humanities	1.0	1.5	0.7
Neuroscience	0.9	0.9	0.5
Medicine	0.9	0.9	0.7
Agricultural_&Bio_Scie	0.9	0.9	0.7
Computer_Science	0.9	1.2	0.7
Energy	0.9	1.0	0.7
Veterinary	0.9	0.9	0.7
Multidisciplinary	0.8	1.0	0.4

From the IN_{ma} indexes, 7 areas presented value lower than 1. On the other hand, 15 areas show a value higher than 1, meaning that the performance is above the average compared with the producer group.

As for IN_{md}, 4 areas had values below 1 and 20 had values above 1, indicating that the majority of areas is above the median behavior.

These results corroborate the data presented by Faria et al (2011), who point out that in this period, in most areas, there was a growth in citations when compared to world performance.

For $IN_{ma_{10\%}}$, it was observed that no area showed a value above 1 and three of them showed values equal to 1.

It was observed that the highest correlation (0.92) was between IN_{ma} and IN_{md}, showing that these indexes tend to exhibit similar behavior. The correlation between $IN_{ma_{10\%}}$ and the other two indexes have moderate intensity values (0.70 with $IN_{ma_{10\%}}$ and 0.51 with IN_{md}).

The two equations of linear regression are presented in Figures 1 and 2.

Figure 1: Linear regression of IN_{md} in function of IN_{ma}.

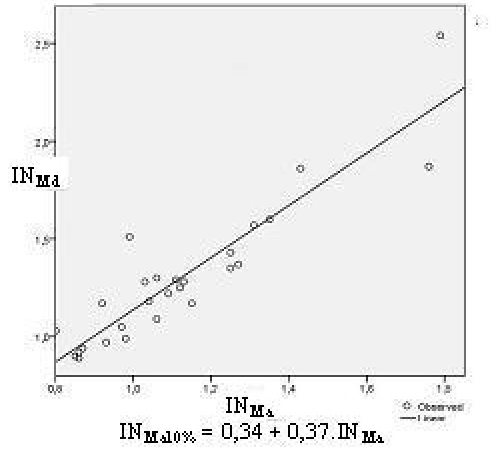
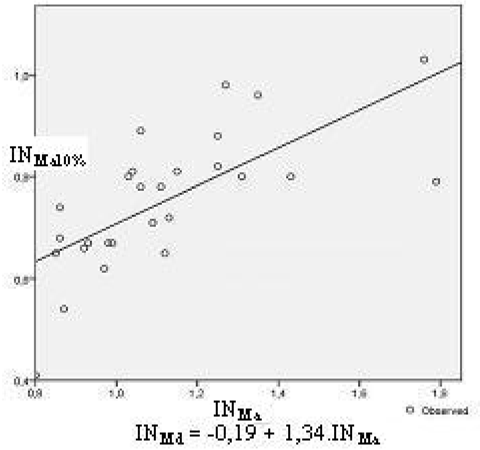


Figure 2: Linear regression of $IN_{ma_{10\%}}$ in function of IN_{ma}.



In Figure 1, out of the 27 areas, in 12 of them the distance between the estimates of IN_{md} in relation to the observed values tended to zero; 3 areas had more significant distance, between 0.3 and 0.4. The remaining areas agglutinated around the line with few significant differences.

For Figure 2, three areas showed a more significant distance, around 0.3. In one area, the distance was very close to zero and the others were evenly scattered around the line.

Final considerations

The model of IN_{md} in function of IN_{ma} presented a better adjustment compared with the model of $IN_{ma_{10\%}}$ in function of the same variable, pointing that IN_{md} and $IN_{ma_{10\%}}$ tend to present a closer behavior, with IN_{md} values slightly higher than those of IN_{ma} at all times. On the other hand, $IN_{ma_{10\%}}$ can be considered a complementary index for explaining the impact of the areas on the scientific community, corroborating Vinkler (2012) observation that the impact of scientific information may not be represented by one single index, given its multifaceted nature.

References

Faria, L. I. L et al. (2011). Análise da produção científica a partir de publicações em periódicos especializados. In: Brentani, R.R.; Brito Cruz, C.H. (Eds.). Indicadores de Ciência, Tecnologia e Inovação em São Paulo 2010. São Paulo: FAPESP, 2011.

Li, Y. et al. (2013). Quantitative evaluation of alternative field normalization procedures. In *Proceedings of* 14th ISSI, 1431-1441.

Moed, H.F. (2009). New developments in the use of citation analysis in research evaluation. *Scientometrics*, 57, 13–18.

Moed, H.F. (2010). *Citation analysis in research evaluation*.Netherlands: Springer.

Persson, O., Glänzel, W. & Danell, R. (2004). Inflationary bibliometric values: The role of scientific collaboration and the need for relative indicators in evaluative studies. *Scientometrics*, 60, 421–432.

Schubert, A., & Braun, T. (1986). Relative indicators and relational charts for comparative assessment of publication output and citation impact. *Scientometrics*, 9, 281-291.

Van Raan, A.F.J. (2004). Measuring science. Capita selecta of current main issues. In H. F. Moed, W. Glanzel, & U.Shmoch (Eds), *Handbook of quantitative science and techinology research. The use of publication and patent statistics of SET systems* (pp. 19-50) Dordrecht: Kluwer Academic Publishers.

Vinkler, P. (1986). Evaluation of some methods for the relative assessment of scientific publications. *Scientometrics*, 10, 157-177.

Vinkler, P. (2012). The case of scientometricians with the “absolute relative” impact indicator. *Journal of Informetrics*, 6, 254-264.

The Impact of Blue Sky Project Based Funding of Academic Research

Nicolas Carayol* and Marianne Lanoe**

* *nicolas.carayol@u-bordeaux4.fr*
GREThA, Université de Bordeaux – CNRS, Avenue Leon Duguit, F-33608 Pessac Cedex.
Observatoire des Sciences et Techniques, 21 Boulevard Pasteur, 75015 Paris.

** *marianne.lanoe@u-bordeaux4.fr*
GREThA, Université de Bordeaux – CNRS, Avenue Leon Duguit, F-33608 Pessac Cedex.

Introduction

In contrast with Anglo-Saxon system of funding, block grants are still funded by the French government but researchers have as well the possibility to apply to and receive competitive grants from funding agencies to perform research projects. Given that these grants are supported by a public agency and that only a small part of selected researchers enjoy extra-funding, we can ask what the real extra output obtained is. Scientific production is of concern because higher education institutions are subject to evaluation and the researchers' productivity inside the laboratories constitutes one of the attentions.

To our knowledge, only a limited number of papers deals with this topic with the aim to explain the selection process of funding agencies and in evaluating the impact of the receipt of a competitive grant on scientific outputs. P. Stephan underscored the essential role of funding for scientific research (Stephan, 1996, 2012) and recently described the working of the main US-funding institutions (Stephan, 2010). The main evaluation impact studies which have been conducted lead to a low positive impact of the grant on the publication output. However the results slightly differ according to the grant characteristics, the country and the impact evaluation method used. Arora and co-authors assess the impact of Italian National Research Council grants in biotechnology and bio-instrumentation fields on the scientific outputs. Their results show a positive impact of the grant on the number of publications and that the impact is greater for researchers with better past performances (Arora et Al, 1998). Arora and al. (2005) focus on economists who obtain NSF support. They find that NSF grant has a positive impact on scientific output for young PI but obtain poorer results for other researcher cohorts.

The impact assessment of a standard research funding (R01) granted by the NIH on the publication outputs also implies a low increase of the scientific productivity (Jacob and Lengfren, 2011). Azoulay et al. (2011) study impact differences of two US-funding institutions with different grant design on scientific outputs and direction of research of accomplished researchers. They find that scientists who obtain a funding which gives more freedom in the orientation of research, get better output results than the one contingent to a given research project. They also explore the research orientation of both scientists groups by the mean of papers keywords and journal citations and show that the former type diversify more their research through novel research lines. Evaluation impact of grants have also been conducted in some developing countries as in Argentina with Chudnovsky and al.'s study which focus on the allocation of grants by the Fund for the Scientific and Technological Research for various scientific fields. They find a positive impact of the grant on researchers' productivity, with a greater effect for youngest scientists as well (Chudnovsky et al., 2008). Finally the impact of the National Science and Technology Research Fund of Chile has also been evaluated by Benavente and co-authors by means of a quasi-experimental design, and their results pointed out an increase of the publications number but no effect on the quality of papers (Benavente et al., 2012).

Our study

We focus on the impact of two competitive projects-oriented grants awarded by the French National Research Agency (ANR) during the period 2005-2007 on the academic productivity. The agency aim associated with the allocation of such fundings is to promote original and quality ideas projects.

The followed methodological approach is based on an impact evaluation method which combines the Propensity Score Matching method with the differences-in-differences estimator to delete bias sources.

We here specifically wonder whether and if so to what extent the allocation of such an ANR subsidy to principal investigators of less than forty years old affects their scientific output both in volume and taking into account impact?

Our sample is composed of 595 fully informed researchers under 40 (which correspond to 611 individual participations in ANR grants between 2005 and 2007) who received a grant from the ANR between 2005 and 2007. For the needs of our methodology, we then made up a control group which include all the French researchers without ANR funding during this period 2. This group is composed of 9,706 researchers or professors under 40 with personal information.

The two lists (funded and controls) were matched with the OST-ISI-Thomson database of scientific publications that allowed us to retrieve their scientific publications from 1999 to 2009. To avoid homonymy issues, we then carried out with the available information a multi-stages process to match as closely as possible researchers to theirs publications. Because of the lack of information with regard to certain publications allowing us to identify the authors reliably, 27 researchers were dropped from the sample that let us with a final sample of 568 funded researchers. For the same reason, the control group is reduced to 7,339 individuals for each year of funding, that is, 44,034 control units for the three years of funding. Since it was not possible to collect direct citations because the upper bound of the publications date is 2009, we refer to a three-year window of journals Impact Factor to consider the visibility of the publications.

We hypothesize that a grant can lead researchers to access research tools and competitive equipments, and allow interactions with skilled partners that might impact positively the scientific output. In a context where good ideas are scarce and scholars face budget constraints to implement their projects, we may expect that fundings modify the scholars' research agenda. It may leads to an increase in autonomy (especially for young scholars), promote the choice of better agenda (the investigation of more original research lines or ideas, larger or more complex problems), encourage collaboration with more and/or better external partners, which could in addition raise the granted effort.

We investigate several dimensions of the research outputs of grantees that may support these assumptions: the number of papers they publish, the prestige of journals, the type of problem they address, and with what kind of co-authors they collaborate. Our study focuses only on the short term impact on the scientific output, though we are also inclined that these fundings allocation may have even stronger long term impact.

Main results

We find that getting a grant has a positive and significant effect on the scientific productivity, both in terms of quantity and of visibility of publications. We estimate an average of 7% increase in quantity of post grant publications and a 11% increase when quality is also considered.

Moreover the results show that the impact varies across scientific disciplines. We obtain as well different inferences according to the year of funding, but it may be caused by the difference in time windows we selected to measure productivity.

We also find that the grant has an effect on the research design of the grantees. Grantees seem to enlarge their research lines and diversify their research interest as shown by the increase of new keywords associated to the publications and by the rise of publications scattering into different specific disciplines. This can suggest that grantees are more inclined to pick more complex research problems at the crossing of several subdisciplines.

Finally, the funding appears to encourage collaboration with new partners and rather skilled scientists (international authors). We can then presume that these consequences take part in the process of increasing the number of publications and their visibility along with the rise of scholars' autonomy. These results are indicative and have to be interpreted with caution. As our data set is quite recent, the study focuses only on the short term impact of some ANR grants and we suppose that these fundings may have stronger long term impact.

References

Arora, A., David, P., and Gambardella, A. (1998). Reputation and competence in publicly funded science: Estimating the effects on research group productivity. Annales d'´economie et de statistique 49/50.

Arora, A. and Gambardella, A. (2005). The Impact of NSF Support for Basic Research In Economics. Annales d'´economie et de statistique, 79/80.

Stephan, P. E. (2012). How economics shapes science. Cambridge, MA: Harvard University Press.

Azoulay, P., Graff Zivin, J. G., and Manso, G. (2011). Incentives and creativity: Evidence from the academic life science. The Rand Journal of Economics 42, 527–554.

Benavente, J. M., Crespi, G., Figal Garone, L., and Maffioli, A. (2012). The impact of national research funds: A regression discontinuity approach to the Chilean FONDECYT. Research Policy, 41(8), 1461 – 1475.

Jacob, B. A. and Lefgren, L. (2011). The impact of research grant funding on scientific productivity. Journal of Public Economics, 95, 1168–1177.

Large scale author name disambiguation using rule-based scoring and clustering

Emiel Caron and Nees Jan van Eck

caroneam1@cwts.leidenuniv.nl, ecknjpvan@cwts.leidenuniv.nl
Centre for Science and Technology Studies (CWTS), Leiden University,
Wassenaarseweg 62A, Leiden, 2333 AL (The Netherlands)

Introduction

Two common questions in bibliometric analysis are:

- Who wrote a particular publication?
- What is the oeuvre of an author?

The answers to these questions should be based on publication information in large bibliographic databases. Unfortunately, the author identification systems in these databases are not fully developed, which makes them difficult to use in bibliometric analysis. For example, if one would query the name “Ding, Y” in the Web of Science[1] (WoS), one would find almost 9000 publications. Obviously, these publications do not belong to a single author “Ding, Y”. The causes for author ambiguity are the fact that many different authors have the same name (i.e., the homonym problem), and the fact that individual authors sometimes publish under multiple names (i.e., the synonym problem). Moreover, manual author disambiguation in these databases is often not feasible if the oeuvre of thousands of authors is studied in a limited timeframe. Therefore, there is need for automatic methods for author disambiguation (Smalheiser & Torvik, 2009).

In this paper, we propose a general author disambiguation method using rule-based scoring and clustering. The method is capable of disambiguating complete bibliographic databases such as the WoS. The results of this method are useful for: academic performance assessment on the author level, research policy-making, the creation of linkages between bibliographic databases, and so on.

Many different solutions are proposed for the author disambiguation problem. See for a comprehensive overview the work of Smalheiser & Torvik (2009) or Ferreira et al. (2012). One solution to the problem would be the establishment of a registry with unique author identifiers. Thomson Reuter’s ResearcherID[2], Open Researcher and Contributor ID[3] (ORCID), and Authorclaim[4] are examples of registries in which where authors are able to register their papers. For such registries to work most authors would need to participate. For the moment this is not the case, however in the future such registries might solve the problem.

The other solution to the problem of author disambiguation is the application of automatic approaches. Supervised or unsupervised learning approaches can be used for this purpose. In supervised learning approaches, a classifier is trained on a data set with pairs of articles,

[1] http://www.webofscience.com
[2] http://www.researcherid.com
[3] http://www.orcid.org
[4] http://www.authorclaim.org

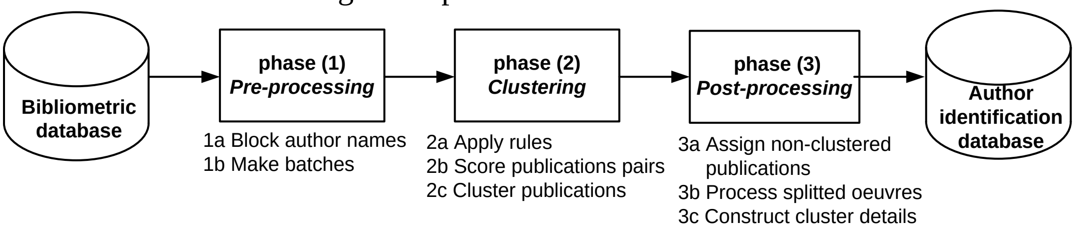
where authors with similar names are classified as being the same person or a different person. Such approaches need a large, manually verified, representative data set for training, which is not easily available. Because we want to disambiguate entire bibliographic databases a supervised approach is not feasible. In unsupervised learning approaches, a similarity metric is defined between pairs of articles and some clustering algorithm is applied (Bolikowski & Dendek, 2011; Levin et al., 2012; Liu et al., 2013; Song et al., 2007). The method that we propose in this paper belongs to the class of unsupervised learning approaches. Inspired by the work of Levin et al. (2012), our method is based on rule-based clustering. An important advantage of our method is, however, that information provided by different rules is combined in a transparent way.

The organization of this paper is as follows. In the next section we explain the phases of our disambiguation method in detail. After that we evaluate the results of disambiguating the WoS database with precision-recall analysis on verified data sets. We close the paper with some concluding remarks.

Methodology

Figure 1 provides a visual summary of the author disambiguation process that is followed by our method. Bibliometric meta data related to authors and their publications is taken as input and clusters of publications likely to be written by the same author is given as output. Our method consists of three main phases: 1. pre-processing, 2. rule based scoring and clustering, and 3. post-processing. The method is developed to disambiguate all authors in the in-house WoS database of the Centre for Science and Technology Studies[5] (CWTS). The total number of publications in the database (version April 2013) is 123,675,056.

Figure 1: Author name disambiguation process.



Pre-processing

Our method starts with the grouping of all author names into blocks (e.g., Levin et al., 2012). These author name blocks are constructed based on the last name and first initial and the removal of all non-alphabetic characters. For example, the author names “Caron, E.” and “Van Eck, N.J.” are respectively assigned to the blocks “carone” and “vaneckn”. The advantage of blocking author names is that the number of pairwise comparisons between publications is greatly reduced and therefore the computational cost. Subsequently, the author name blocks are divided into block size classes 1-6, based on the number of publications within a block. The size class of the block says something about how difficult it is to disambiguate a certain author name. In the scoring mechanism of phase 2, the block size classes are used to adapt the amount of information that is needed to conclude that two publications belong to the same author.

Rule-based scoring and clustering

In this phase, we first detect pairs of publications within blocks that are likely to be written by the same author based on a set of scoring rules. The underlying idea of our scoring rule

[5] http://www.cwts.nl

system is as follows. The higher the number of shared bibliographic elements between two publications, the higher the amount of evidence that these publications are written by the same author, and therefore the higher the score of such a publication pair. The scoring rules that we use are based on four categories of bibliographic meta data (see Table 1): author (rules 1-4), article (rules 5-7), source (rule 8), and citation (rules 9-11).
Rule 1 is defined as pairs of publications with email addresses that match exactly within a block. When two publications relate to the same email address, it is obviously a very strong indicator that the publications are written by the same author. Rules 2a and 2b are defined as pairs of publications with two or three matching author initials in a block, respectively. In general, the more elements are shared between a pair of publications the stronger a rule is. Therefore, rule 2b is stronger than rule 2a and therefore more points are given to it. Rule 2c is a negative rule, it gives a negative number of points to pairs of publications that have conflicting initials. Rule 3a en 3b are specified as pairs of publications with a matching general first name or a matching non-general first name, respectively. A first name is considered to be general when it appears more than 1000 times in the database, otherwise it is a non-general name. Address information directly linked to the author is used in rule 4. Rules 4a, 4b, and 4c find pairs of publications with matching country, city, organization, or department information. The more address information items are shared between two publications the stronger the rule is.

Table 1. Rules, scores and threshold for block size class = 2.

Category	Rule	Field	Criterion	Score
Author	1	email		100
	2a	all initials, more than one	two initials	5
	2b		more than two initials	10
	2c		conflicting initials	-10
	3a	first name	general name	3
	3b		nongeneral name	6
	4a	address (linked to author)	country, city	4
	4b		country, city, org.	7
	4c		country, city, org., dep.	10
Article	5a	shared co-authors	one	4
	5b		two	7
	5c		more than two	10
	6	grant number		10
	7a	address (not linked to author)	country, city	2
	7b		country, city, org.	5
	7c		country, city, org., dep.	8
Source	8a	subject category		3
	8b	journal		6
Citation	9	self citation		10
	10a	bib coupling	one	2
	10b		two	4
	10c		three	6
	10d		four	8
	10e		more than four	10
	11a	co-citation	one	2
	11b		two	3
	11c		three	4
	11d		four	5
	11e		more than four	6
			Threshold	11

Rules 5a, 5b, and 5c score the number of shared co-authors between two publications. The more shared co-authors, the higher the score that is assigned. By applying rule 6, pairs of publications are found that share the same research grant number. In rule 7, elements from the publication's address information that are not directly linked to the author is used.

Publications coupled by rule 8 share the same source, either the subject category (rule 8a) or the journal (rule 8b). In general, if a pair of publications scores on multiple 'sub rules'. The strongest rule is selected. For example, if a publication scores on rule 8b, it does not receive points for rule 8a.

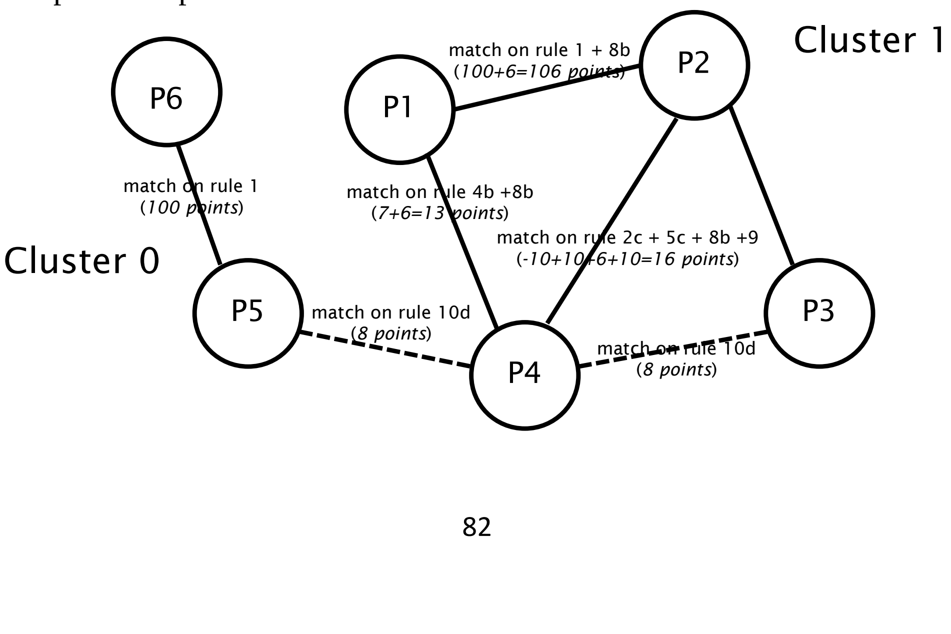
Rules 9-11 are based on citation information. We classify a citation as a self-citation when the citing and cited publication share a common author, based on shared last name and initials. In rules 10 and 11 the concepts of bibliographic coupling and co-citation are used. The stronger the coupling strength the more points are assigned, indicating a higher probability that two publications are written by the same author. And the stronger the co-citation strength, the more likely two publications have the same author.

Furthermore, we also deal with publication specific characteristics as hyper authorship and hyper instituteship. Such type of publications would easily receive too many points, because they have an increased chance, for example, that authors share a number of co-authors, self-citations, or research addresses. In such cases the scores for rules are lowered.

The scoring values of the rules are defined based on expert knowledge of the bibliographic database, and on initial evaluation of the method on a verified data set. First the expert knowledge was used to establish initial values for all the rules. After that the values were fine-tuned by experimental runs on the verified data set.

In step 2b the publication pairs are scored. A publication pair is defined as two publications that have scored on at least one rule. Obviously, two publications can score on multiple rules. In step 2b, first the total score for pairs of publications is computed. For example, see the set of publications (P1-P6) and their scores depicted in Figure 2. For example, P1 and P2 share the same email address (rule 1) and are published in the same journal (rule 8b). Therefore, this publication pair receives 100 + 6 = 106 points in total. The other publications in the sample are scored with the same procedure and are depicted with a connecting line in Figure 2.

Figure 2: Sample set of publications and scores.



Next the pairs with a total score above the threshold are taken into account for clustering (step 2c). The threshold applied is increased dependent on the block size class. In this way, we deal with the increased chance of incorrect coupling of publication pairs when the block size class increases. Basically, the more rules that are active for a pair of publications, the more proof there is that two publications are written by the same author. In general, for a pair of publications always combinations of rules are necessary to exceed the threshold. Only the email matching rule (rule 1) is strong enough to exceed the threshold by itself.

In step 2c, all publication pairs that are above a certain threshold, i.e. matched publication pairs, are clustered by means of single-linkage clustering. For example, when publications P1 and P2 are a matched pair, and publications P2 and P3 are a matched pair, a link between the two initial clusters is made via publication P2, thus the two initial clusters are merged into a new cluster with publications P1, P2, and P3, and so on (see Figure 2). The final cluster represents (a part of) the oeuvre of an author. In Figure 2, for block size class = 2, with threshold >= 11, two clusters are obtained. For the coupling between P4 and P5 and P3 and P4 there is not enough 'proof', denoted by a dotted line. Notice that, in the case of block size class = 3, with threshold >= 13, three clusters would be obtained: P5 and P6 are in cluster 0; P1, P2, and P4 are in cluster 1; and P3 is in cluster 2. Because of the higher threshold more proof is required to connect publications.

Post-processing

The first step in phase 3 deals with publications that are not clustered in phase 2, because there was not enough proof. Non-clustered publications are labelled as separate clusters. In the next step splitted oeuvres are processed. Our method initially clusters publications within name blocks. However, the oeuvre of an author might become dispersed over several blocks because of the synonym problem. For example, a female author might a have a cluster for her maiden name and for her married name. Splitted oeuvres are dealt with by using a correction procedure over the generated clusters on matching email addresses between them. Finally, matching publications are re-assigned to the largest cluster. The final step in this phase is the presentation of the clusters in a useful database for bibliometric analysis.

Evaluation

The results of the clustering are evaluated with metrics as precision, recall, F1, and cosine, which are common metrics in information retrieval research (see for example Levin et al., 2012), on two verified (gold) data sets. The gold datasets, obtained from CWTS' studies at the author level (Van Leeuwen, 2007), are:

- Data set 1, with 133 mainly Dutch researchers with 3,601 journal publications in the period 1990-2011. This data set is used for configuration of the method and for the computation of evaluation metrics on the author level.
- Data set 2, with 1905 mainly Dutch researchers at technical universities with 46,730 journal publication in the period 2001-2010. This data set is used for the computation of evaluation metrics on the aggregated block size class level.

Evaluation on data set 1

The precision and recall values for the best clusters per individual author are depicted in Figure 3, where the author names on the x-axis are ranked based on precision-recall values. The best cluster is defined as the cluster with the highest value for the F1 measure. Moreover, the results in Table 2 show on average a precision of 0.974 and a recall of 0.906 for the best cluster. This shows that the disambiguation method is conservative, it prefers precision over recall. The average total recall for this data set increases to a recall of 0.955 if the 5 best clusters for an author are selected for evaluation. For a number of authors the oeuvre is clearly distributed over several clusters.

Figure 3: Precision-recall analysis data set 1.

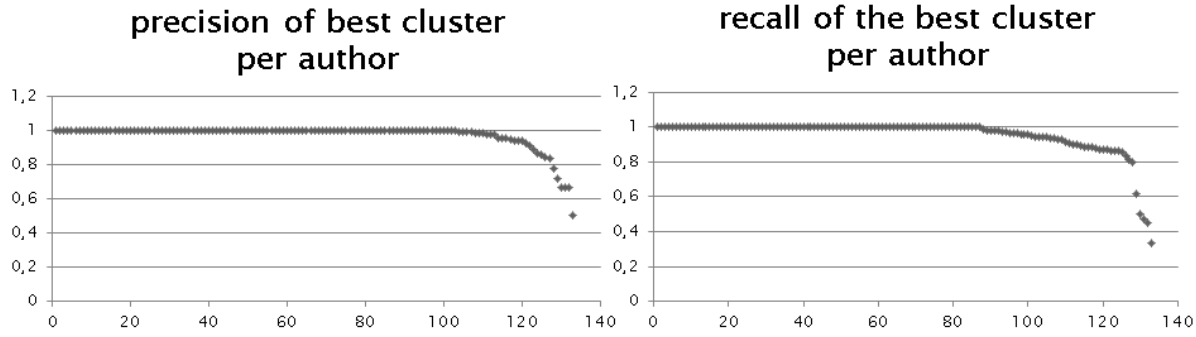


Table 2. Average values of evaluation metrics for the best cluster in data set 1.

	Precision	Recall	F1	Cosine
Best cluster (mean)	0.974	0.906	0.931	0.935
Best cluster (median)	1.000	0.955	0.963	0.964

Evaluation on data set 2

The results in Table 3 show on average a precision of 0.974 and a recall of 0.913 for the best cluster in data set 2. These results are very similar compared to the results in data set 1. The average recall is influenced by the synonym problem of some names.

Table 3. Average values of evaluation metrics for the best cluster in data set 2.

	Precision	Recall	F1	Cosine
Best cluster (mean)	0.974	0.913	0.930	0.936
Best cluster (median)	1.000	1.000	0.992	0.992

From Table 4 it can be concluded that the average number of clusters produced by the method increases when the block size class increases. This means that the oeuvre of an author with a popular name, say block size class 6, will be more splitted on average compared to the oeuvre of an author with a name which is less popular. Because the average precision-recall values for clusters with popular names is still high the clusters are still useful. For such cases it is obviously more work to collect all the relevant clusters. The low average recall of block size class 1, with rare names that are associated with only one publication, is explained by the synonym problem.

Table 4. Results of aggregated evaluation metrics.

Block size class	# Blocks	Avg # clusters	Avg recall	Avg precision
1	80	1,0	0,306	1,000
2	1764	1,4	0,931	0,977
3	122	2,1	0,967	0,949
4	47	3,4	0,976	0,953
5	7	3,8	1,000	0,982
6	4	7,2	1,000	0,944

Conclusion

In this paper we have presented an author disambiguation method for large bibliographic databases that uses rule-based scoring and clustering. The rules are based on bibliographic knowledge and are transparent and easy to understand. Due to the scoring, multiple rules can be combined to link publications. The rules in the system reinforce each other, i.e. the more rules that hold for a pair of publications, the more proof there is that these publications are written by the same author. Erroneous coupling of publications – due to the complexity of popular names, hyper authorship, and hyper instituteship – is partly prevented by lowering the scores for rules and by increasing the threshold values.

The clustering method is conservative, it values precision over recall. This means that if there is not enough proof for joining publications together, they will be put in separate clusters. As a consequence, the oeuvre of an author may be split over multiple clusters. The evaluation of the method shows on average a 95% precision and a 90% recall. The change of errors will increase if an author name is more common. In the future we want to apply the disambiguation method on the Scopus database[6]. In this way it would be possible to make a comparison between our cluster identifiers and the 'black-box' Scopus author identifiers, for which the underlying author disambiguation method is not in de public domain.

References

Bolikowski, Ł & Dendek, P. (2011). Towards a Flexible Author Name Disambiguation Framework. In: Sojka, Petr and Bouche, Thierry (eds.): *Towards a Digital Mathematics Library*. Bertinoro, Italy, July 20-21st, 2011. pp. 27-37.

Ferreira A., Gonçalves M, & Laender A. (2012), A brief survey of automatic methods for author name disambiguation, *SIGMOD Record.* 41, 2, 15-26.

Levin, M., Krawczyk, S., Bethard, S., & Jurafsky, D. (2012). Citation-based bootstrapping for large-scale author disambiguation. *Journal of the Association for Information Science and Technology*, 63(5), 1030-1047.

Liu, W., Islamaj Doğan, R., Kim, S., Comeau, D. C., Kim, W., Yeganova, L., Lu, Z. & Wilbur, W. J. (2013), Author name disambiguation for PubMed. *Journal of the Association for Information Science and Technology*, 65: 765–781.

Smalheiser, N. R. & Torvik, V. I. (2009), Author name disambiguation. *Annual Review of Information Science & Technology*, 43: 1–43.

[6] http://www.scopus.com

Song Y., Huang J., Councill I., Li J., & Giles C. (2007). Efficient topic-based unsupervised name disambiguation. In *Proceedings of the 7th ACM/IEEE-CS joint conference on Digital libraries* (JCDL '07). ACM, New York, NY, USA, 342-351.

Van Leeuwen, T. N. (2007). Modelling of bibliometric approaches and importance of output verification in research performance assessment, *Research Evaluation*, 16 (2): 93 – 105.

Putting citation in its place: Exploring the role of geography in publication impact[1]

John J. Chase *,** and Scott W. Cunningham **

* *john.chase@sri.com*
Center for Science, Technology, and Economic Development,
SRI International, 1100 Wilson Blvd, Arlington, VA 22209, USA

** *s.cunningham@tudelft.nl*
Faculty of Technology, Policy and Management, Delft University of Technology,
Delft 2600 GA, The Netherlands

Theory

The geographic organization of scientific teams is a significant determinant of the impact, and therefore the quality of a scientific paper. Prior knowledge in the field addressed the nature of scientific collaboration (Katz & Martin, 1997). Previous work has also examined the impact of distance on scientific output (Kraut et al., 1988). Despite this previous work, significant questions remain. In this paper we apply spatial measures of scientific collaboration, using a grid to capture the dynamic of both distances as well as places. The resultant output is consequential for understanding interdisciplinary research, regional scientific specialization, as well as the seeding of new research fronts.

In this paper we specifically examine research impact, a comparatively under-examined aspect of scientific geography. Citations are an effective proxy for research impact, albeit a measure which is partial and incomplete (Martin, 1996). The use of citations as a proxy for research impact, or research quality, comes with lots of issues (Bornmann & Daniel, 2008b). Chief among the issues are different patterns or levels of citation across different scientific disciplines. Citation appears to be a form of social capital, accruing over the course of a lifetime of research. Citations can be difficult to fractionate across teams, since team members provide integral capabilities to the research.

Knowledge spillovers in patenting have been examined. Researchers have discovered that knowledge spillovers are closely associated with regions. Some authors have defined regions as metropolitan areas (Jaffe, Trajtenberg & Henderson, 1992). Others have used a looser definition involving geographic distances of less than 300 kilometers (Bottazzi & Peri, 2003). In this paper we apply a novel approach, by creating a multi-resolution grid that spans the earth. Using geo-location databases and newly complete information concerning the location of all scientific researchers, we are able to place publications on this grid. The resultant data enables us to analyze selected fields of research, and specific knowledge regions. Econometric modeling enables us to partial out the effects of distances and region.

The paper confronts explanations of structure and agency in scientific collaboration. Structurally we examine the geographic dispersion of teams, as well as their location in high productivity regions. Nonetheless collaboration is only partly structured by regional and

[1] This work was supported by U.S. Department of Energy, Energy Efficiency and Renewable Energy, SunShot Award DE-EE00613

economic factors. Personal costs and incentives on the part of scientists and institutions must also be considered. We therefore consider search, transaction, and agency costs as an explanation for the findings.

In the analysis that follows we hypothesize that, all things being equal, geographically dispersed teams will show a higher degree of citation impact. The work further hypothesizes that teams located in higher productivity regions will have a smaller geographic collaboration distance, while showing a higher research impact. Smaller research fields will require a larger geographic collaboration distance while not necessarily having a larger research impact in terms of citation.

Analysis

Geocoding and Grid Assignment Procedure

Our analysis is based on a scientific publication dataset related to solar photovoltaic (PV) technologies. Scientific articles were extracted from the Web of Science (SCI-E) with a series of Boolean queries describing three PV technologies: *dye-sensitized solar cells*, *cadmium telluride thin film solar cells*, and *multi-junction solar cells.* The queries, vetted by a panel of PV subject matter experts, generated a data set of 22,924 documents (in XML format), over the time period of 1980 through 2013. A supplemental data set of the 176,897 documents cited by the primary data set was also extracted for future analysis.

For each publication, we follow a geocoding process similar to Waltman, Tijssen, & Eck (2011); we extract author affiliation information from the XML records, ignore all but city, state/province and country, and use a database of geographic names to convert the address information to geographic coordinates. Addresses were processed using the open source *twofishes* geocoder to provide city-level precision latitude and longitude coordinates for each author affiliation in the data set. (Blackman, 2012)

To identify regions of high publication productivity, we partition the globe into a discrete grid and assign each publication to one or more grid locations based on the spatial intersection of the grid boundaries and the geographic location of the publication's author affiliations. For this analysis we employ the Icosahedral Snyder Equal Area Aperture 3 Hexagon (ISEA3H) Geodesic Discrete Global Grid described by Sahr, White, & Kimerling (2003). A level 8 ISEA3H grid composed of 65,612 cells was generated using the public domain software package DGGRID (Sahr, 2013). Each hexagonal cell corresponds to approximately 7,774km^2 – approximately the same area as a 100km diameter circle. To evaluate the effect of grid cell size, a level 9 ISEA3H grid composed of 196,832 cells – each cell corresponding approximately to a 57km diameter circle – will also be generated and used for the analysis. A further advantage of the discrete grid formulation is the ability to further aggregate the grid to create regional or national measures. This enables the equal area projections to be made comparable with planning region approaches, which may aggregate across regions that are very different in area, population and gross domestic product.

Figure 1: Level 8 Icosahedral Snyder Equal Area Aperture 3 Hexagon Discrete Global Grid

To assign publications to appropriate hex cells, the output of the DGGRID software is converted into a series of polygons and stored a spatially enabled relational database for further analysis (PostGIS). Publications, represented as a series of points corresponding to author addresses are inserted into the same database; a *spatial join* operation then associates the publications with grid cells based on the intersection of publication points with grid polygons. With the tables joined, we can identify which publications are associated with a given hex cell and conversely, which hex cells are associated with a given publication.

Indicators

Based on the geocoding results, we calculate two publication-based indicators of collaboration distance. The *geographical collaboration distance* (GCD) described in Waltman et al. (2011) is calculated as the largest geographical distance between any two addresses associated with a publication. The GCD tells us whether or not researchers are engaging in long-distance collaborations, but it doesn't fully reflect the geographic dispersion of collaborators. For example, a publication with one author in Boston and one author in London will have the same GCD as a publication with seven authors in Boston and one author in London, despite the very different overall geographic dispersion of the team.

To supplement the GCD, we calculate the *collaboration standard distance* (SD_C) of each publication. For a set of points (author addresses), the SD_C is defined as the average distance from each point to the mean center of all the points. Standard Distance is calculated as: (Burt, Barber, & Rigby, 2009)

$$SD_C = \sqrt{\frac{\sum_{i=1}^{n}(x_i - \bar{X})^2}{n} + \frac{\sum_{i=1}^{n}(y_i - \bar{Y})^2}{n}}$$

where x_i and y_i are the coordinates for each author affiliation and X and Y are mean center of all author affiliations. Returning to the example above, the publication with one author in Boston and one author in London would have substantially higher SD_C than the publication with seven authors in Boston and one in London.

To analyze the effects of regional concentrations of publication activity and impact, we calculate a series of hex cell-based indicators derived from the results of the spatial join, our first measure, *regional productivity,* is simply the count of the number of publications associated with the region. To characterize *regional impact,* we count the total output of high impact publications in a region. To adjust for the varying number of publications associated with different regions, we borrow a commonly used economic geography concept, the "location quotient." (North, 1955) We define the *high impact publication location quotient* (LQ_{hip}) as the percentage of high impact publications associated with a region divided by the percentage of high impact publications in all regions. LQ_{hip} is calculated as:

$$LQ_{hip} = \frac{x_r/n_r}{x/n}$$

where x_r and x are the number of highly cited publications regionally and globally, and n_r and n are the total number of publications regionally and globally. For a given region, an LQ_{hip} greater than 1 indicates that production of high impact publications is more concentrated in that region than average. Adopting the definition employed by Bornmann & Leydesdorff (2011) and Tijssen, Visser, & Leeuwen (2002), we define the top 10% of publications (by times cited) as "high impact". Because our current analysis is limited to a single domain, solar photovoltaic technology, we do not normalize citation counts by field.

Model

To explore the interaction of geography and citation count, we perform a multiple regression analysis. Based on the distribution of citation counts in the data, and like others before us (Bornmann & Daniel, 2007, 2008a; Davis et al., 2008), we selected the negative binomial regression model, which is well suited to the distribution of our data. (Long, 1997). We model the influence of our independent variables (Table 1) on our dependent variable, citation count.

Table 1. Model Variables

Variable	**Description**
CitationCount (dependent)	Number of times article has been cited
NumAuthors	Number of authors associated with publication
NumCountries	Number of countries associate with publication
GCD	Geographic Collaboration Distance
SD_C	Collaboration Standard Distance
RegionProductivity	Largest of *regional productivity* measures associated with publication
RegionImpact	Largest of *regional impact* measures associated with publication
LQ_{hip}	Largest of LQ_{hip} measures associated with publication

The results of the analysis, to be fully detailed in the forthcoming paper, suggest a complex interplay between the geographical distance and citation impact of scientific work. Geographic dispersion has a non-linear effect on earned citation. This may be because while geographic dispersion is always a cost, the capability of searching across extensive epistemic networks to find exactly the right collaborator is often a boon. This suggests that both geographical distance as well as epistemic distance plays a role in the results.

The analysis also suggests that geographical dispersion has a heterogeneous impact depending on the productivity of the region, and the size of the research field. Here again, we argue that scientific search within epistemic networks provides a potential explanation. Search processes for scientific fields both small and large are fundamentally different in kind. High productivity regions involve a concentration of potentially highly qualified researchers. Distance measures are themselves insufficient to measure the concentration of talent; regional level variables are also needed to model concentration.

Recommendations

Scientific knowledge increasingly requires multi-disciplinary and interdisciplinary work for continued progress. However multidisciplinary work comes with its own costs – of finding the correct people, of balancing the varied epistemic concerns, and of managing scientific production across extensive distances. Another cost involves managing differing rules and incentive systems, a cost which is likely to increase as teams grow geographically more dispersed. Related work has investigated the costs and benefits of various kinds of proximity on research collaboration (Cunningham & Werker, 2013).

The results demonstrate that highly productive regions are also advantaged in the production of high citation work. High productivity regions provide useful knowledge spillovers across a variety of new fronts of knowledge. Researchers seeking complementary knowledge to complete a research agenda, often need not look much further than their local regions. As a result the search costs for finding team members are low, and in addition, the transaction costs associated with maintaining a working scientific relationship are also very low. Furthermore the respective researchers inside a region share a common innovation system, leading to a high degree of systemic proximity.

These results have practical implications for stimulating new scientific discoveries, for the evaluation of research, and for personnel management in scientific teams. Emerging scientific fields may have relatively low search costs, but conversely may have to manage the high transaction costs associated with geographic distribution. Seeding new discovery may involve funding and the smart specialization of districts where relevant knowledge already resides. Alternately there may be concerns of equal access to funding and knowledge, where geographic dispersion is a necessary cost for equity.

In these circumstances science funding agencies may wish to place a higher premium the reduction of transaction costs, through collaboration grants and through the funding of interdisciplinary workshops. A related policy measure might involve reducing systemic distance between team members. Reducing systemic distance may be achieved through mobility grants, a deliberate effort to unify funding requirements across agencies, or the award of grants-in-kind.

Grants-in-kind, where multiple funders collaborate with a common set of application procedures and research incentives, may relieve teams of managing agency problems. Previous research has considered funding for joint ventures, or funding on a team or network basis (Melin, 2000). In particular multiple grant authorities may present teams with a multi-principal problem. As a final note, scientific teams and their managers might think more explicitly about the impact of distance on team formation, and the necessary compromises that distance entails. The result of this paper suggests adopting a healthy degree of skepticism regarding the virtualization of scientific teams.

References

Blackman, D. (2012). Twofishes. *GitHub Repository*. Retrieved from https://github.com/foursquare/twofishes

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. doi:10.1002/asi.20531

Bornmann, L., & Daniel, H.-D. (2008a). Selecting manuscripts for a high-impact journal through peer review: A citation analysis of communications that were accepted by Angewandte Chemie International Edition , or rejected but published elsewhere. *Journal of the American Society for Information Science and Technology*, *59*(11), 1841–1852. doi:10.1002/asi.20901

Bornmann, L., & Daniel, H.-D. (2008b). What do citation counts measure? A review of studies on citing behavior. *Journal of Documentation*, *64*(1), 45–80. doi:10.1108/00220410810844150

Bornmann, L., & Leydesdorff, L. (2011). Which Cities Produce More Excellent Papers Than Can Be Expected? A New Mapping Approach Using Google Maps, *62*(10), 1954–1962. doi:10.1002/asi

Bottazzi, L., & Peri, G. (2003). Innovation and spillovers in regions: Evidence from European patent data. *European Economic Review*, *47*(4), 687–710. Retrieved from http://www.sciencedirect.com/science/article/pii/S0014292102003070

Burt, J. E., Barber, G. M., & Rigby, D. L. (2009). *Elementary Statistics for Geographers* (p. 653). Guilford Press.

Cunningham, S. W. & C. Werker (2013), "Proximity and Collaboration in European Nanotechnology," *Papers in Regional Science, 91*(4): 723-742.

Davis, P. M., Lewenstein, B. V, Simon, D. H., Booth, J. G., Mathew, J., Connolly, L., & Davis, M. (2008). Open Access Publishing, Article Downloads, and Citations: Randomised Controlled Trial. *BMJ: British Medical Journal*, *337*(7665), 343–345. doi:10.1136/bmj.a568

Jaffe, A. B., Trajtenberg, M., & Henderson, R. (1992). Geographic Localization of Knowledge Spillovers as Evidenced by Patent Citations. *The Quarterly Journal of Economics*, *108*(3), 577–598. doi:10.2307/2118401

Katz, J. S. & B. R. Martin (1997), What is Research Collaboration? *Research Policy,* 26(1): 1-18.

Kraut, R., Egido, C. & J. Galegher (2000). Patterns of contact and communication in scientific research collaboration, CSCW '88 Proceedings of the 1988 ACM Conference on Computer-Supported Cooperative Work, 1-12.

Long, J. S. (1997). *Regression models for categorical and limited dependent variables* (7th ed., Vol. 7). Sage.

Martin, B. R. (1996). The use of multiple indicators in the assessment of basic research, *Scientometrics*, 36(3): 343-362.

Melin, G. (2000). Pragmatism and self-organization: Research collaboration on the individual level, *Research Policy*, 29(1):31-40.

North, D. C. (1955). Location Theory and Regional Economic Growth. *Journal of Political Economy*, *63*, 243. Retrieved from http://ideas.repec.org/a/ucp/jpolec/v63y1955p243.html

Sahr, K. (2013). DGGRID (version 6.1). Retrieved from http://discreteglobalgrids.org/software/

Sahr, K., White, D., & Kimerling, a. J. (2003). Geodesic Discrete Global Grid Systems. *Cartography and Geographic Information Science*, *30*(2), 121–134. doi:10.1559/152304003100011090

Tijssen, R., Visser, M., & Leeuwen, T. Van. (2002). Benchmarking international scientific excellence: Are highly cited research papers an appropriate frame of reference? *Scientometrics*, *54*(3), 381–397. Retrieved from http://www.akademiai.com/index/g3j3153567g71w10.pdf

Waltman, L., Tijssen, R. J. W., & Eck, N. J. Van. (2011). Globalisation of science in kilometres. Retrieved February 04, 2014, from http://arxiv.org/pdf/1103.3648v2.pdf

Citations between books and journals in political science

Pei-Shan Chi

peishan.chi@kuleuven.be
Centre for R&D Monitoring (ECOOM), KU Leuven, Waaistraat 6, Leuven (Belgium)
Institute for Research Information and Quality Assurance, Schützenstraße 6a, Berlin, 10117 (Germany)

Background

Citations to and from books are distributed differently from those to and from journal articles (Broadus, 1971). Larivière et al. (2006) analyzed journal articles in SSCI and A&HCI showed that references to journal articles amount to about 40-50% in the social sciences and humanities during the period 1981-2000, or 45% in general. However, Line (1979) found that monographs referenced proportionally fewer journal articles (25%), and more monographs (51%) and other types of literature (24%) compared to journals, which reference to journal articles at 47%, monographs 39%, and 14% to others. These studies show that books reference more books than articles, and journal articles refer more articles than books.

On the side of citations, Samuels (2013) shows books in political science are cited by books (16.3 times on average) more than by SSCI journal articles (6.6 times on average). In another study (Samuels, 2011), SSCI articles are likewise cited more by books than journal articles. This indicates that citations from journal articles are not the largest source of citations obtained by social science publications. Although it is important to point out that the citations from non-journals cannot be measured with the current methodology, these “invisible citations” could increase the overall citations, especially of regional publications, considerably.

These “non-source citations” from non-source items (not indexed in Web of Science (WoS)) exist, but it is difficult to trace them comprehensively. The poor coverage of WoS in the social sciences, which is due to the selection thresholds on high-impact, international and peer-reviewed journals, leads to missed citation links in these fields on a large scale. The Book Citation Index (BCI) in WoS may provide a new opportunity to increase citation coverage, though its publication coverage is not well developed yet. To investigate the citations from outside of WoS, the method applied in this study is to reveal the ratios of citations between books and journals by comparing the citation counts from WoS and BCI, since citations from books are influential in the social sciences, as shown in the previous studies.

Data set

The five-year publication output (2003-2007) from 33 professors in the two top-ranking German political science institutions, the department of Political Science at Mannheim University and the Institute of Political Science at the University of Münster, were collected from the researchers’ official websites, institutional repositories, and SOLIS literature database. The counts of publications of these 1,015 publications in different document types are shown in Table 1. Citations of *ISI journal articles*, *books*, and *edited books* were obtained from the WoS web version on 29th of January 2014.

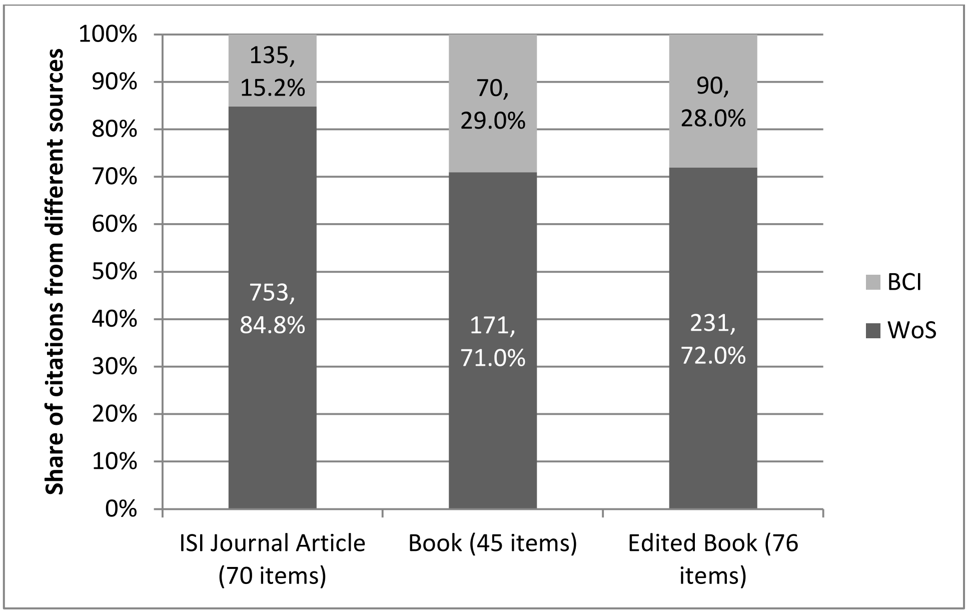
Table 1. Numbers of publications of two German political science institutes.

Document types	No. of publications (%)
ISI Journal Article	70 (6.9)
Non-ISI Journal Article	151 (14.9)
Book	45 (4.4)
Edited Book	76 (7.5)
Book Chapter	396 (39.0)
Conference Paper	151 (14.9)
Others	126 (12.4)
All	1,015 (100)

Results

In order to compare the citations from book to the citations from journals, 70 *ISI journal articles*, 45 *books,* and 76 *edited books* were checked for their citations in WoS and BCI simultaneously. Figure 1 shows that for *ISI journal articles* about 15% of citations are from BCI, while about 30% citations of *books* and *edited books* are from BCI. In general, books have higher percentage of citations from books than ISI journal articles. The result shown in this study is not as strong as those reported by Samuels, but books and journal articles are cited by books more than articles in political science. However, it proves that books receive more citations from books than journal articles do.

Figure 1: Shares of citations from WoS and BCI.

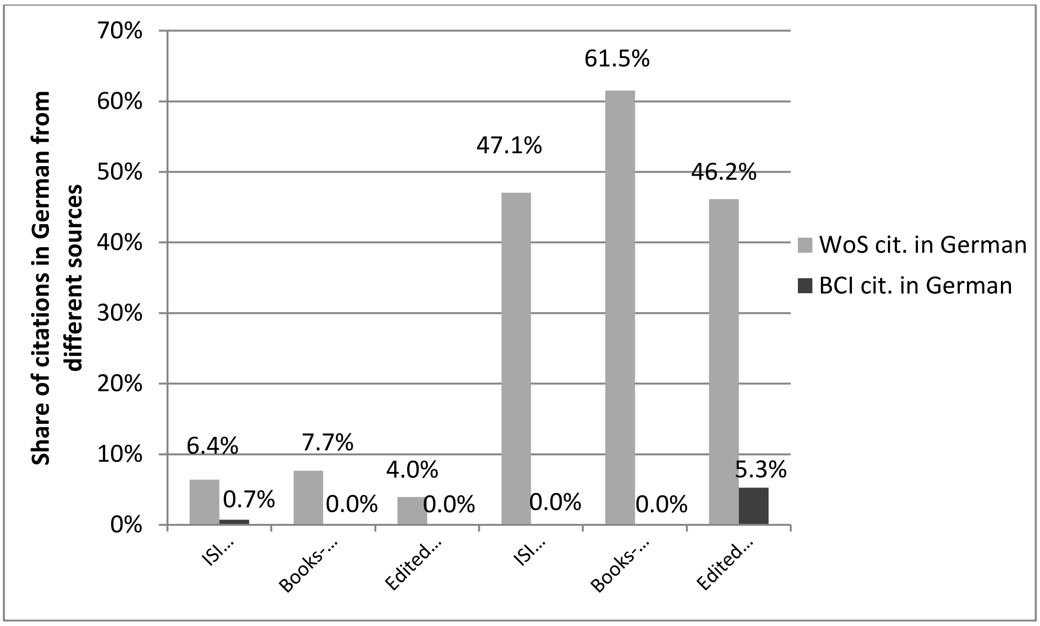


* own search on 29th of January 2014

**without citation window

In Figure 2, items in German (no matter in which document type) were cited by more WoS citations in German than items in English. The *books* in German have 62% of WoS citations in German, while *books* in English have only 8% of WoS citations in German. On the other hand, Figure 2 also shows that BCI does not have a sufficient coverage of books in German in political science, reflecting a very poor percentage of BCI citations in German to all items.

Figure 2: Share of citations in German from WoS and BCI.



References

Broadus, R. N. (1971). The literature of the social sciences: A survey of citation studies. *International Social Sciences Journal, 23*(2), pp. 236-243.

Larivière, V., Archambault, E., Gingras, Y., Vignola-Gagné, E. (2006). The place of serials in referencing practices: Comparing natural sciences and engineering with social sciences and humanities. *Journal of the American Society for Information Science, 57*(8), pp. 997-1004.

Line, M. B. (1979). The influence of the type of sources used on the results of citation analyses. *The Journal of Documentation, 35*(4), pp. 265-284.

Samuels, D. J. (2011). The modal number of citations to a political science article is greater than zero: Accounting for citations in articles and books. *PS: Political Science and Politics, 44*(4), pp. 783-792.

Samuels, D. J. (2013). Book citations count. *PS: Political Science and Politics, 46*(4), pp. 785-790.

The under-representation of developing countries in the main bibliometric databases: a comparison of rice studies in the Web of Science, Scopus and CAB Abstracts[1]

Tommaso Ciarli*, Ismael Rafols** and ÓscarLlopis***

*t.ciarli@sussex.ac.uk
SPRU (Science and Technology Policy Research), University of Sussex, Brighton, BN1 9SL (UK)

**i.rafols@ingenio.upv.es
Ingenio (CSIC-UPV), Universitat Politècnica de València, València, 46022 (Spain) and
SPRU (Science and Technology Policy Research), University of Sussex, Brighton, BN1 9SL (UK)

***osllocor@upvnet.upv.es
Université de Bordeaux (France)

Abstract
Although the main bibliometric databases (Web of Science and Scopus) claim to include journals on the basis of scientific and publication standards, there have long been concerns that its coverage is biased in favour of journals based in industrialised countries. In this article, we investigate this claim in an area of agricultural science, namely rice research, using the database CAB Abstracts. We find unambiguous evidence that for a field such as rice, statistics based on WoS and Scopus may strongly under-represent the scientific production by developing countries, and over-represent that by industrialised countries.

Introduction

Agricultural research has been and remains an important endeavour in developing countries, as it is seen as a potential source of knowledge and innovation crucial for social and economic development. However, given its applied orientation, the local specificity of the topics and the lack of relevance of the topic for developed countries, it is unclear to which extent, research on agriculture-related issues gets published or cited in "international" journals and (even less) gets indexed by main bibliographic databases (Velho, 1986, 1990). Thus, many science policy analysts on developing countries have claimed that their publications are under-represented in main bibliometric databases and that an applied field such agriculture might be further disadvantaged (Royal Society, 2011).

The coverage of the Web of Science is well known to be biased towards English-speaking publications and biomedical publications (Archambault et al., 2006). Scopus has been shown to have a broader coverage, but its overall coverage leads to similar ranks regarding country production over different fields, what has led some analysts to claim that 'indicators of scientific production and citation at the country level are stable and largely independent of the database'(Archambault et al., 2009, p. 1320). Thus, in international benchmarking of science

[1] We are grateful to Diego Chavarro for fruitful discussions. We acknowledge support from the the UK ESRC (grant RES-360-25-0076, "Mapping the Dynamics of Emergent Technologies") and the US National Science Foundation (Award #1064146 - "Revealing Innovation Pathways: Hybrid Science Maps for Technology Assessment and Foresight"). The findings and observations contained in this paper are those of the authors and do not necessarily reflect the views of the funders.

by research agencies, the main databases WoS (e.g UNESCO, 2010) and Scopus (e.g Royal Society, 2011) continue to be used. However, science policy analysts have repeatedly warned of the partial coverage, mainly North-American and European, of the main commercial databases (Chavarro, 2013) and recommendations have been made on the need to improve scientometric indicators in order to "properly evaluate global science" (Royal Society, 2011, p. 107).

In this article we show that for a field such as rice, statistics based on WoS and Scopus may strongly under-represent the scientific production by researchers in developing countries, and over-represent that by researchers in industrialised countries. This is suggested by the acute, though decreasing, differences in coverage and publications counts by country between WoS and Scopus and the database CAB Abstracts which has a much wider coverage (for the sake of brevity CAB Abstracts will be referred to as "CABI" in the text). One may imagine that these differences are likely to apply as well to other fields of agricultural research, particularly those related to crops not important in the temperate climates of industrialised nations.

The agricultural field is an interesting area to investigate since it is “a field in which scientists are under considerable pressure from client groups” so as to solve local problems rather than contribute to the development of “universal” knowledge (Velho, 1990, p. 503). The main reason for focussing on rice is that we would like to monitor a relatively basic crop (although the technology behind research on agricultural crops is far from basic), which serves a large number of people with different needs in different parts of the world. Rice is a crop (i) which feeds a huge number of people around the world, particularly in low and middle income countries; (ii) which was at the core of the *green revolution*, particularly in the 60's and 70's, when high yield varieties of rice where investigated and distributed across the world to reduce the problem of famines in low income countries; (iii) and which, being the symbol of the green revolution is also a controversial technology due to the negative effects such as impoverishment of diets, overuse of water, exhaustion of soils, pollution, etc.

To our knowledge, this is the first bibliometric study using CAB Abstract, though small explorations by practitioners have been reported (e.g. Kawasaki, 2004). This short paper is a preliminary version and thus it only provides a simple first and tentative analysis of the data. The next steps of this investigation will be first to match the articles between databases to check the degree to which they cover different journals, and second to study the research areas (genetics, pest research, rice production, nutrition, etc.) which are over/under-represented in each database. At the current stage, though, this study strongly suggest that CAB Abstracts (CABI) can be a useful complement of WoS and Scopus as a source of information to map socially relevant research in mid and low income countries for issues related to fields such as agriculture, environment, veterinary sciences, applied economics, and food science and nutrition.

Methods and data

Publications on rice were manually downloaded from the WoS (including SCI-Expanded, SSCI, A&HCI, CPCI-S i CPCI-SSH) searching “rice” or “oryza” in the field “topic”. Scopus records were also manually downloaded searching in title, abstract or keywords, i.e. TIT-ABS-KEY ("rice" OR "oryza"). Similarly, documents with “rice” or “oryza” were searched in title and abstract of the database CAB Abstracts. (CAB Abstract allows to retrieve documents classified as strongly related to “rice” in the field “Descriptor”, but for consistency with the search on WoS and Scopus, we did not use this option here).

CAB Abstracts (http://www.cabdirect.org/) is a database focused on environment and agriculture. It is run by CABI, an inter-governmental, not-for-profit organization that was set up by a United Nations treaty, with 48 member countries (many of them belonging to the Commonwealth), with a mission of "providing information and applying scientific expertise to solve problems in agriculture and the environment".[2] Therefore, both CAB Abstract (for agriculture and environment) and Global Health (for public health) are aimed at facilitating the retrieval of relevant information for practitioners, very much as MEDLINE for medical research, but with a focus on development. The data of the three databases was uploaded into the very useful functions for address cleaning and standardisation of the programme VantagePoint.[3].

In this short communication we present a set of descriptive statistics, providing information on the coverage in terms of number of publications by document type, language and year. We then compare the number of publications for the main countries. We use the main author affiliation to retrieve information on the country. An important caveat is that CABI only reports the affiliation of the first author. In the case of WoS and Scopus, on the contrary the affiliation of all authors are included. As a result, the shares of countries will tend to be higher in WoS and Scopus. In this preliminary version, this effect has not been corrected. The error is estimated (using small document samples) on a 10%-30% over-representation, depending on country.

Characterisation of samples

Let us first describe the main differences between the documents retrieved from each database. Given that each database classifies documents into different type categories, we downloaded all the document types, with statistics described in Table 1. It is found that in all cases, journal articles have a dominant share, between 81% (WoS) and 94% (Scopus). Hence, the results that follow will be mainly explained by differences in the coverage of journals used to index articles. The second most important document type is conference proceedings/papers, which make between 3.5% (Scopus) and 7.3% (CABI) of the total publications. The category "Miscellaneous" in CABI (4.7%) deserves further investigation.

Table 1. Share of publications by document type in the three databases investigated.

WoS			**Scopus**			**CAB Abstracts**		
Doc type	**%**	**Cum%**	**Doc type**	**%**	**Cum%**	**Doc type**	**%**	**Cum%**
Article	81.2%	81.2%	JOUR	93.7%	93.7%	Journal article	84.8%	84.8%
Proceedings Paper	7.1%	88.3%	CONF	3.5%	97.2%	Conference paper	6.8%	91.6%
Review	3.4%	91.7%	SER	1.6%	98.8%	Miscellaneous	4.7%	96.3%
Meeting Abstract	2.7%	94.4%	INPR	0.9%	99.7%	Book chapter	2.0%	98.3%
Note	2.4%	96.8%	CHAP	0.3%	99.9%	Book	1.9%	100.2%
Book Review	1.6%	98.4%	BOOK	0.1%	100.0%	Annual report	0.9%	101.1%
Editorial Material	0.7%	99.1%				Bulletin	0.6%	101.7%

[2] http://www.cabi.org/about-cabi/ (Retrieved March 1st, 2014).
[3] https://www.thevantagepoint.com/

Letter	0.6%	99.6%				Conference proceedings	0.5%	102.2%
Correction	0.3%	99.9%				Bulletin article	0.4%	102.7%

Note: Some documents are classified into more than one category.For example, many CABI conference papersa are also Journal articles (this is why cumulative counting is higher than 100%).

In terms of language, as shown in Table 2, CABI is much more comprehensive than WoS (Scopus data temporarily not available), with almost 10% of the documents in Chinese, and 6.7% in Japanese. WoS only covers a few journals in Japanese (2%) and Portuguese (1%). If we consider the actual number of publications rather than the ratio within the data base, the difference is even larger. For exemple, CABI has 7 times the number of publications on rice in Japanese and 5 times the number of publications on rice in Portuguese.

Table 2. Share of publication by original language.

	CABI		WoS	
Language	# docs	%	# docs	%
English	148577	71.84%	92554	94.93%
Chinese	20544	9.93%	490	0.50%
Japanese	13844	6.69%	2032	2.08%
Portuguese	5356	2.59%	1015	1.04%
French	3942	1.91%	560	0.57%
Spanish	3320	1.61%	307	0.31%
Korean	3018	1.46%	31	0.03%
Russian	2396	1.16%	162	0.17%
Italian	1546	0.75%	22	0.02%
German	1462	0.71%	214	0.22%
Persian	501	0.24%	0	0.00%
Dutch	440	0.21%	9	0.01%
Thai	421	0.20%	11	0.01%
Indonesian	285	0.14%	0	0.00%

Note: % documents is computed only over the documents with language reported in the database (98% in WoS, 91% in CABI).

Trends over time show that CABI has had historically a much broader coverage that WoS and Scopus, as shown in Figures 1 and 2. Before the 1980s, coverage by WoS and Scopus of publications on "rice" is very limited. CABI shows a great increase in rice publications from the postwar until the mid 1970s, particularly after the mid 1960s. This in agreement with the diffusion of the "green revolution". The postwar expansion is followed by a period of slow growth from 1975 until 2000, when a renewed growth is observed (perhaps in coincidence with the advent of genomic studies). Since the mid 1990s WoS and Scopus have been catching up with CABI and by 2012 (last year fully indexed), WoS reaches 80% of CABI and Scopus 86%, though with substantial non-overlapping coverage (not shown).

Figure 1: Number of publications on rice per year by database from 1902 until 1975.

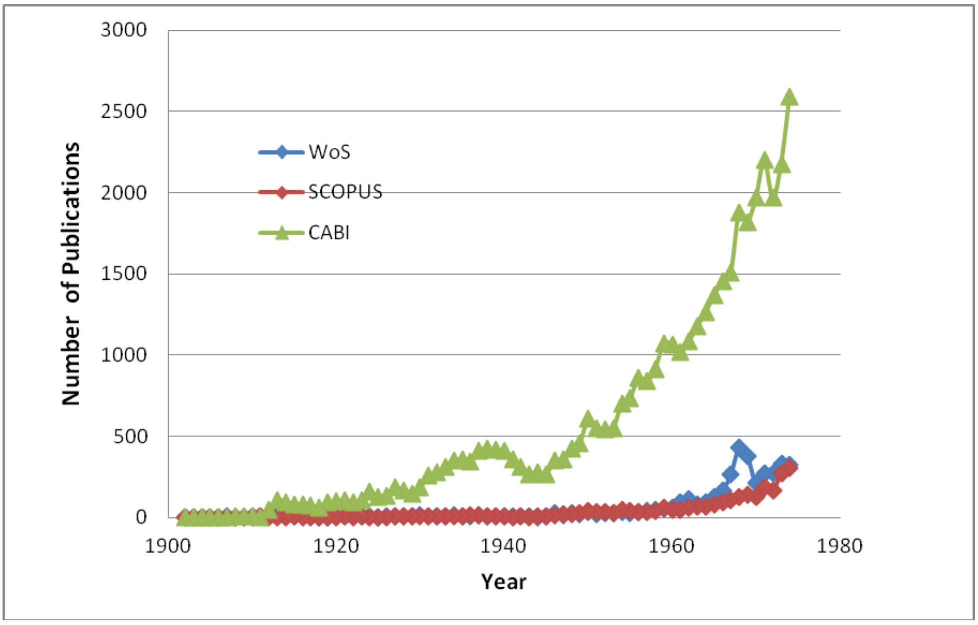
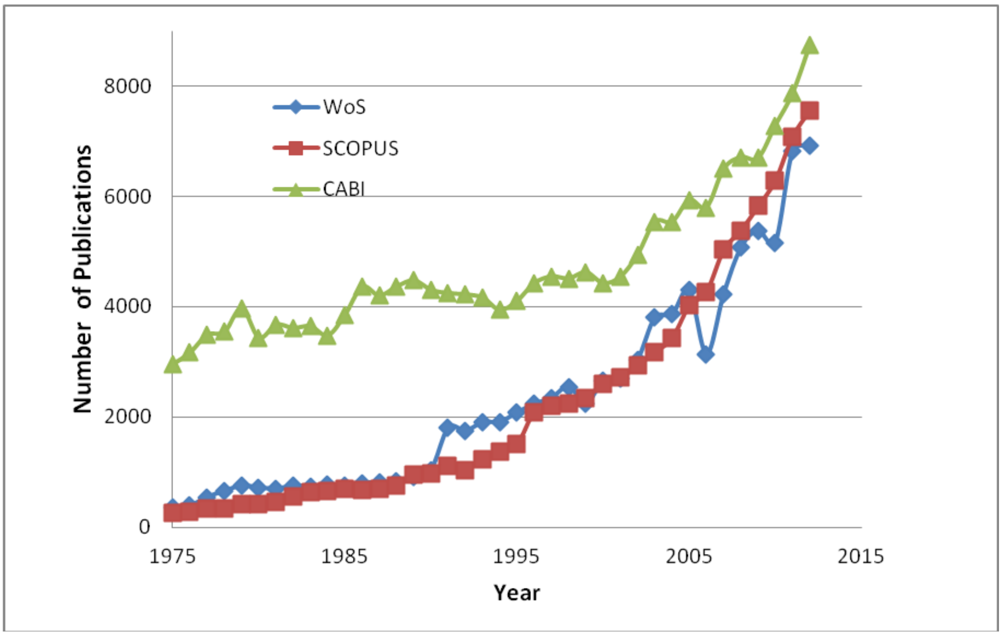


Figure 2: Number of publications on rice per year by database from 1975 until 2012.



Comparison of coverage across countries

The countries publishing the most on rice are India, China, Japan and the US. China's publications on rice have sharply increased in the last twenty years (as expected from global publication trends, cf. Leydesdorff, 2012) whilst the share by India, Japan and the US have decreased, as shown in Figure 3. Interestingly, all three databases agree on these trends.

However, there are major differences in the overall proportion of publication assigned to each country in each database. In the case of CABI, India was the most productive country until the it was caught by China in 2004. But whereas India's publications made 21% of the total in

2000-09 according to CABI, they represent less 9.6% and 8.4% in Scopus and WoS, respectively, as illustrated in Figure 4. Similarly, China's publications were 23% of CABI's publications, but only 16% and 13% according to Scopus and WoS. Oppositely, US publications were only 7% in CABI, but 15% and 16% in Scopus and CABI. Japan stands in the middle, with only a ±1% difference depending on the database used. The differences in coverage between databases have narrowed in recent years, as shown in the right side of figure 4, but there is still a 2-fold difference in the percentage of publications assigned by CABI and WoS for the US and India.

Figure 3: Publications trends by country according to CABI data.

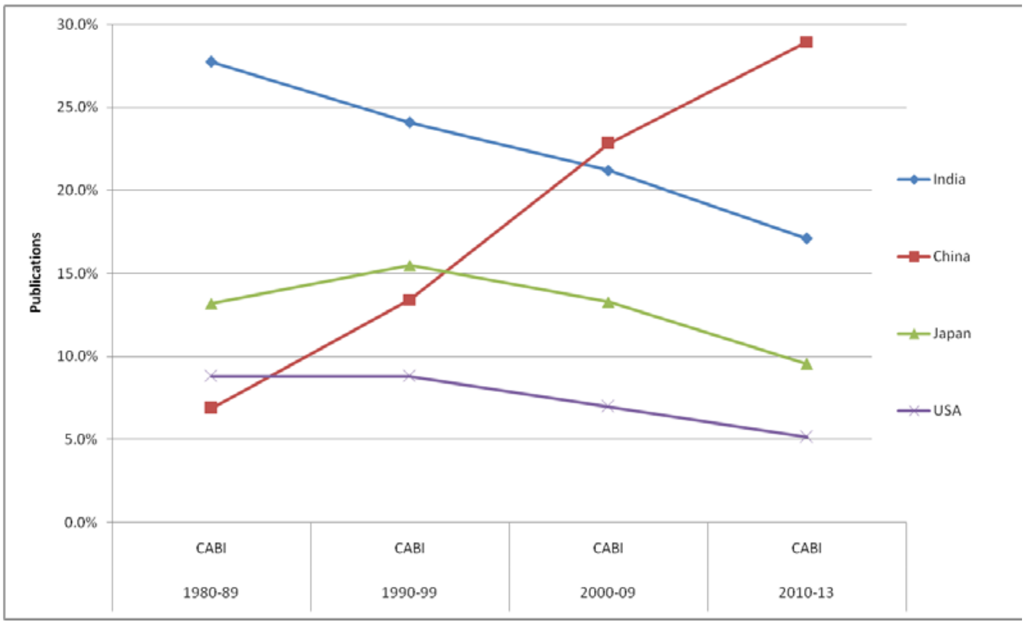


Figure 4: Percentage of publications on rice for large countries for different databases, in two periods.

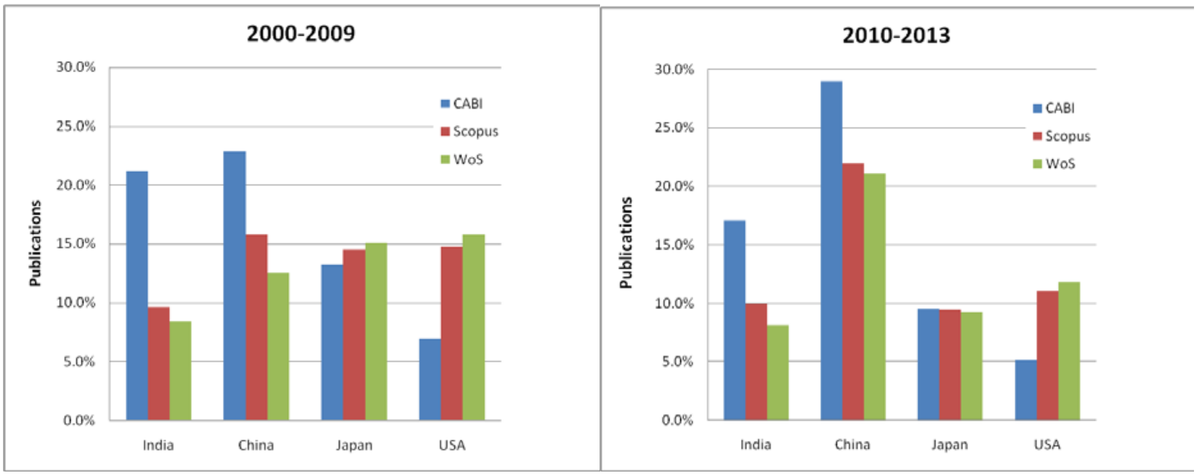
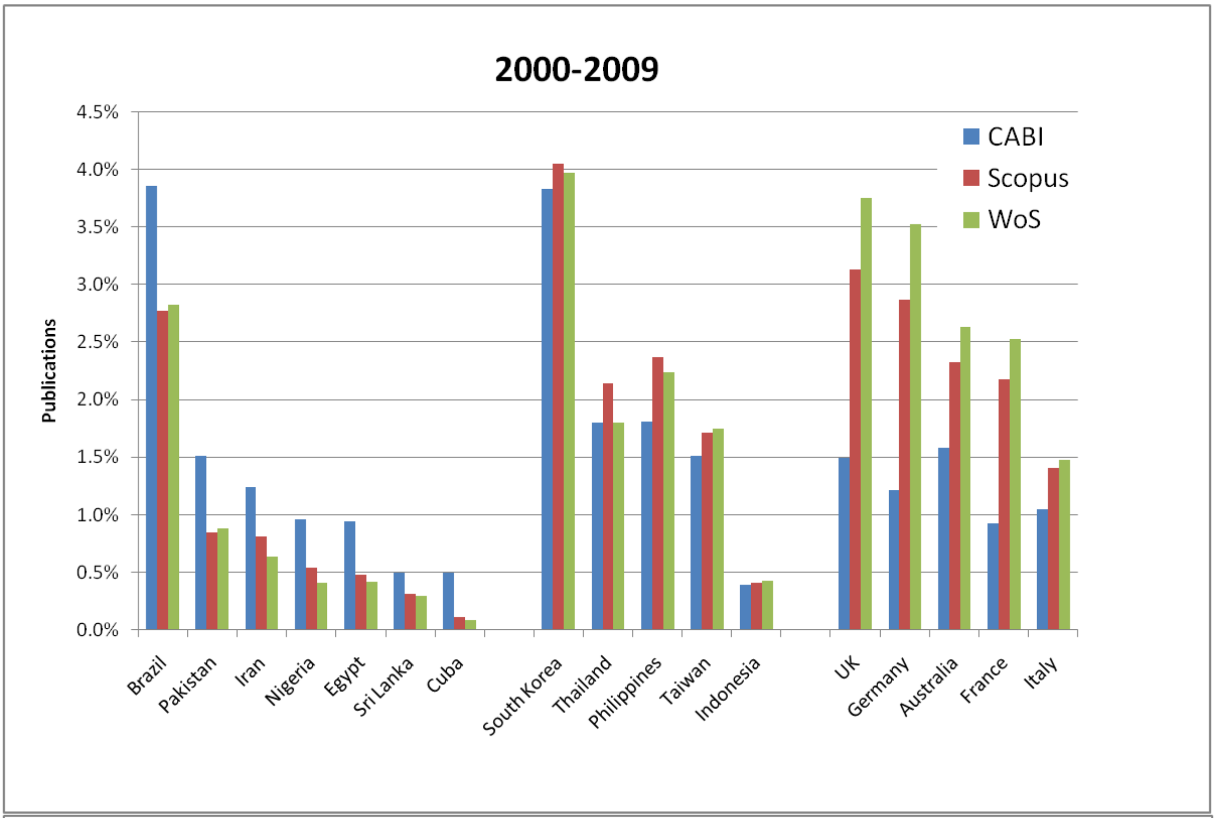


Figure 5: Percentage of publications on rice by countries for different databases, in two periods. Left hand side: countries with a relative higher CABI coverage. Centre: countries with similar coverage. Right hand: countries with higher coverage in WoS.



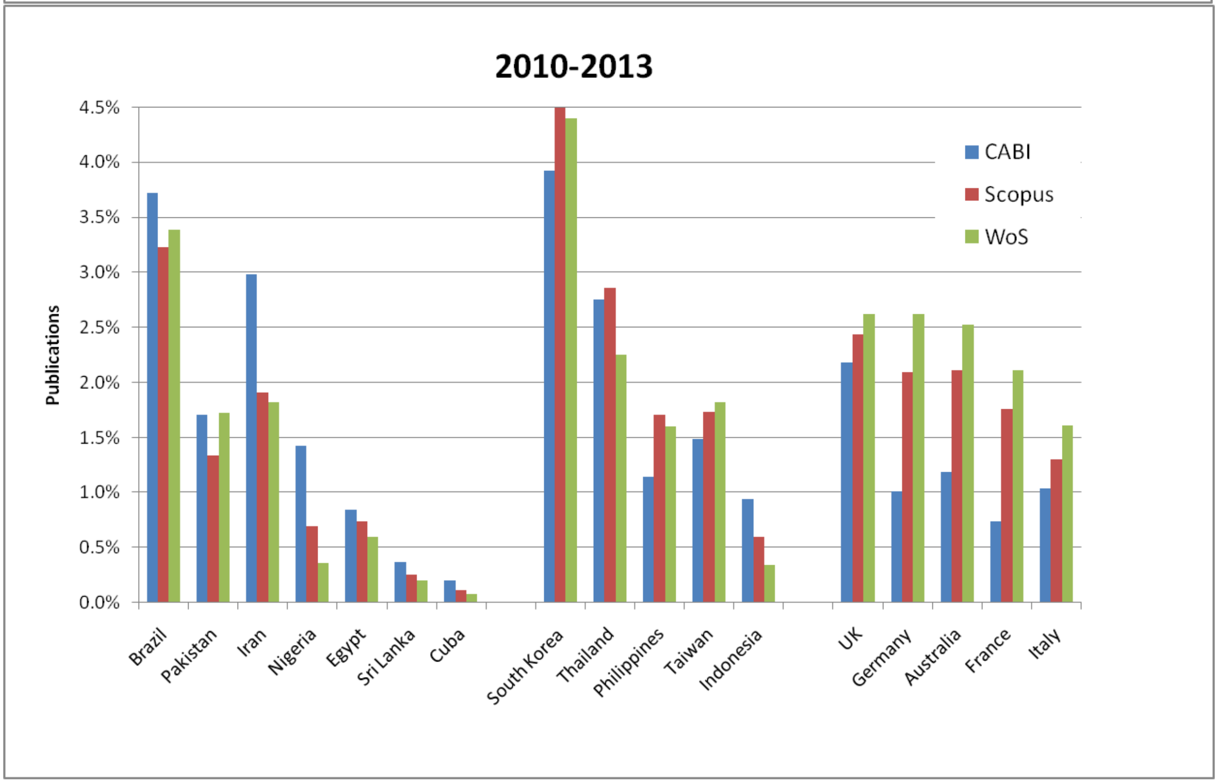


Figure 5 shows the percentage of publications for countries with smaller number of publications of rice for the three databases. As in the case of the large countries, we observe three patterns. On the left hand side, we show developing countries with much higher coverage in CABI. On the right hand side, we present industrialised countries with a much higher coverage in the WoS. In the middle, we have middle income countries from Asia that score similar shares in all three databases. In the latest period (2010-13) the differences between the countries are significantly reduced in some of the countries (e.g. in Brazil, perhaps due to incorporation of Brazilian journals into the commercial databases), but not in others (e.g. Iran and Nigeria). In the case of Western countries, the differences in coverage mostly remain, while in the Asian middle income countries no clear trend is observed.

From the analysis of Figure 4 and 5 it follows that WoS and Scopus cover research published in North-America and Europe, whereas CABI is more comprehensive. As a result of CABI's larger coverage, Western countries relative contribution to scientific production on rice is much smaller than is usually acknowledged when using standard publication databases such as WoS and Scopus.

There are a number of limitations in the empirical strategy adopted here. At this stage, we are not correcting the data for the fact that CABI only reports the affiliation of the first author. Second, although CABI overage of publication is possibly the largest on a subject such as rice (Kawasaki, 2004), publications still represent a subset of the research actually carried out on an applied fields such as agriculture. Particularly in agricultural technologies, many research outputs are not accounted for in publications, such as developments on the field, but also a lot of the research done by private companies and public organisations.

Conclusions

The results of this article suggest that previous assumption on the stability of indicators of scientific production are incorrect (Archambault et al., 2009). Instead, this case study on rice research shows that the indicator of number of publications is very dependent on the database when one analyses low and middle income countries. These preliminary results are potentially important for international organisations such as FAO, IFRI or UNESCO (UNESCO, 2010) that aim to work on human development.

Nevertheless, this finding does not come as a surprise given the proliferation in the last two decades of journal indexing systems at the regional level, such as Scielo[4] or Redalyc[5] that aim to provide visibility to local journals, often in languages other than English (Chavarro, 2013), precisely to compensate for fact that the local science and its journals are not perceived as participating in "international" science (Velho, 1986). A further important issue to address in this study of rice, is whether and how participation in different type of journals (indexed in CABI, Scopus or WoS) has an effect in the choice of research problems (Velho, 1990; Kreimer and Zabala, 2007).

[4] www.scielo.org

[5] http://www.redalyc.org

References

Archambault, É., Campbell, D., Gingras, Y., Larivière, V., 2009. Comparing bibliometric statistics obtained from the Web of Science and Scopus. Journal of the American Society for Information Science and Technology 60, 1320–1326.

Archambault, É., Vignola-Gagné, É., Côté, G., Larivière, V., Gingras, Y., 2006. Benchmarking scientific output in the social sciences and humanities: The limits of existing databases. Scientometrics 68, 329–342.

Chavarro, D., 2013. ¿ Son los sistemas de indexación y resumen un indicador de la buena calidad editorial de las revistas académicas? Presented at the RICyT Conference, Bogotá. Available at http://congreso2013.ricyt.org/files/mesas/1dProduccioncientifica/Chavarro.pdf. Accessed on 1st July 2014.

Kawasaki, J.L., 2004. Agriculture Journal Literature Indexed in Life Sciences Databases. Issues in Science and Technology Librarianship Summer 2004. Available at: http://www.istl.org/04-summer/article4.html. Accessed 1st July 2014.

Kreimer, P., Zabala, J.P., 2007. Chagas Disease in Argentina: Reciprocal Construction of Social and Scientific Problems. Science, Technology and Society 12, 49–72.

Leydesdorff, L., 2012. World shares of publications of the USA, EU-27, and China compared and predicted using the new *Web of Science* interface versus *Scopus*". *El profesional de la información*, 21(1), 43-49.

Royal Society, 2011. Knowledge, networks and nations: Global scientific collaboration in the 21st century. The Royal Society, London.

UNESCO, 2010. UNESCO science report 2010. The current status of science around the world. UNESCO, Paris.

Velho, L., 1986. The "meaning" of citation in the context of a scientifically peripheral country. Scientometrics 9, 71–89.

Velho, L., 1990. Sources of Influence on Problem Choice in Brazilian University Agricultural Science. Social Studies of Science 20, 503–517.

Research quality, characteristics of publications and socio-demographic features of Universities and Researchers: evidence from the Italian VQR 2004-2010 evaluation exercise

Tindaro Cicero*, Marco Malgarini* and Sergio Benedetto*

* *tindaro.cicero@anvur.org; marco.malgarini@anvur.org; sergio.benedetto@anvur.org.*
Italian National Agency for the Evaluation of Universities and Research Institutes (Italy)

Introduction

In July, 2013, ANVUR has published the results of the 2004-2010 Italian evaluation exercise (VQR 2004-2010 or simply VQR in the acronym used hereafter). The VQR Report has presented aggregate results relative to the quality of scientific publications submitted for evaluation by Italian Universities and Research bodies; the final objective of the Report was to rank Italian scientific institutions on the basis of the quality of their research, so as to provide to the Italian Ministry of Education, University and Research (MIUR) information to be used to assign a part of the public funding. The aim of this paper is that of providing a more disaggregated analysis of evaluation outcomes, specifically looking at possible existing correlations among scientific quality and a number of product- and researcher-specific variables. In the following, section 2 will briefly describe the adopted evaluation methods, while section 3 will present the econometric model used to study the relationship among research quality and its possible explicatives, commenting upon the results obtained. Section 4 concludes.

The VQR 2004-2010

The VQR exercise has been kick-started by the publication on the ANVUR Website (http://www.anvur.org/attachments/article/122/bando_vqr_def_07_11.pdf) of the Call for Participation (Bando di Partecipazione) on November 7, 2011. Research outcomes considered for evaluation were: Journal articles; Books, book chapters, conference proceedings (with ISBN codes); critical editions, translations, scientific comments; patents; compositions, designs, performances, work of arts and others. A total number of almost 185,000 research outcomes have been submitted for evaluation by the 61,822 Italian researchers on active duty (either with fixed term or permanent contract) on November 7, 2011 operating in the 14 research Areas defined by the Comitato Universitario Nazionale (CUN, see table 1).

Table 1 – Research outcomes submitted for evaluation by Area and type of publication.

Area*	Outcomes submitted	Journal Articles	%	Books, chapters	%	Conference Proceedings	%	Critical editions, translations	%	Patents	%	Others	%
01	9,682	8,455	87.3	465	4.8	731	7.6	3	0.0	6	0.1	22	0.2
02	19,386	18,105	93.4	181	0.9	934	4.8	1	0.0	22	0.1	143	0.7
03	11,812	11,608	98.3	85	0.7	70	0.6		0.0	46	0.4	3	0.0
04	7,229	6,308	87.3	418	5.8	304	4.2	2	0.0	6	0.1	191	2.6
05	17,298	16,690	96.5	313	1.8	245	1.4	1	0.0	36	0.2	13	0.1
06	27,085	26,266	97.0	521	1.9	242	0.9	1	0.0	39	0.1	16	0.1
07	9,866	8,649	87.7	583	5.9	591	6.0	1	0.0	31	0.3	11	0.1
08	9,657	4,202	43.5	3,943	40.8	1,350	14.0	7	0.1	16	0.2	139	1.4

09	17,654	14,329	81.2	566	3.2	2,590	14.7		0.0	106	0.6	63	0.4
10	13,966	3,707	26.5	7,998	57.3	1,979	14.2	196	1.4		0.0	86	0.6
11	13,158	5,032	38.2	7,295	55.4	720	5.5	60	0.5	3	0.0	48	0.4
12	11,886	3,992	33.6	7,433	62.5	379	3.2	5	0.0		0.0	77	0.6
13	11,765	7,286	61.9	3,964	33.7	467	4.0	1	0.0	1	0.0	46	0.4
14	4,434	1,278	28.8	3,056	68.9	82	1.8	6	0.1		0.0	12	0.3
Total	**184,878**	**135,907**	**73.5**	**36,821**	**19.9**	**10,684**	**5.8**	**284**	**0.2**	**312**	**0.2**	**870**	**0.5**

**01 Mathematics and computer sciences; 02 Physics; 03 Chemistry; 04 Earth Sciences; 05 Biology; 06 Medicine; 07 Agricultural and Veterinary Sciences; 08 Civil Engineering and Architecture; 09 Industrial and information engineering; 10 Antiquities, Philology, Literary Studies, Art History; 11 History, Philosophy, Pedagogy, Psychology; 12 Law studies; 13 Economics and Statistics; 14 Political and Social Sciences.*

Research outcomes have been evaluated by 14 Groups of Experts, one for each area, on the basis of the criteria of relevance (intended as contribution to the advancement of existing literature), novelty and innovation (intended as contribution to creation of new knowledge) and internationalization (intended as positioning of the research output in the international research landscape). On the basis of those criteria, each research product has been assigned to one out of four classes of merit, defined as follows:

A. Excellent: an outcome that falls in the top 20% of the world distribution according to international standards thanks to their originality, methodological rigor and interpretative relevance.
B. Good: an outcome that falls in following 20% of the world distribution according to international standards. Those outcomes have been recognized as relevant in the national and international debate for their contribution to the literature in the field.
C. Acceptable: an outcome that falls in the following 10% of the world distribution according to international standards. Those outcomes have been considered at the national and international level for their – albeit minimal – contribution to the literature in the field.
D. Limited: an outcome that falls in the lower 50% of the world distribution according to international standards. Those outcomes have been considered to provide a modest contribution to the literature in the field.

On the basis of this classification, each outcome was assigned an individual score, respectively equal to 1; 0,8; 0,5 and 0. Publications that were deemed as impossible to evaluate received a -1 score, while those considered as plagiarism or fraud got a -2 score. Missing outcomes with respect to what had to be expected for each researcher counted for -0,5. As for the methods used for evaluation, peer review is generally considered as the main way in which research outputs are evaluated by the scientific community. Peer review is however not immune from criticism: referee may be driven by opportunistic motivations (Frey, 2003) and the procedure may be ineffective in actually controlling for research quality (Baxt et al., 1998); peer review is also considered to be prone to penalize new and innovative theories and scholars in favor of well consolidated approaches, and to favor publications written in English with respect to other languages (Seidl, Schmidt, Grosche, 2005). Starting from those considerations, the VQR adopted a system of "informed peer review", in which for the Areas of Natural and Medical sciences, Mathematics, Engineering and, to some extent, Economics and Statistics, peer evaluation was integrated with the use of quantitative indicators concerning citations and journals' impact, extracted from the ISI/Web of Science and Scopus databases. According to Seglen (1997), the use of citation as a quality measure is based on the assumption that authors select references on the basis of quality. The use of a measure of journals' impact is particularly useful for most recent publications, for which citations count is not always accurate, and in order to reduce the risk of possible distortions caused by self-citations and opportunistic behaviors.

In any case, according to the Call, at least 50%+1 of research outcomes submitted for evaluation had to be peer-evaluated[1]. Evaluation results are presented in Table 2, the analysis being limited to outcomes presented by State Universities. In most Areas the share of excellent outcomes is larger than the top 20% defined as the share of excellent researches (see above). This should not come as a surprise: in fact, here we are not evaluating the overall Italian distribution of research outcomes, but only the three best researches that have been published by each author in the 2004-2010 period. As a consequence, the share of outcomes receiving an "excellent" evaluation is usually larger than the 20% to be expected from an analysis based on the complete distribution of Italian research outcomes.

Table 2 – Evaluation results by scientific Area (% shares)

	Score							
Area‡	**-2**	**-1**	**-0.5**	**0**	**0.5**	**0.8**	**1**	Total number of outcomes
1	0.00	0.50	8.75	15.63	12.64	20.90	41.59	8576
2	0.00	0.22	2.80	6.61	7.44	16.12	66.81	5930
3	0.01	0.04	1.57	9.33	6.86	25.25	56.94	7889
4	0.00	0.41	2.60	27.52	11.07	23.58	34.82	2918
5	0.00	0.94	2.81	22.82	10.02	23.26	40.15	12759
6	0.00	1.88	9.09	27.87	9.22	18.06	33.88	25470
7	0.01	0.66	1.73	29.78	8.65	16.54	42.62	7985
8	0.01	0.09	3.29	25.95	19.26	28.19	23.21	9332
9	0.00	0.24	2.39	14.20	10.27	19.07	53.82	13320
10	0.00	0.56	3.07	11.21	15.82	45.72	23.63	13100
11	0.02	0.58	2.38	20.60	21.23	34.60	20.59	11709
12	0.02	0.55	7.52	18.53	21.76	41.22	10.40	11658
13	0.00	0.33	5.46	50.56	14.61	12.22	16.82	10681
14	0.03	0.15	3.10	29.06	31.12	27.63	8.91	3930
Total	0.01	0.69	4.69	22.41	13.60	25.28	33.32	145257

‡01 Mathematics and computer sciences; 02 Physics; 03 Chemistry; 04 Earth Sciences; 05 Biology; 06 Medicine; 07 Agricultural and Veterinary Sciences; 08 Civil Engineering and Architecture; 09 Industrial and information engineering; 10 Antiquities, Philology, Literary Studies, Art History; 11 History, Philosophy, Pedagogy, Psychology; 12 Law studies; 13 Economics and Statistics; 14 Political and Social Sciences.

[1] See the VQR Report (http://www.anvur.org/rapporto/) and Ancaiani et al., 2014 for a more complete description of the methodology adopted.

Determinants of scientific performance

We assume that the probability of receiving a score equal to $x \in \{-2; 1\}$, may be influenced by three groups of variables, namely the characteristics of researches, researchers and the University:

$$P\ (Score_i = x) = F(Output\ characteristics_i, Researcher\ characteristics, Univerisity\ characteristics_i) \quad (1)$$

In (1), F is the cumulative function of the normal distribution; among the first group of variables we consider the type of outcome (Article; Book or book chapter; Proceeding; Other; Missing), the year and the language of publication (Italian; English; Other), the methodology of evaluation (Bibliometric, peer review or Informed peer review) and a binary variable equal to one when the research is co-authored with a non-Italian author. As for the characteristics of the researcher, we consider age, academic status (Full Professor; Associate Professor; other), gender and a binary variable equal to one when the researcher has been promoted or hired in the period considered. Finally, for the University we consider its location, age (distinguishing among Historical Universities, founded before 1945, Modern Universities, founded between 1946 and 1989 and Contemporary Universities, founded from 1990 onwards), size at the Area level in terms of number of outcomes presented for evaluation and the average amount of non-finalized government financing per capita received in the period considered. We also consider two indicators of academic specialization: the first takes a value comprised between zero and one, being equal to one if in a University all the 14 research Areas have the same weight in terms of research outcomes and being instead equal to zero if all research activity is concentrated in one Area; the second indicator is given in each Area by the ratio between the number of expected research outcomes in the area and the total research outcomes expected for the University. The former indicator is intended to capture the relationship among research quality and the specialization model adopted by the University as a whole, while the latter measures the relationship with academic specialization specific for each Area.

Model (1) is estimated as an Ordered Probit, an extension of the standard binary probit model used when the dependent variable takes the form of a ranked and multiple discrete variable. We normalize with respect to a missing product, evaluated with bibliometric methods, written neither in Italian nor in English, presented by a male Full Professor in Mathematics and informatics, with no mobility and no international co-author, employed in a small contemporaneous University located in the South of the country: i.e. the statistical significance, sign and magnitude of estimated parameters are to be interpreted as differentials with respect to this control group. First of all, the publication year has a different impact across the various areas (Table 3): recent researches are better evaluated in natural sciences and Medicine, while score is higher for outcomes published at the beginning of the period considered in Mathematics, Industrial engineering, Economics and Statistics and in the very heterogeneous Area 10; no effect is found in the remaining areas. Journal articles usually obtain better evaluations in Earth science, Medicine, Agricultural and veterinary sciences, Civil engineering and Architecture, Law and Economic and Statistics; on the other hand, Books and books chapters have a significantly better evaluation in Mathematics, Medicine, Agricultural and veterinary sciences, History and Philosophy, Law and Economics and statistics; conference proceedings obtain more favorable evaluation in Mathematics, Law and Economic and statistics and are instead penalized in Chemistry, Earth sciences and Biology. As for the language of publication, researches published in Italian are usually negatively evaluated in all areas, while English-language publications are usually rewarded with better evaluation, confirming the findings in previous studies both for peer review (i.e. Nylenna et al., 1994) and citation analysis (Poomkottayil et al., 2011). Universities operating in the Center-North usually obtain better evaluations; on the other hand, the effect of the year of foundation of the University is highly differentiated: outcomes from historical Universities are better evaluated in Physics and History and Philosophy and are instead penalized in Chemistry, Earth Science and Agricultural and veterinary sciences; no effect is found in the remaining areas. Similarly, modern Universities have better results in Medicine and History and Philosophy and are instead penalized in Earth science. As for funds received by the government, higher evaluation results are usually associated with higher financing in all natural sciences, while the effect is statistically insignificant in non-bibliometric areas, where indeed research is not in need of special equipment and technologies to be performed. The effect of specialization or de-specialization of research activities is usually insignificant; however, in Civil engineering and Architecture research outputs receive a better evaluation if presented from specialized Universities, while the contrary is true in Earth sciences; if we look more specifically at the specialization in the field under consideration, there is a positive correlation among research quality and specialization in Agricultural and veterinary sciences, History and philosophy and Social sciences; a negative correlation emerges instead in Chemistry and Earth sciences. University size in terms of expected research products is usually statistically insignificant; the only exception are Chemistry, where researches presented by medium and large size Universities receive a better evaluation, and Agriculture and veterinary sciences and Social sciences where the opposite is true. Finally, looking at socio-demographic characteristics of the researcher, being hired or promoted has a positive effect on research quality in Mathematics and Medicine and a negative one in Physics, Earth sciences, Industrial engineering, Humanities and Social sciences (no effect is found in the remaining areas). Researches presented by Full professors are usually better evaluated than those submitted by Associate Professors or Researchers; however, *ceteris paribus*, younger researchers usually receive better evaluations. Significant gender effects also emerge, with researches submitted by women receiving a more negative evaluation in various areas. The latter, rather puzzling, result has already been found in relation to research productivity in various countries (Larivière et al., 2013; Larivière et al., 2011; Frietsch et al., 2009; Mauleón and Bordons, 2006) and may be attributable to various factors mainly linked to the presence of young children in the family and other personal characteristics (Stacks, 2004).

Table 3 –Ordered probit model for research score and its possible correlates.

	Hard sciences								
Explicative variables†	**01**	**02**	**03**	**04**	**05**	**06**	**07**	**08**	**09**
Information about the research products									
PY	-0,018***	0,055***	0,079***	0,035***	0,059***	0,014***	-0,005	0,003	-0,035***
JA	0,357	-0,134	-0,207	-0,670**	0,457*	1,067***	0,901***	0,248**	-0,094
BBC	0,689***	0,252	-0,396	-0,407	0,334	1,224***	0,760***	0,424***	-0,227
P	0,601***	-0,388	-1,137***	-1,676***	-1,162***	-0,216	0,250	-0,098	-0,209
IT	-0,412***	-0,419**	-0,430**	-0,619***	-0,478***	-0,275***	-0,413***	-0,093**	-0,713***
ENG	0,019	-0,135**	0,021	0,100	-0,018	-0,035	0,187***	0,091**	0,051
INT	0,296***	0,694***	0,438***	0,540***	0,539***	0,605***	0,447***	0,294***	0,320***
IPR	-0,632***	-1,376***	-1,050***	-0,494***	-0,725***	-0,804***	-1,056***	-0,948***	-1,287***
PR	-1,435***	-1,890***	-1,483***	-1,830***	-1,434***	-1,664***	-1,415***	-1,329***	-1,913***

Information about the University									
NW	0,465***	0,278***	0,139***	0,004	0,436***	0,326***	0,120***	0,170***	-0,017
NE	0,424***	0,226***	0,362***	0,267***	0,470***	0,326***	0,236***	0,204***	0,149***
C	0,447***	0,222***	0,307***	0,332***	0,353***	0,237***	-0,029	0,236***	-0,001
HU	-0,115*	0,231***	-0,229***	-0,269***	-0,025	0,057	-0,291***	0,065	0,087*
MU	0,003	0,148*	-0,019	-0,294***	0,019	0,109***	-0,127*	0,064	0,037
UF	-0,101	0,393**	-0,080	0,672***	0,320***	0,901***	0,318	-0,022	0,268***
AC	0,299	-0,406	-0,667*	3,055***	0,581*	0,974*	1,965***	-0,393**	-0,337
AS	1,399	1,780	-3,356***	-6,280**	-0,791*	0,008	1,355***	0,102	0,037
MSU	0,028	-0,049	0,421***	-0,121	-0,093	0,006	-0,318***	0,115***	0,066
BU	0,117*	-0,053	0,423***	-0,106	-0,015	0,057	-0,183*	0,053	0,078
Information about the researcher									
WM	0,096***	-0,118***	-0,044	-0,141***	0,014	0,056***	0,039	-0,012	-0,081***
AP	-0,473***	-0,579***	-0,473***	-0,446***	-0,415***	-0,288***	-0,302***	-0,393***	-0,303***
RES	-0,829***	-0,958***	-0,820***	-0,792***	-0,749***	-0,593***	-0,503***	-0,641***	-0,500***
OTH	-1,019***	3,088	0,339	-1,245***	-0,625***	-0,544***	-0,576	-0,530**	-0,508**
AGE	0,047***	0,054***	0,042***	0,043***	0,037***	0,022***	0,022***	0,030***	0,028***
W	-0,178***	-0,136***	-0,133***	0,020	-0,015	0,035**	0,015	0,049*	0,041

	Social sciences				
Explicative variables	**10**	**11**	**12**	**13**	**14**
Information about the research products					
PY	-0,020***	-0,006	0,000	-0,027***	0,001
JA	0,063	0,057	0,667***	1,407***	0,505*
BBC	0,074	0,265**	0,674***	1,286***	0,507*
P	0,020	0,245*	0,629***	0,764***	0,280
IT	-0,113***	-0,232***	-0,091***	-0,721***	-0,279***
ENG	0,162***	0,084**	0,228***	0,337***	0,219***
INT	0,150**	0,420***	-0,041	0,590***	0,286***
IPR		-0,298***		-0,106	
PR		-0,774***		-0,783***	
Information about the University					
NW	0,300***	0,370***	0,257***	0,295***	0,323***
NE	0,300***	0,368***	0,190***	0,336***	0,204***
C	0,244***	0,298***	0,200***	0,223***	0,141***
HU	-0,001	0,201***	0,102***	0,026	0,118
MU	-0,012	0,209***	-0,030	-0,0663	-0,189***
UF	-0,094	-0,056	0,009	-0,078	0,274
AC	-0,383	0,085	0,149	-0,211	0,587
AS	0,086	4,360***	0,442	-0,266	4,968***
MSU	-0,062	-0,014	0,062		-0,267**
BU	-0,015	-0,042	0,000	0,018	-0,245**

Information about the researcher					
WM	-0,082***	-0,016	-0,021	0,014	-0,520***
AP	-0,541***	-0,387***	-0,684***	-0,188***	-0,835***
RES	-0,947***	-0,635***	-1,088***	-0,380***	-0,993***
OTH	-1,495***	-0,812***	-1,172***	-0,237	0,026***
AGE	0,039***	0,025***	0,028***	0,019***	-0,086**
W	-0,121***	-0,025	-0,010	-0,082***	-0,520***

†PY: publication year; JA: journal article; BBC: book, book chapter; PR: proceedings; IT: Italian; ENG: English; INT: International coauthors; IPR: Informed peer-review; PR: peer review; NW: north west; NE: north east; C:center; HU: historical universities; MU: modern universities; UF: university funds; AC: academic concentration; AS: area specialization; MSU: medium-size universities; BU: big universities; WM: work mobility; AP: associate professor; RES: researcher; OTH: other academic ranks; AGE: year born; W:woman.
**** Statistical significant at 1%; ** Statistical significant at 5%; * Statistical significant at 1%.*

Conclusions and further research

The VQR provides invaluable information about scientific research that has been produced in Italian University in the period 2004-2010. In this paper we have related the rating in the assessment exercise to socio-demographic characteristics of the researcher and the University and to intrinsic characteristics of the publication. The analysis shows that ratings crucially depend on language and typology of publication and by the methodology adopted for evaluation; also the personal characteristics of the author submitting the research counts, with younger researchers and Full Professors receiving, *ceteris paribus*, a better evaluation. A negative gender effect for women also emerges, probably attributable to personal characteristics linked to child care and network externalities. Last but not least, outcomes submitted by researchers working in the Centre-North of the country usually obtain better evaluations than those in the South; on the other hand, the effect of size, age and scientific specialization of the University is not clear-cut, being positive in some areas and negative or statistically insignificant in others. Public funding to University research is finally found to be correlated with positive ranking in the assessment exercise, especially in natural sciences and engineering, where there is particular need of appropriate funding for conducing laboratories experiments and research.

References

Baxt, W.G., Waeckerle, J. F., Berlin, J. A., & Callaham, M. L. (1998). Who reviews the reviewers? Feasibility of using a fictitious manuscript to evaluate peer reviewer performance. *Annals of emergency medicine*, 32(3), 310-317.

Frey, B.S. (2003). Publishing as prostitution?–Choosing between one's own ideas and academic success. *Public Choice*, 116(1-2), 205-223.

Frietsch, R., Haller, I., Funken-Vrohlings, M., & Grupp, H. (2009), Gender-specific patterns in patenting and publishing, *Research Policy*, 38(4), 590-599.

Lariviere, V., Vignola-Gagne, E., Villeneuve, C., Gelinas, P., Gingras, Y. (2011), Sex differences in research funding, productivity and impact: An analysis of Quebec university professors, *Scientometrics*, 87(3), 483–498.

Lariviere, V., Ni, C., Gingras, Y., Cronin, B., Sugimoto C.R. (2013), Bibliometrics: Global gender disparities in science, *Nature*, vol. 504 (7479). Available on: http://www.nature.com/news/bibliometrics-global-gender-disparities-in-science-1.14321.

Mauleon, E., & Bordons, M. (2006), Productivity, impact and publication habits by gender in the area of Materials Science, *Scientometrics*, 66(1), 199-218.

Nylenna, M., Riis, P., & Karlsson, Y. (1994). Multiple blinded reviews of the same two manuscripts: effects of referee characteristics and publication language. *JAMA*, 272(2), 149-151.

Poomkottayil, D., Bornstein, M. M., & Sendi, P. (2011). Lost in translation: the impact of publication language on citation frequency in the scientific dental literature. *Swiss medical weekly*, 141, 30.

Schmidt, V., Seidl, C., & Grösche, P. (2005). The performance of peer review and a beauty contest of referee processes of economics journals. *Estudios de economía aplicada*, 23(3), 505-554.

Seglen, P.O. (1997) Citations and journal impact factors: questionable indicators of research quality. *Allergy*, 52, 1050-1056.

Stack, S. (2004). Gender, Children and research Productivity, *Research in Higher Education*, 45(8), 891-920.

From publications to people: bibliometric benchmarking of a selection of countries in the Life Sciences based on individual-level bibliometrics[1]

Rodrigo Costas* and Ed Noyons**

[*] *rcostas@cwts.leidenuniv.nl*
Center for Science and Technology Studies (CWTS), Leiden University, P.O. Box 905, Leiden 2300 AX (The Netherlands)

[**] *noyons@cwts.leidenuniv.nl*
Center for Science and Technology Studies (CWTS), Leiden University, P.O. Box 905, Leiden 2300 AX (The Netherlands)

Introduction

In bibliometrics it is not uncommon to benchmark bibliometric units such as countries or universities based on the output attached to them through the affiliation addresses included in the publications. These data on addresses are directly attached to publications and it is relatively easy to collect the scientific production of a country or any big unit of analysis. However, more difficult is the comparison of different units based on the collective performance of the different scholars affiliated to them (e.g. applying a more bottom-up approach, van Leeuwen, 2007). Research organizations (or even countries) can change or disappear and, although facilitators, they are not the final producers of the new scientific knowledge. It's the scholars working for these organizations who do the research (Bornmann & Marx, 2013). However, one of the most important reasons for the underdevelopment of such studies is the lack of accurate and extensive data on individual scholars.

In spite of these limitations an attempt was done (Zuccala et al, 2010) already showing the potential of this new type of studies. Also, Danell (2013) has shown that different perspectives on the analysis of productivity can provide different results. In this paper we present the results of a broad benchmark study of countries based on the individual performance of the scholars affiliated to them. The main focus is not the "performance of the country" but the "performance of the individuals working in the country". The main objective of this paper is to present the methodological approach and main results of a first explorative extensive benchmark study of countries based on the performance of the scholars that can be attached to them.

Besides the more advanced way of assigning research output to entities (countries, organizations), the approach facilitates statistically more advanced analyses, using distributions.

[1] This work was supported by the Crucell Vaccine Institue. The authors of this report are also very grateful to Jessica Meijer, Dick den Os and Jaco Klap from Crucell Vaccine Institute, Center of Excellence for Immunoprophylaxis, Johnson & Johnson for all their feedback, comments, support and suggestions during the development of this paper.

Methodology

For this analysis we consider all the individual scholars identified in a more extensive study carried out at CWTS of active scholars in the field of Life Sciences (LS) during the period 1980-2011 (Costas & Noyons, 2013)[2]. In that study we took advantage of an algorithm recently developed at CWTS for author name disambiguation of all the author names currently in the Web of Science database (Caron & van Eck, 2014). This algorithm has shown substantial good results in terms of precision and recall values (over 90% in both), and we consider this a suitable tool for our approach. In addition to this author name disambiguation algorithm we have also collected information on the addresses of the individuals identified. This linkage of authors to affiliations is based on 'known' linkages between authors and countries found in scientific publications (e.g. reprint authors, direct linkages of authors with affiliations, e-mail data, publications with only one affiliation, etc.). As a result we have the 'oeuvres' of the different individuals as well as their different affiliations.

For this paper we only took scholars into consideration when they have as their most common certain address (MCAD) any of the countries selected for this study. In other words, we study scholars whose most frequent 'certain affiliation' in their publications is located in that country. An alternative approach could be to use the most recent affiliation of authors to assign them to the different countries. Individuals are assigned to one country only although we realize that they may need to be assigned to more than one (e.g. scholars have sometimes double appointments and affiliations, sometimes they have publish more papers in one affiliation while they spent more time in another one, etc.). For the interpretation this needs to be taken into account but for the purpose of this study and given its high aggregated level we expect that such issues will be cancelled out. More research is necessary to shed more light on this aspect.

The countries selected for this study are the Netherlands, Belgium, United States, China, Germany, Sweden, Denmark, Poland, Finland, Switzerland, Brazil and the United Kingdom.

Results

Descriptive analysis

In this section we provide some general statistics of the population of scholars included in the analysis. In Table 1 we can see a first overview of the identified LS researchers by country (columns total, %country sample and %world scholars). US scholars dominate the analysis as they represent more than 49% percentage of all scholars in the study and more than 22% of all worldwide LS scholars. Germany and UK are second, each representing more than 10% of the scholars in the set of countries and more than 9% (together) of all world's scholars. It should be noted that with the selection of countries we cover almost 45% of all world's scholars active in LS as defined in Costas & Noyons (2013).

Top performance analysis

In this part we focus on the number of top performers in each country. We established several typologies of scholars by means of a classificatory approach (similar to the one implemented by Costas et al, 2010). Thus we were able to identify scholars who can be considered as "top producers", i.e. scholars that are among the 25% most productive scholars worldwide in LS;

[2] For some of the most important methodological details such as the selection of individuals, their linkage of authors to affiliations, bibliometric indicators, citation window and analysis of the different typologies we refer to Costas & Noyons (2013).

"top toppers", i.e. scholars among the 25% most productive, the 25% most productive of highly cited publications, and also among the 25% scholars publishing in the best journals. A third typology of scholars are the "high impact" scholars, i.e., those belonging to the top 25% of the world in terms of publishing highly cited publications and impact of their journals, but belonging to the segment of producers between the median and percentile 25. We consider that these typologies of scholars are relevant and have a research policy value, however we acknowledge that these are not the only possible typologies (e.g. Seiler & Wohlrabe, 2013). In this paper we will focus on these three typologies and leave for further research the exploration of other typologies of scholars.

Top producers

Table 1 and Figure 1 present the results of the analysis of the 'top producers' active across the different countries.

Table 1. Top producers analysis by countries

country	total	%country sample	%world scholars	Top producers	%top producers within country	%top producers within sample	%top producers worldwide
BELGIUM	12,008	2.0%	0.92%	3,511	29%	2.2%	1.1%
BRAZIL	28,798	4.9%	2.20%	4,831	17%	3.0%	1.5%
DENMARK	8,972	1.5%	0.69%	2,622	29%	1.6%	0.8%
FINLAND	9,495	1.6%	0.73%	2,672	28%	1.7%	0.8%
GERMANY	62,515	10.6%	4.77%	18,261	29%	11.4%	5.6%
NETHERLANDS	26,083	4.4%	1.99%	7,561	29%	4.7%	2.3%
PEOPLES R CHINA	46,119	7.9%	3.52%	5,115	11%	3.2%	1.6%
POLAND	10,818	1.8%	0.83%	1,978	18%	1.2%	0.6%
SWEDEN	18,180	3.1%	1.39%	4,709	26%	2.9%	1.4%
SWITZERLAND	13,953	2.4%	1.07%	3,929	28%	2.4%	1.2%
USA	289,494	49.3%	22.11%	87,517	30%	54.4%	26.7%
UK	60,900	10.4%	4.65%	18,150	30%	11.3%	5.5%
Total selected countries	587,335	100.0%	44.85%	160,856	27%	100.0%	49.1%
Total worldwide	1,309,458		100.00%	327,375	25%		100.0%

Figure 1: Share of top producers across countries – red line: international threshold, green line: set of countries threshold.

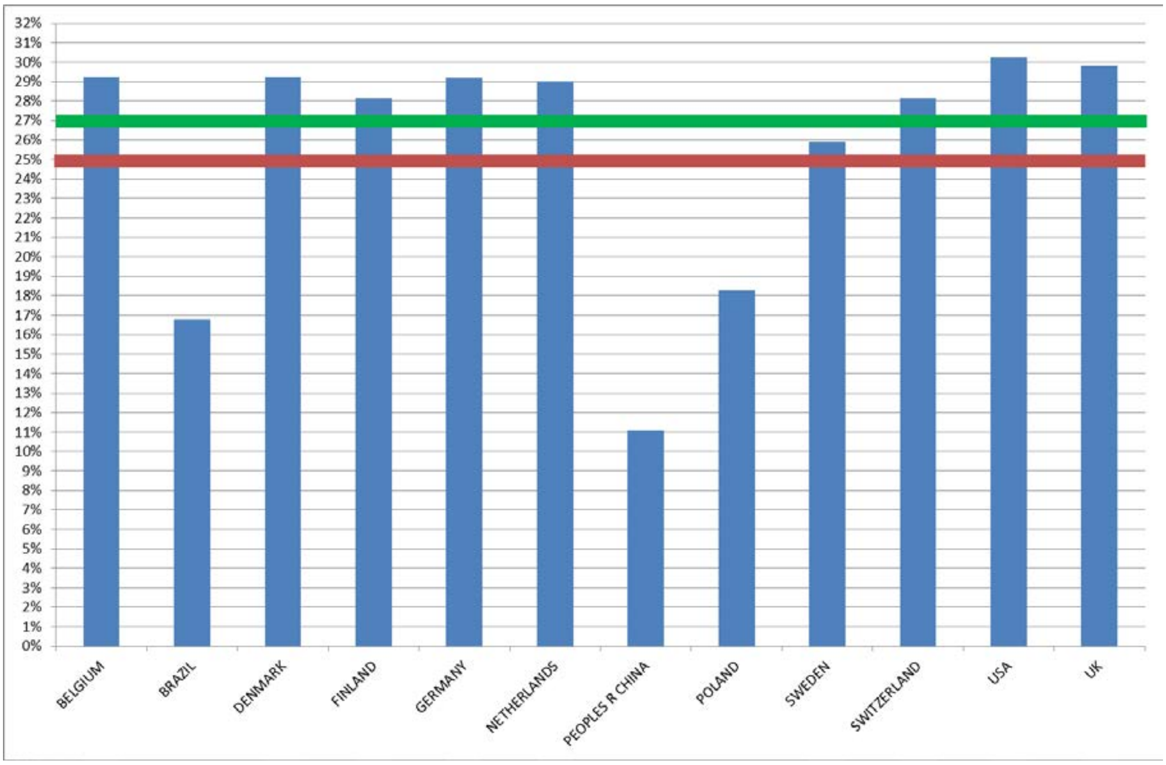


Table 1 shows how 25% of the world scholars are top producers (which is not a surprise, as this is the threshold). In the set of countries selected 27% of the scholars are top producers which means that the selection of countries have proportionally more top producers than would be expected as set by all the researchers included in the analysis. Hence, it is more difficult to become a top producer in our selected set of countries than in LS worldwide. According to Figure 6 we see that most countries in our set (including small countries like the Netherlands or Belgium) have high shares of top producers. The US and the UK have the highest shares of top producers (~30%). China, Brazil and Poland are the countries with the lowest shares. It should be noted that only through analyses as presented here, we are able to investigate productivity, under the assumption that other factors are equal for all researchers in our set.

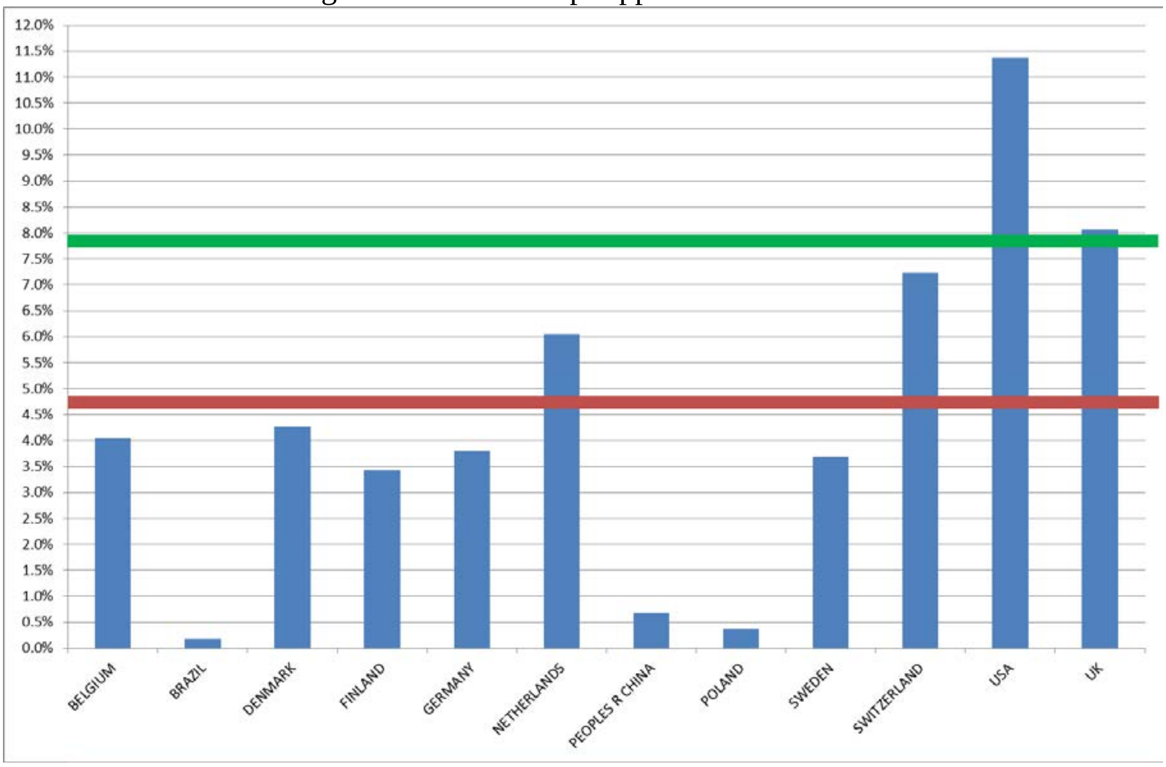
Top toppers

In simple words 'top toppers' are those top producers that also have a high impact. In Table 2 and Figure 2 we analyse the 'top toppers' across the countries. In figure 2, the red line indicates the international share of top toppers and the green line indicates the share of top toppers within the set of the selected countries.

Table 2. Analysis 'Top toppers' by countries

country	total	Top toppers	%top toppers within country	%top toppers within sample	%top toppers worldwide
BELGIUM	12,008	486	4.0%	1.1%	0.8%
BRAZIL	28,798	51	0.2%	0.1%	0.1%
DENMARK	8,972	382	4.3%	0.8%	0.6%
FINLAND	9,495	326	3.4%	0.7%	0.5%
GERMANY	62,515	2,374	3.8%	5.3%	3.9%
NETHERLANDS	26,083	1,576	6.0%	3.5%	2.6%
PEOPLES R CHINA	46,119	310	0.7%	0.7%	0.5%
POLAND	10,818	40	0.4%	0.1%	0.1%
SWEDEN	18,180	670	3.7%	1.5%	1.1%
SWITZERLAND	13,953	1,009	7.2%	2.2%	1.6%
USA	289,494	32,912	11.4%	73.1%	53.5%
UK	60,900	4,914	8.1%	10.9%	8.0%
Total selected countries	587,335	45,050	7.7%	100.0%	73.2%
Total worldwide	1,309,458	61,567	4.7%		100.0%

Figure 2: Share of 'top toppers' across countries.



In table 2 we see that the overall share of top toppers worldwide is 4.7% while the share of top toppers in the sample of the selected countries is 7.7%. The US is the country with the highest share of top toppers, hosting more than 50% of them (as presented in the last column of table 2). Among our set of countries (Figure 2), the US is the country with the highest share, with almost 11% of the researchers of this country being top toppers. The second best countries are the UK and Switzerland, with more than 7% of their scholars within this very competitive type of scholars. The Netherlands comes fourth with 6% but still below the average within our set of countries. These four countries are the only ones (within our set of countries) that present a share of top toppers above the average share top toppers worldwide of 4.7% .

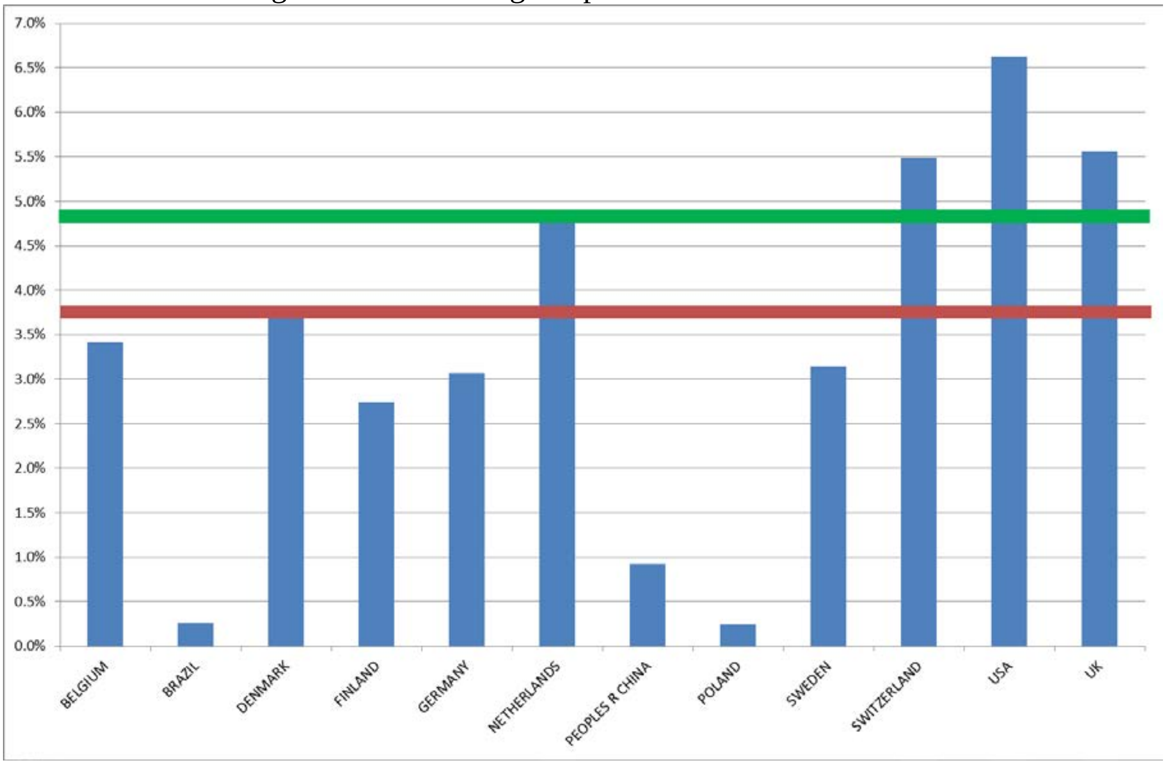
High impact

In this third section we analyze the presence of 'high impact' scholars across countries. Table 3 presents the main scores and Figure 3 the comparison of the countries.

Table 3. Analysis High impact by countries

country	total	High impact	%high impact within country	%high impact within sample	%high impact worldwide
BELGIUM	12,008	410	3.4%	1.4%	0.8%
BRAZIL	28,798	75	0.3%	0.3%	0.2%
DENMARK	8,972	340	3.8%	1.2%	0.7%
FINLAND	9,495	260	2.7%	0.9%	0.5%
GERMANY	62,515	1,919	3.1%	6.7%	3.9%
NETHERLANDS	26,083	1,253	4.8%	4.4%	2.6%
PEOPLES R CHINA	46,119	428	0.9%	1.5%	0.9%
POLAND	10,818	26	0.2%	0.1%	0.1%
SWEDEN	18,180	572	3.1%	2.0%	1.2%
SWITZERLAND	13,953	766	5.5%	2.7%	1.6%
USA	289,494	19,180	6.6%	67.0%	39.1%
UK	60,900	3,385	5.6%	11.8%	6.9%
Total selected countries	587,335	28,614	4.9%	100.0%	58.3%
Total worldwide	1,309,458	49,109	3.8%		100.0%

Figure 3: Share of high impact scholars across countries.



From table 3 and figure 3 we get a similar picture as with the top toppers. We see how the US, UK and Switzerland are the countries with the highest proportions of ‘high impact’ scholars and the Netherlands coming fourth place. These four countries exceed the international average share of high impact scholars, with Denmark just below. Also here the emerging countries (Brazil, China and Poland) are the ones with the lowest rates. In this case the Netherlands reaches the average of selected countries, indicating that this country scores better on this indicator.

Conclusion and Discussion

In this paper we present a novel approach of benchmarking research entities (countries, institutes) based on the performance of the individual scholars linked to them as a complement to the regular analysis of the overall output of countries. This new approach enhances the analytical possibilities and provides multiple and new perspectives and interpretations of the differences among countries regarding their scientific activity and productivity. We move a step forward from the analysis of publications that can be attributed to countries to the analysis of individuals that can with some certainty be linked to those countries.

In general terms, we see countries with different patterns in terms of the performance of individuals. The analysis based on typologies of researchers like for example the ‘high impact’ typology as opposed to top toppers or top producers, may unveil aspects of a country’s or an organization’s performance which cannot be derived from the usual bibliometric analyses. For example, the fact that the Netherlands performs comparatively better in terms of ‘high impact’ scholars than with top toppers may indicate that for this country the focus of scholars is less on production , but more on high impact (i.e. high impact or more ‘selective’ researchers).

All in all, this study opens important and challenging new pathways for research performance analysis. Firstly, it makes possible to move from a publication-based perspective to a more individual-level based approach. This change makes possible the analysis of research questions more targeted to individuals or groups and can help to identify aspects that would be more difficult to grasp from a regular publication-based approach. Secondly, the analysis of different typologies of scholars may help to expand the debate on the research policies across countries, showing that the 'publish or perish' approach is not the only one but that there are other typologies of performance that may be also relevant. In any case, the development of individual-level bibliometric approaches is an open topic. Further research should focus on problems still related with data collection (Smalheiser & Torvik, 2009), the possibilities of improving the approaches here used (e.g. through the 'Characteristics Scores and Scale' approach, Ruiz-Castillo & Costas, 2014) and the more conceptual discussion on the consideration of individuals bibliometrically (e.g. Glänzel & Wouters, 2013).

References

Bornmann, L., & Marx, W. (2013). How to evaluate individual researchers working in the natural and life sciences meaningfully? A proposal of methods based on percentiles of citations, *Arxiv*, http://arxiv.org/abs/1302.3697

Caron, E. & van Eck N.J. (2014). Large scale author name disambiguation using rule-based scoring and clustering", CWTS Working Papers.

Costas, R. & Noyons, E. (2013). Detection of different types of 'talented' researchers in the Life Sciences through bibliometric indicators: methodological outline. CWTS Working papers. http://www.cwts.nl/pdf/CWTS-WP-2013-006.pdf

Costas, R., van Leeuwen, T. N. & Bordons, M. (2010). A Bibliometric Classificatory Approach for the Study and Assessment of Research Performance at the Individual Level : The Effects of Age on Productivity and Impact, *Journal of the American Society for Information Science and Technology*, 61: 1564–1581.

Danell, R. (2013). Productivity differences influence Swedish universities' citation-based university rankings. An analysis of the relationship between work productivity and citation performance. 18th International Conference on Science and Technology Indicators "Translational twists and turns: Science as a socio-economic endeavor". Berlin: iFQ. Pp.96-103.

Glänzel, W. & Wouters, P. (2013). The dos and don'ts in individual-level bibliometrics. 14th International Society of Scientometrics and Informetrics (ISSI). Vienna: Universtät Wien. http://www.slideshare.net/paulwouters1/issi2013-wg-pw

Ruiz-Castillo, J. & Costas, R. (2014). The skewness of scientific productivity. UC3M Working Papers.

http://e-archivo.uc3m.es/bitstream/handle/10016/18286/we1402.pdf?sequence=3

Seiler, C. & Wohlrabe, K. (2013). Archetypal scientists. *Journal of Informetrics*, 7(2): 345-356

Smalheiser, N.R. & Torvik, V.I. (2009) Author name disambiguation. *Annual Review of Information Science and Technology* (B. Cronin, Ed.), pp. 287-313.

Van Leeuwen, T.N. (2007). Modelling of bibliometric approaches and important of output verification in research performance assessment. *Research Evaluation*, 16(2): 93-105

Zuccala, A.; Costas, R. & van Leeuwen, T.N. (2010). Evaluating research departments using individual level bibliometrics. Eleventh International Conference on Science and Technology Indicators. Leiden: University.

http://www.cwts.nl/pdf/BookofAbstracts2010_version_15072010.pdf

The contribution of authors: A study of the relationship between the size and composition of author teams

Rickard Danell

Rickard.danell@soc.umu.se
Department of sociology, Umeå university, SE-90187, Sweden

Abstract
Co-authorship is the norm in most branches of science. With an increasing number of individuals claiming authorship to the same article, however, it is less clear what being an author actually means; i.e. what contributions merit the authorship. In this paper I present the results of an analysis of author contribution statements published in *Nature Neuroscience*; the purpose was to investigate what type of contributions merit authorship, and what happens to the composition of the author team when it increases in size. I classified all tasks according to a scheme suggested by Davenport & Cronin (2001); core author tasks, middle layer tasks, and outer layer tasks. The results show that the largest share of authorship is core authors, but that a substantial share of the authors belong to the outer layer. An analysis of changes in the composition of author teams as they increase in size revealed that the increase in team size was mainly driven by an increase in middle layer authors, indicating a process of functional differentiation in author teams due to the increasing complexity of the work tasks.

Introduction
The size of author teams varies from one to several hundred, or in some extreme cases over a thousand. In most fields of science and technology, papers by a single author are unusual–almost a thing of the past–and in both science and the social sciences, the average number of authors per article is increasing over time (Beaver & Rosen 1978; Endersby 1996; Persson, Glänzel & Danell 2004). There may be multiple factors behind this increase in co-authorship that explain the shift in collaboration patterns. However, a reasonable assumption is that the increasing size of authorship teams is related to the increasing social and cognitive complexity of modern science (McDowell & Melvin 1983; Jeong, Choi & Kim 2011; Nowell & Grijalva 2011).

In biomedicine the increase in the number of authors per publication is a source of concern and debate among editors and scientists (Smith 1997; Flanagin et al. 1998; Rennie & Yank 1998; Shewan & Coats 2010). The intense discussion about authorship concerns ethics and accountability–or the lack thereof–in research, and motivated the International Committee of Medical Journal Editors (ICMJE) to formulate an official position regarding criteria for authorship (Biagioli 1999). The ICMJE authorship standard is an attempt to strengthen the idea of an author as the individual responsible for the content of a text, and can be interpreted as a defence of the traditional view. Others have argued that this idea of an author is obsolete, and that it should be replaced with a notion of a contributor; and that the list of authors should be replaced with a list of contributors resembling the list of contributors found in other complex cultural products such as films or plays (Rennie et al. 1997; Rennie & Yank 1998).

In science studies and information science there is increasing research concerning scientific authorship. Notable examples are survey studies describing the prevalence of authors not conforming to ICMJE criteria for authorship (Flanagin et al. 1998); pioneering empirical and theoretical work done by Blaise Cronin and his co-authors (Cronin 2001; Cronin, Shaw & La Barre 2003; Cronin 2004; Cronin 2005); and Birnholtz' (2006) investigation into what it mean to be an author in high energy physics. It should also be mentioned that Cronin (1995) extends the notion of an author when he goes beyond individuals listed as authors, and emphasizes the individuals listed in the acknowledgement as a kind of sub-author. There is a need for research, however, into what type of contributions motivate authorship, and how author teams are composed with regard to different types of contributions.

The results presented in this paper are based on an analysis of author contributions published in *Nature Neuroscience*. The aim of this study was to investigate what kind of contributions merit authorship, and what happens to the composition of the author team when it increases in size.

Data and method

The data consist of author contribution statements retrieved from articles published in *Nature Neuroscience* in 2012 and 2013. A total of 275 articles were checked for author contribution statements; statements were found in 237 of them. Among the publications where author contribution statements were missing were articles by one author, or specific document types such as brief communication, commentary, perspectives, or reviews.

Nature Neuroscience, like all journals from the *Nature* publishing group, requires authors to include a statement that specifies the contribution of every author. One example of a contribution statement:

"This study was designed, directed and coordinated by F.C.d.A. and L.-H.T. L.-H.T., as the principal investigator, provided conceptual and technical guidance for all aspects of the project. F.C.d.A. planned and performed the in utero electroporations and analyzed the data with A.L.R. and O.D. F.C.d.A. performed and analyzed the immunohistochemistry experiments. K.M. generated and characterized the shRNA constructs. K.M., A.L.R. and D.R. contributed to the neuronal cultures. T.T. performed and analyzed the data from the neuronal cultures of Nrp1Sema- mice. A.L.R., O.D., J.G. and R.M. contributed to the biochemistry experiments. T.S. generated the lentiviral shRNA construct and produced the virus particles. D.D.G. and A.L.K. provided the Nrp1Sema- mouse brains and suggested and commented on the design of the experiments. The manuscript was written by F.C.d.A. and L.-H.T. and commented on by all authors." (de Anda et. al 2012)

The descriptions of the authors' contributions varies between manuscripts. *Nature* allows two co-authors to be specified as having contributed equally to the work, but prefers clear statements of author contributions. The length of the descriptions produced by the authors varies from 45 to 1323 characters, with an average length of 378 characters per statement. As expected, more specific information concerning the contribution of each author is provided in articles with more authors.

Classifying authors according to their contribution was carried out in several steps. In order to classify author contributions, the contribution statements divided into specific work tasks; I used regular expressions to extract the authors from the descriptions of the work tasks. After that, I attributed the work task to all authors mentioned as performing the task. After constructing a database consisting of authors linked to work tasks, I classified all tasks according to a scheme suggested by Davenport & Cronin (2001); core author tasks, middle layer tasks, and outer layer tasks. Table 1 describes the classification scheme used.

Table 1. A three-tier author contribution taxonomy

Type of contribution	Examples
Core task	Conception and design Writing the manuscript
Middle layer task	Conducting experiments Data analysis Interpretation of data Project management
Outer layer task	Obtaining funding Providing samples Providing technical assistance Collecting data

Authors were classified into core authors, middle layer authors, and outer layer authors according to the scheme above. The classification was hierarchical: authors performing core author tasks were classified as core authors; authors to whom middle layer contributions could be attributed were classified as middle layer authors, if they had not already been classified as core authors; and authors to whom neither core author tasks nor middle layer tasks were attributed were classified as outer layer authors.

In the core layer we find work tasks typically attributed to research leaders, such as conception and design, and drafting and revising the manuscript. Almost all the articles mention who designed the study and who wrote the manuscript. In the original taxonomy suggested buy Davenport and Cronin, however, they listed final approval of the manuscript as a core contribution. It was not possible to use this work task. Almost all the articles stated that all authors had read and approved the manuscript. It is a matter of discussion whether all the authors really read the manuscript or not. However, the journal demands that this is the case, so it therefore cannot be used as a criteria to distinguish the authors from each other.

Results

The composition of the author team

In this first part of the results section I will present the distributions of authors by contribution type. It should be noted that most core authors perform middle-layer tasks as well, but they are only classified as core authors. Middle-layer authors are those who perform middle range tasks, but no core tasks. Outer-layer authors are those who perform only outer-layer tasks.

Table 2. Distribution of authors according to contribution

Layer	No authorships	Share of authorships (%)	Avg. authors per article	St.d. authors per article
Core authors	815	47.3	3.5	2.3
Middle layer authors	671	38.9	2.8	3.1
Outer layer authors	238	13.8	1,1	2.0
All authors	1724	100.0	7.4	4.9

Table 2 displays authorships distributed by type of author. The core authors–the authors who conceptualized the studies, designed the experiments, and/or wrote the articles–constitute the most frequent type in this sample of articles published in *Nature Neuroscience*. All documents contain core authors, and this layer of authors comprises 47.2 percent of the author collective; the 815 core authors average out to approximately 3.5 core authors per article. The 671 authors classified as middle layer authors comprise 38.9 percent of the author collective. These authors did not perform core author activities, but they were involved in the research process–conducted experiments, analysed and interpreted the data, and so on. The description of middle layer tasks is often very specific in contrast to the description of core author tasks, and varies depending on whether the content of the articles represent fMRI studies, biochemical research, or clinical research. It is the outer layer of the author collective, however, that has sparked debate and concern, particularly in the biomedical research community. The outer layer is defined negatively: authors are classified as other layer authors if they were not involved in the planning, analysis, or writing phases, i.e. they did no work directly connected to the research presented in the article. The 238 outer layer authors in this study comprise 13.8 percent of the author collective. The percentage of outer layer authors observed confirms results from a survey study of author contributions in biomedicine (Flanagin et al. 1998).

It can be concluded that outer layer authors constitute a substantial part of the author collective. This in spite of their contributions, or lack of contribution, are described in the article. Since the content of outer layer contributions is not stated in the classification scheme used, taking a closer look at what type of contributions authors in the outer layer made in order to be listed as authors is justifiable. Table 3 displays the distribution of authorships by contribution.

Table 3. Author contributions attributed outer layer authors

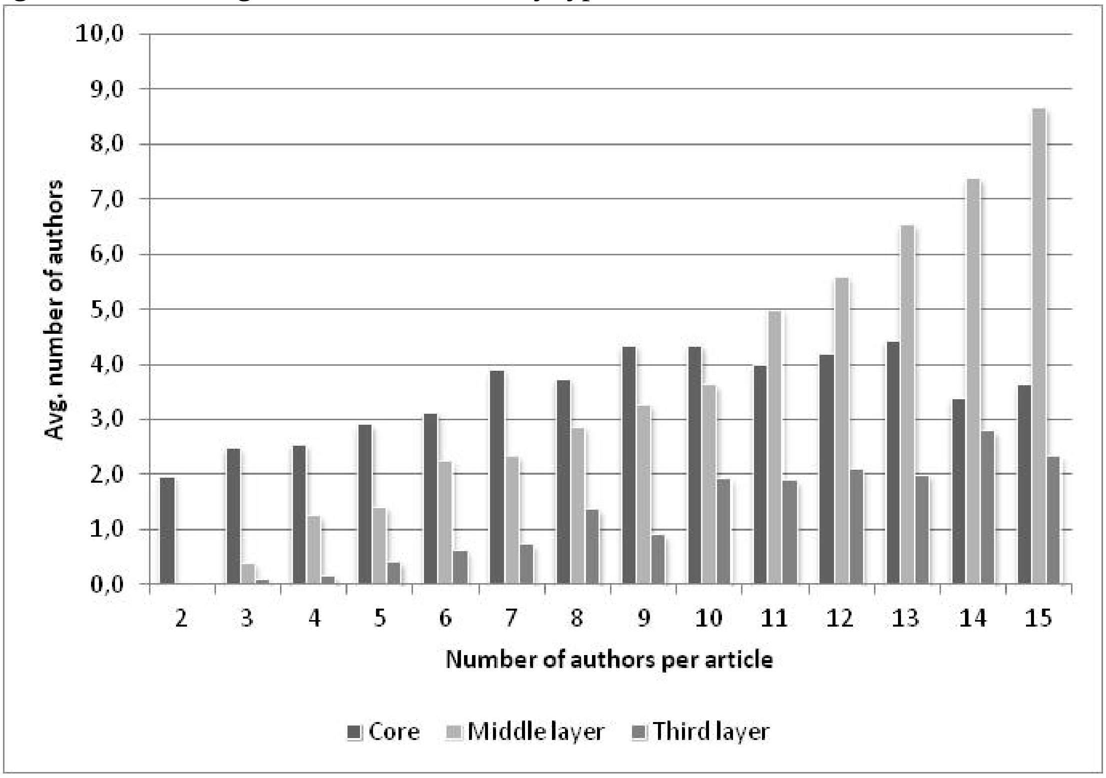
Author contribution	Number of authorships	Percent of outer layer
Created and/or provided samples (knockout mice, reagents, stem cells, etc.)	129	54.2%
Data acquired or collected	50	21.0%
Supervision	16	6.7%
Advice	13	5.5%
Organised studies	12	5.0%
Equipment	6	2.5%
Programming	3	1.3%
Manuscript preparation	3	1.3%
Assisted the project	2	0.8%
Technical assistance	2	0.8%
Financial support	1	0.4%
Institutional support	1	0.4%
Total	238	100,0%

The results presented in Table 3 show that the most common contributions among outer layer authors consist of providing samples or collecting data. One interpretation is that the co-authorship network in part expresses an underlying exchange network in which expensive and hard-to-produce samples are traded between researchers and research groups in exchange for a place among the authors. The most common commodity traded for authorship is genetically engineered mice. These so called ‘knock in’ and ‘knockout’ mice are mice in which specific genes have been either introduced or knocked out. Another commodity traded for authorship is unpublished reagents–compounds that start specific chemical reactions. Stem cells are also valuable enough to guarantee their owner a place among the authors. Another large group in the outer layer of authors is those collecting data–a contribution that is most common in clinical research, where the studies include large patient groups. If we combine those providing and creating samples with those collecting data, we find that 75 percent of the authors in the outer layer have made contributions related to material or data.

The fact that so many of the outer layer authors contribute samples illustrates the dilemma that research groups face when including researchers among the authors. Although not acknowledged by ICMJE as a reasonable justification for acquiring authorship status, denying this group status as authors would probably negatively affect the progress of science, as it would hinder the exchange of very expensive samples.

Changes in the composition of the author team with an increase in the number of authors
In this section I will present the results of an analysis of the same sample regarding how the composition of the author collective changes as the number of authors per article increases. This analysis will provide insight into the process behind the proliferation of authors.

Figure 1. The average number of authors by type of contribution and size of author team



In Figure 1 the average number of core, middle layer, and outer layer authors has been plotted against the author teams of different sizes. The x-axis represents the size of the author team, and the y-axis represents the average number of authors in different contribution categories. It is evident that the number of core authors–the authors who conceptualized, designed experiments, and/or wrote the articles–increases as the size of the author collective increases, but levels off when the size of the team is seven authors. For teams larger than seven authors, the number of core authors remains stable with an average of four core authors per article. In comparison to the number of core authors, the average number of middle layer authors in the author team increases linearly with the size of the team. The average composition of the author teams for articles with 15 authors consists of 3.7 core authors, 8.7 middle layer authors and 2.3 outer layer authors. The rapid increase in middle layer authors indicates a functional relationship to the type of research conducted.

Conclusions

The aim of this study was to investigate what type of contributions merit authorship, and what happens to the composition of the author team when it increases in size. The conclusions drawn from our analysis are limited by the use of author contribution descriptions published in *Nature Neuroscience*. I would like to note the usefulness of the three-layer classification scheme used to divide the author collective into core authors, middle layer authors, and outer layer authors.

Authors classified as core authors constitute the largest group in the sample (47.3 percent). In our study the core authors are those authors that conceived the idea, designed the experiments, or wrote the paper. Although not a criteria for being classified as core author, these authors are in general involved in most core activities, and are also usually involved in the experimental work. In our analysis of the changes to the composition of the author collective with regard to the type of author, it was clear the number of core author stabilized at an average of four core authors per paper, independently of the size of the author collective.

Middle layer authors are almost as numerous as core authors (38.9 percent). In terms of contributions, this group is the most complex part of the author collective. Middle layer authors are those who have not been attributed any core tasks, but tasks such as conducting experiments and analyses, analyzing data, preparing samples or collecting data. If the description of the core author contributions is general, and not specific to any type of neuroscientific research, the description of middle layer tasks is often very specific, and varies depending on whether the content of the articles represents fMRI studies, biochemical research, or clinical research. The analysis of how the author collective changes when the number of authors increases clearly shows that it is this group of middle layer authors that increases, indicating that it is in fact the complexity of the work tasks that drives the increase in authors per article. It should also be noted that there are indications that articles with large author teams report on several experiments. This aspect has not been analyzed in this study, but I propose a hypothesis stating that part of what explains the size of the author collective is the number of experiments reported in the article.

The portion of the author collective that has caused most concern in the medical community is the outer layer of authors (13.8 percent). The outer layer consists of authors that have not contributed either core of middle layer tasks, have not been part of the research team in any normal sense of the word, and therefore cannot be held responsible for the content of the article. The most common contribution attributed to authors in the outer layer is that they have provided samples, unpublished reagents, or knockout and knock-in mice. This is a clear indication that authorship is currency in an exchange system, and part of the co-authorship network is something other than pure research collaboration. The analysis of changes to the composition of the author collective shows that the number of outer layer authors tends to increase when the number of authors increases.

The results presented in this study demonstrate that the issue of accepting outer layer authors is a dilemma. According to guidelines for authorship in most medical journals, this is a type of author that should not be accepted, since these authors have little influence over the research conducted and can therefore not take responsibility for the content. The rationale behind efforts to ban this type of author is understandable, if we accept a traditional definition of what an author is and the responsibilities associated with authorship. If we consider the findings presented here, however, intervening in this process could be problematic. If this type of authorship were prohibited, researchers would lose their incentive to exchange data and scientific progress would probably be negatively affected.

References

Beaver, D. de B. & Rosen, R. (1978). “Studies in scientific collaboration. Part I: The professional origins of scientific co-authorship”, *Scientometrics,* 1, pp. 65-84.

Biagioli, M. (1999) “Aporias of scientific authorship: Credit and responsibility in contemporary Biomedicine”, in ed. M. Biagioli, *The Science Studies Reader,* Routledge: New York, pp. 12-30

Birnholtz, J.P. (2006) “What does it mean to be an author? The intersection of credit, contribution, and collaboration in science”, *Journal of the American Society for Information Science and Technology,* 57, pp. 1758-1770

Cronin, B. (1995) *The scholar’s courtesy: The role of acknowledgement in the primary communication process*, London: Taylor Graham.

Cronin, B. (2001) “Hyperauthorship: A postmodern perversion or evidence of a structural shift in scholarly communication practices?”, *Journal of the American Society for Information Science and Technology*, 52, pp. 558-569

Cronin, B., Shaw, D. & La Barre, K. (2003) “A cast of thousands: Coauthorship and subauthorship collaboration in the 20th Century as manifested in the scholarly journal literature of psychology and philosophy, *Journal of the American Society for Information Science and Technology*, 54, pp. 855-871

Cronin, B. (2004) “Bowling alone together: Academic writing as distributed cognition”, *Journal of the American Society for Information Science and Technology*, 55, pp. 557-560

Cronin, B. (2005) *The hand of science: Academic writing and its rewards,* Oxford: Scarecrow Press Inc.

Davenport, E. & Cronin, B. (2001) “Who dunnit? Metatags and hyperauthorship”, *Journal of the American Society for Information Science and Technology*, 52, pp. 770-773

de Anda F.C. et al (2012) “Autism spectrum disorder susceptibility gene TAOK2 affects basal dendrite formation in the neocortex”, *Nature Neuroscience,* 15, pp. 1022-1021

Endersby, J.W. (1996) “Collaborative research in the social sciences: Multiple authorship and publication credit”, *Social Science Quarterly*, 77, pp. 375–392.

Flanagin, A. et al. (1998) “Prevalence of articles with honorary authors and ghost authors in peer review medical journals”, *Journal of the American Medical Association*, 208, pp. 222-224

Jeong, S., Choi, J.Y. & Kim, J. (2011) “The determinants of research collaboration modes: exploring the effects of research and researcher characteristics on co-authorship”, *Scientometrics*, 89, pp. 967-983.

McDowell, J. M. and Melvin, M. (1983) “The determinants of co-authorship: an analysis of the economics literature”, *Review of Economics and Statistics,* 65, pp. 155-60.

Nowell, C. & Grijalva, T. (2011) “Trends in co-authorship in economics since 1985”, *Applied Economics*, 43, pp. 4369-4375

Persson, Olle, Wolfgang Glänzel & Rickard Danell (2004) “Inflationary bibliometric values: the role of scientific collaboration and the need for relative indicators in evaluative studies”, *Scientometrics,* 60, pp 421-432.

Rennie, D. & Yank, V. (1998) “If authors became contributors, everyone would gain, especially the reader” *American Journal of Public Health*, 88, pp. 828-829

Rennie, D., Yank, V., & Emmanuel, L. (1997). “When authorship fails: A proposal to make contributors accountable”, *Journal of the American Medical Association,* 20, pp. 579–585.

Sher, A. & DeVries S.H. (2012) “A non-canonical pathway for mammalian blue-green color vision”, *Nature Neuroscience,* 15, pp. 952-953

Shewan, L.G. & Coats, A.J.S. (2010) "Ethics in the authorship and publishing of scientific articles", *International Journal of Cardiology*, 144, pp. 1-2

Smith, R. (1997) "Authorship: Time for a paradigm shift", *British Medical Journal*, 314, p 992

Towards Concordance Tables of Different Subject Classification Systems. A literature review with policy implications

Cinzia Daraio*, Flavia Di Costa** and Henk F. Moed***

*daraio@dis.uniroma.it (corresponding author); **dicosta@dis.uniroma.it
Department of Computer, Control and Management Engineering Antonio Ruberti, Sapienza University of Rome, (Italy)

***h.moed@elsevier.com
Informetric Research Group, Elsevier, Radarweg 29, 1043 NX Amsterdam (The Netherlands)

Abstract
Our paper focuses on subject classification systems, and aims to systematically analyse which attempts have been made to develop concordance tables between different subject classifications; which concordance tables have actually been created; which methods were used to create these, and how successful these methods were, in terms of the degree of validity of the proposed concordance.
The efficiency and effectiveness of intensive-based selection of papers vs manual search is compared. The policy relevance of the concordance generation between different types of classification in science, technology and economy is discussed.

Introduction
The field of quantitative science and technology studies is more and more becoming a "big data" science in which large datasets on different aspects of the science, technology and innovation (STI) system are being combined. Combining different datasets is especially important in studies analysing the relationship between the various parts of the STI system or between the various components within each part. At the same time, international organisations such as the OECD, UNESCO, and EUROSTAT generate standardised statistics on R&D activities, both in terms of input and output.

For instance, studies on the science system seek to capture the relationship between funding and scientific-scholarly output in the various domains of science and human scholarship. But statistics from funding organizations may use subject classifications that are different from publication- or journal-based scientific subject classifications. The description and evaluation of teaching and research activities of staff members in academic departments or institutions is confronted with the problem that teaching and research subject classifications do not coincide.

Luwel (2004) noted that attempts to calculate per scholarly field productivity measures relating these input measures to output indicators are hampered by the two types of statistics giving aggregate measures based on different subject classification systems. Studies on the science-technology-industry interface are confronted with the need to create concordance tables between technology and industry subject classifications (e.g., Schmoch et al., 2003) and between patent (IPC) and technology classifications (Schmoch, 2008; Lybbert and Zolas (2014).

The research-in-progress presented in this paper relates to a project that is based on the notion that the emergence of "big data" scientometrics, the increasing emphasis on multi-dimensional assessment, and the increasing interest of research institutions and their funding organizations in valid, reliable and useful indicators, lead to the need to analyse, further develop and – if possible- align a series of relevant classification systems.

The current project focuses on subject classification systems, and aims to systematically analyse which attempts have been made to develop concordance tables between different subject classifications; which concordance tables have actually been created; which methods were used to create these, and how successful these methods were, in terms of the degree of validity of the proposed concordance. This research-in-progress paper presents the outcomes of the first step in the project: the retrieval of relevant articles on the subject based on a literature search in Scopus, and a content analysis of these articles, using VosViewer software combined with a manual approach.

A search of the literature

As a first step a manual search was carried out to identify a core of relevant "seed" articles with essential keyword like "classification", "taxonomy", "concordance table". Subsequently, from their titles keywords were extracted keywords to build up a query with which an automatic search was conducted in Scopus. A wide-ranging analysis was performed using TITLE-ABS-KEY search in the combined field that searches abstracts, keywords, and article titles. Table 1 presents the terms from the "seed articles" that were used as argument in the query.

Table 1: List of terms used in the search

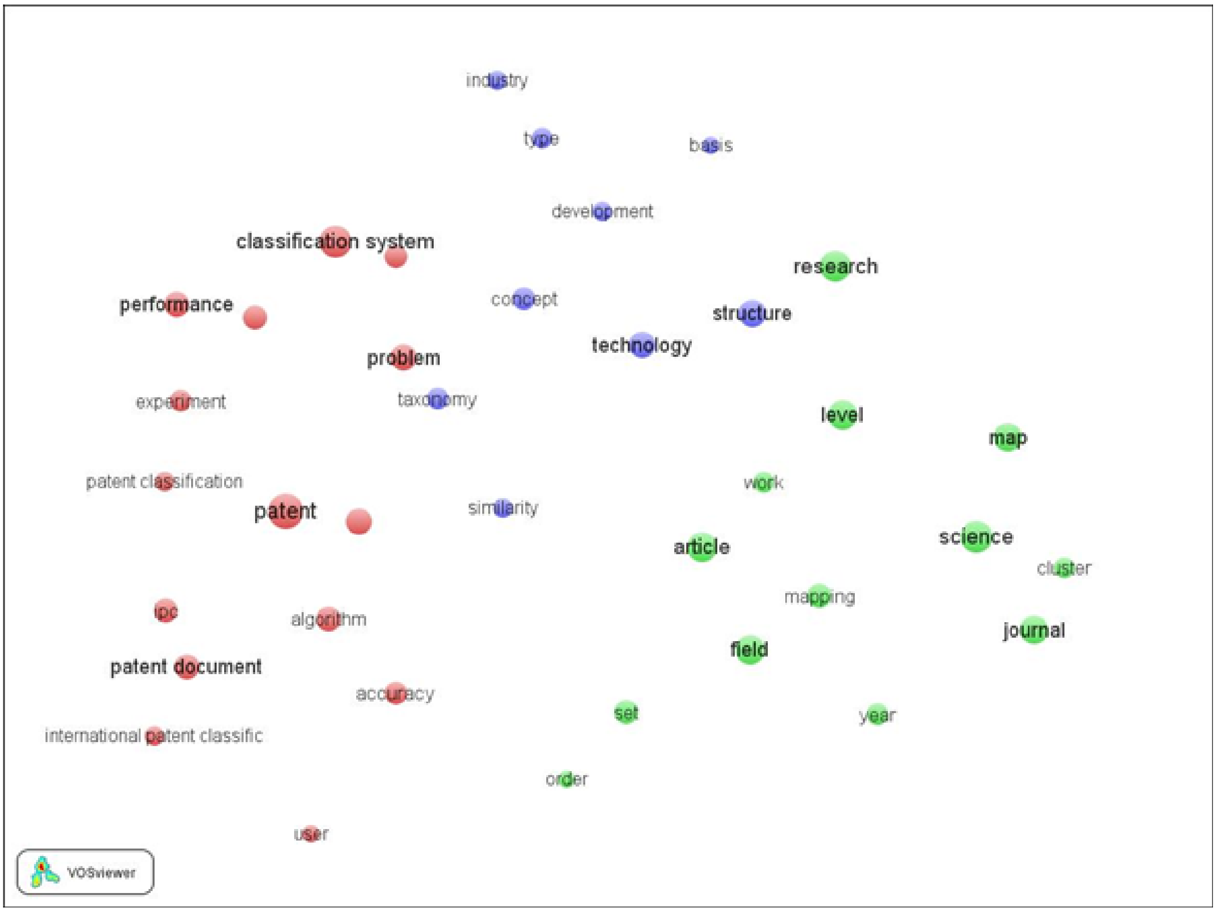
"industr* classification"	"patents AND paper*"	MeSH AND classification
"statistical classification*"	"Classification Systems"	IPC
"Classification of Industries" AND Technology	"Subject clustering" AND "ISI category classification"	subfields AND publications
hierarchical AND taxonom*	Hierarchical AND classification	keyword AND classification
Taxonomy AND Classification	"Research literature" AND maps	clustering AND "scientific texts"
"patent classification*"	"Patent Categorization" AND IPC	"document categorization"
"subject-classification schemes"	"classification scheme" AND "science fields"	categorisation AND patent
"Technology classification" AND indexing"	"Context*aware systems"	"map of science"
"medical data classification"	"automatic classification" AND "scientific literature"	"Hybrid Clustering" AND classification
Classification AND articles	"publication*classification"	Patent AND Categorization
Taxonomy AND Mapping	" hybrid mapping"	Patent AND Science
"terminology mapping"	"classification* AND journals"	patent AND "classification system"
"Technology Concordance"	"Structure AND literature"	coding and "classification systems"
"cross*classification table*"	Manufacturing AND classification	"concordance table*"

435.855 records were obtained. In this set the following five additional selections were made:
1) Articles, reviews and conference papers only (number of records was reduced to 415.485);
2) Documents written in English only (reduction to 338.593 recs);
3) Documents included in the following Subject Area: Engineering, Computer Science, Social Sciences, Mathematics, Economics, Econometrics and Finance, Business, Management and Accounting, Decision Sciences (reduction to 50.656 recs);
4) Given a large amount of records remaining in the set, its number was further reduced by selecting those that contained least one of the following keywords: "patent", "classification", "science", "field", "map*", "taxonom*". The remaining set contained circa 900 records

5) These records were then manually analyzed, on basis of their title. After removing duplicates they were added to the list of those found in the initial manual search.
The final work dataset consisted of 167 records indexed in Scopus.

Content map
The software tool VosViewer (www.vosviewer.com) was used to analyse the contents of the set of 165 articles. A map was created based upon a text analysis of the abstracts using the modele "Create a map based on a text corpus". The text analysis resulted in 3,470 terms. Setting a frequency threshold at 10, 62 terms were selected. Figure 1 shows a map of the 37 most relevant ones. The graphical representation of the keyword structure reveals basically two worlds of work: "technology" and "science". The technology word consists of two clusters, one related to technology-industry concordance and another cluster on patent versus technology classifications. The science world relates to subject classifications based on scientific articles or journals.

Figure 1: VosViewer map of 37 most relevant keywords



A detailed, paper-by-paper analysis reveals from the science cluster a series of papers by Glanzel and co-workers on citation- and text-based subject classifications of scientific journals indexed in Web of Science, and the development of hybrid classifications, including Janssens et al. (2009). A second series, in the technology domain, relates to the concordance between technology and industry classifications (e.g., Schmoch et al., 2003) and between patent (IPC) and industry classification (ISIC) systems, including Schmoch (2008), and Verspagen, Van Moergestel and Slabbers (1994).

Conclusions and further steps

The approach adopted in this paper has generated a useful list of key papers with great relevance for the issue as to which studies have dealt with the concordance of classifications either in science or in technology. This set constitutes the basis for further research into this issue. A recently published article by Lybbert and Zolas (2014) appears to be the most valuable in terms of describing and validating a methodology for creating concordance tables.
The paper will be developed along the following lines:

i) Compare the efficiency and effectiveness of the extensive database search with the one of a manual identification of relevant documents, looking for citing and cited documents in Scopus and/or in Google scholar;
ii) Discuss the advantages and disadvantages of the approach proposed by Lybbert and Zolas (2014) in light of the most relevant studies identified in the literature;
iii) Investigate the policy relevance of the concordance generation between different types of classification in science, technology and economy.

References

Janssens, F., Zhang, L., de Moor, B., Glänzel, W. (2009). Hybrid clustering for validation and improvement of subject-classification schemes. Information Processing & Management, 45(6), 683–702.

Luwel, M. (2004). The use of input data in the performance analysis of R&D systems. In: Moed, H.F., Glänzel, W., and Schmoch, U. (2004) (eds.). Handbook of quantitative science and technology research. The use of publication and patent statistics in studies of S&T systems. Dordrecht (the Netherlands): Kluwer Academic Publishers, 315–338.

Lybbert, T.J., Zolas, N.J (2014). Getting patens and economic data to speak to each other: An 'Algorithmic Links with Probabilities' approach for joint analyses of patenting and economic activity. Research Policy, 43(3), 530-542.

Schmoch, U., LaVille, F., Patel, P. and Frietsch, R. (2003). Linking technology areas to industrial sectors. Final report to the European Commission.

Schmoch, U. (2008). Concept of a Technology Classification for Country Comparisons. Final Report to the World Intellectual Property Office (WIPO), Karlsruhe: Fraunhofer ISI.

Verspagen, B, van Moergastel, T and Slabbers, M. (1994). MERIT concordance table: IPC - ISIC (rev. 2). MERIT Research Memorandum 2/94-004.

"Intentions and strategies for evaluating the societal impact of research: Insights from REF2014 evaluators"

Derrick, GE.*

*gemma.derrick@brunel.ac.uk,

Health Economics Research Group, Brunel University, Uxbridge, London, UB8 3PH (United Kingdom)

Introduction

This research in progress paper describes the initial results of a long-term, large-scale analysis of the operationalization of evaluation of the societal impact of research. Results from the first stage of qualitative interviews are used to illustrate the strength of the methodological design of the study.

The study will eventually include two sets of interviews with the REF2014 Main Panel-A Evaluators. The first set of interviews (The Pre-evaluation interviews) was conducted between January-March 2014, prior to the REF2014 Impact evaluation taking place. The second set of interviews (The Post-evaluation interviews) will be conducted after the evaluation process is completed (estimated time December 2014). The aim of this paper is to present the results of the first round of Pre-evaluation interviews with REF2014 Main Panel-A Evaluators. In particular, the results presented here are concerned with the Evaluators' experience in evaluating impact prior to the REF2014, as well as the variety of strategies evaluators intend to apply in order to guide their assessment.

There are currently many conflicting debates about how the wider societal application of research can be formally evaluated, however no study has empirically studied the formal evaluation of impact. This is mainly due to lack of opportunity, where evaluation frameworks that incorporate the formal evaluation of impact have not existed to provide a formal methodological design. The UK Research Excellence Framework (UK REF) is a world-first framework that dedicates a major proportion of its overall criteria (20%) to the assessment of ex-post impacts. This is done by assessing 4-page case study descriptions submitted by each HEI. The incorporation of this criterion provides a unique opportunity to investigate how evaluators assess research impact.

Broad definitions of societal impact include concepts concerned with the social, cultural, environmental and economic returns from publicly funded research (Bornmann, 2012). The absence of a firm definition of societal impact emphasises the difficulties encountered by evaluators when assigning value to the variety of ways that research can achieve a societal impact. For health research, societal impact can manifest itself by: informing policy and practice development (Kuruvilla et al, 2006), contributing to public debate (Davies et al, 2005), attracting media coverage (Chapman et al, 2014), improving understanding of health risks or determinants (Armstrong et al, 2009), changing ways of thinking about health problems and/or solutions (Weiss, 1986), or being used as evidence in legal proceedings (Cwik & North, 2001). Currently, it is unclear to researchers how to achieve a societal impact that is evaluated positively by peer review panels. Research policies in the UK (REF) and in Europe (Horizon2020) emphasise the importance of societal impact but there is little understanding about what aspects are valued during evaluations. Proposing the inclusion of this criterion has stimulated heated debate among researchers (Nolan et al, 2008) and, as such, the research policy community is struggling with how to recognise and reward research that has both scientific and a societal impact (Smith, 2001; Cooksey, 2006).

Critics argue that quantitative measures of scientific impact are poorly related to applied utility and socioeconomic outcomes (Macilwain, 2009). This is amplified by a poor correlation between more traditional measures of scientific impact with the societal value of research (Nightingale & Scott, 2007); and the realisation that achieving a societal impact may involve factors not under the researcher's control or within the ability of evaluators to predict, such as political will and economic need. These considerations, along with the possibility that evaluators may have differing preferences for rewarding societal impact, make the equitable evaluation of societal impact extremely challenging. Despite these difficulties, a number of tools have been proposed to help evaluate societal impact. These include frameworks such as the HERG payback model (Hanney et al, 2003) and the Research Impact framework (RIF) (Kuruvilla et al, 2006) to guide the evaluations. In contrast, Nutley et al (2007) and Spaapen & van Drooge (2011) promote a focus on behaviours that are considered a precondition for achieving societal impact. This concept is confirmed by public health policy research that identifies a number of researcher behaviours positively associated with societal impact (Derrick et al, 2011; Haynes et al, 2011). This focus on evaluating behaviours acknowledges that the research to societal impact pathway is rarely linear and instead is a maze of complex social and political interactions that is rarely controlled by the researcher nor causally related to the quality of the research (Bowen et al, 2009; Humphreys & Piot, 2012). In fact, many accepted models of public health research utilisation have likened the process to a complex dance (Edwards, 2001), a garbage can of ideas waiting to be needed (Cohen et al, 1972) and as parallel streams awaiting a social, political or economic reason to stimulate a convergence and create an impact (Kingdon, 2003).

In regards to assessing research impact via peer review, while the involvement of experts and peers brings status and credibility to the evaluation process (Boaz et al, 2009), evaluating societal impact can be highly subjective. The incorporation of "societal impact" can be described as a Kuhnian revolution for research evaluation criteria (Luukkonen, 2012). As such, in order to achieve a revolutionary change towards including considerations of societal impact, the idea must be constantly debated, re-defined and reformed before the new paradigm is adopted. An important implication of using peer review is, therefore, that during a period of time in which paradigm shift is occurring, there are multiple scientific contenders who support highly variable viewpoints, making it challenging to achieve consensus within peer review committees (Luukkonen, 2012). The broad REF 2014 definition of societal impact further complicates its evaluation (Kearnes & Wienroth, 2011). In addition to navigating the assessment hurdles of causality, attribution and time lags, when assessing societal impact, evaluators are required to step out of their role as research-peers and instead assess the value of the wider impact of the research, using a different perspective as a public stakeholder. Therefore, without a clear precedent or prior experience of effective impact evaluation, differences in what is believed to constitute a societal impact are likely to be more pronounced where there are already conflicting viewpoints about what constitutes excellent research, such as health (Derrick et al, 2011). This lack of precedents and experience in its evaluation can further hamper government goals of reaping the *"full economic, health and social benefits of public investment in health research"* (Cooksey, 2006).

This research will outline how evaluators intend to overcome the barriers to impact evaluation that have been discussed extensively in the literature (causality, time-lags and attribution). In addition, this research will identify any further barriers anticipated by the REF2014 Main Panel-A evaluators. The research will also describe the strategies the evaluators have developed to overcome these barriers, as well as describe more generally, their approaches to the formal evaluation of ex-post impacts.

This research provides the first, large scale, mixed methods investigations of the formal, ex-post evaluation of impact during a national evaluation framework. This research will provide a unique perspective of how the evaluation of societal impact is constructed by evaluators, as well as provide systematic guidelines of how to evaluate research impact.

Methods

The UK REF2014

The UK REF2014 will dedicate 20% of its overall assessments of university research to how research has had *"…an effect on, change or benefit to the economy, society, culture, public policy or services, health, the environment or quality of life, beyond academia"* (HEFCE, 2011). Within this model, peer review evaluation panels will review 4-page case studies of how research conducted within the last 15 years has benefitted the economy and society (HEFCE, 2011).

Interview Sampling

The Main Panel-A within the REF2014 is comprised on one overarching Main Panel, and 6 sub-panels each divided into different fields under the auspices of Health and Medical Research. The breakdown of each of the panels with the corresponding number of evaluators is included below in Table 1.

All unique REF Main Panel-A evaluators (n=215) were identified and invited to participate in the project. In total, 64 evaluators agreed to participate in the interview, representing a 28.2% response rate. Specific care was taken to obtain a representative sample of evaluators who were evaluating impact and outputs (n=47), Outputs only (n=8), and Impact only (n=9).

Interview questions

The interview schedule was designed to include one, main, overarching question designed to explore a certain theme, followed by a series of prompts to further investigate this theme. The prompts were used to keep the interviewee on topic, while also serving as a method to entice less forthright interviewees to address the theme, without leading.

The interview themes were based around the common issues currently discussed in the academic literature about the evaluation of research impact and peer review. These themes included: Interviewees personal definition of impact; Implicit bias in research impact evaluation; Productive interactions as indicators of impact; Intentions and strategies for assessing impact and overcoming difficulties (including causality, attribution and time lag issues); Anticipated difficulties and power relationships; The role of different types and levels of impacts; and Indicators of impact, attribution and causality. Interview questions also drew on the interviewee's previous research and peer-review research evaluation experience and the influence of research impact in these situations. Finally, the past experience of the interviewees with impact was also used as a prompt to explore their opinions about the importance of evaluating research impact, and its inclusion as a formal criterion in the REF2014.

Table 1: The number of interviews conducted with REF2014 Main Panel A and its 6 sub-panels.

Panel name	Total	Academic evaluators (AEs)	User evaluators (UEs)	Total interviewed
Main Panel A	19	14 (73.7%)	5 (26.3%)	8 (42.1%)
Sub-panel 1 – Clinical Medicine	39	32 (82.0%)	7 (18.0%)	10 (25.6%)
Sub-panel 2 – Public Health, Health services and Primary care	27	23 (85.1%)	4 (14.9%)	13 (48.1%)
Sub-panel 3 – Allied Health Professions, Dentistry, Nursing and Pharmacy	51	42 (82.3%)	9 (17.7%)	14 (27.5%)
Sub-panel 4 – Psychology, Psychiatry and Neuroscience	35	28 (80.0%)	7 (20.0%)	9 (25.7%)
Sub-panel 5 – Biological Sciences	35	30 (85.7%)	5 (14.3%)	6 (17.1%)
Sub-panel 6 – Agriculture, Veterinary and Food Science	29	16 (55.1%)	13 (4.9%)	4 (13.8%)
TOTAL	235	185 (78.7%)	50 (21.3%)	64 (29.7%)

Analysis

The analysis of the qualitative data collected in the interviews is based on two, interlinked rounds of coding and analysis (Round 1: In-depth memo making and analysis; and Round 2: Full qualitative analysis using a cognitive-based grounded theory design. At this stage of the study, the results presented here are based on the first round of coding and analysis: Round 1: In-depth memo making and analysis.

Round 1: In-depth memo making and analysis.

At this stage of the study, the analysis was performed by reviewing the in-depth, post interview notes made by the interviewee immediately after completing the interviews. Extensive memo-making was employed by the interviewer directly after each interview. This allowed for the interviewer to reflect and note the emergence of different themes for analysis, as well as to draw parallels between interviewees as the interviews progress. This recording of themes analysed as they emerged, was noted within the memos and is used to provide theme description within this research in progress paper. In addition, this initial, first-glance stage of coding allows for the further testing and coding of the full transcripts during the second round of coding, described below.

Further analysis of the interviews will be based on a full transcription of each interview (Round 2: Full qualitative analysis using a cognitive-based grounded theory design) using the software Nvivo, and employing Morse's outline of the cognitive-basis of qualitative research (Morse, 1994) and Charmaz's outline of data analysis in grounded theory (Charmez, 2006). However, due to the time restraints, the brief results presented below are based on the first round of coding of the memos made directly after each interview. Future versions of this research-in-progress paper will be based on the results of the second round of analysis which involves a systematic qualitative analysis described above.

Results

Past experience with assessing impact

The Evaluators described their previous experiences with evaluating "impact", but stressed that this had been done informally, and that the REF2014 was to be one of the first occasions when they were asked to evaluate it formally. In this regard, their previous experiences with assessing impact, in general, fell into two categories: (1) Any consideration of "impact" being disregarded totally (Disregarded totally); and (2) Indirect considerations of impact being incorporated into the assessment of the scientific quality of submissions (Indirect consideration).

Disregarded totally

Evaluators described how their previous experience evaluating impact has been done with ex-ante impacts only. In this, they described that impact during previous peer-review had been disregarded totally, and that in many senses impact evaluation was applied after the more *"scientific"* evaluation of proposals was complete. One evaluator described how the evaluation of impact in these circumstances was a *"tick-box criteria"*, rather than an in-depth discussion. The reasons for this were not explored fully in this round of interviews, but one emerging theme was in relation to the seriousness in which these evaluators themselves put together their own impact statements. For the UK research councils, many grant applications are requested to be accompanied by a "Pathways to impact" statement. Within this statement, applicants are asked to describe how their research will be influence and be translated to non-academic audiences. It is an essential component of all grant applications, but one that many evaluators stated was not taken as seriously as the other, more scientific application components. As such, one evaluator stated that with their previous experience in evaluating these impact statements that they saw this as *"impact as rhetoric"* and therefore was not inclined to formally, and seriously consider it as part of the overall grant evaluation.

Indirect consideration

The other, other way that researchers had evaluated impact in the past is what is described here as indirect consideration. Here, evaluators were aware of the importance of research application to non-academic questions, but incorporated this into other, more traditional considerations such as *"the importance of the questions"*, or the *"originality"* of the research. In this way, researchers were not opposed to aspects of impact being evaluated in grant proposals, but they felt more *"comfortable"* and *"experienced"* in evaluating these as part of more traditional, scientific peer review processes, rather than as a separate, formal criterion. Further analysis of the pre-evaluation interviews, as well as combination of the Post-evaluation interviews, will reveal to what extent these traditional avenues of impact evaluation are used as proxies in the evaluation of ex-post impacts under the REF2014.

Intended strategies for assessing impact

Evaluators described the prospect of evaluating impact as *"one big experiment"* where they felt they had had little experience that they could use to evaluate the research impact formally. Nonetheless, the evaluators felt comfortable that by using the traditional processes inherent in peer-review, that the evaluation of research impact would succeed and that *"it will be done well"*. The evaluators had little understanding of what characteristics of impact they would be valuing over others, but emphasised that they felt confident that *"they would know it [impact] when they saw it."*

Despite this optimism, evaluators identified a number of anticipated barriers to the evaluation of the research impact associated with the causality of the impact. In general, evaluators exhibited a preference for impact within the case studies that could show a *"clearly defined link*", between the underpinning research and the impact being claimed. It is unclear, at this stage of the analysis, whether this preference was related to the evaluator's lack of prior experience when evaluating impact; misunderstanding of the role of salient factors and productive interactions that play a role in ensuring impact, implicit bias towards a linear model of achieving impact, or something else. Further analysis of the interviews using the in-depth, coding of interview described above, will reveal more insights regarding any unintended implicit bias towards impact case studies that exhibit a high degree of causality in their description.

The role of the Case studies

As mentioned above, evaluators expressed a preference towards evaluating impact case studies that will demonstrate a strong causal link. As such, some evaluators suggested that a positive evaluation outcome will be as dependent on "how well they tell a story" as with the impact outcome. In this way, a positive evaluation outcome will be as dependent on how "convincing" the case study constructs the strong causal link between the underpinning research and the impacts being claimed. However, evaluators also noted that the advantage that came from clearly written case studies would be no different from the advantage good writing has in all peer-review evaluations.

In addition, in contrast to a number of studies that denigrate (Bornmann & Marx, 2014) the use of case studies in impact evaluation, evaluators felt confident that the case studies would prove beneficial in facilitating the evaluation process. Many evaluators had also been involved in preparing their own organisation's REF2014 Impact case study submissions, and reflected, the wide range of way that research impact could be defined. This had served to demonstrate the enormity of the task in front of them in terms of evaluating the impact within the constructed case studies. However, some evaluators expressed a preference for quantitative measures of impact and felt that translating "these words" in the case study, into quantitative appreciation of the value of the impact, would be their strategy in assessing the case studies.

Discussion

The preliminary results presented in this research in progress paper suggest that despite evaluators having some experience related to the evaluation of ex-ante impacts in the past, that they have limited experience when formally considering the ex-post impacts for evaluation. This reflects the views by Cozzens et al (2002), where it was suggested that the challenge in measuring societal impact, was due to the lack of well-developed models explaining the processes leading from innovation to impact. Indeed, this sentiment is equally echoed in evidence-based policy making studies that emphasise that the strength of the evidence is rarely directly linked to its implementation and that more nuanced factors were involved (Weiss, 1986).

Many researchers reported that they anticipated difficulties in the evaluation of the impacts for the REF2014 that were associated with the causality and the attribution of the claimed impacts. Time lags were not considered a problem, as the clarity of the REF2014 rules regarding claimed impact, and time since the original research was conducted were seen to be sufficient and fair for health and medical research. However, evaluators had little strategy in mind on how to deal with these issues of impact evaluation, admitting that the majority of these issues would primarily be worked out during the evaluation process. Evaluators also felt confident that the range of evaluators on the panel, including evaluators representing stakeholder or user organisations, would help to facilitate the impact evaluation as part of a *"big experiment"*.

Evaluators, however, seemed to exhibit an implicit bias towards impacts that will exhibit a strong causal link between the underpinning research and the impacts claimed. This evaluator-led preference for strong causality suggests that the value of the supporting evidence supplied, as well as the ease with which the case studies are constructed, may play an influential role in facilitating and favour its assessment by REF2014 impact evaluators. Further analysis of the interviews will reveal whether there is a role for productive interactions in demonstrating the strength of this causal link to evaluators, and whether and how this may facilitate its evaluation.

Further insights will be drawn from the complete, in-depth analysis of the complete transcripts of these Pre-evaluation interviews. In addition, the cross-reference of the results of these pre-evaluation interviews, with the results of the post-evaluation interviews (expected in December 2014), as well as determining the differences between academic-based and user-based evaluators, will provide further insights into the evaluation of research impact.

References

Armstrong R, Doyle J, Walters E (2009) Cochrane Public Health Review Group update: incorporating research generated outside of the health sector. Journal of Public Health 31: 187-189.

Bornmann L (2012) Measuring the societal impact of research. EMBO Reports 13: 673-676.

Bornmann L, Marx, W (2014) How should the societal impact of research be generated and measured? A proposal for a simple and practicable approach to allow interdisciplinary comparisons. Scientometrics 98: 211-219.

Boaz A, Fitzpatrick S, Shaw B (2009) Assessing the impact of research on policy: a literature review. Science and Public Policy 36: 255-270.

Bowen S, Zwi AB, Sainsbury P, Whitehead M (2009) Killer facts, politics and other influences: what evidence triggered early childhood intervention policies in Australia. Evidence & Policy 5: 5-32.

Chapman S, Haynes A, Derrick GE, Sturk H, Hall WS (in press) Reaching "an audience that you would never dream of speaking to": influential public health researchers' views on the role of news media in influencing policy and public understanding. Journal of Health Communication.

Charmaz K (2006) Constructing grounded theory. London: Sage.

Cohen MD, March JG, Olsen JP (1972) A Garbage Can Model of Organizational Choice. Administrative Science Quarterly 17: 1-25.

Cooksey D (2006) A Review of UK Health Research Funding. Norwich.

Cozzens S, Bobb K, Bortagaray I (2002) Evaluating the distributional consequences of science and technology policies and programs. Research Evaluation. 11(2): 101-107.

Cwik CH, North JL (2001) Scientific evidence review: Monograph 5. Chicago: American Bar Association. Section of Science and Technology Law.

Davies P, Bowen S (2005) Getting research into policy and practice. NSW Public Health Bull 16: 177-180.

Derrick GE, Haynes A, Chapman S, Hall WD (2011) The Association between Four Citation Metrics and Peer Rankings of Research Influence of Australian Researchers in Six Fields of Public Health. PLoS ONE 6: e18521.

Edwards M (2001) Social policy, public policy: from problems to practice. St Leonards, N.S.W: Allen & Unwin.

Hanney SR, Gonzalez-Block MA, Buxton M, Kogan M (2003) The utlisation of health research in policy-making: concepts. examples and methods of assessment. Health Research Policy and Systems 1: available online (http://health-policy-systems.com/content/1/1/2).

Haynes AS, Derrick GE, Chapman S, Redman S, Hall WD, et al. (2011) From "our world" to the "real world": Exploring the views and behaviour of policy-influential Australian public health researchers. Social Science & Medicine 72: 1047-1055.

HEFCE (2011) REF2014: Assessment framework and guidance on submissions.

Humpreys K, Piot P (2012) Scientific evidence alone is not sufficient basis for health policy. British Medical Journal 344: e1316.

Kearnes M, Wienroth M (2011) Tools of the trade: UK Research Intermediaries and the Politics of Impact. Minerva 49: 153-174.

Kingdon J (2003) Agendas, alternatives, and public policies (2nd Edition). New York: Longman.

Kuruvilla S, Mays N, Pleasant A, Walt G (2006) Describing the impact of health research: a Research Impact Framework. BMC Health Services Research 6: available online (http://www.biomedcentral.com/1472-6963/1476/1134).

Luukkonen T (2012) Conservatism and risk-taking in peer review: Emerging ERC practices. Research Evaluation 21: 48-60.

Macilwain C (2009) The RAE: An Assessment Too Far? Cell 139: 643-646.

Morse JM (1994) Emerging from the data: the cognitive processes of analysis in qualitative inquiry. In: Morse JM, editor. Critical issues in qualitative research methods. Thousand Oaks: Sage. pp. 23-42.

Nightingale P, Scott A (2007) Peer review and the relevance gap: ten suggestions for policy-makers. Science and Public Policy 34: 543-553.

Nolan M, Ingleton C, Hayter M (2008) The Research Excellence Framework (REF): A Major Impediment to Free and Informed Debate? International Journal of Nursing Studies 45: 487-488.

Smith R (2001) Measuring the social impact of research - Difficult but necessary. British Medical Journal 323: 528-528.

Spaapen J, van Drooge L (2011) Introducing 'productive interactions' in social impact assessment. Research Evaluation 20: 211-218.

Weiss CH (1986) The circularity of enlightenment - diffusion of social-science research to policy-makers. Knowledge-Creation Diffusion Utilization 8: 274-281.

Unwrapping "impact" for evaluation: A co-word analysis of the UK REF2014 policy documents using VOSviewer[1]

Derrick, GE.*, Meijer I** and van Wijk, E**

* *gemma.derrick@brunel.ac.uk*
Health Economics Research Group, Brunel University, Uxbridge, London, UB8 3PH (United Kingdom)

** *i.meijer@cwts.leidenuniv.nl; wijk@cwts.leidenuniv.nl*
Center for Science and Technology Studies (CWTS), Leiden University (The Netherlands)

Introduction

This paper analyses the conceptualisation of "impact" for the UK's REF2014, in order to gain an understanding of its definition for the 2014 evaluation process. This research provides a quantitative analysis of themes and words associated with the development of the concept of "impact" within UK research policy documents. The aim of this analysis was to use an objective, quantitative method to investigate the overarching impact evaluation, policy implementation and its adoption as well as concepts pertaining to the evaluation of the "societal impact" of research at the UK level. The results, therefore, may contribute to a more precise understanding of underlying policy intentions in relation to research "impact".

During 2014, the UK will be running the Research Excellence Framework (REF2014). The framework, part of a series of research evaluation programs previously known as the Research Assessment Evaluations (RAE) currently run in the UK every 5 years. Whereas in the past, the evaluation has primarily involved the peer review, of academic outputs such as research articles, the REF will incorporate a new criterion known as "impact". For the REF, impact is defined as research that has had *"…an effect on, change or benefit to the economy, society, culture, public policy or services, health, the environment or quality of life, beyond academia."* (HEFCE, 2011) This criterion will constitute 20% of a university's overall assessment which will then be used to allocate UK government research funding. Within this model, peer-review evaluation panels will review 4 page case studies of how selected research conducted in the last 15 years, has benefited the economy and society. As such, the REF2014 provides the world's first, formal, ex-post assessment of how research has influenced the wider, non-academic community.

While the involvement of research "experts" brings status and reliability to the evaluation process, evaluating the societal "impact" of research can be highly subjective. Subjectivity on impact has also been shown in the evaluation of the 'broader impact' criterion of the National Science Foundation in the USA, even though this dealt with ex ante assessment instead of ex post (Holbrook & Frodeman 2011). In addition, the uniqueness of the criteria and the lack of a firm precedent for the evaluation of "impact", raises the risk that researchers will resort to other methods to evaluate impact that have previously been described as problematic in peer-review evaluations. In order to address the subjectivity of the impact evaluation, researchers may resort to one or more of the following peer review problems: conservative bias (researchers translating their own values of convictions in evaluations); implicit bias (a

[1] This work was supported by the UK Economic and Social Research Council (ESRC).

positive or negative mental attitude at an unconscious level); quantitative bias (positive evaluations of impacts that are "countable"); and assuming research excellence evaluation as a proxy for impact excellence.

Although the definition of "excellent, research impact" associated with the "impact" of research have been thoroughly discussed in the research policy literature, blogosphere and UK media, it is not yet clear how impact will refer to the impacts on the economy, society, culture and/or health. This confusion makes it difficult to interpret the results or to anticipate what characteristics of impact will perform well under this world-first formal criterion. This also makes it difficult for other countries (Australia, New Zealand and Europe) to adopt similar frameworks and guidelines for its evaluation.

By providing an objective, quantitative view of the words associated with "impact", this research will allow for a clearer understanding of what the UK government intends the impact criterion to reflect, as well as a more transparent interpretation of the REF2014 results, available in December 2014. A secondary aim of this paper is to investigate the suitability of the software VOSviewer, for the analysis of themes within policy documents.

Methods

Document identification

Policy and other related documents to the development of the impact evaluation criterion were sourced via the REF2014 website (www.ref.ac.uk). Only documents directly related to the REF2014, written, endorsed or commissioned by the Higher Education Funding Council of England (HEFCE) which runs the REF process, were included in the analysis. By only including documents underlying the UK's REF2014 evaluation process, it was assured that all results listed were relevant to the conceptualisation of impact for its submission and evaluation under this framework.

A total of policy 41 documents were identified from the REF2014 website (www.ref.ac.uk) as underpinning the development of the tools and procedures of the evaluation framework. To increase the validity of the results obtained, a number of checks were carried out on the documents to assess their validity. In particular, all documents were individually analysed in order to ensure that they were similar in structure (Executive summary/recommendations, Introduction, Aims/Objectives, Background, Outcomes and Annexes), and style (Policy documents, Consultation responses, Evaluation guidelines etc). This pre-analysis assessment of the documents ensured that the analysis run by VOSviewer 1.5.2 (Van Eck & Waltmann 2010, 2011) would be both robust and yield representative results. All 41 original underpinning REF documents were determined to be comparable in the analysis.

Analysis

VOSviewer 1.5.2 was selected to visualise the noun phrases associated with the word "impact" within these policy documents. Using VOSviewer, as opposed to more qualitatively focused software can be useful in determining frequently used noun phrases within these documents without having to manually perform such searches, or rely of expert judgement of associations which can be dangerously subjective (Van Eck, 2011).

For the analysis all 41 policy documents were all converted to plain text style and converted into a text corpus for analysis. The analysis focused on the noun phrase 'impact' and the words surrounding impact, taking into account the distance to 'impact' and to each other. The distance set in the initial analysis was set at five, and cutting off at an interpunction, meaning that only words in the same sentence were analysed.

The total number of words analysed was 446. Meaningless words were removed, and the resulting data matrix was visualised in the VOSviewer. Only words that appeared 10 times or more were displayed. In the display the word "impact" was removed and the size of the circles reflects the number of occurrences of that particular term (word) in connection with impact, and the relative distance between the terms (shown graphically) reflects the relatedness of the terms.

Cluster analysis grouped linked words together according to the strongest occurrence. These words were therefore mapped in closed proximity with each other and each cluster was grouped by colour. In addition, link strength was calculated for each of the words with the highest co-occurrence. The link strength is a measure of the proximity of the words to each other so that higher link strength indicates the words that are in closer proximity to each other. The degree of link strength between words can be used to investigate the conceptualisation of the term "impact" in relation to these closely associated words within the text of the policy documents being analysed.

Results

The initial analysis of the REF policy documents yields a visual display of the words in the vicinity of the keyword "impact".

The map created and shown on Figure 1 consists of words indicating the process of assessment, such as 'research', 'assessment', 'case study', 'criteria' and 'outputs' and 'evidence', which are the most frequent. When zooming in more procedural terms appear that are related to the REF framework and process, such as 'pilot', 'panel', 'template', 'submissions', 'statements' and 'indicators'. These words could be expected in guidance documents that are part of the set of 41 documents.

The cluster analysis results yielded 7 distinguishable clusters of noun phrases shown in Figure 1. Of these 7 clusters, 5 were related to the **process** of applying under the REF impact criterion, rather than elaborating on the **concept** and definition of impact. These 5 process clusters were labelled as (1) Case, which contained noun phrases related to the submission of the 4-page impact case studies; (2) Criteria, where the noun phrases related to the weighting of and the sections involved with each REF criteria; (3) Assessment, which was clustered with other noun phrases such as "quality" and "academic"; (4) Economic, which was a small cluster along with "councils" and referred to the investment associated with the UK Research Councils; and (5) Submitted, which referred to the units of assessment (UoA), and other nouns relating to the units assessed under the framework.

More importantly, were the 2 clusters associated with the content of impact; (6) Evidence; and (7) Research. A table containing the nouns identified within these clusters, as well as their link strength to the primary cluster word where this exists (i.e. Evidence or Research) is shown in Table 1. For Cluster (6) Evidence, the clustered words included noun phrases such as "appropriate" and "claimed", but these were not directly linked to the source word "Evidence". Only 2 words were linked to "Evidence", with the noun phrase "indicators" showing the strongest link strength at 124. On the other hand, a larger set of noun phrases

Table.1. Noun phrases grouped under Cluster 6: Evidence; and Cluster 7: Research and corresponding link strength

Cluster	**Noun phrases**	**Link Strength**
Evidence	Appropriate	-
	Claimed	-
	Confirmed	-
	Examples	**42**
	Impacts	-
	Indicators	**124**
	Menu	-
	Statement	-
	Supporting	-
	Table	-
Research	Key	-
	Focus	-
	Report	-
	Workshop	-
	Achievement	-
	Arising	-
	Time	-

	Areas	-
	Broad	-
	Support	-
	Groups	-
	Council	-
	Nature	-
	institution	-
	Australia	-
	Activity	**46**
	Beneficiaries	-
	Need	-
	Enabling	-
	Wider	**42**
	Applied	-
	Relevant	-
	Individual	-
	Collaboration	-
	Data	-
	Users	**54**
	Underpinned	-
	Identify	-

	Sector	-
	Disciplines	-
	Staff	-
	Distinct	-
	Researchers	-
	Humanities	-
	Exploring	-
	Citation	-
	Future	-
	Contributed	-

Figure.1. Cluster analysis of the REF2014 Policy documents

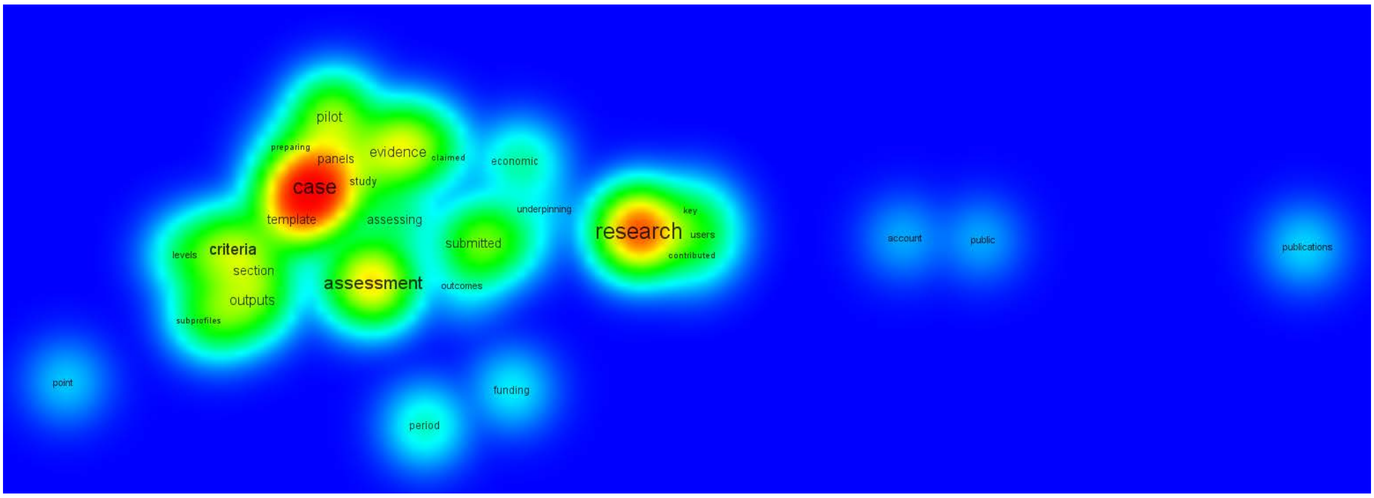
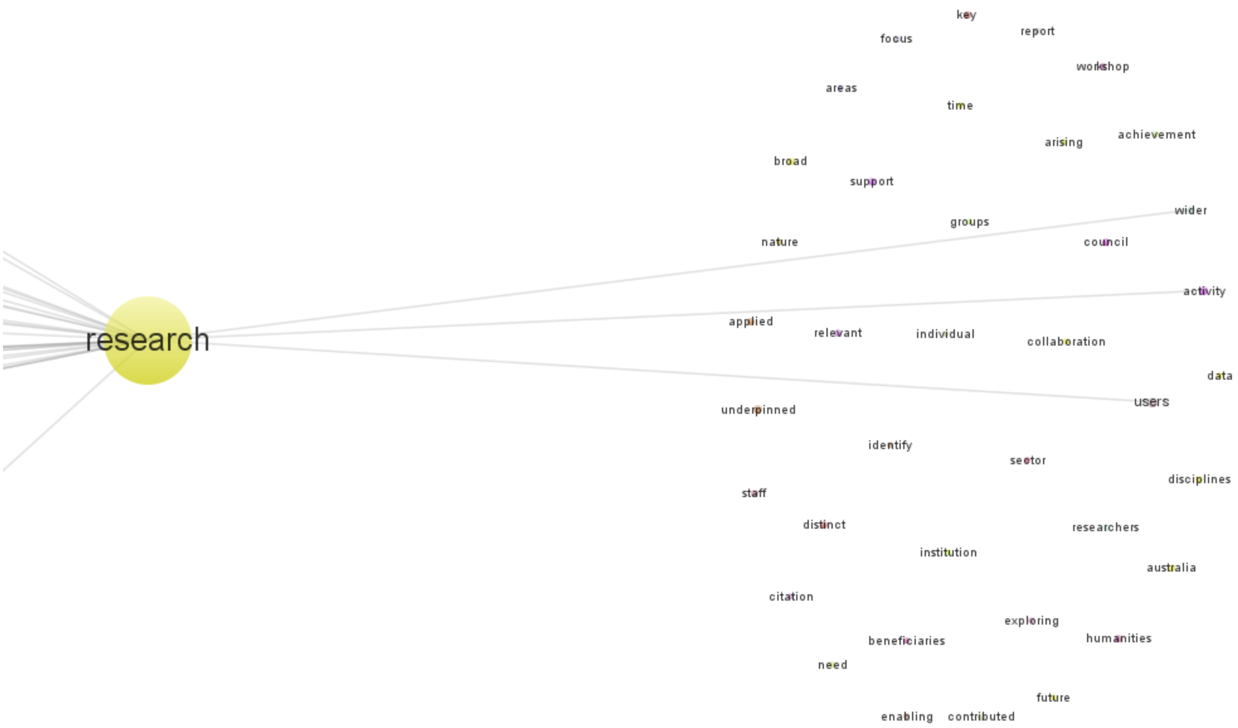


Figure.2. Closer analysis of Content Cluster 7: Research



were associated within Cluster (7) Research. Three words were found to be directly linked with "Research" all with similar link strengths. These were found to be "Wider", "Activity" and "Users". Interestingly, these words describe the concept of "impact" but were found to be clustered around "research", rather than in the other 6 clusters described above. These associations and the wider Cluster (7) Research are shown below in Figure 2.

Indeed, an interesting result for publications pertaining to the REF2014 policy documents is that words that could be taken to elaborate on the meaning of the word 'impact' are scarce and, apart from the 3 noun phrases linked to the word "research", do not appear within the other clusters. Indeed, the words "social" and "societal" do not appear at all in the analysis. This is surprising considering the broad definition of impact adopted by the REF2014 (HEFCE, 2011; Grant et al, 2009).

Only two other noun phrases appear frequently enough to be displayed: "economic" and "publications". They appear not in each other's vicinity, indicating that they are not related. Economic refers to the connotation that research should eventually benefit society by creating more jobs, and more turnovers, based on linear thinking. Publications were not to be expected as part of societal impact, but the appearance can be explained because in the case statements also key publications are included. It may also refer to the widely accepted idea that excellent research is a prerequisite for impact. More likely, however, it is related to the REF2014 Impact template that required submissions to nominate up to 4 publications as "underpinning research", in each impact case study. However the word "underpinning" was seen as separated from the word "publications", so this is unlikely to be the case.

Discussion

The results described above suggest that there is a lack of transparency related to the content of the definition of impact within REF2014 policy documents. This result is interesting in the policy documents of an evaluation framework that will be a world first in formally evaluating the impact of research (20% of the overall evaluation). A possible limitation of this study may be in the availability of policy documents. Indeed, the focus of a cluster associated with the **process** of the REF, suggests that many of the documents underpinning the UK REF201 may be primarily focused on procedural information, rather than larger discussions on the conceptualisation of "impact". Instead, the results suggest that the majority of the REF2014 policy documents are concerned with the "process" of Impact submission. This is indicated by the strong clusters around the noun phrases "criteria" and "case". The lack of this information within these primary REF documents is worrying in light of these documents representing the only government guide for the submission and evaluation of "impact" as part of the REF framework. However, the lack of this conceptualisation in these policy documents may suggest that the debate about what constitutes "impact" is occurring elsewhere, perhaps in the academic literature or perhaps within the REF2014 evaluation panels during the assessment process. Future research will employ alternative, qualitatively focused methods, to explore this conceptualisation of "impact" for research evaluation more thoroughly. Indeed, further analysis of the academic literature using VOS viewer, as well as interviews with REF2014 (See Derrick, (submitted to STI-ENID 2014), for more information), will reveal further information regarding the content of "impact" as well as its "content" within the REF2014 evaluation process.

Furthermore, there is a suggestion of an implicit bias within the text through the association found between the noun phrases "research"-"economic" and "evidence"-"indicators". In addition, there was a lack of wider impact associated words such as “social” and “societal” that further indicate an implicit bias within the policy text towards impacts that are measurable and economic in nature. This reflects a restricted, linear definition of ex-post impacts that may manifest itself in its assessment, favouring those impacts that can be described in quantitative measures. However, this is in contrast to the association between the word “research” and the words “wider”, “activity” and “users”, which suggest a more broad definition of “impact”. Nonetheless, in absence of noun phrases associated with the “content” of impact, there is little information to guide HEIs in the submission of their impact case studies. Indeed, the implicit bias indicated with the policy text, suggests that HEIs may have prioritised impact case study submissions that contain impacts that are quantifiably measureable or economic in nature. This may mean that a number of more nuanced impacts, including research influencing policies or those containing valuable, non-measureable, productive interactions, would not have been submitted as part of the REF2014 as the HEIs may have deemed them too risky. In addition, the absence of these more salient impact case studies has the potential to further bias the impact assessment by REF2014 evaluators towards those impacts that can be measured by quantitative indicators.

Finally, the above results show that VOSviewer can successfully be applied to the mapping and identification of noun phrases within government policy documents. VOSviewer has successfully been applied to identifying noun phrases in academic articles (Mingers & Leydesdorff, 2013; Rodrigues, et al, in press; Romo-Fernandez, et al 2013) and editorials (Waaijer, 2013; Waaijer, et al 2011; Waaijer, et al 2010), but this study is the first to extend the analysis to government policy documents.

Indeed, the use of VOSviewer for the analysis of policy documents is unique in this study and there are indications in the results that the approach adopted within this study is appropriate. Indeed, a strong link was found between the noun-phrases “quality”-“research” (link strength = 158) and “quality”-“environment” (link strength = 40). Although the noun phrase “research” is flexible, its association with “quality” along with the parallel association between “quality” and “environment” reflect the other two REF2014 assessment criteria where “outputs” will constitute 65% of the overall score, and “environment” 15%. This finding acts as a methodological control and provides evidence to suggest that our approach is justifiable.

In general, government policy documents are distinct from the academic literature as they, (Problem 1) are more prone to be laden with values and language designed to “convince” the readers; and (Problem 2) may use noun phrases differently to how they are accepted to be used as jargon in the academic literature. The advantages of analysing government policy documents empirically using software such as VOSviewer therefore is that a more removed, and therefore objective, approach to analysing noun phrases can be employed. This can identify key phrases independent of being swayed by the argument constructed within the text (Problem 1). This issue also demonstrates why a methodology using VOSviewer, is preferable in this study to a more traditional, content analysis of the policy documents. In addition, it is possible to discern how key noun phrases are defined and utilised within the text (Problem 2) by observing those words clustered together, and/or closely associated.

Future research in this area will include combining the policy documents from the impact policy supported within the European Horizon2020 programmes, as well as a comparison of the noun phrase clusters identified in the academic literature pertaining to the content of research impact. The addition of these policy and academic documents would mirror the methodologies of previous studies that have used VOSviewer to identify clusters in the academic literature (Mingers & Leydesdorff, 2013; Rodrigues, et al, in press) over time (Romo-Fernandez, et al 2013), while extending the methodology to include a larger sample of policy documents.

References

Derrick, G.E. (submitted to STI-ENID 2014) Intentions and strategies for evaluating the societal impact of research: Insights from REF2014 evaluators.

Grant, J., Brutscher, P-B., Kirk, S, E., Butler, L., Wooding, S. (2009) Capturing research impacts: A review of international practice.: RAND Europe.

HEFCE (2011) REF2014: Assessment framework and guidance on submissions.

Holbrook, J.B., and Frodeman, R. (2011). Peer review and the *ex ante* assessment of societal impacts. *Research Evaluation 20 (3): 239-246.*

Mingers, J., & Leydesdorff, L. (2013). Identifying research fields within business and management: A journal cross-citation analysis. – Available http://arxiv.org/abs/1212.6773.

Rodrigues, S.P., Van Eck, N.J., Waltman, L., & Jansen, F.W. (in press). Mapping patient safety: A large-scale analysis of the literature using bibliometric visualization techniques. BMJ Open.

Romo-Fernández, L.M., Guerrero-Bote, V.P., & Moya-Anegón, F. (2013). Co-word based thematic analysis of renewable energy (1990-2010). Scientometrics, 97(3), 743-765.

Van Eck, N.J., & Waltman, L. (2010). Software survey: VOSviewer, a computer program for bibliometric mapping. Scientometrics, 84(2), 523-538.

Van Eck, N.J., & Waltman, L. (2011). Text mining and visualization using VOSviewer. ISSI Newsletter, 7(3), 50-54.

Waaijer, C.J.F. (2013). Careers in science: Policy issues according to Nature and Science editorials. Scientometrics, 96(2), 485-495.

Waaijer, C.J.F., Van Bochove, C.A., & Van Eck, N.J. (2011). On the map: Nature and Science editorials. Scientometrics, 86(1), 99-112.

Waaijer, C.J.F., Van Bochove, C.A., & Van Eck, N.J. (2010). Journal Editorials give indication of driving science issues. Nature, 463(7278), 157.

Factors influencing PhD students' scientific productivity

Hanne Derycke*, Katia Levecque**, Noëmi Debacker***, Karen Vandevelde****, and Frederik Anseel*****

*hanne.derycke@ugent.be
Centre for Research & Development Monitoring (ECOOM), Department of Personnel management, work and organizational psychology, Ghent University, Henri Dunantlaan 2, Ghent, 9000 (Belgium)

**katia.levecque@ugent.be
Centre for Research & Development Monitoring (ECOOM), Department of Personnel management, work and organizational psychology, Ghent University, Henri Dunantlaan 2, Ghent, 9000 (Belgium)

***noemi.debacker@ugent.be
Centre for Research & Development Monitoring (ECOOM), Department of Research Affairs, Ghent University, Sint-Pietersnieuwstraat 25, Ghent, 9000 (Belgium)

****karen.vandevelde@ugent.be
Centre for Research & Development Monitoring (ECOOM), Department of Research Affairs, Ghent University, Sint-Pietersnieuwstraat 25, Ghent, 9000 (Belgium)

*****frederik.anseel@ugent.be
Centre for Research & Development Monitoring (ECOOM), Department of Personnel management, work and organizational psychology, Ghent University, Henri Dunantlaan 2, Ghent, 9000 (Belgium)

Abstract

Scientific productivity is a major topic for all those who aspire to an academic career and it is an important precondition to obtain a doctoral degree. In this study, we investigate factors influencing both the research quantity and quality of PhD students at a large public university in Flanders (Belgium). Bibliometric data included in the Thomson Reuters Web of Knowledge database were gathered for all 1112 PhD students who completed the Survey of Junior Researchers II. Multivariate logistic regression analyses were performed. With regard to research quantity, our results show that scientific discipline, phase of the PhD process, funding situation, family situation and organizational culture within the research team are important factors predicting the number of publications. Similar findings were obtained with regard to research quality, although a more competitive culture in the research team could not predict the likelihood of publishing in high ranked journals.

Introduction

Scientific productivity is a critical aspect of academic achievement and is an important issue for researchers even at the early career stages. Junior researchers at the start of their career become rapidly familiarized with the prevailing publish or perish culture in academia and are stimulated to join in. However, while the research performance of faculty has received considerable attention, little is known about the scientific productivity of junior researchers (Cardoso, Guimaraes, & Zimmermann, 2010). In addition, the knowledge of which factors relate to their scientific productivity is limited. Gender differences dominate the scientific productivity literature (Duffy, Jadidian, Webster, & Sandell, 2011): a recurrent finding is that men are more productive than women (e.g. Leahey, 2006; Stack, 2004; Symonds, Gemmell, Braisher, Gorringe, & Elgar, 2006). However, the relation between gender and research productivity is complex and is likely moderated by a number of variables including work

experience, family characteristics and research specialization (Duffy, et al., 2011). Academic tenure can be considered an important influencing factor, as it is evident that more experienced researchers are likely to be more productive than researchers at the start of their scientific career. Family characteristics may affect research performance as well, although mixed results were obtained with regard to the influence of marriage and childcare responsibilities on women's scientific productivity (Fox, 2005). Another factor that can affect research productivity is the specific scientific field in which the research is conducted. Different scientific cultures and accordingly other norms and practices, for instance with regard to publishing strategies and funding resources may differ considerably according to the specific scientific discipline (Manana-Rodriguez & Gimenez-Toledo, 2013). Previously, studies have revealed that the research performance of young PhD graduates is determined by the quality of the academic training they receive (Cardoso et al., 2010; Ruane & Tol, 2009). The quality of the research training environment, in turn, depends on the quality of the institution awarding the PhD, and the research productivity and guidance provided by the PhD supervisor (Duffy et al., 2011; Ruane & Tol, 2009).
Overall, the aim of this study is to investigate which factors influence the scientific productivity of junior researchers, both in terms of publication quantity and quality.

Methods

Study sample

The present study is based on a subsample of the Survey of Junior Researchers II which was organized in 2013 among the total population of junior researchers at Flemish universities. The subsample relates to the PhD students at Ghent University. Among other topics, respondents were asked about their education, their research and well-being, and the coaching they received from their supervisor(s). In addition, questions addressed working conditions, team climate, and organizational culture. From the 3830 junior researchers at Ghent University who were invited to participate (February-May), 1313 researchers completed the questionnaire (34.3%). Our study sample consists of 1112 doctoral students who were officially enrolled in a PhD program. For all participants, the survey data were matched to the bibliometric data that were included in the *Thomson Reuters Web of Knowledge* database. More specifically, all publications that were published between 2004 and 2013 in journals included in the Web of Knowledge of which a doctoral student was author or co-author were taken into account.

Productivity variables

In this study, we measured the scientific productivity of the junior researchers in different ways focusing on both publication quantity and quality. Publication quantity was measured by the number of publications and the number of first-author publications. The measurement of research quality can be measured in many ways (Bornmann & Marx, 2014). However, taken into account the turnaround time for publications in Web of Knowledge and the fact that the participating junior researchers have an average work experience as doctoral researcher of 2.6 years (SD 1.6) we did not include citation counts. In the present study, publication quality measurement is based on the number of publications in top 25% journals. This ranking is based on the ranking of the specific journal within its subject category according to its impact factor. The Journal Impact Factor was retrieved from the Journal Citations Report in Web of Knowledge. All outcome variables were dichotomized (0= no publications; 1= ≥ 1 publication).

Statistical analysis
Multivariate logistic regression analyses were performed in order to assess which factors influence the research quantity and quality of junior researchers. Three logistic regression models were constructed, one for each of the research productivity variables outlined above. Factors included in each of these models were gender, age, scientific discipline in which they are active, type of funding, phase of the PhD process, family situation (i.e. having a partner and/or children), satisfaction with the guidance provided by the main PhD supervisor, and the organizational culture within the research group. Organizational culture was operationalized as a continuum between a very supportive and a very competitive culture. Examples of a competitive culture include a strong focus on individual results, considering colleagues as competitors, and decision-making processes that only involve a limited number of people, whereas the emphasis of a supportive climate is on good relationships with colleagues (Hofstede, Neuijen, Ohayv, & Sanders, 1990). Analyses were conducted using SPSS 21.0 software.

Results
Descriptives
The sample consists of 45.2% (N=501) male and 54.8% (N=607) female PhD students. The average age is 28.1 years (SD 4.4), while about 4 in 5 are younger than 30. The majority of PhD students (72.4%; N=805) have a partner but only 15.2% (N=169) have childcare responsibilities. In total, 20.1% (N=222) of the doctoral students are in the initial planning phase of their PhD project, 58.5% (N=646) are in the executing phase, and 21.4% (N=237) are in the finishing phase. Doctoral students in biomedical science form the largest group (30.1%; N=332), followed by students in the social sciences (24.8%; N=274), applied science (17.8%; N=196), natural science (17.5%; N=193), and humanities (9.8%; N=108). While 39.6% (N=440) of the doctoral students have a competitive scholarship, 24.6% (N=269) have an appointment based on project funds, 18.3% (N=204) have an assistant lectureship and 10.7% (N=119) are registered as doctoral student but receive no funding from the university. On average the participating PhD students publish 0.76 (SD 1.6) publications. The majority of them (67.6%) have no publication yet, 14.7% have one publication, and 1 in 10 researchers have three or more publications (Table 1). In total, 32.4% (N=360) of the respondents were (co-)author of at least one publication, 19.4% (N=216) had at least one first-author publication and 21.5% (N=239) were (co-)author on at least one publication in a top 25% journal.

Table 1. The frequency of junior researchers in terms of their scientific production

	N	%
Total number of publications		
No publications	752	67.6
1 publication	163	14.7
2 publications	98	8.8
3 or more publications	99	8.9
First-author publications		
No publications	896	80.6
1 publication	138	12.4
2 publications	49	4.4
3 or more publications	29	2.6
Top 25% journal publications		
No publications	873	78.5
1 publication	129	11.6
2 publications	65	5.8
3 or more publications	45	4.0

To assess whether the scientific productivity differed between various groups, chi-square tests were performed. The number of publications does not differ significantly between male and female PhD students (Table 2). Research output differences exist between scientific disciplines (χ^2=32.53, df=4, p<.001): 4 in 10 doctoral students in the biomedical science have at least one publication, while ‘only’ about 2 in 10 doctoral students in the social sciences and humanities have already one or more publications. The scientific productivity also differs according to the type of funding a PhD student receives (χ^2=26.61, df=4, p<.001), and the phase of the PhD project (χ^2=215.48, df=2, p<.001). Assistants and researchers in the finishing phase of their PhD project were most likely to have one or more publications.
Differences in research output also exist between single researchers and those who have a partner, suggesting that single researchers and doctoral students without childcare responsibilities are less productive. However, this could possibly be explained by the fact that researchers with childcare responsibilities are more likely to be in the finishing phase of their PhD (χ^2=54.116, df=2, p<.001).

Table 2. Prevalence of having publications according to various socio-demographic characteristics

	0 publications		≥ 1 or more publications		$\chi^2(p)$
	N	%	N	%	
Gender					1.15 (ns)
Men	347	69.3	154	30.7	
Women	402	66.2	205	33.8	
Scientific discipline					32.54 (<.001)
Humanities	85	78.7	23	21.3	
Social sciences	214	78.1	60	21.9	
Science	128	66.3	65	33.7	
Biomedical science	196	59.0	136	41.0	
Applied science	125	63.8	71	36.2	
Type of funding					23.61 (<.001)
Assistant	117	57.4	87	42.6	
Personal grant	285	64.8	155	35.2	
Researcher of a project	196	72.9	73	27.1	
No funding	91	76.5	28	23.5	
Other	63	78.8	17	21.3	
Phase of PhD project					215.48 (<.001)
Planning phase	212	95.5	10	4.5	
Executing phase	458	70.9	188	29.1	
Finishing phase	77	32.5	160	67.5	
Having a partner					10.25 (<.001)
No	226	75.1	75	24.9	
Yes	523	65.0	282	35.0	
Childcare responsibilities					26.76 (<.001)
No	595	70.7	246	29.3	
Yes	85	50.3	84	49.7	

ns: not significant

Logistic regression models

Table 3 presents the results for the three multivariate logistic regression models. Gender is no significant predictor of PhD students' research productivity, neither is age. The likelihood of publishing more than one publication either as (co-)author or first author and of having one or more publications in top 25% journals is related to the scientific discipline. Our results show that compared to their colleagues in the biomedical sciences, PhD students in social sciences and humanities are less likely to have a publication in a peer reviewed journal and to have published in high ranked journals. Moreover, junior researchers in social sciences have a lower probability of being the first author of their publications in comparison with doctoral students in the biomedical sciences. However, no differences were observed with regard to publication quantity and quality between PhD students in the biomedical sciences and those in natural and applied sciences. The type of funding or scholarship a PhD student receives is also related to the odds of having more than one publication. Assistants are significantly more likely to have a publication as (co-)author or first-author and of having one or more publications in top 25% journals, compared to researchers who have an appointment based on project funds. Researchers who have obtained a personal grant are also more likely to have at least one first-author publication compared to researchers on project funds. As expected, the odds ratios for executing phase and finishing phase suggest that junior researchers in these phases of their PhD are substantially more likely to have one or more publications compared to doctoral students in the planning phase. No significant associations are found between having a partner, and the three productivity outcomes. However, having childcare responsibilities increases both the probability of having publications, either as (co-)author or first author, and of having publications in high ranked journals. Being satisfied with the guidance provided by the main supervisor is not associated with a higher likelihood of being more productive. The more competitive the culture within the research team, the more likely doctoral students are to be productive in terms of publication quantity.

Table 3. Logistic regression models for the 3 productivity outcomes, including odds ratio (OR), b coefficient (B), Wald, and significance of the included variables

	Number of publications (≥ 1 publication)				**Number of first-author publications (≥ 1 publication)**				**Number of publications in top 25% journals (≥ 1 publication)**			
	OR	**B**	**Wald**	***p***	**OR**	**B**	**Wald**	***p***	**OR**	**B**	**Wald**	***p***
Gender												
Male	1				1				1			
Female	1.05	0.05	0.09	.765	0.93	-0.08	0.15	0.93	0.96	-0.04	0.02	0.879
Age	0.99	-0.01	0.38	.539	0.99	-0.01	0.10	.755	0.97	-0.03	1.21	.271
Discipline												
Biomedical science	1				1				1			
Humanities	0.22	-1.50	19.25	<.001	0.61	-0.49	2.00	.157	0.04	-3.26	19.07	<.001
Social sciences	0.32	-1.12	23.48	<.001	0.40	-0.91	11.43	<.001	0.40	-0.92	13.52	<.001
Science	0.86	-0.15	0.38	.539	0.62	-0.48	2.72	.099	1.09	0.08	0.11	.737
Applied science	1.08	0.08	0.10	.748	1.16	0.15	0.30	.581	1.09	0.08	0.11	.740
Type of funding												
Researcher of a project	1				1				1			
Assistant	2.74	1.01	15.76	<.001	3.24	1.18	15.71	<.001	2.23	0.80	8.76	.003
Personal grant	1.48	0.39	3.53	.060	1.82	0.60	5.49	.019	1.41	0.34	2.28	.131
No funding	0.68	-0.39	1.49	.222	0.92	-0.09	0.05	.818	0.51	-0.66	3.31	.069
Other	0.76	-0.28	0.51	.476	1.13	0.12	0.07	.791	0.24	-1.44	6.25	.012
Phase of PhD project												
Planning phase	1				1				1			
Executing phase	8.90	2.19	32.92	<.001	7.09	1.96	13.84	<.001	11.71	2.46	22.17	<.001
Finishing phase	54.16	3.99	92.57	<.001	47.10	3.85	50.39	<.001	46.99	3.85	49.99	<.001
Having a partner												
No	1				1				1			
Yes	1.21	0.19	0.94	.332	0.97	-0.03	0.02	.897	1.09	0.09	0.15	.697
Childcare responsibilities												
No	1				1				1			
Yes	1.59	0.46	3.93	.047	1.75	0.56	5.08	.024	2.04	0.71	8.37	.004
Organizational culture	1.36	0.31	4.32	.038	1.62	0.48	8.00	.005	1.03	0.03	0.04	.845
Satisfaction with coaching by supervisor	0.97	-0.03	0.13	.720	1.18	0.17	3.92	.068	1.05	0.05	0.45	.501

Discussion

The present study investigates whether personal characteristics, specific features of the PhD project, and research environment-related factors influence the publication performance of junior researchers, using a unique dataset combining bibliometric data of PhD students in a public research university and data obtained from the Survey of Junior researchers II. We considered both PhD students' research quantity and quality.

Men have been consistently shown to be more productive than women (Leahey, 2006; Stack, 2004; Symonds et al., 2006), although in our study no gender gap in research productivity is found. One possible explanation might be that the population studied is on average quite young and the large majority of respondents have no children yet. Gender difference may be more common at a later stage in the academic career when more researchers take up childcare responsibilities, for instance during the postdoc phase (Stack, 2004). Nevertheless, our results reveal that PhD students who have childcare responsibilities are more productive, both in terms of publication quantity and quality. These group of PhD students might be more eager to finish their doctoral degree within a reasonable time span due to their family obligations and in order to be able to consider new more long-term career prospect on the (non-)academic labour market.

Not surprisingly the phase of the PhD project, reflecting the work experience of PhD students, is related to a higher probability of having a higher number of publications.

As expected, the total number of publications, the number of first-author publications, and the publication quality based on scientific output in the Web on Knowledge are significantly lower in social sciences and humanities compared to the other scientific disciplines. There were no substantial differences in publication quantity and quality between doctoral students in the biomedical, natural and applied sciences.

PhD students who are embedded in a research team characterised by a competitive culture, with a strong focus on assertiveness, power, and ambition, have a significant higher probability of being more productive than students in research teams where a more supportive culture, placing more value on relationships and quality of life is prevailing.

PhD students who have been awarded a scholarship are more likely to have at least one first-author publication, whereas assistants seem to be more productive in terms of both quantity and quality than researchers who have an appointment on a project. For researchers with a prestigious personal grant this is in line with the expectations, as they generally have a full-time appointment to focus on their research, but for assistants this is more surprising as they have to spend on average 50% of their time on teaching activities.

So, scientific discipline, phase of the PhD process, having childcare responsibilities, funding situation and organizational culture are important factors predicting the total number of publications and first-author publications. Similar results are obtained with regard to publication quality, although the organizational culture in the research group could not predict the probability of publishing in top 25% journals. Possibly other factors may play a role here, for example the intrinsic motivation of the PhD student and the overall focus of the PhD supervisor on high quality publications.

However, when interpreting our results, we need to take into account the turnaround time for articles to get published in the Web of Knowledge and the short time period or our data. Due to the fact that only 32.4% of the PhD students had already one or more publications, we opted to use whole counts instead of fractional counts that adjust for the total number of co-authors on a publication. If we could repeat the same exercise within a couple of years we might get more detailed information concerning doctoral students' publication process. We could expect that investing in high quality research takes considerable more time than publishing in lower ranked journals, but further research is needed to support this assumption.

Although we controlled for scientific discipline in our analyses, we cannot rule out that it might be more difficult to get a publication in a top 25% journal in a small field as compared to a larger field with more journals in its subject category in Web of Knowledge. A further step would be the creation of an optimal quality measurement.

References

Bornmann. L.. & Marx. W. (2014). How to evaluate individual researchers working in the natural and life sciences meaningfully? A proposal of methods based on percentiles of citations. Scientometrics. 98(1). 487-509.

Cardoso. A. R.. Guimaraes. P.. & Zimmermann. K. F. (2010). Comparing the early research performance of PhD graduates in labor economics in Europe and the USA. Scientometrics. 84(3). 621-637.

Duffy. R. D.. Jadidian. A.. Webster. G. D.. & Sandell. K. J. (2011). The research productivity of academic psychologists: assessment. trends. and best practice recommendations. Scientometrics. 89(1). 207-227.

Fox. M. F. (2005). Gender. family characteristics. and publication productivity among scientists. Social Studies of Science. 35(1). 131-150.

Hofstede. G.. Neuijen. B.. Ohayv. D. D.. & Sanders. G. (1990). Measuring organizational cultures - a qualitative and quantitative study across 20 cases. Administrative Science Quarterly. 35(2). 286-316.

Leahey. E. (2006). Gender differences in productivity - Research specialization as a missing link. Gender & Society. 20(6). 754-780.

Manana-Rodriguez. J.. & Gimenez-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(3). 893-91.

Ruane. F. P.. & Tol. R. S. J. (2009). A Hirsch measure for the quality of research supervision. and an illustration to trade economists. Scientometrics. 80. 613–624.

Stack. S. (2004). Gender. children and research productivity. Research in Higher Education. 45(8). 891-920.

Symonds. M. R. E.. Gemmell. N. J.. Braisher. T. L.. Gorringe. K. L.. & Elgar. M. A. (2006). Gender Differences in Publication Output: Towards an Unbiased Metric of Research Performance. Plos One. 1(2).

Visser. M. S.. Luwel. M.. & Moed. H. F. (2007). The attainment of doctoral degrees at Flemish Universities: a survival analysis. Higher Education. 54(5). 741-757.

Exploring internationality and collaborative behaviour of scientists in Social Sciences and Humanities[1]

Adrián A. Díaz-Faes*, Maria Purificación Galindo** and María Bordons*

*adrian.arias@cchs.csic.es; maria.bordons@cchs.csic.es
Quantitative Analysis in Science & Technology Group (ACUTE), IFS, Centre for Human and Social Sciences (CCHS), Spanish National Research Council (CSIC), Albasanz 26-28, Madrid 28037 (Spain)

** pgalindo@usal.es
Statistics Department, University of Salamanca, Alfonso X El Sabio s/n, Salamanca 37007 (Spain)

Introduction

Research is increasingly collaborative and interdisciplinary. Collaboration allows scientists to face larger and more complex problems and to optimize resources; and its positive effects on the quantity and quality of scientific contributions have been extensively described (Glänzel, 2001). Moreover, interdisciplinarity has been associated with the most innovative breakthroughs in science.

Research organizations have to adapt to the changing conditions of research and take steps to facilitate the development of high quality and highly innovative research. Spanish National Research Council (CSIC), the largest public institution devoted to research in Spain, develops four-year action plans to define strategies, establish priorities and assign resources. Strategic planning is developed to promote international presence, gain visibility and remain competitive (CSIC, 2009). Interdisciplinary collaborative research is promoted as a way to enhance quality and scope of the research and the dissemination of scientific output in high prestige international channels is recommended.

CSIC comprises more than 4,000 scientists and 125 institutes spread all over the country which are organized in eight scientific areas. In this paper the study of the Social Sciences and Humanities area (SSH), which is the least collaborative one, is addressed from a micro-level perspective. The aim of this paper is to study the publication behaviour of SSH scientists at CSIC with special attention to specific aspects that characterize modern science, such as collaboration, interdisciplinarity and international orientation. Differences by age, gender, and academic rank are studied, as well as between Social Sciences and Humanities.

Methods

This paper focuses on 261 active scientists in the SSH area of CSIC in 2007. Personal data and publication-based indicators were analysed.

1. Personal data: age, professional rank (P=postdoc, TS=tenured scientist, RS=research scientist and RP=research professor) and gender of scientists were provided by CSIC.
2. Publications of CSIC's scientists in Web of Science (WoS) during 2007-2011. Publications in SSH were identified (Gómez et al., 2012) and assigned to scientists by means of specific algorithms that normalize author's names. The following indicators were obtained for every researcher:

[1] Financial support from a JAE predoctoral fellowship from CSIC is acknowledged.

a. Number of papers (only articles and reviews).
b. Impact, measured through the percentage of papers in high impact journals (journals in the first quartile), percentage of non-cited papers, and relative citation rate (RCR) (citations normalised to the average citation rate of the country in the category of the publication journal).
c. Collaboration. The collaboration profile of each scientist (% of single-address papers, % national collaboration and % international collaboration) and the percentage of single-authored papers were obtained.
d. Local/international orientation, which is analysed through the trend of scientists to publish in Spanish, English or another language.
e. Interdisciplinarity. Pratt index (Pratt, 1977) is used to study the concentration of papers by subject categories according to the assignment of the publication journals to WoS categories. It ranges from 0 (low concentration) to 1 (high concentration).

The relationship between personal data and publication behaviour is explored through Non-linear Principal Components Analysis (NLPCA) (SPSS v.20), which allows reducing a set of variables to a smaller number of non-correlated underlying variables. Discrete variables as professional rank and gender are quantified assigning numeric values that maximize association between quantified variables through optimal scaling.

Findings

A total of 1,612 papers were published by CSIC scientists in SSH during 2007-2011. 173 scientists out of 261 (66.3%) had at least one publication in WoS. Large differences in the bibliometric profile of social and humanities scientists are observed, being on average social scientists more collaborative, interdisciplinary and internationally oriented (Table 1).

Table 1. Average behaviour of SSH scientists.

		HUM (n=117)	SOC (n=56)
Personal data	Age	50.30	48.36
	Professional rank (P/TS/RS/RP)(%)	7/50/20/23	7/52/21/20
	%Women	43.6	48.2
Production	No. papers	2.89	5.41
Impact	%Papers in Q1 journals	11.94	34.76
	%Non-cited papers	75.92	36.37
	RCR	0.67	0.99
Collaboration	%Single authored papers	66.23	13.99
	%Collaboration papers	22.11	48.06
	%International coll. papers	6.92	22.57
	%National coll. papers	17.62	32.00
Interdisciplinarity	Pratt Index	0.55	0.49
International orientation	%Spanish	72.17	35.15
	%English	24.74	64.40

NLPCA including all scientists and variables is developed and a two-dimension solution which accounts for 47% of the variance is obtained. Graphically, the scatterplot for component loadings (Figure 1) shows that nationally oriented activity (written in Spanish) is positively correlated with single-authored papers, low interdisciplinarity and non-cited papers; while internationally oriented activity (written in English) is associated with higher number of papers, higher impact and greater collaboration. Secondly, the behaviour of Social Sciences and Humanities scientists is analysed separately. The national/international activity dimension can be observed in both cases, but in Social Sciences academic rank is positively related with production and impact and older scientists show higher tendency to get involved in nationally oriented research.

Figure 1. Component loadings.

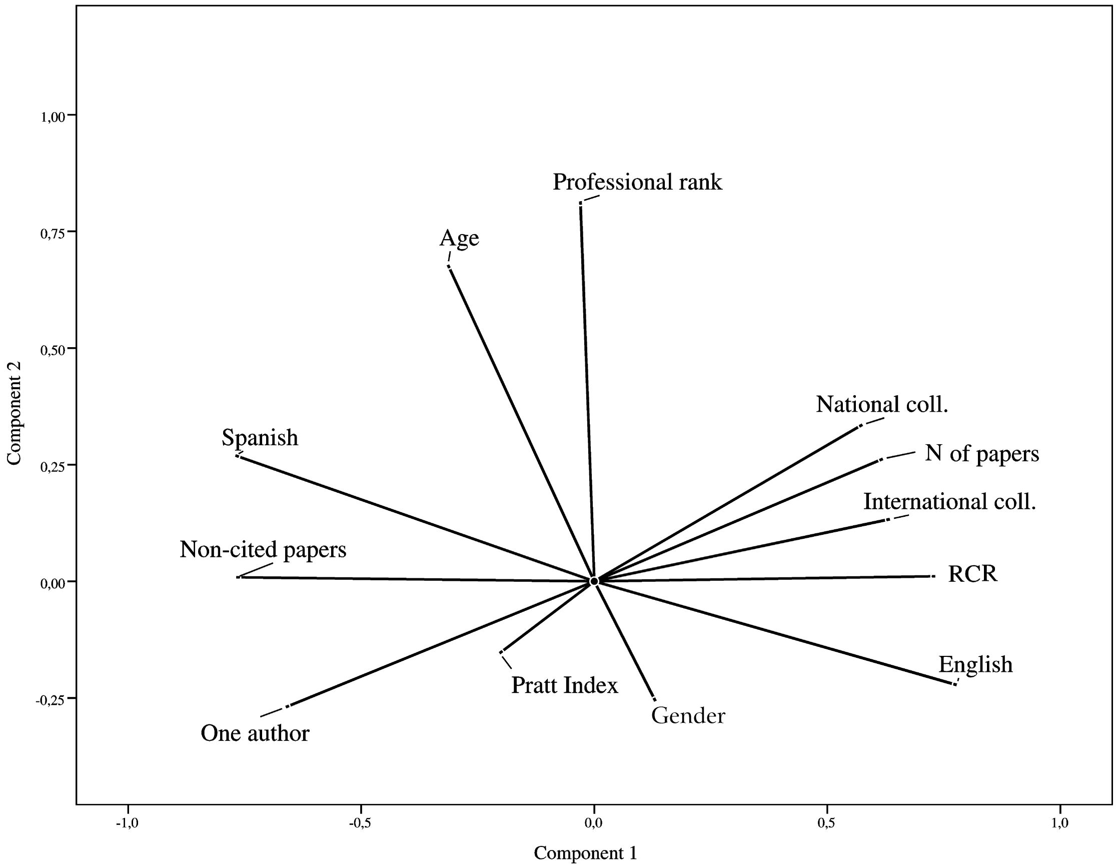
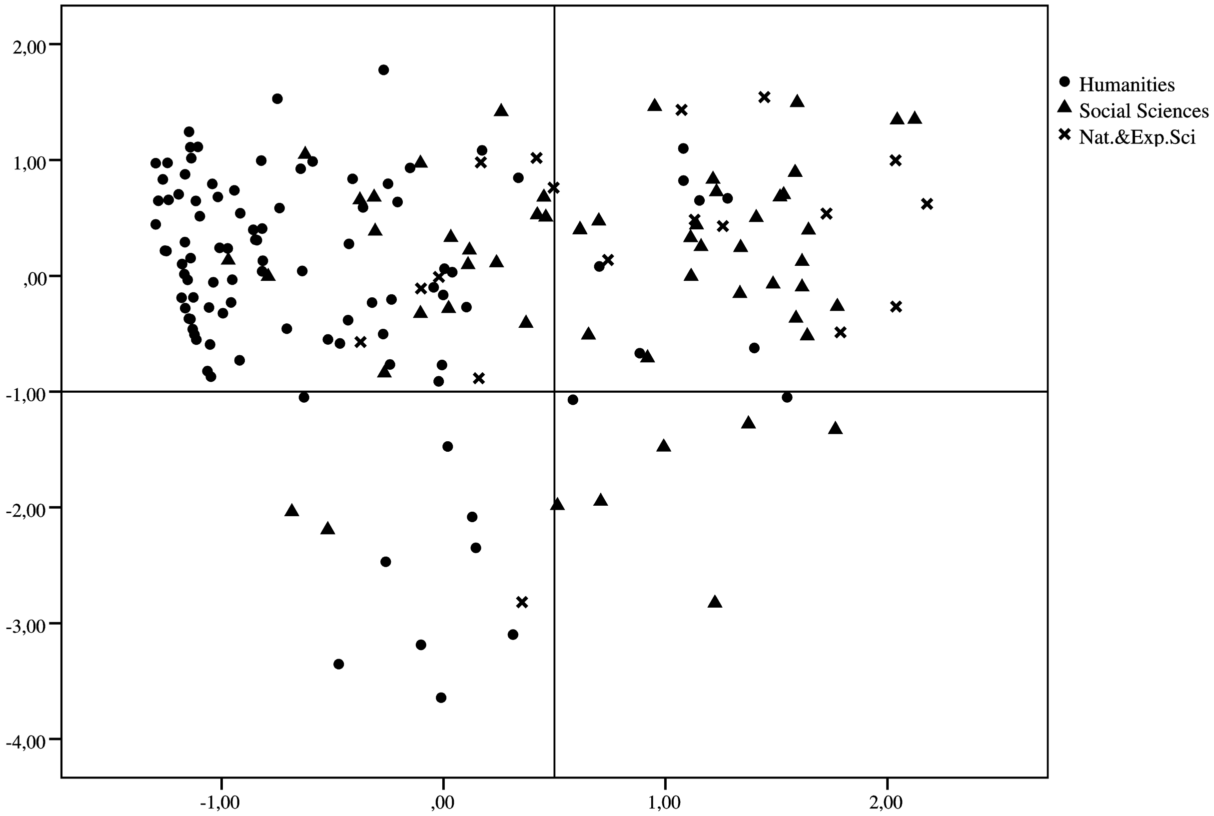


Figure 2. NLPCA scores for author's behaviour[2].



Discussion and conclusions

The research output of SSH scientists is only partly considered in this study since non-WoS publications (books, book chapters or a number of national journals), which are important in this area (Huang & Chang, 2008; Ossenblonk et al. 2014), are not taken into account. However, we considered it was not a major limitation since we had descriptive purposes rather than evaluative ones.

Humanities scientists show lower values of collaboration, interdisciplinarity and internationality -measured through the presence of foreign co-authors or English language in papers-. Moreover, higher rates of collaboration and interdisciplinarity are observed for the most internationally-oriented scientists in both Social Sciences and Humanities.

The low international orientation of most humanities scientists can be due to the local and regional concerns of their research, whose target audience is also local (Bordons & Gómez, 2000). Their low collaborative activity suggests the predominance of a "traditional" conception of research as opposed to a "modern" one more cooperative and interdisciplinary (Oschner et al., 2013).

Our analysis shows that SSH is a heterogeneous area. The large differences observed in the behaviour of social and humanities scientists point to the need to analyse separately both communities of scientists. The fact that academic rank is positively related with production and impact in Social Sciences but not in Humanities is probably due to the best WoS coverage

[2] Author's field was assigned based on WoS category publication profile.

of the output of scientists in Social Sciences. Our next objective is to include non-Wos output in the analysis in order to achieve a more comprehensive picture of the area.

References

Bordons, M. & Gómez, I. (2000). Collaboration networks in science. In B. Cronin & H.B. Atkins (Eds.), The Web of Knowledge: A Festschrift in Honor of Eugene Garfield (pp. 197–213). Medford, NJ: Information Today, Inc. & ASIS.

CSIC. Plan de Actuación 2010-2013. http://www.csic.es/web/guest/plan-de-actuacion-2010-2013.

Glänzel, W. (2001). National characteristics in international scientific co-authorship relations. Scientometrics, 51(1), 69-115.

Gómez, I., Bordons, M., Morillo, F., Moreno, L., Aparicio, J. & González-Albo, B. (2012). La actividad científica del CSIC a través de indicadores bibliométricos. Madrid: CSIC.

Huang, M.H. & Chang, Y.W. (2008). Characteristics of Research Output in Social Sciences and Humanities: From a Research Evaluation Perspective. JASIST, 59(11), 1819–1828.

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.

Ossenblok, T.L.B., Verleysen, F.T. & Engels, T.C.E (2014). Coauthorship of journal articles and book chapters in the Social Sciences and Humanities (2000-2010). JASIST, 65(5), 882-897.

Pratt, A.D. (1977). A measure of class concentration in bibliometrics. JASIS, 28(5), 285-292.

Article Properties Associating with the Citation Impact of Individual Articles in the Social Sciences

Fereshteh Didegah* - Mike Thelwall**

*f.didegah@wlv.ac.uk - **m.thelwall@wlv.ac.uk
Statistical Cybermetrics Research Group, School of Mathematics and Computing, University of Wolverhampton, Wulfruna Street, Wolverhampton WV1 1LY, UK

Abstract
This study investigates a range of metrics available when an article is published to see which metrics associate with its eventual citation count. The purposes are to contribute to developing a citation model and to inform policymakers about which predictor variables associate with citations in the social sciences. Despite the complex nature of reasons for citation, some attributes of a paper's authors, journal, references, abstract, field, and country and institutional affiliations are known to associate with its citation impact. This study investigates some common factors and some new factors. Using negative binomial hurdle models, journal and reference factors were found to be the most effective determinants of future citation counts. Individual and institutional teamwork give a citation advantage in the social sciences but international teamwork seems not to contribute to citation impact.

Introduction

This study investigates a range of properties that are available when a social science article is published to see which of them associate with its eventual citation count. The purposes are to contribute to developing a statistical model of the key factors that associate with citations and to inform policymakers about which predictor variables associate with increased citations in the social sciences. This study examines some common article properties that have been previously assessed and some new factors (see Table 1).

Although the content and quality of a research paper is presumably the main reason for its citations, other factors are known to statistically associate with article citation counts. For example, collaboration increases citation impact (Gazni & Didegah, 2010; Sooryamoorthy, 2009). Journal impact, reference impact, author impact, and institution and country impact can also increase the impact and visibility of articles (Boyack & Klavans, 2005; Leimu & Koricheva, 2005). Papers with more references will be cited more (Peters & Van Raan, 1994), as will funded research (Levitt, 2011) and papers with less readable abstracts, at least in the five top institutions in the world (Gazni, 2011). Articles with longer abstracts have also been found to receive more citations and longer papers may be cited more too (Kostoff, 2007).

Research Questions

No previous studies have examined many document properties that may associate with citations for the broad social sciences; only a few factors in specific sub-fields of the social sciences have been investigated. Moreover, previous studies have used statistical tests (e.g., correlations) that are powerful enough to differentiate between correlating factors (Chen, 2012; Leimu & Koricheva, 2005). Using a negative binomial-logit hurdle model, this study conducts a simultaneous analysis of multiple social science document properties that may associate with future citations. The following research questions are (simultaneously) addressed.

1. Which types of research collaboration (individual, institutional and international) associate with increased citation impact?
2. Do author, institution and country impact associate with increased citation impact?
3. Do journal and reference characteristics (journal impact and internationality, reference impact and internationality, and total references) associate with increased citation impact?
4. Which field size and article size attributes (article, abstract, and title length) associate with increased citation impact?
5. Do articles with more readable abstracts receive more citations?
6. *To what extent* do the above factors associate with increased citation counts?

Methods

The publications selected for this study were taken from Thomson Reuters' Web of Science (WoS). A sample of 16,096 journal articles and conference papers was taken from the Social Sciences Citation Index from 2000 to 2009, with each year randomly sampled in proportion to the number of articles in that year. The year 2009 was selected as the end date in order to allow all articles sufficient time (3 years) to accrue a substantial number of citations. The dependent variable is the citation counts for papers and the independent variables and their measures are listed in Table 1.

In this study, the Gini coefficient is a value between 0 and 1 where 0 represents the highest level of internationality (i.e., all authors from the different countries) and 1 represents the least internationality (i.e., all authors from the same country) (Buela-Casal, Perakakis, Taylor & Checa, 2006). All Gini coefficients were multiplied by -1 so that higher values associate with increased internationality in order to make the results easier to interpret. The Gini formula is:

$$\mathrm{Gini} = \left|1 - \sum_{i=1}^{N} (X'_i - X'_{i-1})(Y'_i + Y'_{i-1})\right|$$

Where:

N = Number of countries contributing to the journal;

Xi = Cumulative proportion of X where X=1/N;

Yi = Cumulative proportion of authors publishing in or citing the journal from countries 1 to i, where the countries are arranged in descending order of the number of authors contributing to the journal.

Note: When i=1, X'_{i-1} and Y'_{i-1} equal zero.

Table 1. Independent variables and measures.

Main factor	Sub-factors (Independent variables)	Measure
Research collaboration	Number of authors	Number of authors in the WoS AU field.
	Number of institutions	Number of different institution names in the WoS C1 field.
	Number of countries	Number of different country names in the WoS C1 field.
Impact of the paper	Author impact	Maximum h-index of the publishing authors in the WoS AU field.
	Journal Impact Factor (JIF)	Journal Impact Factor retrieved from JCR for the journal in the WoS SO field.
	Reference impact	Median citation impact of references

	Institution impact	Maximum Mean Normalized Citation Score (MNCS) of the institutions in the WoS C1 field.
	Country impact	Maximum Mean Normalized Citation Score (MNCS) of the countries in the WoS C1 field.
Internationality of the paper	Journal author internationality (J. auth. internationality)	Gini coefficient of the journal named in the WoS SO field.
	Journal citing author internationality (J. citer internationality)	Gini coefficient of the journal in the WoS SO field.
	Cited journal author internationality (Ref. auth. internationality)	Average Gini coefficient of the references in the WoS CR field.
	Cited journal citing author internationality (Ref. citer internationality)	Average Gini coefficient of the references in the WoS CR field.
Paper size	Paper length (Number of pages)	Number of pages in the WoS PG field.
	Abstract length	Number of words in the WoS AB field.
	Title length	Number of words in the WoS TI field.
	Number of references	Number of references in the WoS CR field.
	Field size	Number of publications in the related sub-field
Readability of the paper	Abstract readability	Flesch readability score of the WoS AB field. Flesch scores range between 0 to 1 and a score closer to 1 means that the abstract is more readable.

A Negative Binomial (NB)-logit hurdle model was fitted to the data (Hilbe, 2011). This analyses the zeroes separately from the other citation counts on the assumption that different processes underlie them. The hurdle model is suitable because it seems reasonable to assume that it is a significant hurdle for a paper to receive its first citation but after this it is more likely to be cited in the future, for example by others finding the paper with a citing references search or by the paper being ranked higher in search systems because of its citation(s). The overdispersion parameter is significant in this model, further justifying the use of the NB model (p for alpha < 0.01) rather than a simpler one. To diagnose multicollinearity between the variables the Variance Inflation Factor (VIF) was used with the rule of thumb that VIFs over 4 indicate severe multicollinearity (Chatterjee, Hadi & Price, 2000). The largest VIF value was 2.09, suggesting the absence of serious multicollinearity among the variables. The results also report the change in the expected number of citations between an article with a value on the 25th percentile for a property to an article with a value on the 75th percentile for the property. This indicates the importance of the property for citations in the sense of the increase in citations obtainable from a reasonable change in its value.

Instead of normalizing the citation counts by year of publication or using a citation window, the publication year was entered into both the logit and NB models to control for the effect of the publication year.

Results

With respect to the NB model (Table 2), field size and number of countries do not significantly associate with citation counts. Except for journal citer internationality, reference citer internationality and title length, all the other factors significantly associate with increased citation counts to articles. Title length significantly associates with decreased citation counts and a unit increase in the factor (i.e., an extra word in the title) associates with a 2% decrease in citation counts. Moreover, an increase in title length from the 25th quartile to the 75th quartile decreases the expected number of citations by only 11.3%. Journal citer internationality and reference citer internationality also significantly associate with decreased citation counts and a unit increase in each factor associates with 36.8% and 42.1% decrease in citation counts, respectively. A unit increase in reference impact, number of references, number of pages, abstract length, and abstract readability gives a small increase in citations (less than 1%) but the percentage increase in the citation counts for the change between 25th quartile and 75th quartiles of each factor is considerable (20.4%, 16.2%, 10.1%, 8.8%, and 5.5%, respectively).

With respect to the logit model (Table 3), field size, number of countries, reference author internationality, reference citer internationality, and title length are not significant factors for zero citations. All the other factors except for the number of institutions and journal citer internationality contribute to decreased zero citations.

Discussion and conclusions

Research Collaboration: Individual and institutional collaborations significantly associate with increased citations. Multi-author research is known to receive more citations than does solo research (Gazni & Didegah, 2010) except perhaps in library and information science (Hart, 2007), economy and finance (Medoff, 2003), social and personality psychology (Haslam, et al., 2008), and chemistry (Bornmann, Schier, Marx, & Daniel, 2012). International collaboration is not significant in the social sciences, however. The contradiction between this finding and previous studies of international collaboration may be due to the limited geographical and institutional coverage of previous research or may be due to the simpler statistical models used in most previous studies, which mostly did not analyse multiple factors simultaneously. In other words, the key factor may be collaboration rather than specifically international collaboration. Alternatively, given the significance of the impact of the affiliated countries (see below), it could be that only collaboration with high impact countries is beneficial for future citations.

Article properties: Higher JIFs significantly associate with substantially increased citations and decreased zero citations. An increase from the lower to the upper JIF quartile results in a 51% increase in expected citations and a 70% decrease in the number of zero citations. This is unsurprising because JIFs are based upon average citations and so article citations should be mathematically related to the publishing journal JIF and this is in agreement with previous findings for a range of scientific fields (e.g., Boyack & Klavans, 2005). Higher impact references also associate with increased citations and decreased zero citations. A change between quartiles associates with a 20% increase in positive citation counts and a 13% decrease in zero citations. This finding concurs with the results of previous studies (Bornmann, Schier, Marx, & Daniel, 2012; Didegah & Thelwall, 2013ab). Author impact significantly associates with increased citation counts and decreased zero citations. A unit increase in author impact associates with a 3% increase in positive citation counts whereas an increase between lower and upper quartiles gives a 30% increase in the expected number of citations. This is in agreement with previous studies (Vanclay, 2013). Institutional impact is also a significant but minor determinant of increased citation counts. The percentage increase in the mean parameter for citations due to a unit increase and a change between quartiles both give only a 4% increase, however. The factor contribution to decreased zero citations is 28% from an increase from the lower to the upper quartile. Country impact is a significant factor for both positive citation counts and zero citations. A unit increase in country impact associates with a 29% increase and a change from the lower to upper quartile associates with a 61% increase in the positive citation counts. With respect to the logit model, an increase from the lower to the upper quartile associates with a 57% decrease in zero citations.

Journal and reference internationality: An increase from the lower to the upper quartile in journal author internationality associates with a 5.3% increase in expected citations. With respect to the logit model, increased journal author internationality associates with decreased zero citations. Increased journal citer internationality significantly associates with decreased citation counts. Journal citer internationality gave contrary results: a significant association with decreased citation counts, showing that being cited from across the world does not matter for citations whereas having authors from different countries publishing in the journal helps articles' citations. Reference internationality was also measured in two ways with the Gini coefficient, the geographic dispersion of the cited journal authors and that of the cited journal citers. Reference author internationality significantly associates with increased citation counts. The results show that articles with more international references in terms of the geographic dispersion of their authors receive more citations than do articles with less international references. The increase in citation counts for an increase between the lower and upper quartiles in reference author internationality is 20.2%. Reference citer internationality is a significant determinant of decreased citation counts, showing that articles with more international references in terms of the geographic dispersion of their citers received fewer citations.

Article size attributes: More references significantly associates with increased citation counts and each additional reference associates with a 0.5% increase in expected citations. Moreover, an increase from the lower to the upper quartile increases the expected citation counts by 16%. It is known that articles with more references are cited more (Mingers & Xu, 2010). Field size in terms of the number of publications in the social sciences WoS sub-field is not a significant factor. Title length significantly associates with decreased citation counts showing that articles with shorter titles receive more citations and concurring with Ayres and Vars (1999). Abstract length significantly associates with increased citation counts. The same result was found in Kostoff (2007) for medical articles. Paper length is a significant factor for increased citation counts. A number of micro-studies in different subject areas have also confirmed this (Vanclay, 2013; Mingers & Xu, 2010).

Abstract readability: More readable abstracts significantly associates with increased citation counts although an increase from the lower to upper quartile (from difficult to easier abstracts) increases the mean parameter for citation counts by only 6%. Moreover, abstract readability associates with decreased zero citations in the social sciences and a change between the lower and upper quartiles results in a 18% decrease in the probability of zero citations.

To summarize, and using the inter-quartile changes in Table 2 as the main guide, journal and reference characteristics, and particularly journal and reference impact, are the main extrinsic properties of articles that associate with their future citation impact in the social sciences.

Journal and reference internationality can also help with the prediction of future citation impact. Research collaboration, and particularly individual and institutional collaboration, can help to predict citation counts for articles but international collaboration alone is not important, unless it is with a high impact nation. Paper length, abstract length and abstract readability are also significant determinants of citation counts, but not all make a substantial difference. In the world top institutions, articles with more readable abstracts (i.e., easier to read) were less cited but in the social sciences more readable abstracts are more cited. These conclusions have been obtained using a method that minimises the chance that spurious factors have been identified due to their correlations with genuine factors.

References

Ayres, I. & Vars, F.E. (1999). Determinants of citations to articles in elite law reviews. In: *Conference on Interpreting Legal Citations*, Evanston, Illinois.

Bornmann, L., Schier, H., Marx, W. & Daniel, H-D. (2012). What factors determine citation counts of publications in chemistry besides their quality? *Journal of Informetrics*, 6, 11-18.

Boyack, K. W. & Klavans, R. (2005). Predicting the Importance of Current Papers. *Proceedings of ISSI 2005*, 335–342. In P. Ingwersen & B. Larsen (eds.), July 24-28, Stockholm, Sweden.

Chatterjee, S., Hadi, A. S. & Price, B. (2000). *Regression analysis by example*. John Wiley and Sons, New York, New York, USA.

Chen, C. (2012). Predictive effects of structural variation on citation counts. *Journal of the American Society for* Information *Science & Technology*, 63(3), 431–449.

Didegah, F. & Thelwall, M. (2013a). Determinants of research citation impact in Nanoscience and Nanotechnology. *Journal of the American Society for Information Science & Technology*, 64(55), 1055-1064.

Didegah, F. & Thelwall, M. (2013b). Which factors help authors produce the highest impact research? Collaboration, journal and document properties. *Journal of Informetrics*, 7, 861-873.

Gazni, A. & Didegah, F. (2010). Investigating different types of research collaboration and citation impact: A case study of Harvard University's publications. *Scientometrics*, 87(2), 251-265.

Gazni, A. (2011). Are the abstracts of high impact articles more readable? Investigating the evidence from top research institutions in the world. *Journal of Information Science*, 37 (3), 273-281.

Hart, R.L. (2007). Collaboration and article quality in the literature of academic librarianship. The Journal of Academic Librarianship, 33(2), 190–195.

Haslam, N. et al. (2008). What makes an article influential? Predicting impact in social and personality psychology. *Scientometrics*, 76(1), 169-185.

Hilbe, J. M. (2011). Negative Binomial *Regression* (2ed). Cambridge, UK: Cambridge University Press.

Leimu, R., & Koricheva, J. (2005). What determines the citation frequency of ecological papers? *Trends in Ecology & Evolution*, 20, 28-32.

Levitt, J. M. (2011). Are funded articles more highly cited than unfunded articles? A preliminary investigation. *Proceedings of ISSI 2011*, 1013-1015. E. Noyons, P. Ngulube, & J. Leta, (eds.), July 4-7, Durban, South Africa.

Kostoff, R. (2007). The difference between highly and poorly cited medical articles in the journal Lancet. *Scientometrics*, 72, 513-520.

Medoff, M.H. (2003). Collaboration and the quality of economics research. *Labour Economics*, 10, 597–608.

Mingers, J. & Xu, F. (2010). The drivers of citations in management science journals. *European Journal of Operational Research*, 205, 422–430.

Peters, H.P.F. & Van Raan, A.F.J. (1994). On determinants of citation scores: A case study in Chemical Engineering. *Journal of the American Society for Information Science*, 45(1), 39-49.

Sooryamoorthy, R. (2009). Do types of collaboration change citation? Collaboration and citation patterns of South African science publications. *Scientometrics*, 81(1), 177–193.

Vanclay, J.K. (2013). Factors affecting citation rates in environmental science. *Journal of Informetrics*, 7(2), 265-271.

Appendix: Tables 2-3

Table 2. Analysis of hurdle model results for 16,096 WoS social sciences journal articles and conference papers from 2000 to 2009.

Logit model	Significance	Decreasing/Increasing probability of zero citations	% change in the probability of zero citations	% change between lower and upper quartiles	Unit
Field size	Insignificant				
No. authors	Significant	Decreasing	-17.3	-37.7	Citations per extra author
No. institutions	Significant	Increasing	15.0	15.0	Citations per extra institution
No. countries	Insignificant				
JIF	Significant	Decreasing	-19.2	-69.9	Citations per extra IF
Ref. impact	Significant	Decreasing	-0.4	-13.2	Citations per extra median citations
Author impact	Significant	Decreasing	-3.4	-35.8	Citations per extra h-index
Institution impact	Significant	Decreasing	-28.4	-0.6	Citations per extra MNCS
Country impact	Significant	Decreasing	-22.9	-56.8	Citations per extra MNCS
J. auth. internationality	Significant	Decreasing	-88.7	-12.4	Citations per extra GINI
J. citer internationality	Significant	Increasing	0.4	30.2	Citations per extra GINI
Ref. auth. internationality	Insignificant				
Ref. citer internationality	Insignificant				
No. refs	Significant	Decreasing	-0.8	-27.1	Citations per extra reference
No. pages	Significant	Decreasing	-3.0	-41.6	Citations per extra page
Title Length	Insignificant				
Abs. length	Significant	Decreasing	-0.2	-18.3	Citations per extra word
Abs. readability	Significant	Decreasing	-0.9	-17.5	Citations per extra Flesch Score
NB model	**Significance**	**Decreasing/Increasing citations**	**% change in the mean parameter of positive citation counts**	**% change between lower and upper quartiles**	**Unit**
Field size	Insignificant				
No. authors	Significant	Increasing	10.6	22.4	Citations per extra author
No. institutions	Significant	Increasing	5.0	5.0	Citations per extra institution
No. countries	Insignificant				
JIF	Significant	Increasing	61.1	50.9	Citations per extra IF
Ref. impact	Significant	Increasing	0.6	20.4	Citations per extra median citations
Author impact	Significant	Increasing	2.9	29.8	Citations per extra h-index
Institution impact	Significant	Increasing	4.4	0.1	Citations per extra MNCS
Country impact	Significant	Increasing	28.6	61.1	Citations per extra MNCS
J. auth. internationality	Significant	Increasing	32.0	5.3	Citations per extra GINI
J. citer internationality	Significant	Decreasing	-36.8	-80.4	Citations per extra GINI
Ref. auth. internationality	Significant	Increasing	63.1	20.2	Citations per extra GINI
Ref. citer internationality	Significant	Decreasing	-42.1	-19.8	Citations per extra GINI
No. refs	Significant	Increasing	0.5	16.2	Citations per extra reference
No. pages	Significant	Increasing	0.8	10.1	Citations per extra page
Title Length	Significant	Decreasing	-2.0	-11.3	Citations per extra word
Abs. length	Significant	Increasing	0.1	8.8	Citations per extra word
Abs. readability	Significant	Increasing	0.3	5.5	Citations per extra Flesch Score

Table 3. The results of the hurdle model for 16,096 WoS social sciences journal articles and conference papers from 2000 to 2009.

Logit model	Coef.	Exp (Coef.)	Std. Err.	z	P>z	[95% Conf. Interval]	
Field size	0.000	1.000	0.000	0.780	0.433	0.000	0.000
No. authors	0.160	1.173	0.016	10.080	0.000	0.129	0.191
No. institutions	-0.162	0.851	0.057	-2.850	0.004	-0.273	-0.050
No. countries	0.038	1.039	0.107	0.350	0.724	-0.172	0.248
JIF	0.785	2.192	0.032	24.260	0.000	0.721	0.848
Ref. impact	0.004	1.004	0.002	2.390	0.017	0.001	0.007
Author impact	0.034	1.034	0.007	4.910	0.000	0.020	0.047
Institution impact	0.250	1.284	0.067	3.730	0.000	0.119	0.382
Country impact	5.445	231.611	36.766	0.690	0.489	-4.615	7.505
J. auth. internationality	0.635	1.887	0.181	7.170	0.000	0.546	1.653
J. citer internationality	-5.503	0.004	0.161	-34.150	0.000	-5.819	-5.187
Ref. auth. internationality	-0.018	0.982	0.167	-0.110	0.912	-0.345	0.308
Ref. citer internationality	0.015	1.015	0.293	0.050	0.960	-0.559	0.589
No. refs	0.008	1.008	0.003	2.990	0.003	0.003	0.013
No. pages	0.029	1.030	0.008	3.880	0.000	0.014	0.044
Title Length	0.004	1.004	0.011	0.350	0.727	-0.017	0.024
Abs. length	0.002	1.002	0.001	2.670	0.007	0.001	0.004
Abs. readability	0.009	1.009	0.004	2.460	0.014	0.002	0.017
Constant	-2.341	0.096	0.541	-4.330	0.000	-3.402	-1.280
NB model	Coef.	Exp (Coef.)	Std. Err.	z	P>z	[95% Conf. Interval]	
Field size	0.000	1.000	0.000	-0.620	0.534	0.000	0.000
No. authors	0.101	1.106	0.013	7.760	0.000	0.075	0.126
No. institutions	0.049	1.050	0.024	2.060	0.040	0.002	0.095
No. countries	0.052	1.053	0.045	1.130	0.256	-0.038	0.141
JIF	0.477	1.611	0.024	19.490	0.000	0.429	0.525
Ref. impact	0.006	1.006	0.001	10.590	0.000	0.005	0.007
Author impact	0.029	1.029	0.002	14.120	0.000	0.025	0.033
Institution impact	0.043	1.044	0.013	3.280	0.001	0.017	0.069
Country impact	5.401	221.560	6.081	0.890	0.374	-6.518	17.320
J. auth. internationality	0.278	1.320	0.113	0.910	0.007	-0.119	0.324
J. citer internationality	-1.001	0.368	0.111	-32.790	0.000	-0.848	-1.414
Ref. auth. internationality	0.489	1.631	0.121	2.710	0.007	0.091	0.564
Ref. citer internationality	-0.865	0.421	0.228	-9.360	0.000	-2.580	-0.687
No. refs	0.005	1.005	0.001	5.640	0.000	0.003	0.006
No. pages	0.008	1.008	0.003	3.230	0.001	0.003	0.013
Title Length	-0.020	0.980	0.004	-5.470	0.000	-0.027	-0.013
Abs. length	0.001	1.001	0.000	4.470	0.000	0.001	0.002
Abs. readability	0.003	1.003	0.001	2.100	0.035	0.000	0.006
Constant	-1.283	0.277	0.235	-5.470	0.000	-1.743	-0.823
Alpha	0.118	1.125	0.037	3.220	0.001	0.046	0.190

Enhanced self-citation detection by fuzzy author name matching

Paul Donner

donner@forschungsinfo.de
Institute for Research Information and Quality Assurance (iFQ), Schützenstr. 6a, 10117 Berlin

Introduction

With this study we seek to gain some understanding into the empirical magnitudes of errors in self-citation detection. We use fuzzy string distance measures to select highly similar names from a set of possible candidates to increase recall of relevant names and compare the results to those of a simple name comparison method. The automatic methods' results are benchmarked against manually curated samples. We arrive at estimates of the proportion of latent to overt self-citations for several detection methods.

Data

This study was performed on scholarly publication and citation data obtained from Elsevier Scopus through a data subscription by the Competence Center for Bibliometrics (KB). It has been processed, cleaned and enhanced and subsequently loaded into the KB's internal bibliometrics database. The data provider's supplied citation linkage information was used. All publications from the year 2009 regardless of document type or publication type and all articles citing them up to the date of May 2013 were taken into account.

Methods

A framework was programmed within the database in PL/SQL to run and evaluate different author self-citation detection routines. All documents citing a given work are collected and the author information of the cited work is compared against those of the citing works.
The methods are:
1) fullname-field method: The "fullname" field is created from the author given name and surname fields in the original data when loaded into the database. This is most often used for self-citation matching up to now and used as a baseline here. Exact string match on the cited and citing author fullname field is performed.
2) complete name method: A method taking into consideration as much author name information as available. Surnames are compared directly. If both names have full given names, those are compared. If both contain initials, they are aligned to the common minimum number of initial letters should they differ, then they are compared. If one should contain just initials and the other a full given name, the latter is reduced to its initials and two are compared (a longer chain of initials would likewise be truncated to be as long as the shorter one).
3) Enhanced algorithms: the surnames are first compared by regular exact string match. If they do not match, we subject the two surnames to one of two string distance measures, either Levenshtein distance (Levenshtein 1966) or Jaro-Winkler (JW) distance (Winkler 1990).
If the computed similarity score is equal or greater than a threshold value, the pair is a match. Should the given names be the same or similar above the defined threshold (for not abbreviated given names) or be the same initials (if at least one given name is abbreviated), the match is kept, else it is discarded.

The algorithm methods were evaluated with samples randomly drawn from the population of candidate pairs described above. The two author names in each pair were manually assessed as to whether they denominate the same person. Apart from the names, the institutional affiliations were taken into account, as well as personal online publication lists. Decisions on identity from this ground truth data set were then compared to the results of runs of the automatic methods. Algorithmic decisions were classified into true/false positive and true/false negative. From those, precision, recall, and the F-measure were computed.

Results and discussion

Random sample: 1012 author pairs were randomly drawn from the candidates and the methods evaluated. The ground truth showed that 1009 of those pairs consist of the same authors and are thus true self-citations. See Table 1 for the results. The complete name method outperforms the fullname field method. Fullname field is easily beaten by both of the surname similarity based methods at any examined threshold, while the complete name method works almost as well as them. The advantage from the increased recall is not mitigated by the slightly lower precision.

From these results we estimate that an elaborate fuzzy string based method can find close to all author-author self-citation (AASC) with little if any error, while a method based on surname and initials (baseline) will miss about 2.6% of author-author self-citations. The method making use of full given name information will miss only an estimated 0.7% of AASC.

Table 1: Results of the evaluation of methods for a true random sample, n = 1012

	Threshold	Precision	Recall	F-Measure
fullname field		1.000	0.974	0.987
complete name		1.000	0.993	0.996
Levenshtein name similarity	0.70	0.999	1.000	0.999
	0.80	1.000	0.999	0.999
	0.90	1.000	0.994	0.997
Jaro-Winkler name similarity	0.70	0.997	1.000	0.998
	0.80	0.998	1.000	0.999
	0.90	0.999	1.000	0.999

Sample of not-equal candidate names: We analyzed sample of edge cases. Name pairs that were similar but not identical were taken into consideration. From all pairs that were found to have Levenshtein distance similarities between 0.6 and 0.99 we randomly selected 2750 pairs. As would be expected, the plain methods perform much worse than in the sample discussed previously. Results are reported in Table 2.

Table 2: Evaluation of methods for similar but not equal names, n = 2750

Method	Threshold	Precision	Recall	F-Measure
fullname field		1.000	0.014	0.028
complete name		1.000	0.005	0.010
Levenshtein name similarity	0.70	0.865	0.999	0.927
	0.80	0.973	0.776	0.863
	0.90	1.000	0.169	0.289
Jaro-Winkler name similarity	0.70	0.843	1.000	0.915
	0.80	0.853	0.999	0.920
	0.90	0.913	0.908	0.910

Conclusion

The standard method for detecting AASC is found to miss about 2.6% of AASCs. Should this be considered non-trivial, using complete name information and string similarity methods can satisfyingly alleviate the problem. Preferable thresholds are 0.70 to 0.75 for Levenshtein distance similarity and 0.80 to 0.90 for Jaro-Winkler similarity.

References

Levenshtein, V. I. (1966). Binary codes capable of correcting deletions, insertions and reversals. Soviet Physics Doklady, 10: 707-710.

Winkler, W. E. (1990). String Comparator Metrics and Enhanced Decision Rules in the Fellegi-Sunter Model of Record Linkage. Proceedings of the Section on Survey Research Methods (American Statistical Association): 354–359.

CitNetExplorer: A new software tool for analyzing and visualizing citation networks

Nees Jan van Eck and Ludo Waltman

ecknjpvan@cwts.leidenuniv.nl; waltmanlr@cwts.leidenuniv.nl
Centre for Science and Technology Studies, Leiden University, The Netherlands

Introduction

A lot of work has been done on the analysis and visualization of many different types of bibliometric networks (e.g., Van Eck & Waltman, in press). The most frequently studied types of bibliometric networks are based on citation relations. Examples include networks of co-citation relations or bibliographic coupling relations between journals, authors, or individual publications.

The analysis and visualization of direct citation networks has received relatively limited attention. Interest usually focuses on co-citation and bibliographic coupling networks rather than direct citation networks. Many techniques have been developed for analyzing and visualizing co-citation and bibliographic coupling networks, and a number of software tools are available to support the study of these networks (Cobo, López-Herrera, Herrera-Viedma, & Herrera, 2011; Van Eck & Waltman, in press). One of these tools is our own VOSviewer (Van Eck & Waltman, 2010; www.vosviewer.com).

Important work on the analysis and visualization of direct citation networks has been done by Eugene Garfield. Garfield emphasizes the value of direct citation networks for studying the history and development of scientific fields. He refers to this as algorithmic historiography (Garfield, Pudovkin, & Istomin, 2003). Garfield has developed a software tool called HistCite (www.histcite.com) that can be used to construct and visualize direct citation networks based on data downloaded from the Web of Science (WoS) bibliographic database.

In this paper, we introduce CitNetExplorer, a new software tool that we have developed for analyzing and visualizing direct citation networks. CitNetExplorer, which is an abbreviation of ‘citation network explorer’, can be downloaded from www.citnetexplorer.nl. It has been developed in Java and therefore should run on any system that offers Java support. CitNetExplorer builds on Garfield’s work on algorithmic historiography. Compared with HistCite, it can handle much larger citation networks, possibly including millions of publications and citation relations. Moreover, CitNetExplorer offers sophisticated functionality for drilling down into a citation network, for instance allowing users to start at the level of a full network consisting of several millions of publications and to then gradually drill down into this network until a small subnetwork has been reached including no more than, say, 100 publications, all dealing with a specific topic of interest. CitNetExplorer borrows various ideas from VOSviewer, especially features related to visualization and user interaction.

CitNetExplorer may be used for many different types of applications. Below we give three examples.

Studying the development of a scientific field over time. This is what Garfield refers to as algorithmic historiography. The idea is that by showing the most important publications in a field, ordered based on the year in which they appeared, and the citation relations between these publications, one obtains an overview of the development of a field over time.

Identifying research areas. Suppose one wants to identify all publications on a certain research topic or in a certain research area. One may attempt to identify publications using keywords or based on the journal in which they have appeared, but this will usually yield incomplete results. Publications that do not contain the right keywords or that have not appeared in the right journal will be missed. Starting from a core set of relevant publications, CitNetExplorer can be used to identify publications based on citation relations. One may for instance select all publications that have at least a certain minimum number of citation relations with publications in the core set.

Literature reviewing. Literature reviewing can be a time-consuming task, especially when one attempts to be exhaustive in one's overview of the literature. To make sure that no relevant publications are overlooked, large numbers of publications need to be checked, often by going through the reference lists of publications that have already been identified as being relevant. Or the other way around, relevant publications need to be identified by checking all publications that cite one or more publications already identified as being relevant. CitNetExplorer simplifies literature reviewing in various ways, in particular by making it possible to easily select all publications that cite, or are cited by, a given set of publications.

Below we first discuss how citation networks are constructed, visualized, and analyzed in CitNetExplorer. We then provide a demonstration of the tool.

Construction, visualization, and analysis of citation networks

In this section, we discuss the approach taken by CitNetExplorer to construct, visualize, and analyze citation networks.

Construction of citation networks

CitNetExplorer is able to directly construct citation networks based on WoS output files. To identify citation relations between publications, the cited references in a WoS output file are matched with the publications in the file. CitNetExplorer first attempts to match based on DOI. If DOI data is not available, matching is done based on first author name (last name and first initial only), publication year, volume number, and page number. A perfect match is required for each of these data elements. Data on the title of the cited journal usually is available as well, but because journal titles often are not written in a consistent way, this data is not used.

CitNetExplorer assumes citation networks to be acyclic. This for instance means that it is not allowed to have both a citation from publication A to publication B and a citation from publication B to publication A. Likewise, it is not allowed to have a citation from publication A to publication B, a citation from publication B to publication C, and a citation from publication C to publication A. In practice, citation networks are not always perfectly acyclic. CitNetExplorer therefore removes citation relations that cause a citation network to have cycles. CitNetExplorer also removes citation relations that point forward in time, for instance from 2013 to 2014.

Visualization of citation networks

If a citation network consists of more than 50 or 100 publications, displaying all publications and all citation relations is typically of little use. There will usually be lots of citation relations, and many of them will cross each other, leading to a visualization that is hard to interpret. In the case of larger networks, CitNetExplorer therefore displays only a selection of all publications, by default the 40 most frequently cited ones. For simplicity, below we assume that we are dealing with a small network and that all publications in the network are displayed. Larger networks are visualized in the same way as described below, with the exception that only a selection of all publications are included in the visualization.

In the visualization of a citation network, CitNetExplorer uses the vertical dimension to represent time, with more recent years being located below earlier years. Publications are positioned in the vertical dimension based on the year in which they appeared. The vertical dimension is organized into layers, each of equal height. Each year is represented by at least one layer. If there are citation relations between publications from the same year, the year is represented by multiple layers. The horizontal dimension in the visualization of a citation network is used to provide an indication of the closeness of publications to each other in the citation network. Examples of visualizations of citation networks can be found in Figures 1 and 2.

Positioning the publications in a citation network in the horizontal and vertical dimensions of a visualization is known as a hierarchical graph drawing problem. Following the literature (e.g., Healy & Nikolov, 2013), CitNetExplorer first assigns each publication to a layer in the vertical dimension. This is done based on the year in which a publication appeared. In addition, it is ensured that citations always flow in an upward direction. So for each citation relation, the layer to which the citing publication is assigned must be located below the layer of the cited publication. After each publication has been assigned to a layer, CitNetExplorer positions the publications in the horizontal dimension. In general, the closer two publications are to each other in the citation network, the closer to each other they are positioned in the horizontal dimension.

To optimize the visualization of a citation network, CitNetExplorer uses similar techniques as VOSviewer. CitNetExplorer labels publications by the last name of the first author. To prevent labels from overlapping, labels may sometimes be displayed only for a selection of all publications. Like VOSviewer, CitNetExplorer offers zooming and panning (scrolling) functionality.

Analysis of citation networks

CitNetExplorer offers the following options for analyzing citation networks:

- Extracting connected components.
- Clustering publications.
- Identifying core publications.
- Finding shortest and longest paths between publications.

Publications are clustered based on their citation relations. The clustering methodology of Waltman and Van Eck (2012) is used combined with the smart local moving algorithm of Waltman and Van Eck (2013). The level of detail of the clustering is determined by a resolution parameter.

The identification of core publications is based on the concept of *k*-cores introduced by Seidman (1983). A core publication is a publication that has at least a certain minimum number of citation relations with other core publications. The identification of core publications makes it possible to get rid of unimportant publications in the periphery of a citation network.

Demonstration

We now offer a demonstration of CitNetExplorer. We use the tool to analyze the literature on two scientometric topics: The *h*-index and science mapping. The demonstration aims to give a general idea of the possibilities offered by CitNetExplorer. The demonstration is not intended as a step-by-step tutorial. A tutorial is available online at www.citnetexplorer.nl/gettingstarted/.

Data collection

Bibliographic data for all 25,242 publications in the 13 journals listed in Table 1 was downloaded from the WoS database[1]. To select these journals, we started with *Scientometrics* and *Journal of Informetrics*, the two core scientometric journals. We then used the Journal Citation Reports to identify closely related journals. We took all journals listed among the five most closely related journals to either *Scientometrics* or *Journal of Informetrics*, excluding journals that seem to be mainly nationally oriented. For each selected journal, we also added possible predecessors to the selection. The 25,242 publications in the 13 selected journals cover the period 1945–2013.

Table 1. Journals included in the data collection.

Journal	No. of pub.
American Documentation	796
ASLIB Proceedings	2,697
Information Processing and Management	3,036
Information Scientist	254
Information Storage and Retrieval	372
Journal of Documentation	3,778
Journal of Information Science	1,855
Journal of Informetrics	399
Journal of the American Society for Information Science	2,995
Journal of the American Society for Information Science and Technology	2,486
Research Evaluation	383
Research Policy	2,596
Scientometrics	3,595

Analysis

We first analyze the *h*-index literature. We then consider the literature on science mapping.

h-index

After loading the data downloaded from the WoS database into CitNetExplorer, we obtain a citation network consisting of 28,482 publications and 158,292 citation relations. The citation

[1] Data collection took place on November 7, 2013.

network includes 3,240 publications that are not among the 25,242 publications included in the data collection. These are publications that are cited at least ten times in the 25,242 publications. This includes classical scientometric publications, such as the work by Bradford and Lotka, it also includes books, for instance by Garfield, Kuhn, and De Solla Price, and it includes scientometric publications in multidisciplinary journals such as *Nature*, *PNAS*, and *Science*.

A visualization of the citation network is presented in Figure 1.[2] The scientometric literature is located in the center of the visualization. Information science and information retrieval publications are located in the left part, and publications on technology and innovation studies in the right part.

We now drill down into the subnetwork consisting of the publication by Hirsch in 2005 in which the *h*-index was introduced and all publications that directly or indirectly cite this publication. These publications are called successors in CitNetExplorer. A visualization of the subnetwork is shown in Figure 2. The subnetwork includes 1,371 publications.

Many publications that cite the publication by Hirsch in 2005 may be only weakly related to the topic of the *h*-index. Among the 1,371 publications in our subnetwork, we therefore make a selection of core publications. We define a core publication as a publication that has citation relations with at least ten other core publications. Based on this criterion, 230 core publications are identified. We drill down into the subnetwork consisting of these 230 publications. After drilling down, we cluster the publications. Using the default value of 1.00 for the resolution parameter, two clusters are identified. The visualization that we obtain is shown in Figure 3. Based on our knowledge of the scientometric literature, it is immediately clear that the blue cluster consists of publications on the *h*-index and its variants. The publications in the green cluster are not directly about the *h*-index but instead deal with the closely related topic of advanced citation-based indicators.

Drilling down into the blue cluster yields the visualization presented in Figure 4. The visualization displays the citation network of the most frequently cited publications on the *h*-index, starting with the publication by Hirsch in 2005 and ending with the publication by Waltman on the inconsistency of the *h*-index in 2012.

In CitNetExplorer, we can navigate back and forth between different subnetworks of a citation network in a similar way as we can navigate back and forth between web pages in a web browser. After moving back to our subnetwork consisting of 230 publications, we drill down into the green cluster. The visualization that we obtain is shown in Figure 5. The visualization offers an overview of the development of the literature on advanced citation-based indicators after the introduction of the *h*-index, starting with well-known publications by Bollen, Lundberg, and Zitt and ending with recent work by for instance Glänzel, Leydesdorff, and Waltman.

[2] In all visualizations shown in this paper, the 70 most frequently cited publications are displayed. Furthermore, only citation relations included in the so-called transitive reduction of a citation network are displayed. This for instance means that if A cites B and C and if B cites C, the citation relation between A and C is not displayed. This relation is considered non-essential, since one can also get from A to C via B.

Science mapping
We now look at the literature on science mapping. We first move back to the full network, and we then cluster the 28,482 publications using a relatively high value of 5.00 for the resolution parameter. Publications by Small, Wasserman, and White are assigned to the same cluster, suggesting that this cluster covers the topic of science mapping. We drill down into the cluster, which consists of 1,105 publications. We then again create a clustering of publications, this time using the default value of 1.00 for the resolution parameter. This yields the visualization shown in Figure 6.

We observe four clusters in the visualization in Figure 6. The blue cluster can be considered to cover the core of the science mapping literature, in particular the work on co-citation and bibliographic coupling analysis. The orange cluster mainly covers the topic of co-word analysis. The purple cluster covers the topic of (social) network analysis. The green cluster is a bit more difficult to label. On the one hand it covers the topic of interdisciplinarity, but on the other hand it also covers a large number of publications from a single author (of the 248 publications in the cluster, 57 are authored by Leydesdorff), suggesting that to some degree the cluster may represent the oeuvre of an author rather than a scientific topic.

We drill down into the blue cluster. This gives us the visualization presented in Figure 7. We select the publication by Kessler in 1963, in which the concept of bibliographic coupling was introduced, and the publication by Van Eck in 2010, in which the VOSviewer software was presented. We then identify the longest path in the citation network between these two publications. There turn out to be multiple longest paths, as shown in Figure 8.

Conclusion
We have introduced CitNetExplorer, a new software tool for analyzing and visualizing citation networks. The most important functionality of CitNetExplorer has been discussed, and a demonstration of the tool has been given. Because of space limitations, we have not been able to discuss all possibilities offered by CitNetExplorer. We also have not been able to show how CitNetExplorer can be applied to very large citation networks, including millions of publications and citation relations.

The scientometric community is still relatively inexperienced in the types of analyses made possible by a tool such as CitNetExplorer. As we have discussed, the community has focused more on the analysis of co-citation and bibliographic coupling networks than on the analysis of direct citation networks. Given the limited experience with the analysis of direct citation networks, it remains to be seen for what types of applications CitNetExplorer is most useful. Based on our own experience with CitNetExplorer and the feedback we hope to receive from others, we plan to continue the development of the tool. Among other things, we will consider the possibility of including additional options for analyzing citation networks, for instance related to the idea of main path analysis (Hummon & Doreian, 1989). We especially hope that CitNetExplorer, probably in combination with other tools, will contribute to a better understanding of the evolution of scientific fields.

Figure 1. Citation network of the field of scientometrics and closely related fields.

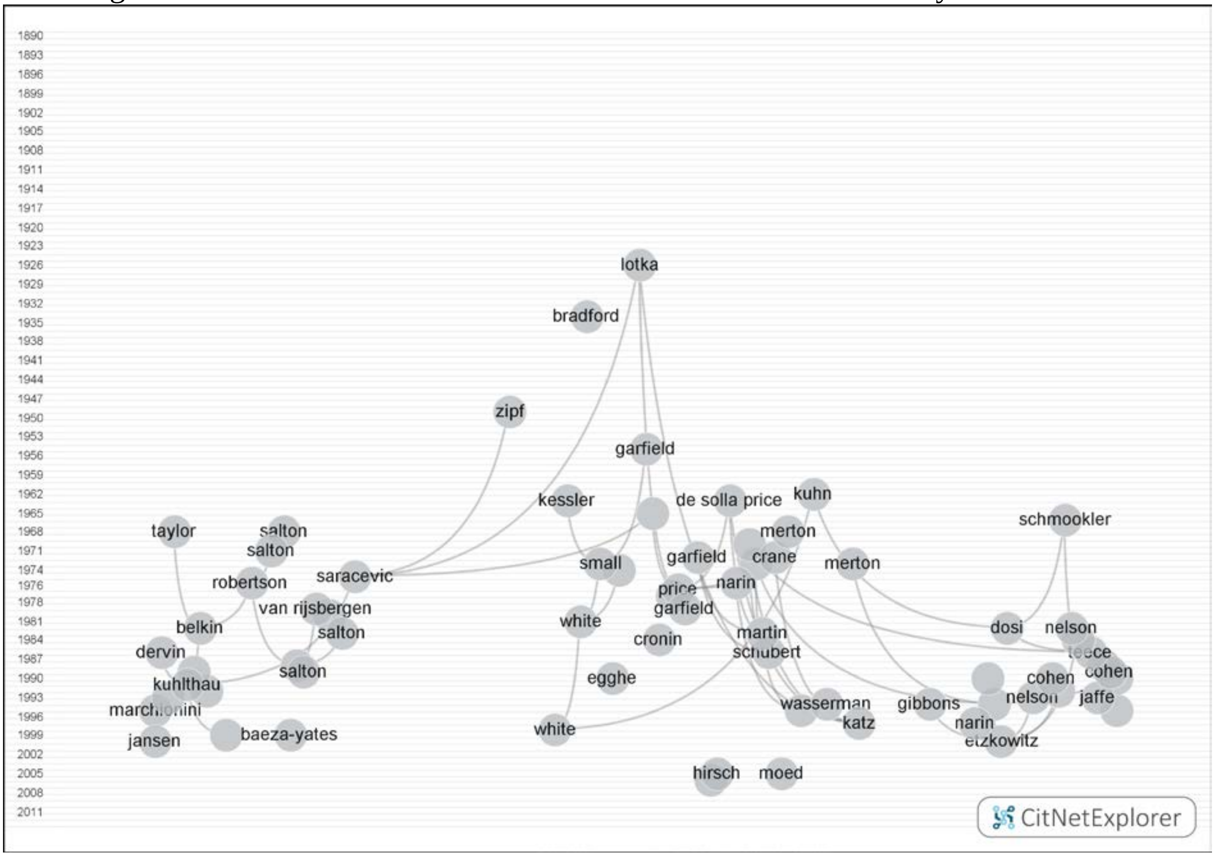


Figure 2. Citation network of the paper by Hirsch in 2005 and all its direct and indirect successors.

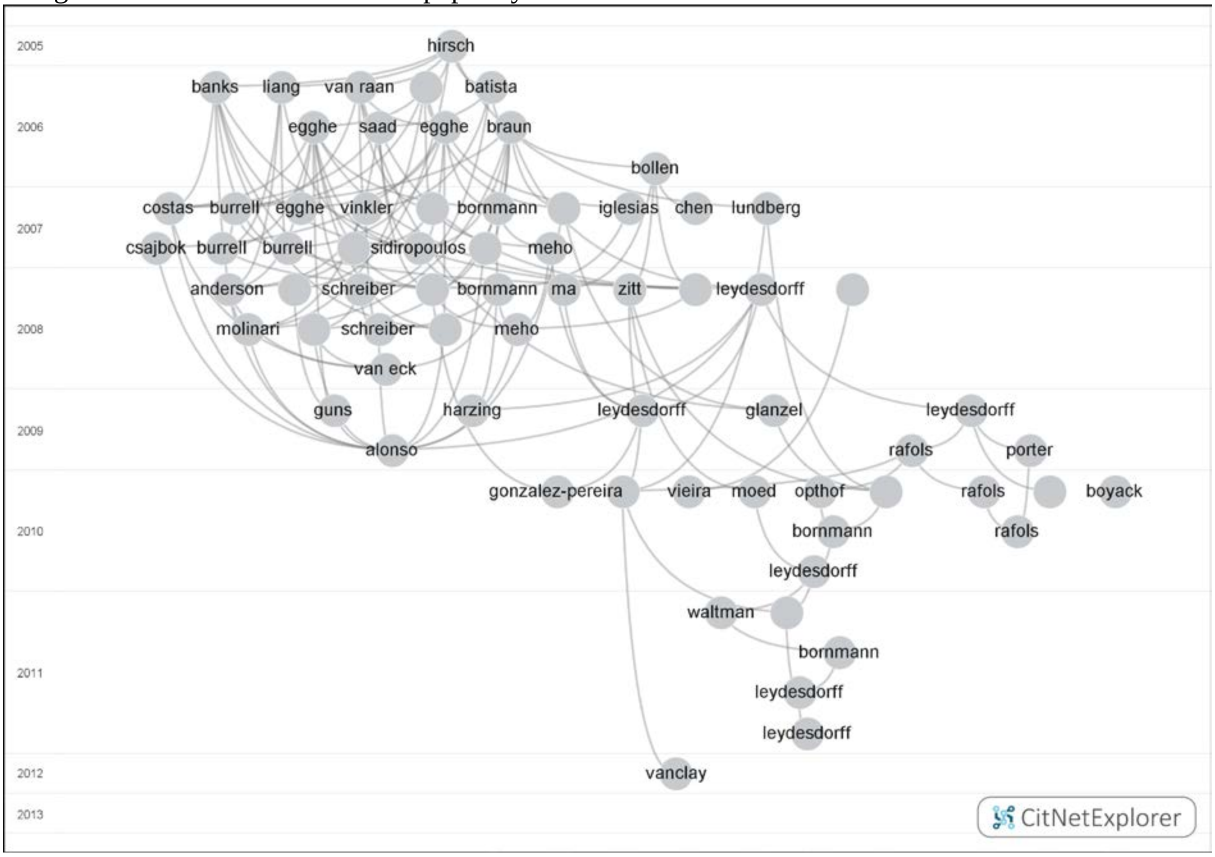


Figure 3. Citation network of the paper by Hirsch in 2005 and the core of its direct and indirect successors.

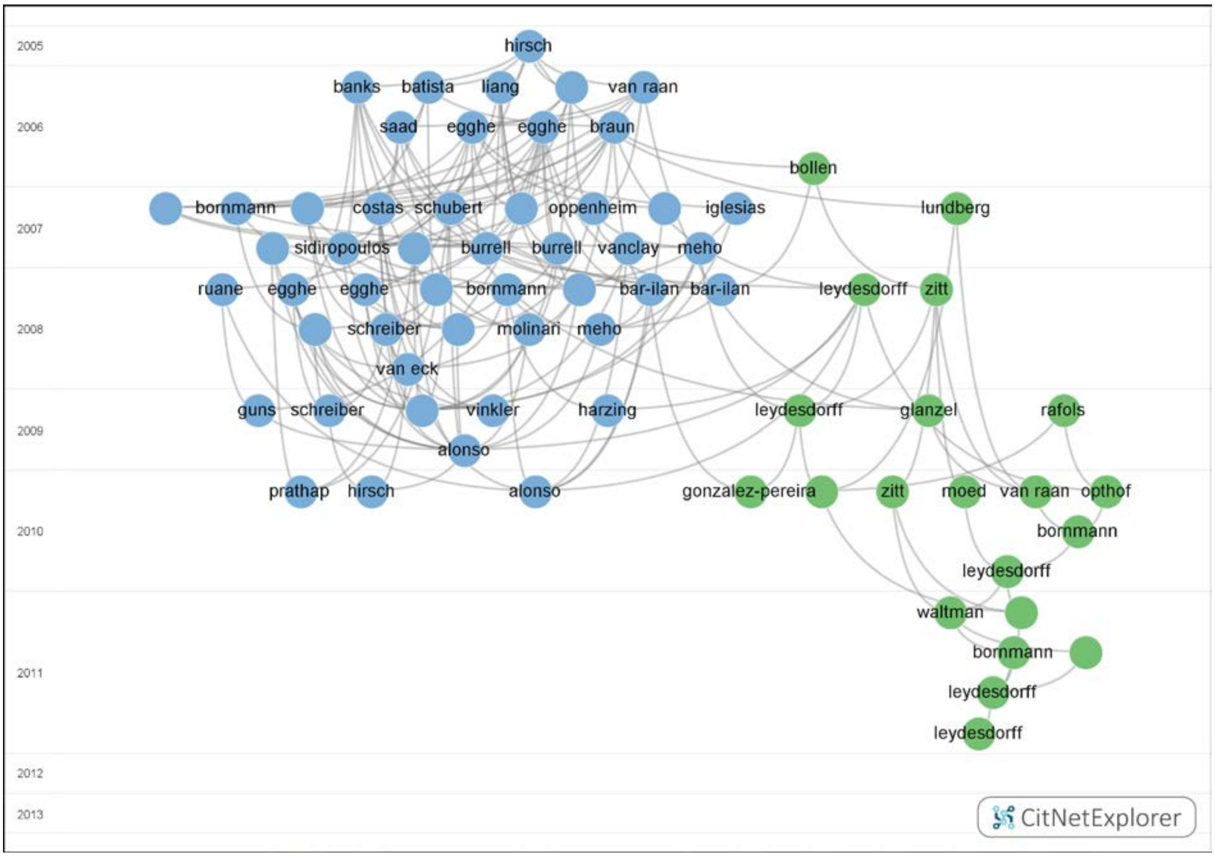


Figure 4. Citation network of the literature on the h-index and its variants.

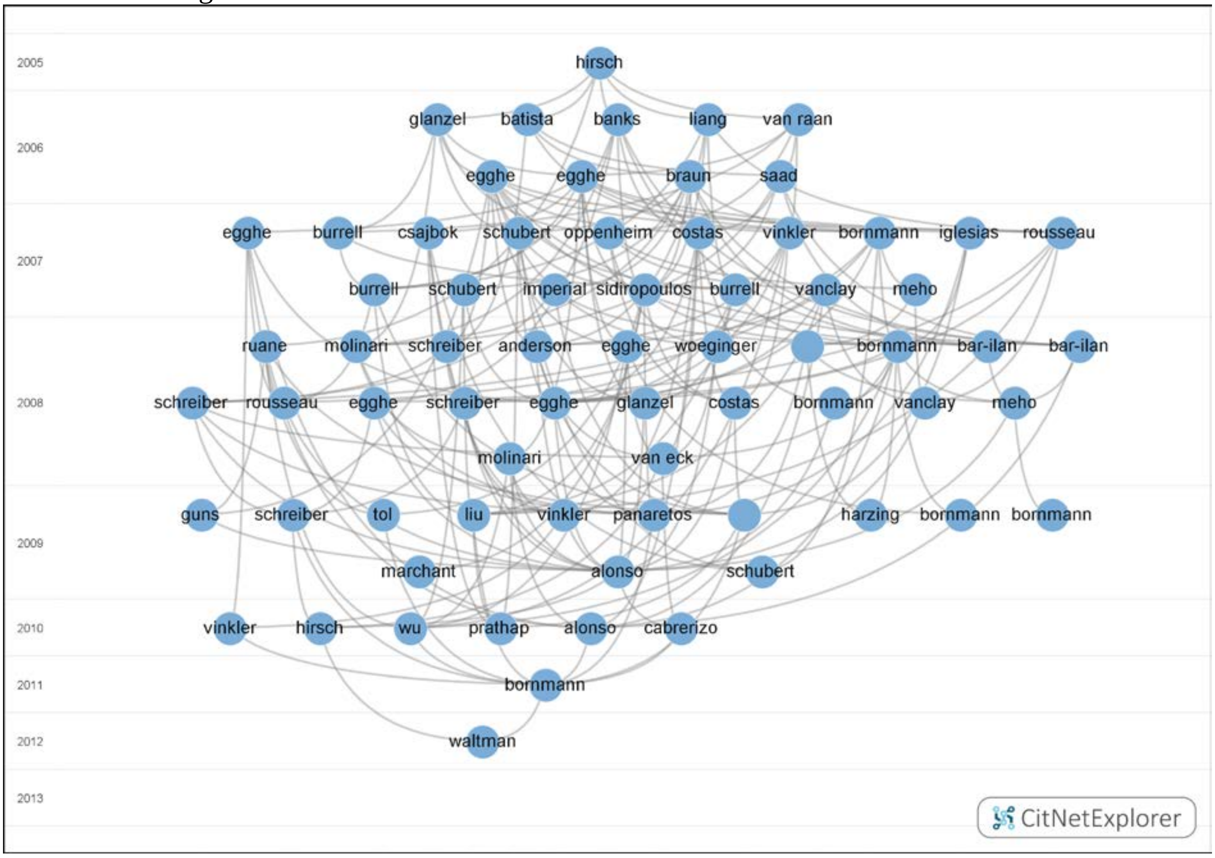


Figure 5. Citation network of the literature on advanced citation-based indicators.

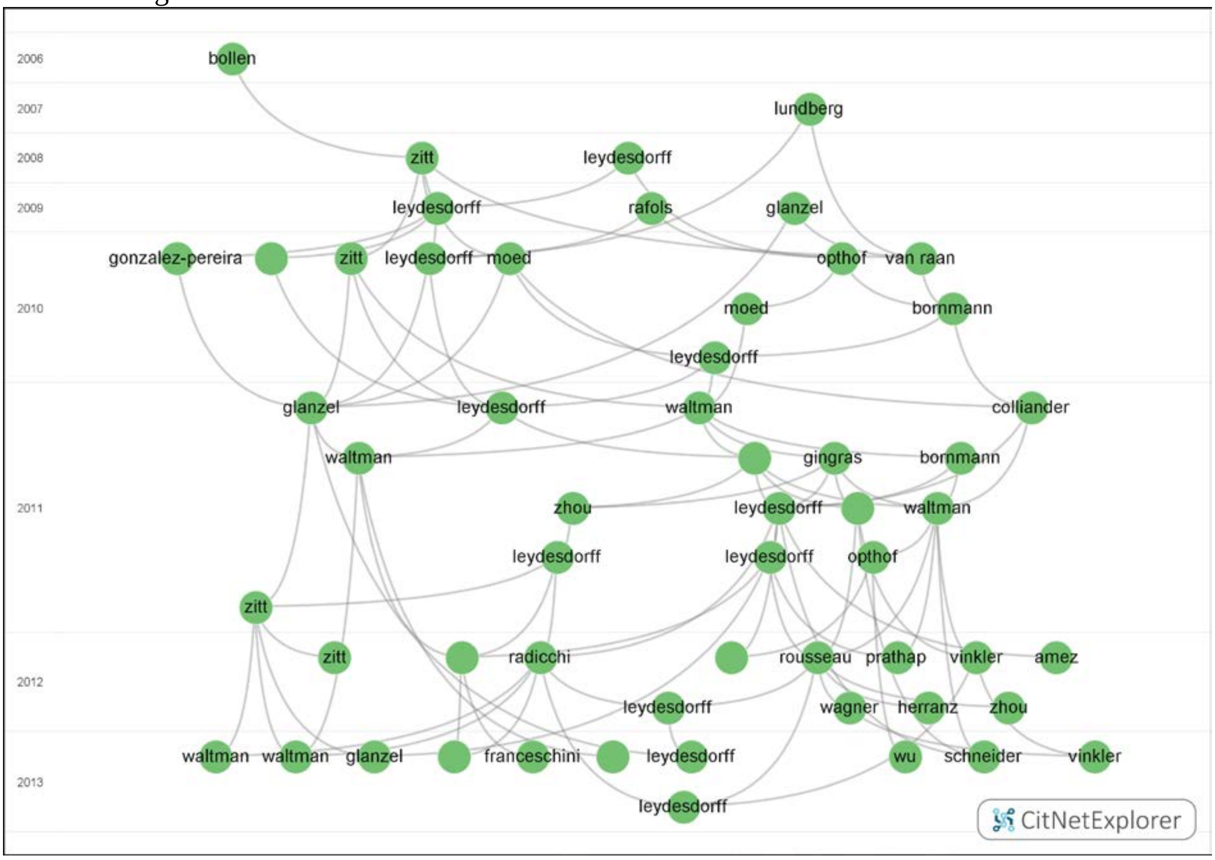


Figure 6. Citation network of the literature on science mapping.

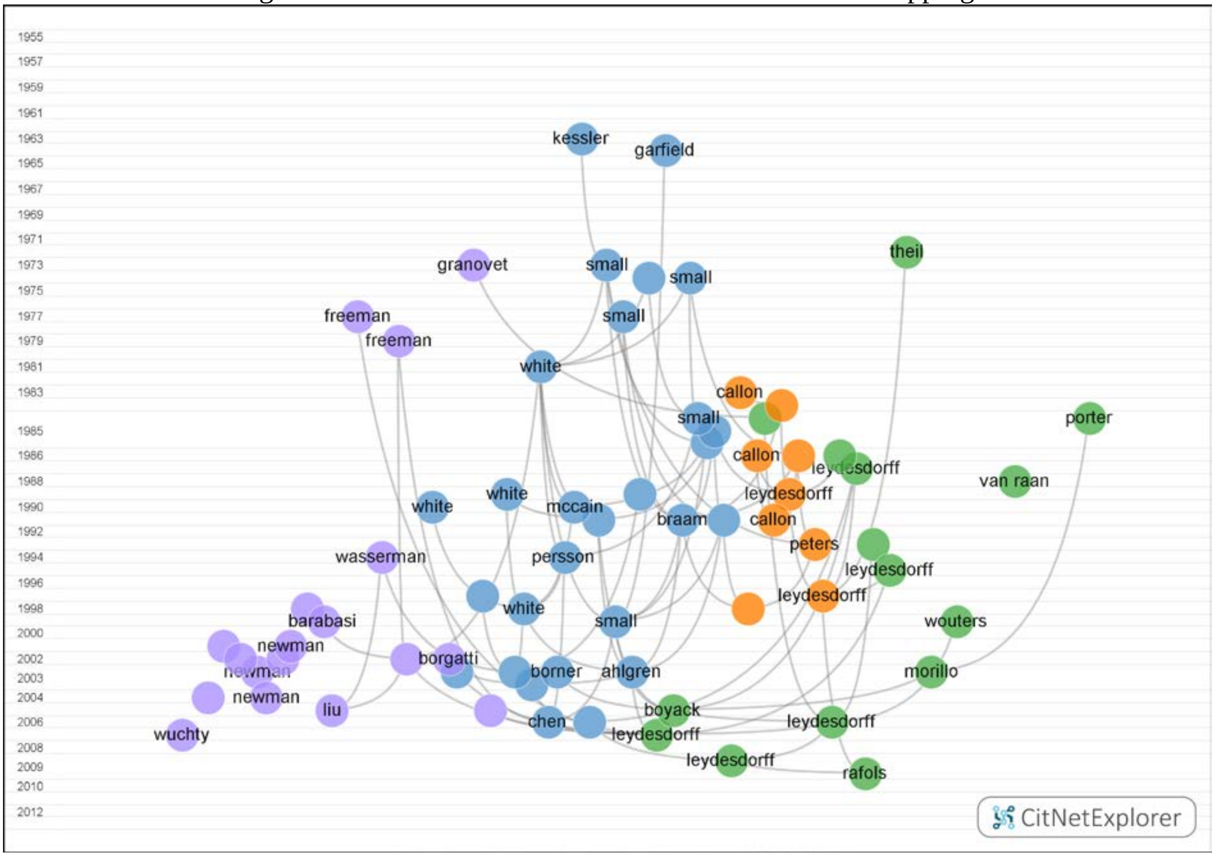


Figure 7. Citation network of the literature on co-citation and bibliographic coupling analysis.

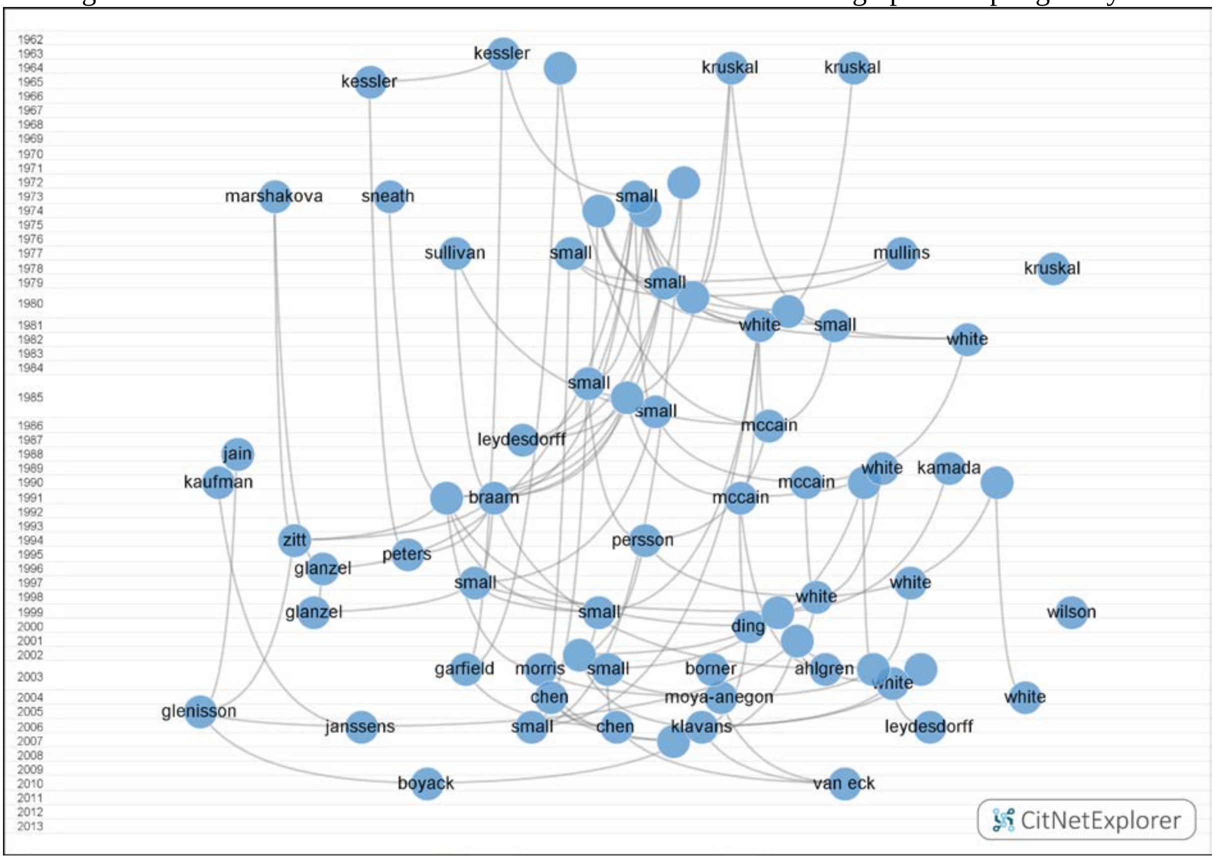
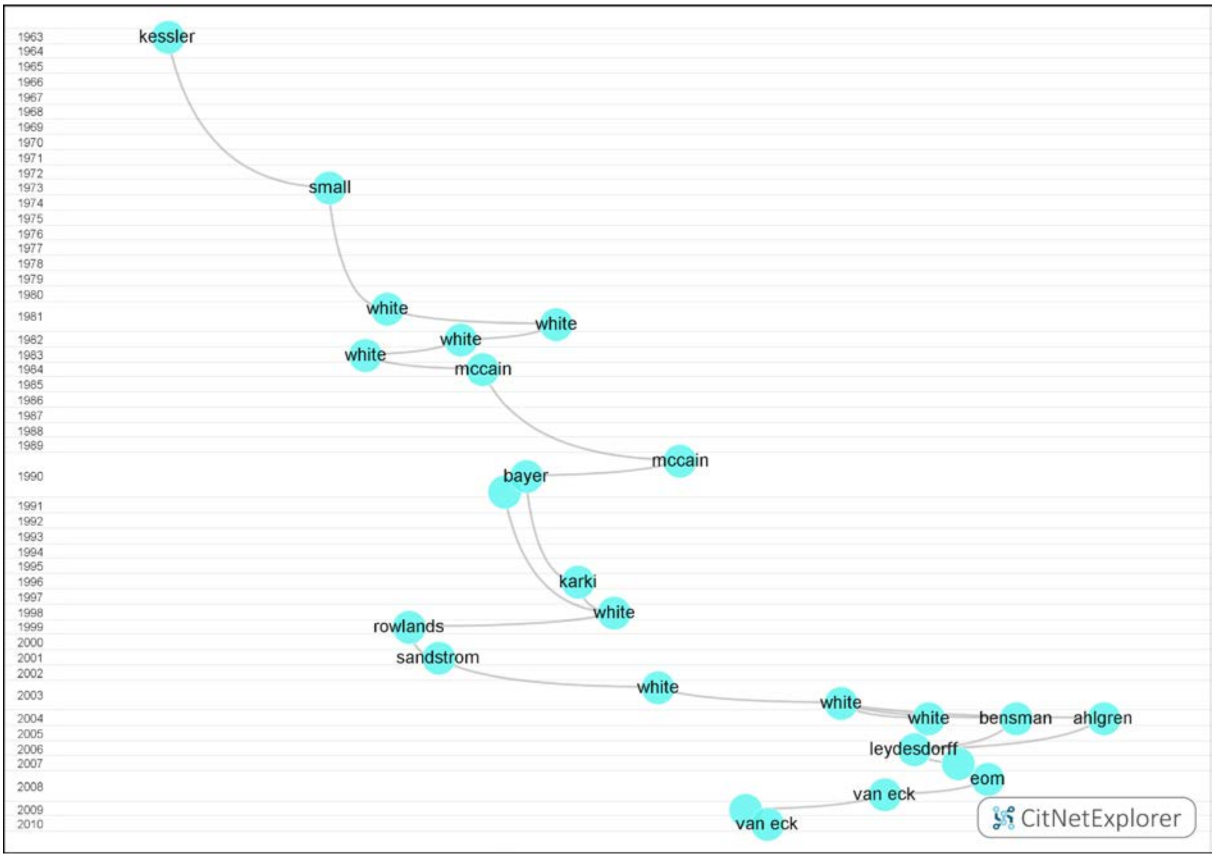


Figure 8. Longest paths between the publication by Kessler in 1963 and the publication by Van Eck in 2010.



References

Cobo, M.J., López-Herrera, A.G., Herrera-Viedma, E., & Herrera, F. (2011). Science mapping software tools: Review, analysis, and cooperative study among tools. *Journal of the American Society for Information Science and Technology*, *62*(7), 1382–1402.

Garfield, E., Pudovkin, A.I., & Istomin, V.S. (2003). Why do we need algorithmic historiography? *Journal of the American Society for Information Science and Technology*, *54*(5), 400–412.

Healy, P., & Nikolov, N.S. (2013). Hierarchical drawing algorithms. In R. Tamassia (Ed.), *Handbook of graph drawing and visualization* (pp. 409–453). CRC Press.

Hummon, N.P., & Doreian, P. (1989). Connectivity in a citation network: The development of DNA theory. *Social Networks*, *11*(1), 39–63.

Seidman, S.B. (1983). Network structure and minimum degree. *Social Networks*, *5*(3), 269–287.

Van Eck, N.J., & Waltman, L. (2010). Software survey: VOSviewer, a computer program for bibliometric mapping. *Scientometrics*, *84*(2), 523–538.

Van Eck, N.J., & Waltman, L. (in press). Visualizing bibliometric networks. To appear in Y. Ding, R. Rousseau, & D. Wolfram (Eds.), *Measuring scholarly impact: Methods and practice*. Springer.

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.

Waltman, L., & Van Eck, N.J. (2013). A smart local moving algorithm for large-scale modularity-based community detection. *European Physical Journal B, 86*, 471.

Beyond traditional metrics at the University of São Paulo: scientific production in the PLOS journals

Sibele Fausto*, Rogério Mugnaini**

*sifausto@usp.br, **mugnaini@usp.br
Escola de Comunicações e Artes, University of São Paulo,
Av. Prof. Lúcio M. Rodrigues, 443, São Paulo, SP, CEP 05608-020 (Brazil)

Introduction

In the evolution of the Open Access Movement, there have been pioneering initiatives to provide free and open access to published scientific content, such as the creation of the BioMed Central (BMC) and the Public Library of Science (PLOS) in early 2000. These set out a new funding model for journals where the liability for publishing costs is transferred from the readers to the authors. The PLOS was also innovative since it led to the creation of Article-Level Metrics - ALM (Fenner & Lin, 2013) and expanded by issuing a series of new titles - PLOS Biology in 2003; PLOS Medicine in 2004, PLOS Computational Biology, PLOS Genetics and PLOS Pathogens in 2005, PLOS ONE in 2006 and more recently PLOS Currents. As a result, today the set of PLOS journals is widely recognized and prestigious.

Recently (on June 18, 2013), PLOS introduced a new search engine - PLOS-ALM Reports (http://almreports.plos.org/) (Allen, 2013) which allows more detailed investigations to be carried out in all the PLOS journals showing consolidated alternative measures of visibility and impact earned by published articles. Since the University of São Paulo is considered to be "Brazil's leading academic institution in research and graduate education" (Schwartzman, 2006), this study is an attempt to find evidence of the USP performance that goes beyond traditional metrics, by using the alternative indicators provided by PLOS-ALM and making a comparison with other articles in the PLOS journals that come from Brazil.

Methods

We carried out a search in PLOS-ALM Reports for [Author affiliation country: Brazil; Period: January 01, 2005 to December 31, 2012; Journal: All Journals]. The data were downloaded in a CSV file, so that a comparison of the PLOS-ALM indicators[1] could be made between the USP and non-USP articles from Brazil.

Results

A total number of 481 articles from USP are represented within the analyzed date range in the PLOS journals, which is 0.69% of the overall number of PLOS publications (n=69,306) and 30.1% of the publications where Brazil is the authorship country (n=1,598) in the same period. The PLOS ONE journal has published almost all the items where Brazil features as the "author affiliation country" (n=1,303), followed by PLOS Neglected Tropical Diseases (n=213), PLOS Pathogens (n=27), PLOS Medicine (n=18), PLOS Genetics (n=17), PLOS Biology and PLOS Computational Biology (10 items each). With regard to ALM, Table 1 shows articles from Brazil with the four best PLOS-ALM indicators (viewed, cited, saved and

[1] The PLOS-ALM set of relevant indicators for the impact made by the articles is described in: http://www.plosone.org/static/almInfo/#static-content-wrap

discussed), and makes clear that USP has the most “saved” items with 438 bookmarks in the Mendeley reference management service.

Table 1: PLOS articles from Brazil with the four best ALM indicators

ALM indicator	Brazil origin	DOI	Title	Total reach
most viewed	Non-USP	10.1371/journal.pone.0019881 Non-USP	What Is New for an Old Molecule? Systematic Review and Recommendations on the Use of Resveratrol	PMC: 7,869
most cited	Non-USP	10.1371/journal.pmed.0020059	A Space-time Permutation Scan Statistic for Disease Outbreak Detection	Scopus: 270
most saved	USP	10.1371/journal.pone.0013666	Beyond the Fragmentation Threshold Hypothesis: Regime Shifts in Biodiversity Across Fragmented Landscapes	Mendeley: 438
most discussed	Non-USP	10.1371/journal.pone.0043007	Glass Shape Influences Consumption Rate for Alcoholic Beverages	Facebook: 947

Table 2 shows the set of PLOS-ALM indicators that compare the reach of articles from Brazil of USP and non-USP origin, where the Total is the sum of all the ALM indicators and the Average per Article is the mean of each indicator for the USP and non-USP articles. The Table is arranged in descending order for the ALM Classification, to show the respective indicators.

Table 2: ALM indicators for the USP and non-USP articles from Brazil

ALM Classification	Indicator	Total			Average per article		
		non-USP	USP	% USP	non-USP	USP	USP/non-USP
Viewed	PMC Total	818.945	353.351	30,1%	733,16	734,62	1,00
	PMC views	551.442	235.661	29,9%	493,68	489,94	0,99
	PLOS PDF downloads	518.595	214.778	29,3%	464,27	446,52	0,96
	PMC PDF Downloads	267.728	117.690	30,5%	239,68	244,68	1,02
	PLOS XML downloads	48.228	19.753	29,1%	43,18	41,07	0,95
	Figshare	7.679	3.713	32,6%	6,87	7,72	1,12
Cited	PMC Europe Database Citations	66.515	68.406	50,7%	59,55	142,22	2,39
	Scopus	12.511	4.567	26,7%	11,20	9,49	0,85
	PMC Europe Citations	8.844	3.264	27,0%	7,92	6,79	0,86
	CrossRef	8.807	3.222	26,8%	7,88	6,70	0,85
	Web of Science	6.132	2.436	28,4%	5,49	5,06	0,92
	PubMed Central	5.725	2.045	26,3%	5,13	4,25	0,83
Saved	Mendeley	24.600	9.133	27,1%	22,02	18,99	0,86
	CiteULike	407	169	29,3%	0,36	0,35	0,96
	DataCite	5	-	-	0,00	0,00	-
	Reddit	2	-	-	0,00	0,00	-
Discussed	Facebook	10.278	4.172	28,9%	9,20	8,67	0,94
	Twitter	1.669	615	26,9%	1,49	1,28	0,86
	Wikipedia	132	80	37,7%	0,12	0,17	1,41
	Wordpress.com	64	15	19,0%	0,06	0,03	0,54
	Research Blogging	23	19	45,2%	0,02	0,04	1,92
	Nature Blogs	3	2	40,0%	0,00	0,00	1,55
	Science Seeker	-	-	-	0,00	0,00	-
Recommended	F1000Prime	52	14	21,2%	0,05	0,03	0,63

It is clear that the number of views and downloads is the most significant PLOS-ALM indicator, in absolute numbers, and USP represents about 30% of the total figure for Brazil. When account is taken of the average figure per article, USP stands out in virtually all the "Viewed" ALM indicators, but especially in Figshare.

When all the "Citations" are analyzed, the overall percentage of USP is 43.6% compared with the rest of Brazil (data not shown), with the citations in PMC Europe Citations Database being highlighted. As well as representing 50.7% of the citations made for Brazilian articles, this also has 2.39 times the average number of citations for articles by authors from other national institutions.

With regard to the indicators that show how many times articles were "Saved" through bookmark, Mendeley is most prominent with 27.1% for USP. Finally, with regard to the "Discussion" of the articles in social media outlets, those from USP stand out particularly in Facebook and Twitter. Although to a lesser extent, there is a significant difference in favour of USP in blogs (Nature Blogs and Blogging Research).

Conclusions

The presence of USP in the PLOS journals collection, as measured by its share of publications, reflects the considerable size of its physical structure and personnel, in comparison with that of other Brazilian institutions. Our results show how far this is the case, when the range of indicators is analyzed (although there are a couple of exceptions). ALM provides a wider range of indicators related to published articles that goes beyond the traditional citations, with analytical methods that involve alternative metrics for determining article usage and dissemination.

This initial investigation in the PLOS-ALM Reports is now under way with this case study based on the production by USP published in the PLOS journals which seeks to determine the different types of analysis that can be conducted with the data obtained by the tool.

References

Allen, L. Announcing ALM Reports – a new tool for analyzing Article Impact (18 jun. 2013). PLOS Blogs. Retrieved February 20, 2014 from: http://blogs.plos.org/plos/2013/06/announcing-alm-reports-a-new-tool-for-analyzing-article-impact/.

Lin, J. & Fenner, M. (2013). Altmetrics in Evolution: Defining and Redefining the Ontology of Article-Level Metrics. *Information Standards Quarterly*, 25 (2), 20.

Schwartzman, S. (2006). A universidade primeira do Brasil: entre *intelligentsia*, padrão internacional e inclusão social. *Estudos Avançados*, 20 (56), 161-189.

Extracting and sharing data citations from Google Scholar for collaborative exploitation

Sibele Fausto*, Tiago Rodrigo Marçal Murakami**

*sifausto@usp.br
Escola de Comunicações e Artes, University of São Paulo,
Av. Prof. Lúcio M. Rodrigues, 443, São Paulo, SP, CEP 05608-020 (Brazil)

**tiago.murakami@dt.sibi.usp.br
Departamento Técnico, Sistema Integrado de Bibliotecas, University of São Paulo
Rua da Biblioteca, S/N, Complexo Brasiliana, Piso Embasamento, São Paulo, SP, CEP 05508-050 (Brazil)

Background and purpose

There are studies that have drawn attention to the lack of indexing for the titles of scientific journals in the Social Sciences, Applied Social Sciences and Humanities in large commercial databases (Frandsen & Nicolaisen, 2008; Neuhaus & Daniel, 2007). This lack is even more acute when it comes to journals concerned with these areas published in languages other than English and published in developing countries (Archambault & Larivière, 2010), which makes it difficult to carry out an investigation of the importance and impact of these journals.

This situation is changing as a result of the new opportunities provided by the emergence of Open Access (OA) and tools as the search engine Google Scholar (GS) and software for data processing such as Publish or Perish - PoP (Harzing, 2007). The increasing shift of Social Sciences and Humanities journals to the Web - including those of Library and Information Science (LIS) is making them more widespread. This is allowing detailed searches to be conducted through GS and the recovery of citations of articles, which can be regarded as an alternative to traditional databases in bibliometrics studies on the impact of scientific production published in these areas. In addition it highlights the fact that GS is a free access source, in contrast with expensive commercial databases. It has a broad coverage of other kinds of material, even in the Social Sciences and Humanities (SSH), such as books, book chapters, conference materials, etc. which are not normally covered by traditional databases and hence it is able to make a comprehensive recovery of open access journals, in languages other than English, some of which come from emerging countries.

However, this apparently favorable context for research into bibliometrics in these areas still faces challenges owing to questions about the reliability of the GS as a data source (Jacsó, 2010). This criticism regarding to GS is a restatement of the need for more research into the tool to finds a rational basis for understanding the full potential of Google Scholar for bibliometrics studies, especially in areas not covered by commercial databases (Caregnato, 2011).

This situation stimulated our attempt to share citation data from Brazilian LIS journals as a pilot scheme to allow further investigation by the Brazilian scientometrics community in employing Google Scholar with the aim of encouraging its greater use for bibliometric purposes.

Methods

This pilot scheme adopted the following procedures:

a. Conducting a survey of LIS journals titles through compiling lists of those that exist on the web;
b. Carrying out searches using PoP software for Windows, with the journal title as a parameter, and confirming the official titles and abbreviations, in the period from January 28, 2014 to March 02, 2014;
c. Displaying the results in Google Drive spreadsheets, one for each retrieved journal title;
d. Creating a spreadsheet that brings together all the spreadsheets with the articles that had at least one citation;
e. Carrying out statistical tests using Excel and Tableau Public[1].

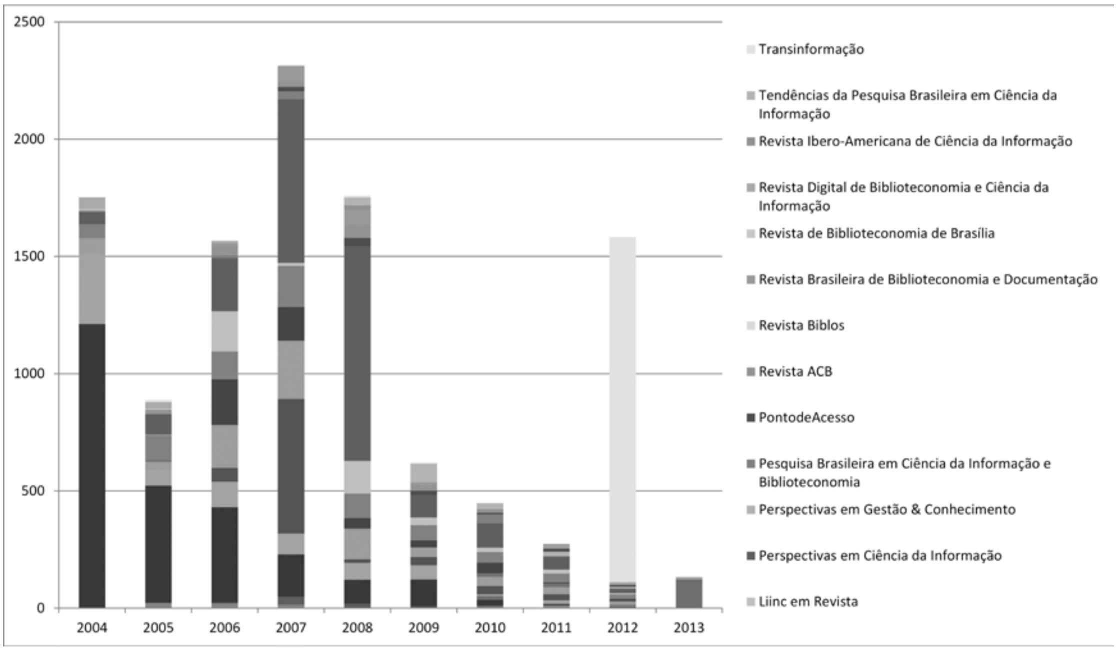
Google Drive allows its contents to be shared publicly, and the extracted data to be made available through the following link:
https://docs.google.com/spreadsheets/d/19kcMMnfi_5Ohe60_mev-myFc85FkppqRJy-HhXpfB_Q/edit.

Preliminary findings

Data extraction from the GS with PoP resulted in a total of 24 Brazilian LIS journals, all in open access. However, the searches recovered some inaccurate data which were then analyzed article by article and those with inconsistencies were withdrawn. The data obtained allowed some exploratory exercises to be conducted with Tableau Public, by various categorizations such as the received citations for each journal, including citations per year and the articles cited, among others. These preliminary exercises were also publicly shared through the following link:
http://public.tableausoftware.com/views/EstudodascitaesrecebidasporperidicosdaCI/Citaesrecebidasporperidicos?:embed=y&:display_count=no, e.g. as shown in Figure 1.

[1] Tableau Public: http://www.tableausoftware.com/public/ .

Figure 1. Number of Citations per journal and per year



Final considerations

Citation studies are an important subject research in Bibliometrics and their sources of reliable data were, until recently, a prerogative of restrictive and expensive commercial databases, despite these sources still continue to show inconsistencies as is widely discussed in the literature. Google Scholar provides an alternative source to these studies, particularly in the areas of the SSH, where many journals are not considered by the large databases.

The emergence of tools that facilitate the extraction and data processing from GS, such as PoP and tools like Google Refine, Google Drive and Tableau Public help to simplify the task of validating these data. In our view, the public sharing of pretreated citation data can stimulate more collaborative investigations by the community of Brazilian scientometricians with the aim to demonstrate the capacity of Google Scholar to act as an alternative and reliable data source in the metrical studies of national journals and thus enable better measures of the SSH results in the context of scientific evaluation in Brazil.

References

Archambault, E. & Larivière, V. The limits of bibliometrics for the analysis of the social sciences and humanities literature (2010). In UNESCO (Ed.), *2010 World Social Science Report*: *Knowledge Divides* (pp. 251-254). Paris: UNESCO, International Social Science Council. Retrieved February 20, 2014 from: http://unesdoc.unesco.org/images/0018/001883/188333e.pdf.

Caregnato, S. E. (2011). Google Acadêmico como ferramenta para os estudos de citações: avaliação da precisão das buscas por autor. *Ponto de Acesso*, 5 (3), 72-86.

Frandsen, T.F. & Nicolaisen, J. (2008). Intradisciplinary differences in database coverage and the consequences for bibliometric research. Journal *of the American Society for Information Science and Technology*, 59 (10), 1570-1581.

Harzing, A.-W. *Publish or Perish* (2007). Retrieved February 20, 2014 from: http://www.harzing.com/pop.htm.

JACSÓ, P. (2010). Metadata mega mess in Google Scholar. *Online Information Review*, 34 (1), 175–191.

Neuhaus, C.; Daniel, H-D. (2007). Data sources for performing citation analysis: An overview. *Journal of Documentation*, 64 (2), 193-210.

Effects of cuts in R&D investment in the quality and quantity of the research output of Spanish universities.

Gaizka Garechana*, Rosa Río-Belver** Ernesto Cilleruelo-Carrasco*** and Javier Gavilanes-Trapote****

* *gaizka.garechana@ehu.es*
Department of Industrial Engineering, University of the Basque Country (UPV/EHU), C/Elcano 21, Bilbao, 48008 (Spain)

** *rosamaria.rio@ehu.es*
Department of Industrial Engineering, University of the Basque Country (UPV/EHU), Calle Nieves Cano 12, Vitoria, 01006 (Spain)

*** *ernesto.cilleruelo@ehu.es*
Department of Industrial Engineering, University of the Basque Country (UPV/EHU), Alameda Urquijo s/n, Bilbao, 48030 (Spain)

**** *javier.gavilanes@ehu.es*
Department of Industrial Engineering, University of the Basque Country (UPV/EHU), Calle Nieves Cano 12, Vitoria, 01006 (Spain)

Introduction

Academic institutions are not invulnerable to economic recessions. Both public and private universities see their disposable incomes falling, mainly due to public budget cuts, the fall in enrolments (sometimes coupled with increasingly inverted population pyramids) because of unaffordable fees and cuts on the public grant system, and a reduced investment in public R&D by firms. Investments on new facilities, equipment or intangible assets are cancelled because of budget cuts. Salaries are usually a very significant share of expenditures, so employee downsizing is often the first choice for account balancing.

Spain has passed through two recession periods during “Great Recession” that forced government to run into considerable deficits. Several measures have been adopted by national and regional administrations to reduce spending, and cuts in public budgets have had immediate effects on R&D expenditures in higher education (HE) sector.

This paper addresses the question of the effects of R&D investment cuts in the quantity and quality of the scientific output of HE institutions in Spain. Budget reduction has not stopped: year 2012 has been the year of the biggest reduction in public national spending in R&D since the first national R&D programs starting in 1988, so this research must be further enhanced and complemented by future available data.

Data sources

Spanish National Institute of Statistics (INE, 2014) offers some of the key variables in HE R&D expenditures. Intramural expenditure is the main variable, reflecting all expenditures for R&D performed within a statistical unit or sector of the economy. Evolution in capital expenditures (expenditures on fixed assets), and in full-time equivalent (FTE) employees in R&D will also be considered in order to get a more complete picture.

SCOPUS database has been chosen to obtain the data corresponding to Spanish HE institutions, due to its wide coverage (21,000 titles in February 2014) (Elsevier B.V, 2014) and the availability of SCImago Journal & Country Rank portal (SCImago, 2014), which offers a full system of journal and country scientific indicators.

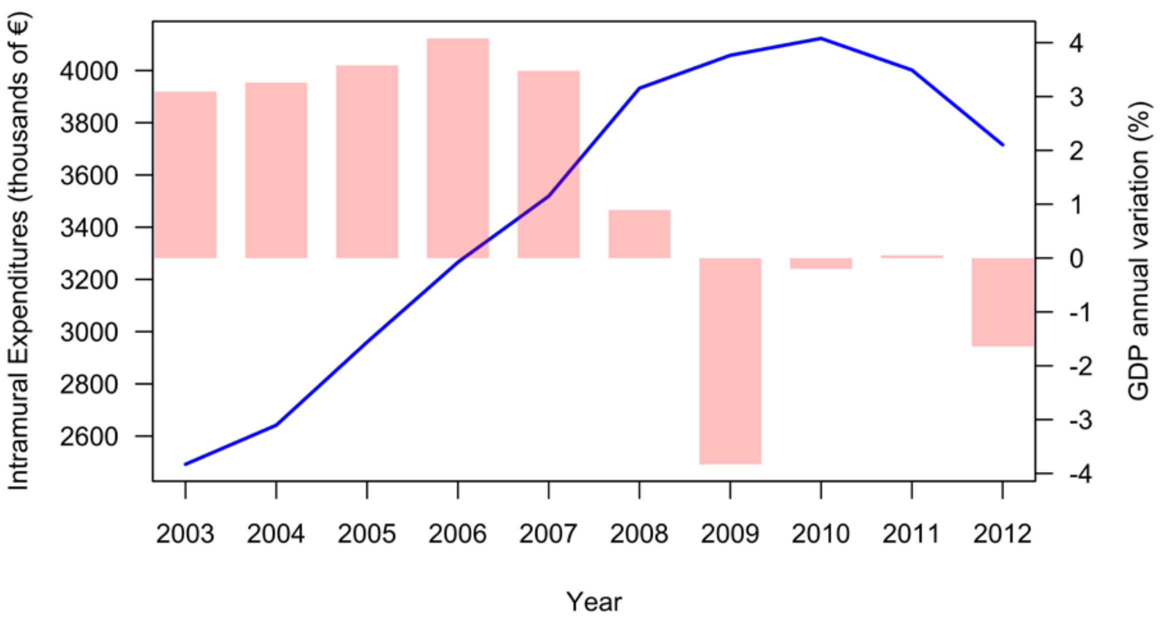
Method

A query looking for the co-occurrence of "Spain" in "Affiliation Country" field and "Universi*" or "Univ" (acronym) in "Affiliation Name" has been run on SCOPUS. Every publication with at least one Spanish affiliation was incorporated to the dataset. The Conference of Spanish University Rectors estimates the scientific production of Spanish universities in roughly 39,000 articles indexed in Web of Science for year 2010, and this query returns 38,640 journal articles for that year (query run on 05/07/14), so a high recall can be expected from this information retrieval technique. Publication and citation data will be obtained from SCOPUS and SJR ranking will be used to determine the quantiles, normalized by field. This study covers interval 2003-2012.

Results

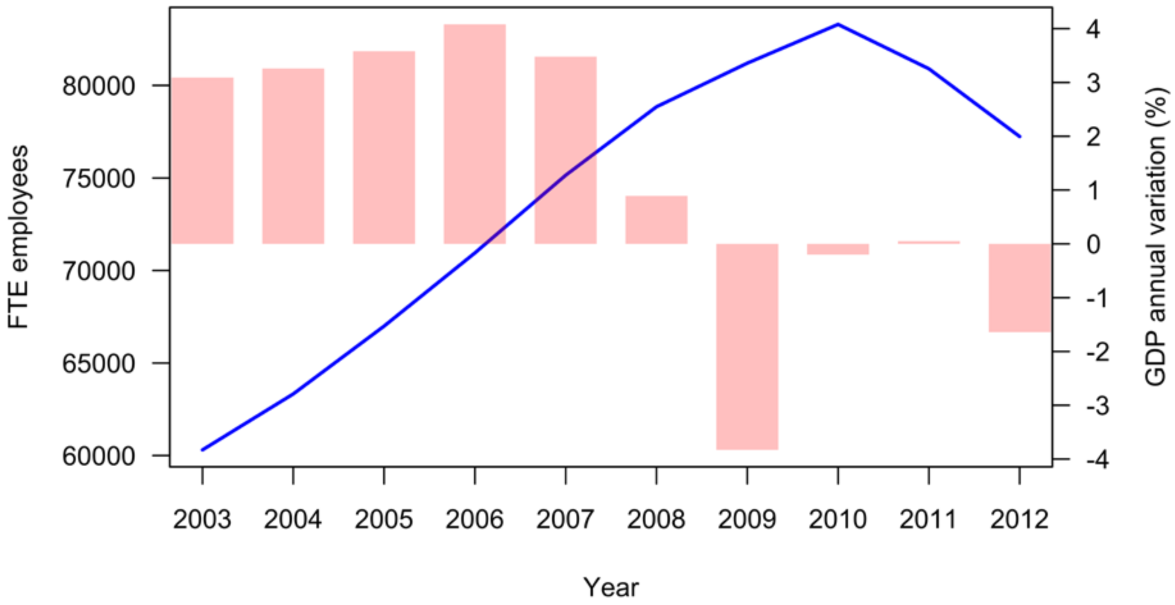
Looking at intramural expenditure data in HE sector (Fig 1), cuts have been going on since 2010, with a sharp, unprecedented decrease in 2012 (7.16%). Bars show the annual percentage growth rate of Spanish economy in those years.

Figure 1: R&D intramural expenditure evolution in Spanish HE.



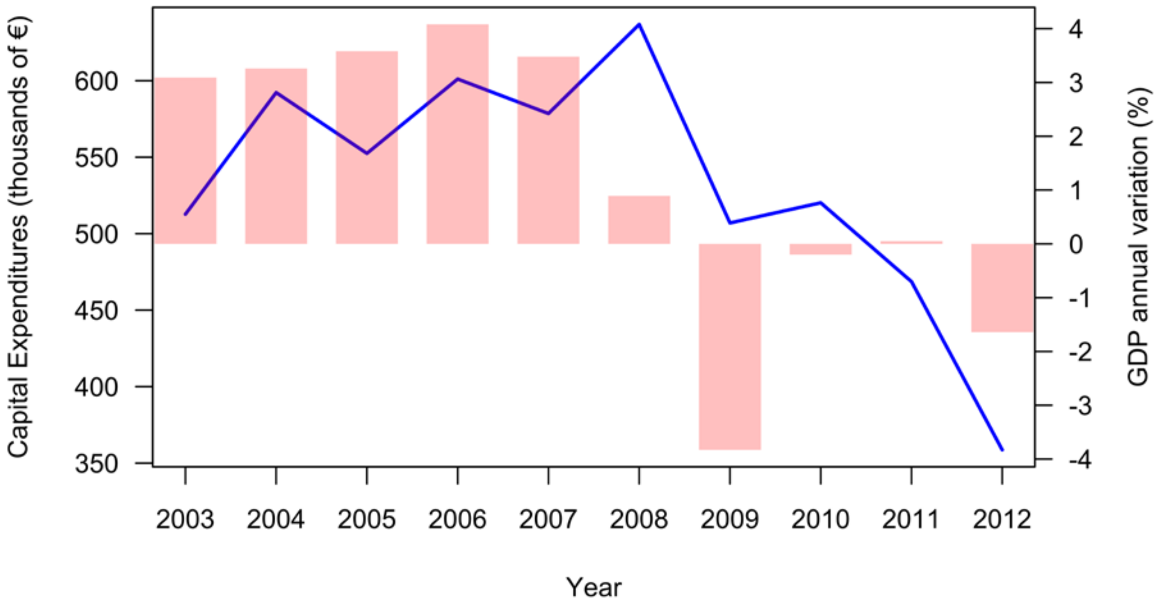
The lion's share of total internal expenditures in HE are the workforce salaries (58% of total expenditures in 2010), so both the size and salaries of workforce will likely be affected by cuts. Figure 2 shows roughly 6,000 FTE researchers disappearing from HE R&D workforce since 2010.

Figure 2: FTE R&D employees in Spanish HE.



An eye should be kept on capital expenditures (Fig 3), since different areas of research may have different needs of fixed assets, and cuts in these expenditures may have different effects, for example, on the productivity of social scientists or material science researchers.

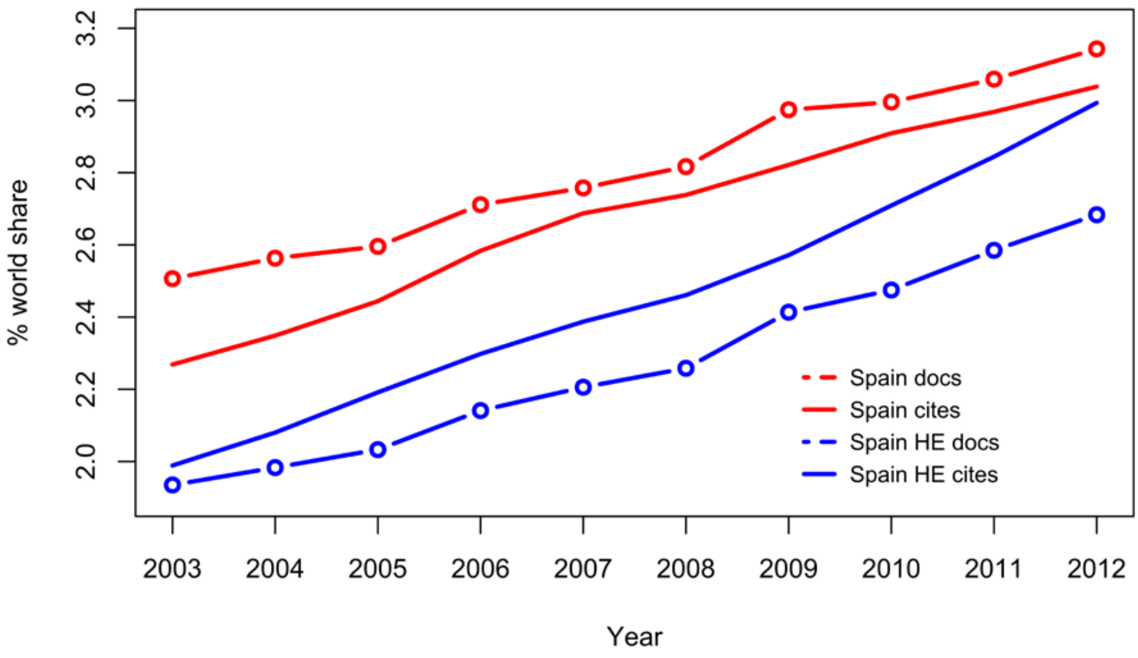
Figure 3: Capital expenditure evolution in Spanish HE.



The downfall in capital expenditures started with the first shadows of the forthcoming economic crisis, and has experienced a 50% decrease from its maximum value in year 2008.

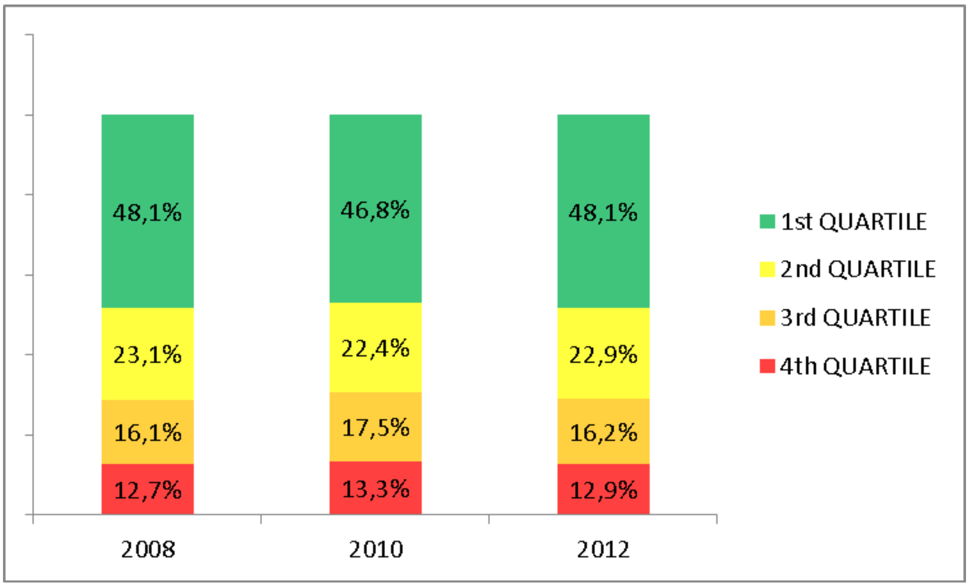
Expenditure cuts do not usually have an immediate effect on the quantity and quality of scientific production. This can be partly explained by the delay between expenditures in R&D and the scientific papers that could come as a result. Total amount of Spanish papers has been continuously rising by an average yearly 10% since 1991 until 2013, where slightly decreased by a -0.4%. Not by chance, Spanish CSIC expects to notice the first drop in their scientific production in year 2014, due to “budget cuts and employee downsizing”. Fig. 4 shows the evolution of Spanish scientific document production and citations received (until 2012) as % of world share and compares it with data corresponding to Spanish HE. Both indicators show an upwards trend and the increasing relevance of HE institutions in Spain’s received citations.

Figure 4: Spain's scientific production and citations received, both at country and HE levels.



Finally, a quartile analysis of SJR values for Spanish publications has been conducted for years 2008-2010-2012, with the aim of studying the evolution of the "excellence" in Spanish scientific production. Figure 5 shows the results.

Figure 5: Quartiles of Spanish publications..



The quartile proportions remain fairly stable if we compare the before and after situation (cuts started in 2010) therefore we cannot state that the excellence of Spanish publications has decreased, at least looking at quartile analysis.

Conclusions

This work sets a method and indicators to evaluate the effects of budgetary cuts on the quantity and quality of Spanish HE scientific output. The authors will continue complementing this work as new data is available, and further studies focused on the effect of cuts in particular subject areas will be conducted.

- There are not clear signs of any decrease in the quantity or quality of scientific production, but early signs of deceleration have been detected.
- The analysis of forthcoming data can be the key to determine the influence exerted by expenditure cuts in the quality of scientific production, since typical research project's length in Spain takes 3 to 4 years.

References

Instituto Nacional de Estadística. http://www.ine.es/ (05/07 accessed)

Elsevier B.V. Content Overview. http://www.elsevier.com/online-tools/scopus/content-overview (05/07 accessed)

SCImago. SCImago Journal & Country Rank. http://www.scimagojr.com/ (05/07 accessed)

Retrieved March 14, 2014, from http://www.scimagojr.com

Cross-national preferences and similarities in downloads and citations of scientific articles: A pilot study[1]

Wolfgang Glänzel* and Sarah Heeffer**

[*] *Wolfgang.Glanzel@kuleuven.be*
Centre for R&D Monitoring (ECOOM) and Dept. MSI, KU Leuven, Leuven (Belgium)
Dept. Science Policy & Scientometrics, LHAS, Budapest (Hungary)

[**] *Sarah.Heeffer@kuleuven.be*
Centre for R&D Monitoring (ECOOM) and Dept. MSI, KU Leuven, Leuven (Belgium)

Introduction

Much has already been written about the relationship between the download of publications and their citation impact. In particular, specific issues such as the role of open access have been studied in this context (e.g., Harnad & Brody, 2004; Brody et al., 2006; Gaule & Maystre, 2011). Also the question of causality has been raised since the download process usually starts up earlier than the corresponding citation process and the frequency of downloads exceeds that of citations by one or more orders of magnitude (Moed, 2005). Download statistics are already celebrated as the true measure of usage and on the basis of their objectivity and independence of particular databases seen as a real alternative to and predictor for citation impact (Brody et al., 2006). However, the situation is not so simple. The interrelation between the two processes is much more complex as might appear since frequent citation might imply or affect the download of full-text papers (Moed, 2005). In particular, careful reading of scholarly publications might also draw the reader's interest to cited references and thus stimulate downloading further related work.

Before we study several aspects of download frequency and citation rates, we would clarify that downloads, which have already superseded the formerly popular photocopying of printed documents, has become one important contemporary electronic form of gaining access to the full text of scientific publication. However, it is by far not the only one. (Electronic) reprint requests of otherwise unavailable documents as well as acquiring hard copies of printed matter still remain important sources of scientific information. This is still the easiest and most favoured form of gaining access whenever downloading the complete article is for any reason difficult, too slow or even impossible.

In the following study we intend first to analyse the interrelation between downloads and citations of a large sample set of about 80,000 documents put online in 2008 and downloaded/cited till June 2013. The second part of the analysis will be devoted to cross-national information flow in the sense of the notion proposed by Glänzel & Schubert (2006). In particular, the analysis will be conducted along the following research questions.

1. Do the findings confirm earlier observations made in previous studies (e.g., Moed, 2005; Brody et al., 2006; Thelwall, 2012), concerning the correlation between the two statistics?

[1] The authors wish to acknowledge the use of data provided by Elsevier in the framework of the Elsevier Bibliometrics Research Programme (EBRP). We also thank Marc Luwel and Henk Moed for their initiative and their valuable input. The Elsevier data set was supplemented by data sourced from Thomson Reuters Web of Science used for analysis of the cross-national preferences in citations.

2. Is there a deviation of cross-national download patterns from cross-national citations flows?

Download and citation data used to answer the first and part of the second question has been provided by Elsevier. The data set comprises monthly download and citation counts for papers published in journals from various fields in the sciences, social sciences and humanities. Along with these statistics also the uniquely identified location (country) of download was provided. These data have been used as they are, that is, they have not undergone any further cleansing process. In order to clean noise caused by the superposition of field-specific peculiarities in download and citation behaviour one individual journal has been chosen to obtain more specific results. We decided to use the journal *Physica A – Statistical Mechanics and Its Applications.* In addition, cross-national citation preferences according to the 2nd research question have been analysed on the basis of data extracted from Thomson Reuters Web of Science for publications in the same journal indexed in the 2008 volume and cited in the period 2008–2012. This was necessary to determine the countries of authors who have cited the papers in question.

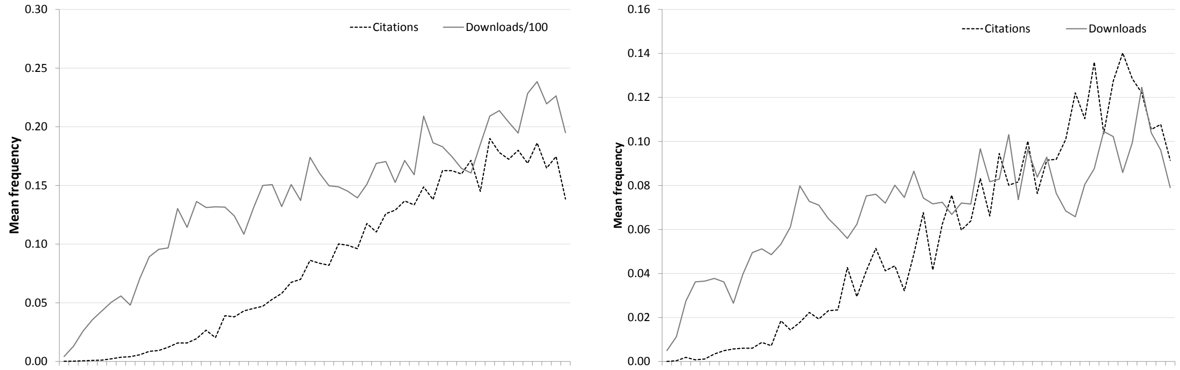
Methods and results

As mentioned in the introduction, we have used a sample set provided by Elsevier for the analysis. The data set comprises documents from different journals that were online available in 2008. Download and citation statistics were drawn till June 2013 on monthly base. In addition to the complete data aggregation, *Physica A* was used to break down data to one specific journal. In a first step the monthly evolution of citation rates was plotted against the corresponding download frequencies. Thereafter a regression analysis of the conditional expectation of citations vs. the scaled number of downloads is conducted for the complete observation period. Both analyses were applied to the complete set (N = 77,887) as well as to the selected journal (N = 2,646). This part of the study refers to the first research question.

Regression analysis

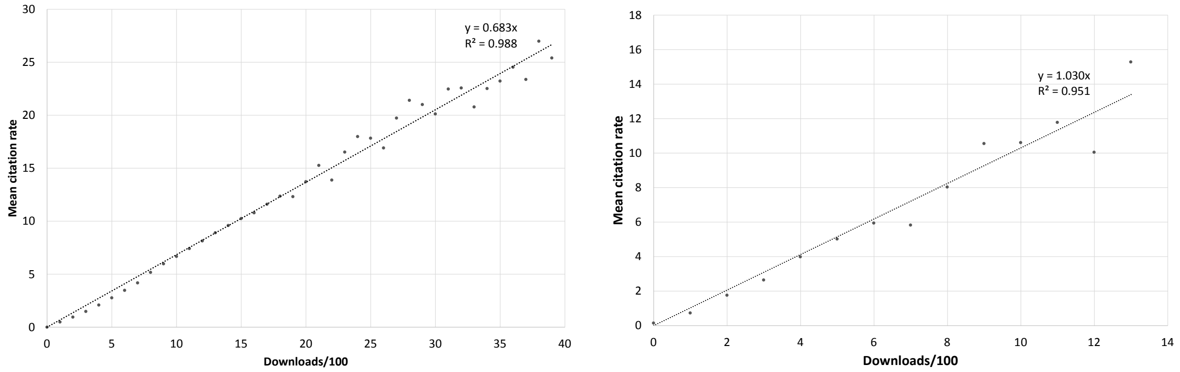
Figure 1 shows the evolution of downloads and citation for the complete set (left) and the journal *Physica A* (right). For this representation the mean values over documents has been used. In either case, the number of downloads has been divided by 100. This 'scale factor' is in line with the findings by Moed (2005), who concluded based on a strong rank correlation between downloads and citations that during an initial period of two years in a papers' life cycle one citation corresponds to about 100 downloads. Indeed, this resulted in the same order of magnitude in our set as well. Both curves show a parallel evolution, where downloads start and increase earlier than the corresponding citation rates. This observation, which confirms conclusions drawn in a study by Schloegl et al. (2010) according to which the correlation between *Usage Impact Factor* and *Journal Impact Factor* in the field of oncology was rather moderate because of the different obsolescence of downloads and citations, is important for possible conclusions from the following analysis.

Figure 1: Monthly evolution of downloads vs. citations for the complete set (left) and *Physica A* (right). The x-axis represents the time elapsed from publication by month.



We have kept the transformation of download frequencies by dividing by 100 for the following regression analysis as well. Unlike in previous studies by Moed (2005), Brody et al. (2006), O'Leary (2008), Liu (2011) and Thelwall (2012), we use for the regression analysis and the visualisation the same method that has been proposed and applied by Glänzel et al. (2004). The mean citation rate of documents have been calculated under the condition that those have been downloaded a given number of times. In order to calculate these conditions the integer part of the transformed download, i.e., the number of downloads divided by 100, has been calculated. In verbal terms, the condition "0" stands for less than 100 downloads, "1" form at least 100 but less than 200 downloads, etc. Ideally, the conditional means should follow a (not necessarily linear) function of the number of downloads; otherwise, if the means are constant, citations and downloads are not correlated. We have to stress that the sample sizes underlying conditional means dramatically decreases with growing number of download thus resulting in huge fluctuations at the high end of the download scale. We have therefore truncated at a point beyond which the number of underlying documents drops below ten. Since the mean citation rate was slightly below zero for documents downloaded less than 100 times, we decided to put the intersection zero. The results for the complete paper set and the journal *Physica A* about five years after the papers were available online are presented in Figure 2. The correlation is rather strong ($r > 0.95$) and the slope substantiate that download frequency is roughly by two orders of magnitude larger than citation rates. The "exchange rate" in the case of the physics journal is about 100 (i.e., one citation corresponds to 100 downloads), while this amounts to about 70 for the complete set (which includes life sciences, engineering and mathematics as well).

Figure 2: Conditional mean citation rates as a function of downloads five years after online availability for the 'complete set (left) and *Physica A* (right)



The results shown in Figures 1 and 2 partially confirm results of earlier studies Moed (2005), Brody et al. (2006), O'Leary (2008), Liu (2011) and Thelwall (2012), who found significant positive correlation between citations and downloads in several fields in the sciences. In this context, it should be mentioned that using the number of downloads as the condition, by no means implies causality. This is in line with observations by Moed (2005) and Schloegl et al. (2011), who stressed in contrast to the results of other authors (e.g., Brody et al., 2006, who suggested the use of downloads as early predictors of citation impact) that no conclusions might be drawn on the possible effect of early downloads on later citation rates of a paper.

In conclusion, we would like to mention that a similar regression model as applied above but using citations as the condition and calculating mean download frequencies would, of course, provide comparable results.

Finally we have to mention one general limitation of this type of analysis: Since articles might already be available as online-first versions or accessible via institutional or individual repositories, downloads of or citations to early or 'in-press' versions can indeed affect *response indicators*. However, there is no evidence of systematic or serious distortion of download and citation processes.

Cross-national preference in downloads and citations

The second part of the analysis, which refers to the 2nd research question, aims at providing completely new insights in what downloads and citations stand for in terms of information use. Taking up the idea of analysing country-by-country cross-reference and cross-citation networks (Glänzel & Schubert, 2005; Schubert & Glänzel, 2006), this section aims at a comparison of the patterns found in citation analysis with similar results from download statistics. In order to be able to assign citing articles to the country of co-authors a citation database that records affiliation information is needed. Since corporate addresses of citing papers were not included in the dataset and for reasons of comparability with previous results (cf. Glänzel & Schubert, 2005; Schubert & Glänzel, 2006), Thomson Reuters Web of Science is used here again. While the previous studies on cross-citation patterns revealed that scientific collaboration, geopolitical location, cultural relations and language are determining factors in shaping the national preference, one might assume that downloads are rather subject to phenomena reflecting globalisation and rather general patterns of electronic communication. In fact, before information flow is manifested by proper citations, publication of the results in a scientific journal and indexing the document in question in a bibliographic database is required. By contrast, downloading scientific documents is not necessarily or directly linked to the production of own research results.

Figure 3: National shares in all downloads/citations of the 15 most active countries based on the journal *Physica A*

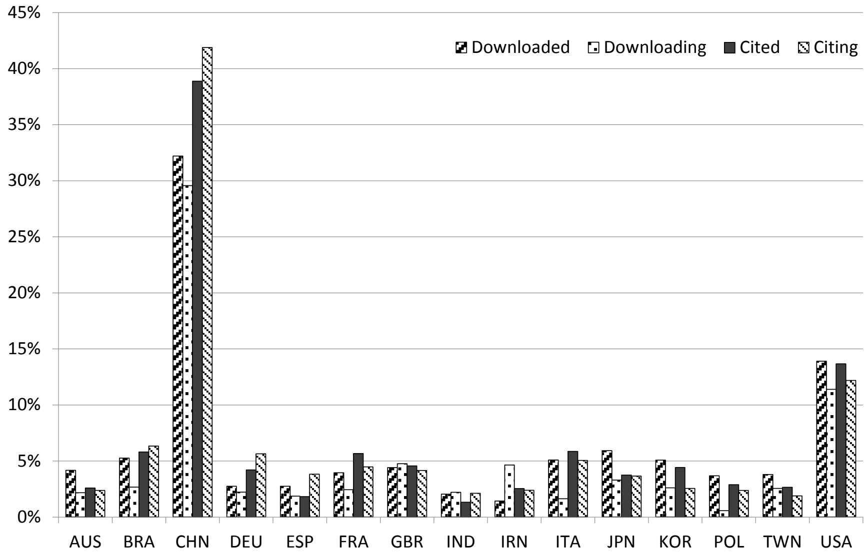


Figure 3 shows the national share in all downloads of and citations to articles published in *Physica A* in 2008. In the figure the 15 most active countries in terms of downloads and citations are displayed. Downloads and citations are not based on exactly the same data set since for the download statistics those papers were used that were online available in 2008 and downloaded till June 2013, while for the citation statistics the 2008 volume of the Web of Science has been used and citations were counted till end 2012. This was necessary to be able to assign citations to the countries of citing authors. Counts have not been fractionated. There is also a basic difference between the country of download and citation: while the first one, i.e., the location of download, is unique, the country (countries) of citation depends on possible international co-authorship. Nevertheless, the basic features of the two processes should be captured by these counting schemes as well. Data have been normalised by the total number of downloads and citations respectively. In this manner, a direct comparison across countries and between national download and citation patterns is possible.

Besides the enormous share of both downloads and citations of/in China and the US also the share of Brazil and Korea is worth mentioning. The high share of cited/citing articles in Germany, France, UK and Italy are not unexpected as those are in line with citation patterns known from this science field (e.g., Glänzel et al., 2002). However, the relatively low share of downloads compared with the citation patterns in Germany and France is striking. Even more interestingly, the low share of downloads in Brazil, Italy, Poland, Japan and Korea is contrasted by pronouncedly higher share of 'downloadedness'. A common pattern in China, Germany and France is that citations are apparently more important than downloads in these countries. Finally, the strikingly high share of downloads in Iran – compared with both, downloadedness and references and citations – is worth mentioning.

In order to have a look at (cross-)national preferences and similarities in downloads and citations we break down statistics by downloaded, downloading, cited and citing countries. Again we have used the journal *Physica A* to demonstrate the model. Unlike the separate normalisation by publications, references and citation for capturing cross-national preferences proposed by Schubert and Glänzel (2006), we use a simple cosine measure to capture similarities. In particular, download and citation frequencies are divided by the square root of the national total according as downloaded/downloading and cited/citing direction is needed.

In this manner possible asymmetry (downloading vs. downloaded and citing vs. cited direction) is kept. In particular the following measure is used.

$$r_{ij} = \frac{p_{ij}}{\sqrt{p_i \cdot p_j}},$$

where p_{ij} denotes the number of downloads (citations) in country j of papers published in country i. p_i denotes the number of all downloaded (cited) papers published in country i, p_j is the number of all downloading (citing) papers in country j. By contrast to the previous statistics that where normalised by the world total, this measure is sensitive to the activities of the corresponding partner country.
In what follows, similarities for the 25 most active countries in terms of downloads and citations are shown in Table 1. The value 0.05 was used as the lower threshold. Only similarities r_{ij} stronger than this threshold value are displayed. As we have expected, the domestic activities represented by the main diagonal are quite dominant. The enormous downloading activities of China and the US confirm the finding shown in Figure 3, where frequencies had been normalised by the world total. Downloading in Iran is spread over almost all other countries; the maximum value is taken for Iran itself. The relationship is – unlike for China and the US – a completely asymmetrical one. This is contrasted by Iran's low cross-national similarities in citations (both directions). The most striking pattern has been found for Israel. This country has very weak cross-national similarities in downloads (both directions). Even the domestic strength is low (about 0.01). For most of the selected countries, however, the two types of similarities express alike patterns. Germany's and France's overall property of being more active in citation flows than in download activities is also reflected the cross-national similarities. Here again we find a slight deviation from China, where cross-national preferences in downloads and citations are, in contrast to the corresponding national shares, quite similar (cf. Table 1 and Figure 3). On should notice that data in Table 1 are based on similarities, i.e., on relative indicators, while the bars in Figure 3 stand for shares of absolute numbers. This explains the dominance of China in the chart which is contrasted by its more moderated figures in the table.

Table 1: Cross-national similarity in downloads and citations for the 25 most active countries based on *Physica A* (2008).
Download similarities (top; row vector: downloaded country, column vector: downloading country) based on source data provided by Elsevier), citation data (bottom; row vector: cited country, column vector: citing country) sourced from Thomson Reuters Web of Science

Country	ARG	AUS	BEL	BRA	CAN	CHE	CHN	DEU	ESP	FRA	GBR	HUN	IND	IRL	IRN	ISR	ITA	JPN	KOR	MEX	POL	PRT	TUR	TWN	USA
ARG	0.13	0.06		0.07	0.06		0.15	0.07	0.11	0.07	0.08		0.07		0.15			0.05					0.07		0.12
AUS		0.11					0.15				0.08				0.11										0.07
BEL			0.06				0.08				0.07				0.06										0.09
BRA				0.23			0.11				0.08				0.13								0.06		0.09
CAN					0.06										0.06										
CHE						0.09	0.15				0.05				0.09										0.07
CHN		0.05			0.06		0.49	0.05			0.09		0.07		0.19			0.07	0.05				0.07	0.08	0.15
DEU							0.07	0.07							0.05										0.07
ESP							0.09		0.09						0.06										0.07
FRA							0.09	0.06		0.12	0.07				0.07										0.10
GBR							0.10				0.10				0.06										0.08
HUN							0.07					0.08													
IND							0.17	0.06			0.06		0.34		0.11			0.05					0.06		0.10
IRL					0.05		0.18	0.06			0.11		0.06	0.06	0.14								0.08	0.06	0.10
IRN													0.05		0.34										0.07
ISR																									
ITA				0.05	0.06		0.15	0.06	0.06	0.06	0.08		0.06		0.15		0.10	0.06	0.05				0.08		0.14
JPN							0.14				0.06				0.09			0.14							0.08
KOR							0.14				0.11				0.12				0.20				0.06		0.08
MEX					0.08		0.09	0.06		0.06	0.06		0.07		0.09					0.14					0.12
POL		0.05		0.06	0.08		0.19	0.07	0.05	0.05	0.10		0.06		0.28						0.09		0.11	0.09	0.17
PRT		0.06			0.07		0.20	0.06		0.05	0.11		0.06		0.15			0.06	0.05			0.11	0.07	0.06	0.15
TUR																							0.20		
TWN							0.10				0.06				0.15								0.05	0.11	0.06
USA					0.06		0.19	0.06		0.06	0.12				0.12				0.07				0.05		0.18

Country	ARG	AUS	BEL	BRA	CAN	CHE	CHN	DEU	ESP	FRA	GBR	HUN	IND	IRL	IRN	ISR	ITA	JPN	KOR	MEX	POL	PRT	TUR	TWN	USA
ARG	0.35	0.09		0.13					0.23				0.06												0.05
AUS	0.19	0.24		0.12			0.09		0.08																0.12
BEL			0.32					0.08		0.09	0.11					0.11	0.07				0.06				0.26
BRA				0.46			0.06	0.06	0.11														0.09		0.19
CAN					0.31																				
CHE					0.13	0.36	0.10																		0.07
CHN		0.10			0.09	0.05	0.78	0.08		0.05	0.06		0.06					0.06	0.06				0.06		0.25
DEU				0.05	0.05		0.08	0.23			0.07		0.07												0.16
ESP									0.33																
FRA			0.09			0.07		0.21	0.15	0.41	0.11						0.14	0.06			0.06				0.21
GBR		0.05					0.12	0.06		0.06	0.24		0.11			0.05									0.30
HUN			0.07		0.08	0.08	0.13		0.17		0.05	0.17										0.12			0.16
IND													0.30												0.07
IRL				0.08			0.06	0.09	0.08					0.16					0.05				0.24		0.18
IRN							0.09								0.42										0.06
ISR				0.05	0.07		0.08	0.16								0.32				0.05					0.15
ITA	0.07			0.09		0.06	0.06	0.07	0.12								0.40	0.05							0.43
JPN				0.08														0.43					0.09		0.12
KOR	0.12	0.09		0.08	0.08		0.16		0.09		0.07						0.05		0.28		0.06		0.08		0.22
MEX					0.08			0.08	0.05			0.07								0.45					0.18
POL	0.05		0.06	0.07			0.05	0.05	0.06							0.11	0.06				0.44				0.12
PRT	0.09			0.07	0.07		0.06	0.08	0.17		0.05						0.06					0.21			0.14
TUR																							0.33		
TWN					0.09		0.09	0.08		0.06	0.06				0.07			0.06						0.39	0.14
USA	0.07			0.11	0.09		0.15	0.12	0.12	0.07	0.11						0.16	0.06	0.10				0.12	0.05	0.95

Geopolitical location, cultural relations and language being determining factors in shaping preferences in cross-citations (cf. Schubert & Glänzel, 2006) seem to be somewhat less pronounced in cross-download relations (cf. Table 1).

Conclusions

While the first part of this pilot study referred to research questions, which aimed at checking our data set against results reported in earlier studies (e.g., Moed, 2005; Brody et al., 2006; O'Leary, 2008; Schloegl & Gorraiz, 2010 and 2011; Liu, 2011 and Thelwall, 2012), confirmed those findings in terms of the scale factor, i.e., that downloads are of two orders of magnitude more frequent than citations in an initial phase, and also showed a rather strong correlation between the two analysed statistics in both the complete sample of about 80,000 papers and the selected journal *Physica A*, the second part attempted to depict patterns of cross-national information flow as measured by downloads and citations. Downloads are not closely related to documented scholarly communication as citations are by nature. Documents might be downloaded by anybody who has access without using or incorporating downloaded information in own publishable research. This is also substantiated by the asymmetries we have found. These asymmetries refer to both the deviation of citation patterns from downloads and the different patterns of the two directions within downloads and citations, respectively. The case of Iran might serve as the most striking example of this phenomenon. Iran is among the most active downloading countries with respect to the selected physics journal; it is downloading information from many other countries but is not significantly downloaded by others, and has rather weak citation links with other countries in both directions.

Deepening the results of this pilot study and applying methodology to larger samples and other disciplines as well as analysing the dynamic aspects of the two analysed processes will be tasks of future research.

References

Brody, T., Harnad, S. & Carr, L. (2006), Earlier web usage statistics as predictors of later citation impact, *JASIST*, 57(8), 1060-1072.

Gaule, P. & Maystre, N. (2011), Getting cited: Does open access help?, *Research Policy*, 40(10), 1332-1338.

Glänzel, W., Schubert, A. & Braun, T. (2002), A relational charting approach to the world of basic research in twelve science fields at the end of the second millennium. *Scientometrics*, 55(3), 335-348.

Glänzel, W. & Schubert, A. (2005), Domesticity and internationality in co-authorship, references and citations, *Scientometrics*, 65(3), 323-342.

Glänzel, W., Thijs, B. & Schlemmer, B. (2004), A bibliometric approach to the role of author self-citations in scientific communication, *Scientometrics*, 59(1), 63-77.

Harnad, S. & Brody, T. (2004), Comparing the impact of Open Access (OA) vs. non-OA articles in the same journals. *D-Lib*, 10(6).

Liu, X.L., Fang, H.L. & Wang, M.Y. (2011), Correlation between Download and Citation and Download-citation Deviation Phenomenon for Some Papers in Chinese Medical Journals , *Serials Review*, 37(3), 157-161.

Moed, H.F. (2005), Statistical relationships between downloads and citations at the level of individual documents within a single journal, *JASIST,* 56(10), 1088-1097.

O'Leary, (2008) The relationship between citations and number of downloads in Decision Support Systems, *Decision Support Systems,* 45(4), 972-980.

Schloegl, C. & Gorraiz, J. (2010), Comparison of citation and usage indicators: the case of oncology journals. *Scientometrics*, 82(3), 567-580.

Schloegl, C. & Gorraiz, J. (2011), Global Usage Versus Global Citation Metrics: The Case of Pharmacology Journals. *JASIST*, 62(1), 161-170.

Schubert, A. & Glänzel, W. (2006), Cross-national preference in co-authorship, references and citations, *Scientometrics*, 69(2), 409-428.

Thelwall, M., (2012), Journal impact evaluation: a webometric perspective, *Scientometrics,* 92(2), 429-441.

Humanities in the bibliometric spotlight - Research output analysis at the University of Vienna and considerations for increasing visibility

Christian Gumpenberger*, Johannes Sorz**, Martin Wieland* and Juan Gorraiz*

*christian.gumpenberger@univie.ac.at , *martin.wieland@univie.ac.at , *juan.gorraiz@univie.ac.at
Library and Archive Services (Bibliometrics Department), University of Vienna, Boltzmanngasse 5, Vienna, 1090 (Austria)

**johannes.sorz@univie.ac.at
Rector's Office, University of Vienna, Universitätsring 1, Vienna, 1010 (Austria)

Abstract
A quantitative analysis of the longitudinal research output development in the humanities at the University of Vienna was performed for a six years interval (2007-2012). According to target agreements between the rectorate and the faculties, the language requirement was met successfully with an increase of the non-German (particularly English) output. The results also show an increasing trend line regarding the percentage of peer reviewed articles and of publications indexed in world-renowned databases like Arts & Humanities Science Index. The number of publications with a Digital Object Identifier or in Gold Open Access journals is very low.
Further strategies are recommended in order to increase the international visibility of the research output in the humanities.

Introduction
Without doubt the humanities can be regarded as the Achilles' heel for a successful application of scientometric and bibliometric methods in these disciplines' research assessment. It is therefore one of the biggest challenges in these fields. However, evaluations are increasingly based on quantitative bibliometric indicators primarily designed for the sciences, which are either of limited value or even inappropriate in the humanities and therefore criticized.

Previous studies have already pointed out that the humanities follow a different set of rules than the sciences (Nederhof, 2006; Hammarfeldt, 2013). Three major differences can be identified:

1) different publication habits and channels (e.g. importance of monographs and edited books, preference of single-authored publications, language, etc.),
2) different audiences not only restricted to the scholar community, and
3) lack of globally available data sources for bibliometric purposes.

Since research assessment is commonly based on Web of Science (WoS) and Scopus as data sources, which are knowingly not representative for the humanities, alternative data sources need to be taken into account. That is why a few initiatives (like e.g. "The Excellence in Research for Australia" (ERA) or the European Reference Index for the Humanities (ERIH)) have come into being. Furthermore it has been suggested to use alternative data sources like Libcitations, Google Scholar, Google Books (Kousha and Thelwall, 2009), Book Citation Index (Gorraiz et al, 2013), institutional information systems (Ossenblok et al. 2012), etc.

Language biases concerning the coverage of the Science Citation Index and its consequences for international comparisons of national research performance have also been reported) for the non-English language countries in Europe (e.g. van Leeuwen et al, 2001, van Leeuwen, 2013).

However, further and/or even better criteria, sources and indicators are still required for meaningful quantitative research assessment in the humanities (Li & Linmans, 2010; Hug et al., 2013).

One of the major attempts of quantitative research assessments is to measure the impact of research output. Some have already tried to do so in the humanities, but have insisted in using citations (Leydesdorff et al., 2011). But these can only be an acceptable proxy for impact measurement in disciplines, where the „publish or perish" community is actually the most relevant target group (Gorraiz et al., 2014).

Since e-media and the social media have gained momentum, emerging metrics, like usage metrics (Kurtz and Bollen, 2010; Gorraiz and Gumpenberger, 2010) and altmetrics open up new vistas for alternative approaches (Kousha and Thelwall, 2009; Priem et al., 2012; Tang et al., 2012; Wouters and Costas, 2012). These new approaches sound promising, since they might have the potential to overcome the inadequacy of conventional bibliometric methods. Unfortunately first results of recent studies are sobering and have clearly shown that the research output in the humanities is still light-years away from the digital era and therefore has a very low online visibility (Hammarfeldt, 2014).

While these new approaches might not be appropriate for impact measurement, they will certainly enhance visibility, which is of major importance for these disciplines. In order to learn which measures should be taken into account to increase the visibility of the quantity and quality of the research output in the humanities, it is crucial to gain deeper insight into its development in the last years. The University of Vienna was exemplarily used for this case study.

The University of Vienna is the oldest university in the German-speaking world and one of the largest in Central Europe. With 92.000 students, 9.500 employees, 6.700 of who are scientific staff, it is the largest teaching and research institution in Austria.
The humanities are well represented at the University of Vienna which makes it an ideal case for this study. In 2012 approximately 28% of the total FTE scientific staff (incl. professors and researchers funded by third-party funds) could be attributed to the humanities. In the same year approximately 31% of all publications published in the same year could be attributed to the humanities

The humanities at the University of Vienna are organized in three faculties and a center for translational studies. For this study two of these humanity faculties were selected. Faculty 1 comprises mainly of language and area studies. It is one of the largest faculties of the university with 484.3 FTE scientific staff (2012). Faculty 2 comprises the historical sciences incl. archeology and art history with 258.4 FTE scientific staff (2012).
Periodically, target agreements are negotiated and signed between the rectorate and the faculties, which define the corresponding budget allocations. Although the research output of the faculties is not an indicator for budget allocation, it is constantly monitored by the

rectorate and discussed in the negotiation process. This allows the rectorate to take steps ("negotiate targets") if the research output of a faculty is exceptionally low or does not correspond to the university's overall research strategy. According to the University's research strategy and in order to increase international visibility, the Viennese research output in the humanities should not exclusively be written in German. Peer review is considered as the most essential quality indicator for all types of publications. According to further target agreements, both faculties were encouraged to increase the total number of publications in high-quality journals, particularly in ERIH-listed journals. At the same time monographs were stressed as important document types for both faculties. Although the subsequent aspects have been frequently discussed with the faculties, so far the target agreements did not include any recommendations regarding Arts & Humanities Science Index (A&HCI) or Social Sciences Citation Index (SSCI) coverage, co-authorship or Open Access publishing.

Aims of the study

The main research questions are:

- What can be learned from the longitudinal research output development in the humanities at the University of Vienna?
- How can the observed trends be explained? Are influences of the "target agreements" somehow reflected?
- Which strategies can be developed in order to increase the visibility of the research output in these disciplines?

Methodology

A quantitative analysis was performed based on data drawn from the research documentation system of the University of Vienna called RAD (Research Activity Documentation)[i].
The longitudinal research output development in the humanities was studied by means of two selected faculties. A six years interval (2007-2012) was applied.

The two obtained data pools were analysed according to the following criteria:

- total output, distinguished by following document types:
    - books
    - editions (abstract volumes, art catalogues and collections, encyclopaedias, edited books, proceedings, abstracts, etc.)
    - articles in journals
    - letters
    - articles in proceedings
    - contributions in editions
    - other publications (including book reviews, reprints, translations, working papers, articles in newspapers, reports, internet publications, annotations, audiovisual contributions, meeting abstracts, etc.)
- percentages of single authored publications
- percentages of publications affiliated at least to one non-Austrian affiliation (percentages of team authorship international)
- language (in percentages): English, German, not-only-German publications
- percentages of peer-reviewed publications[ii]
- coverage in different databases and indexes: SSCI, A&HCI, ERIH
  SCI was not considered in the analysis of the humanities due to insignificant coverage. Coverage data could have been overlapping in some cases, thus the same publication

could have been indexed in more than one database. The available RAD data did not allow for further distinction. ERIH lists were periodically updated, the last update was made in 2011.
The coverage analysis was expanded for benchmark purposes including the two most related German-speaking universities, i.e. Humboldt University of Berlin and University of Zurich.

- percentages of open access publications in journals indexed in the Directory of Open Access Journals (DOAJ). DOAJ journal lists were periodically updated, the last update was made in 2012.
- percentages of publications containing a Digital Object Identifier (DOI)[iii]

In order to cross-check and validate data collected from RAD, complementary searches (affiliation: University of Vienna or University Wien) were performed in SSCI and A&HCI. The obtained results for each analysed faculty in Vienna were finally compared to the target agreements for the corresponding time interval.

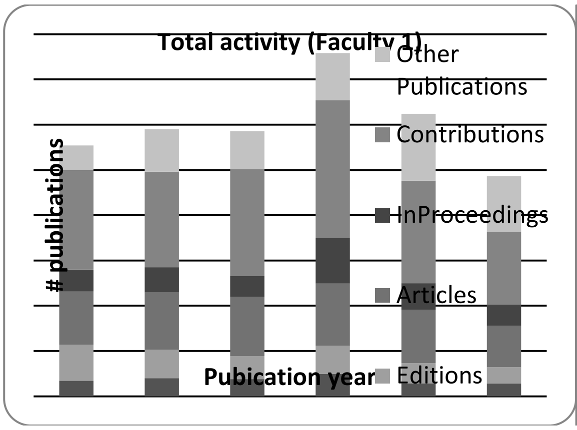
Results and Conclusions

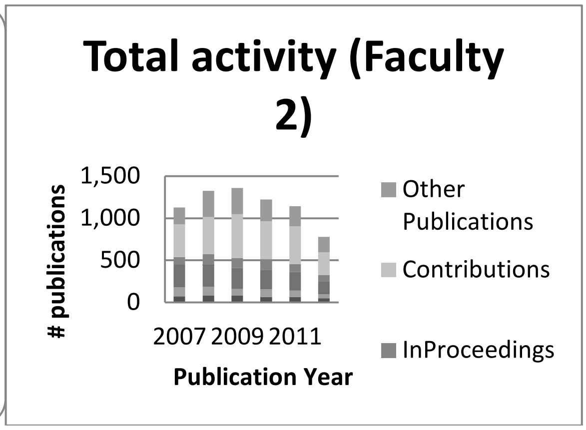
Document types and total activity

Overall neither the number of monographs nor the number of journal articles is significantly increasing. Document types like “Other publications” and “Contributions” are responsible for almost half of the research output. It therefore seems very unlikely that journal articles will become the preferred communication channel for the humanities in the near future. Currently new types of publications (mostly internet-based) are gaining momentum worldwide.

The total activity is comparatively shown for both faculties distinguished by document type in Figures 1 a-b.

Figures 1 a-b: Total activity (faculties 1 and 2).



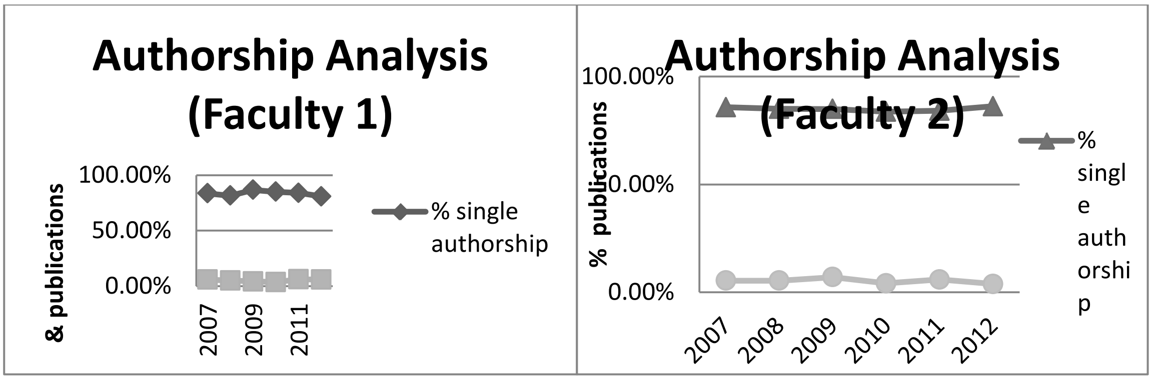


Authorship

The percentage of single-authored publications remains quite constant ranging from 80 to 90%. International collaboration is static with values below 10% for faculties 1 and 2.
Single-authored publications are still most common in the humanities.
Of course well-tried discipline-specific publication cultures are not likely to change as rapidly as all ambitious digital initiatives might hope, particularly not in the humanities.

Single authorship is opposed to international team authorship comparatively for both faculties in Figures 2 a-b.

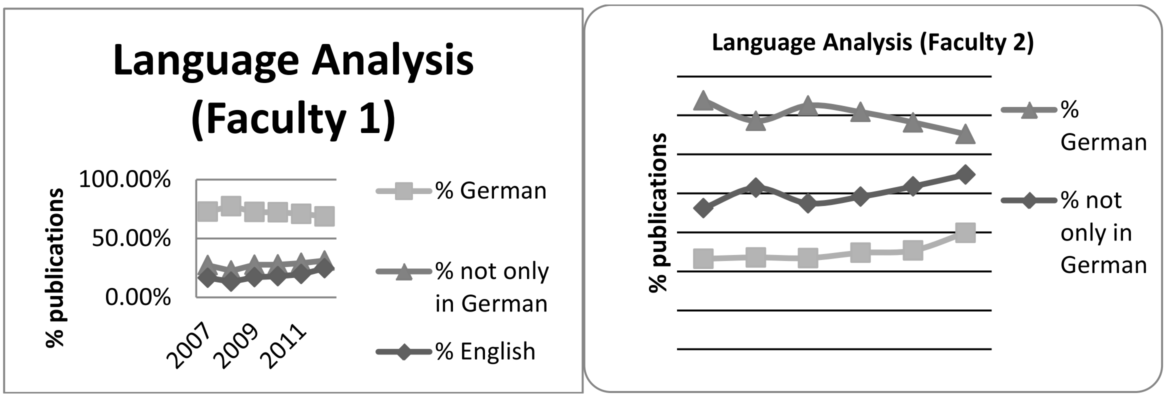
Figures 2 a-b: Authorship Analysis (faculties 1 and 2).



Language

In both analysed faculties, the number of English language publications is slightly increasing. Restricted to research articles only this trend is even more obvious. This development is in accordance with the need for a higher degree of internationalisation and more visibility.
Moreover the number of “not only in German” language publications shows an upward trend as well. Additional languages are certainly connected to the particular research focus.
The number of simultaneous publications in German and English is low and fluctuating between 0 and 24 publications with no clear trend observed.
The results are comparatively shown for both faculties in Figures 3 a-b.

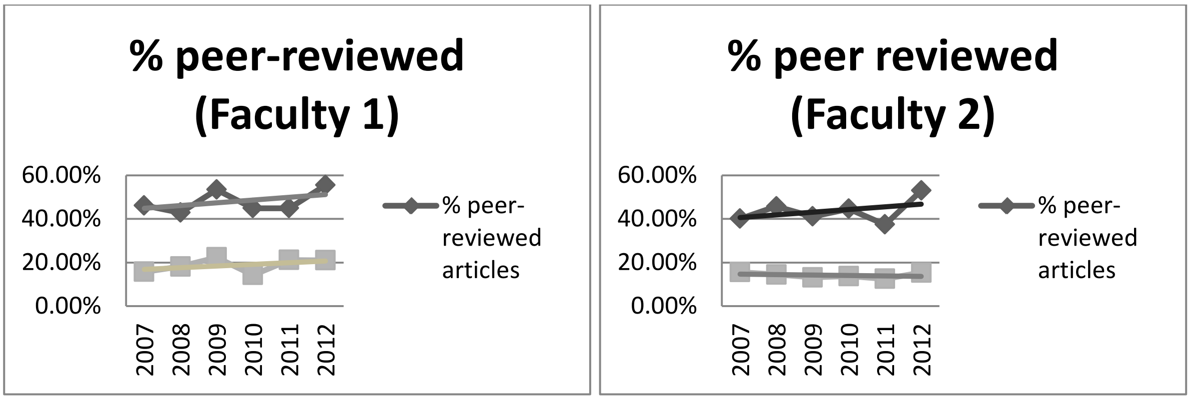
Figures 3 a-b: Language Analysis (faculties 1 and 2).



Peer-reviewed output

The percentage of peer-reviewed output ranges from 10 to 20% for both faculties resp. 40 to 60% restricted to journal articles only. Figures 4 a-b show an increasing trend line for articles.

Figures 4 a-b: Percentage of peer-reviewed contributions (faculties 1and 2).



Coverage in databases and indexes

The percentage of articles indexed in A&HCI is low, but at least increasing from less than 10 to 15% in both faculties.

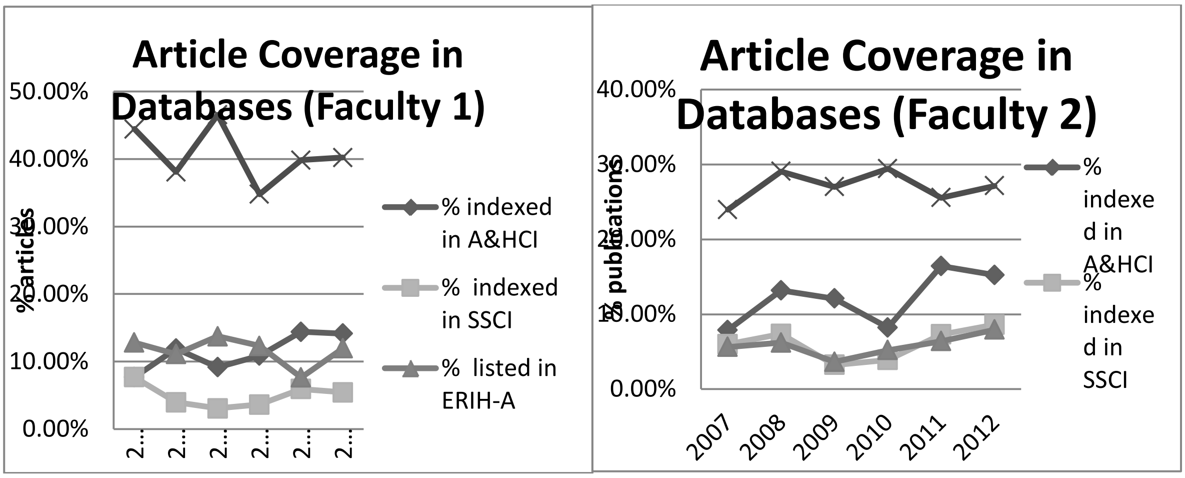
The values for SSCI indexed articles are even lower (<10%). While for faculty 2 the trend line is slightly increasing, it is rather static for faculty 1.

Apparently the coverage in ERIH is low as well (35-45% for faculty 1, 25-30% for faculty 2). Coverage in ERIH-A is comparable to the coverage in A&HCI for faculty 1 (<15%) and to the coverage in SSCI for faculty 2 (<10%).

Considering all the efforts taken and time invested worldwide to generate this alternative Europe-centric index for the humanities, one could certainly question its value when looking at the results.

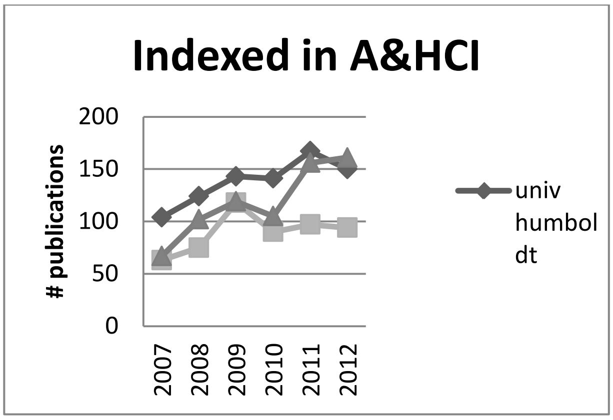
The evolution of articles indexed in A&HCI and SSCI as well as listed in ERIH and ERIH-A is comparatively shown in Figures 5 a-b.

Figures 5 a-b: Article coverage in databases and journal lists (faculties 1 and 2)



The results of the expanded coverage analysis suggest that the increasing coverage trend lines in A&HCI are very similar for the University of Vienna, for Humboldt University of Berlin and University of Zurich (see Figure 6).

Figures 6: Coverage in A&HCI – benchmark analysis

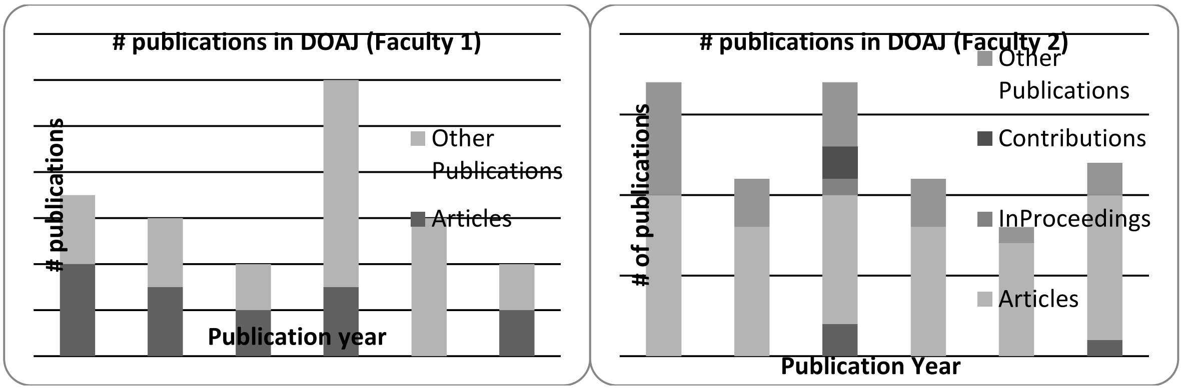


Output in Gold Open Access Journals

The number of items published in Open Access (OA) journals indexed in DOAJ is negligibly low (approximately 1% of the whole publication output) for both faculties. An increasing trend is definitely missing, despite of the fact that OA has been very actively promoted at the University of Vienna. There are a couple of reasons that can explain the low uptake. First of all, in the reporting period, the University of Vienna has had no official OA policy, which would certainly have helped to encourage researchers to make up their minds accordingly. The official OA policy was published in June 2014 and will hopefully help to increase OA awareness at the University of Vienna in all scientific fields. Second, OA journals are still scarcely available in the humanities. Third, if available, their quality might not meet the required standards. Fourth, the popular author-pay model is difficult to understand and accept in these less-funded disciplines. It can still happen in the humanities to pay authors for their submissions, thus the other way round feels rather alien.

The number of items published in Gold OA Journals and indexed in DOAJ is comparatively shown in Figures 7 a-b.

Figures 7 a-b: DOAJ indexed publications (Faculties 1 and 2)



Percentages of publications containing a DOI (Digital Object Identifier)

The results show that almost 60% of the articles affiliated to the University of Vienna indexed in A&HCI have no DOI, while the percentage in SSCI decreases to 30%.

The high percentage of publications without DOI in A&HCI seems not to be solely distinctive for the University of Vienna. Very similar values were found for the Humboldt University of Berlin (64%) and the University of Zurich (53%)[iv].

Discussion & Outlook

The data shows an increase in international visibility in both faculties in the reporting period, indicated by the increase in international (non-German) publications, and the percentages of peer-reviewed articles and of indexed articles in A&HCI. In the view of the university, the instrument of target agreements proved to be a valuable tool to foster this positive trend especially by raising awareness for international visibility and by offering a forum for critical reflection of publication strategy. The surprisingly low ERIH coverage implicates, that despite the efforts of the university management made in the past to promote this project, it did not find great acceptance among the university's scientists. Although the scope of this case study is limited to one university, publications indexed in A&HCI seem much more promising to increase overall visibility in the humanities. However, despite the positive results of this case study, it is obvious that further strategies are needed to increase international visibility of publications in the humanities. Certainly the researchers need further encouragement to opt for peer-reviewed publication channels as a major quality criterion. However, peer-review in the humanities itself needs to be strengthened, expanded and adjusted to the desiderata of new communication habits. Furthermore, in order to increase visibility non-English publications should at least always provide title, abstract and keywords information in English for international indexing purposes.

Living in the digital era, scientists in the humanities should embrace digital media whenever they are available. Online presence can be augmented by means of permanent identifiers like Open Researcher and Contributor ID (ORCID) on researcher level or DOI on single publication level.
When switching to online, the humanities should simultaneously aim for OA. Existing traditional journals should gradually be transferred to the new publication model, whereas the launch of new OA journals should be supported, provided that they meet certain quality criteria (peer review) and can therefore be indexed in prestigious databases.
But even the non-existence of appropriate OA journals is no obstacle for making one's research openly accessible, since self-archiving in institutional and/or disciplinary repositories can easily complement publication in a traditional journal.

Since research in the humanities often targets society, researchers are well advised to become more internet-savvy and take advantage of social media and tools like Wikipedia, Google Scholar Citations, Academia.edu, etc., which are often customizable and allow setting up individual profiles.

This case study gives a rough idea of the slow uptake of digital humanities, which is probably not only a local but a global phenomenon. Other studies performed in other countries as in Flanders and Norway arrived at similar conclusions and show only a slow change of publication patterns into the direction of internationalisation (Ossenblok et al, 2012).

Our quantitative analysis is currently complemented by semi-structured interviews of researchers. All gained insight will hopefully find its way into future target agreements with more specific requirements and recommendations.

References

Australian Research Council. (2012) The Excellence in Research for Australia (ERA) Initiative, http://www.arc.gov.au/era/ accessed 11 March 2014

European Science Foundation. (2011) European Reference Index for the Humanities (ERIH), http://www.nsd.uib.no/nsd/english/erih_announcement.html, accessed 11 March 2014

Gorraiz, J., Gumpenberger, C., Schlögl, C. (2014) Usage Versus Citation Behaviours in Four Subject Areas. *Scientometrics*, in Press.

Gorraiz, J., Purnell, P.J. & Glänzel, W. (2013) Opportunities for and limitations of the book citation index. *Journal of the American Society for Information Science and Technology*, 64 (7),1388-1398

Gorraiz, J., & Gumpenberger, C. (2010). Going beyond citations: SERUM – a new tool provided by a network of libraries. *Liber Quarterly*, 20, 80–93

Hammarfelt, B. (2014). Using altmetrics for assessing research impact in the humanities. Scientometrics (in press)

Hug, S., Ochsner, M. and 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, 369–383

Kousha, K., & Thelwall, M. (2009). Google book search: Citation analysis for social science and the humanities. *Journal of the American Society for Information Science and Technology*, 60(8), 1537-1549

Kurtz, M. J., & Bollen, J. (2010). Usage bibliometrics. *Annual review of information science and technology*, 44(1), 1-64.

Leydesdorff, L., Hammarfelt, B., & Salah, A. (2011). The structure of the Arts & Humanities Citation Index: A mapping on the basis of aggregated citations among 1,157 journals. *Journal of the* American *Society for Information Science and Technology*, 62(12), 2414-2426

Li, X., Linmans, J. A. M. (2010). Why with bibliometrics the humanities does not need to be the weakest link. Indicators for research evaluation based on citations, library holdings and productivity measures. Scientometrics, 83(2), 337-354

Nederhof, A. J. (2006). Bibliometric monitoring of research performance in the social sciences and the humanities: A review. Scientometrics, 66(1), 81-100

Ossenblok, T.L.B.; Engels, T.C.E.; 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

Price, D. J. de Solla (1963), Little science, big science, Columbia Univ. Pr., New York.

Priem, J., Piwowar, H. A., & Hemminger, B. M. (2012). Altmetrics in the wild: Using social media to explore scholarly impact. *arXiv preprint*:1203.4745.

Tang, M., Wang, C., Chen, K. & Hsiang, J. (2012). Exploring alternative cyber bibliometrics for evaluation of scholarly performance in the social sciences and humanities in Taiwan. *Proceedings of the ASIS&T Annual Meeting*, Vol. 49. Retrieved, December 12, 2012, from www.asis.org/asist2012/proceedings/openpage.html

van Leeuwen, TN; Moed, HF; Tijssen, RJW; Visser, MS; van Raan, AFJ (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; DOI: 10.1023/A:1010549719484

van Leeuwen, TN. (2013). Bibliometric research evaluations, Web of Science and the Social Sciences and Humanities: a problematic relationship? *Bibliometrie-Praxis & Forschung, 2, http://www.bibliometrie-pf.de/article/viewFile/173/218*

Wouters, P. & Costas, R. (2012). Users, narcissism and control – Tracking the impact of scholarly publications in the 21st century. SURF-foundation. Utrecht.

[i] In November 2014 RAD was replaced by u:cris based on Pure. While the RAD information is not online, data from u:cris will be available online in the future. RAD journal information (index, peer review) was supplemented with data from the Thomson Reuters Master journal list (http://ip-science.thomsonreuters.com/mjl/) and data from ulrichsweb (https://ulrichsweb.serialssolutions.com/).

[ii] Information is taken from RAD and only includes journal articles and other items explicitly published in peer-reviewed journals. The percentage of total publications with peer review also includes the journal articles with peer review.

[iii] Due to the lack of available information in RAD this analysis was done in A&HCI and SSCI for journal articles only clearly affiliated to the University of Vienna.

[iv] Further analyses at macro and meso-levels will be conducted in this regard.

Pruning cooccurrence networks

Raf Guns

raf.guns@uantwerpen.be
University of Antwerp, Institute of Education and Information Sciences, Venusstraat 35,
2000 Antwerpen (Belgium)

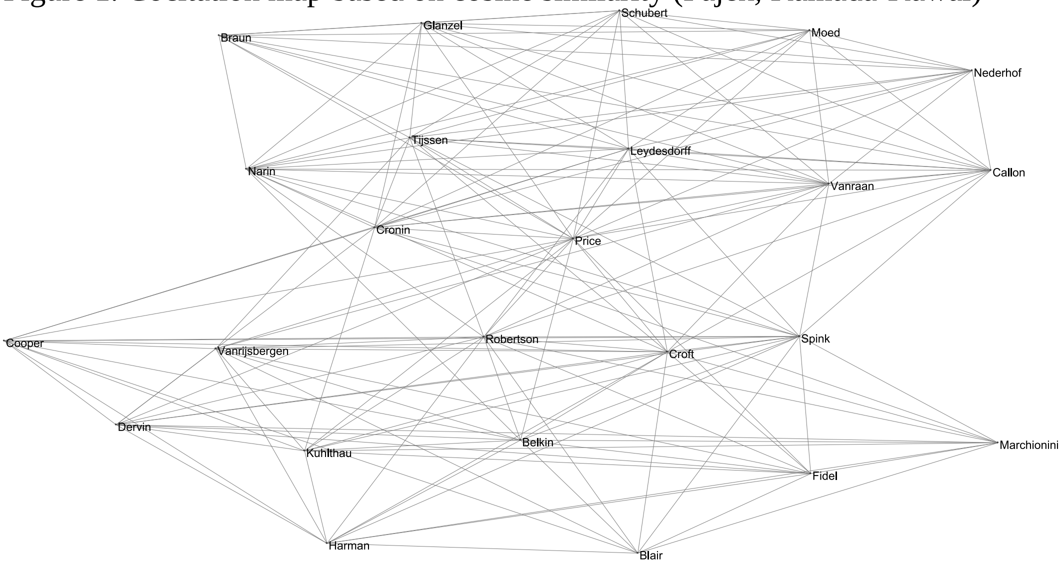
Abstract
Cooccurrence (e.g., cocitation) networks tend to be dense, which renders them hard to visualize and interpret. This paper presents a new method for pruning cooccurrence networks. Every coocurrence network is derived from a two-mode network (e.g., authors and citing papers). Starting from this two-mode network, we apply Markov Chain Monte Carlo sampling to determine the statistical significance of each link and only retain links between nodes (e.g., authors) with probability $p < 0.001$.
This procedure accounts for large variations in the degrees of both top and bottom nodes, which is not the case for other pruning techniques. The feasibility and usefulness of the method is illustrated on two empirical examples.

Introduction
Bibliometric research often employs networks to study the relations between authors, journals or other entities. Many bibliometric indicators are most easily interpreted in a network context (e.g., González-Pereira et al., 2010). Moreover, networks form the input to different approaches to bibliometric mapping, whereby entities are positioned to represent their relative proximities, as reflected in publications and citations.

Networks can range from sparse (few links) to dense (many links). Dense networks are hard to visualize, interpret and work with. The author cocitation network in Figure 1 provides an example. The density of this author cocitation network is 0.57. It is unclear by visual inspection which links are most important and whether the network consists of any cohesive subgroups. This kind of problem has led researchers to prune dense networks, that is, to discard less important links and retain only the most important ones. The current paper presents a new method to prune cooccurrence networks.

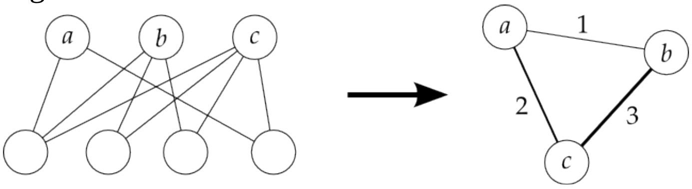
Different pruning techniques have been described in the literature. The Pathfinder algorithm (Schvaneveldt, 1990) prunes a network by eliminating links that are not needed to preserve shortest paths between pairs of nodes. De Nooy, Mrvar, and Batagelj (2005) discuss deleting low-weight links “to obtain a clear picture.” Persson (2010) proposes to prune article citation networks by only retaining those links between articles that are simultaneously cocited and bibliographically coupled. In several studies, Loet Leydesdorff (e.g., Leydesdorff, 2007; Leydesdorff et al., 2013) prunes cocitation or bibliographic coupling links between nodes whose cosine values are below a certain threshold (often 0.2). Egghe and Leydesdorff (2009) introduce an approach to determine a threshold value for the cosine measure, such that the corresponding Pearson correlation coefficient is guaranteed to be positive.

Figure 1. Cocitation map based on cosine similarity (Pajek, Kamada-Kawai)



A kind of network that is especially prone to the density problem are cooccurrence networks, such as cocitation, coauthorship, or bibliographic coupling networks. Cooccurrence networks are derived from two-mode (bipartite) networks (Figure 2), which explains why they are more likely to be dense. For instance, if one article cites 20 authors, then all 20 authors are linked in the corresponding author cocitation network, leading to 190 links. This article presents a new pruning method that aims to retain only the important relations in a cooccurrence network while discarding the trivial ones. The method does not apply to non-cooccurrence networks, such as citation networks.

Figure 2. Cooccurrence network derived from a two-mode network



Formally, cooccurrence relations can be interpreted as two-mode (bipartite) networks with two sets of nodes: the top nodes represent the entities of interest (e.g., words, authors or journals) and can only be linked to bottom nodes (usually, but not necessarily, articles). Two top nodes co-occur (are indirectly connected) if they have one or more neighbouring bottom nodes in common. Another way to interpret a two-mode network is as a matrix $\mathbf{O}$ with dimensions $n \times m$ (n top nodes, m bottom nodes). A cell in $\mathbf{O}$ is 1 if the top node is related to the bottom node (e.g., if the keyword is used in the article) and 0 otherwise. One can obtain a cooccurrence matrix by multiplying the matrix $\mathbf{O}$ with its transpose: $\mathbf{C} = \mathbf{O} \times \mathbf{O}^T$. Values in the cooccurrence matrix indicate the number of times two nodes (e.g., two keywords) occur together. For purposes of bibliometric mapping, one typically does not work with the 'raw' cooccurrence matrix $\mathbf{C}$ but instead applies a similarity measure, such as the cosine measure (Ahlgren, Jarneving & Rousseau, 2003), to row vectors in $\mathbf{O}$. The matrix of all pairwise similarities is a normalized cooccurrence matrix.

We propose another approach, that starts from the two-mode network and only retains node pairs whose cooccurrence can be interpreted as statistically significant (Zweig & Kaufmann, 2011). This typically results in less dense networks, whose structure is easier to interpret and better reflects reality.

Methods

Our method is based on the work of Zweig and Kaufmann (2011), who propose a systematic approach to the projection of bipartite networks. In the remainder of this section, we will use the terminology of author cocitation analysis. That is, top nodes (entities under study) are authors, who are connected with a bottom node (citing article) if the latter cites work by the former. Note, however, that the method is general and can equally well be applied to bibliographic coupling between journals, cowords between articles etc.

We have a bipartite network $G = (A \cup C, E)$, where the node set is the union of the set of authors $A = \{a_1, .., a_t\}$ and the set of citations $C = \{c_1, \ldots, c_b\}$, and the link set is $E \subseteq A \times C$. We denote the set of neighbours of node x by $nbr(x)$. Each node in A and C has a certain degree: $\deg(x) = |nbr(x)|$. Hence, we can determine the degree sequence $T = [\deg(a_1), \deg(a_2), \ldots, \deg(a_t)]$ of A. Likewise, $B = [\deg(c_1), \deg(c_2), \ldots, \deg(c_b)]$ denotes the degree sequence of C.

Essentially, the method can be summarized as follows:

1. First, we define a pattern or motif (Milo et al., 2002) of interest. In our case, this is the cooccurrence of two authors. We denote $coocc_G(a_i, a_j) = |nbr(a_i) \cap nbr(a_j)|$.
2. Next, we determine the *interestingness* of each cooccurrence. Zweig & Kaufmann (2011) discuss several measures of interestingness. We propose using the z-score:
$$z(x, y) = \frac{Obs - Exp}{\sigma} \tag{1}$$
Here, Obs denotes the observed number of times a motif is found ($Obs = coocc_G(x, y)$), Exp denotes the expected number of cooccurrences and σ is the standard deviation; these are discussed below. If a cooccurrence is found significantly more than expected, it is considered *interesting*.
3. Finally, the resulting projection is constructed by only linking those authors whose cooccurrence is considered interesting.

Before we move on, we want to make it more explicit why interestingness is important. Why can we not, for instance, simply single out those author pairs whose cocitation strength is high? Consider the extreme case where an author is cited by *all* citing articles (e.g., because they are an authority in a certain field). Consequently, this author has many citations in common with all other authors, although the other authors are likely to be more specialized on specific problems in the field. Slightly less extremely, if an author is cited many times, we could say that their high value of cooccurrence with some other authors is a natural consequence of their high degree. In other words, a high cooccurrence value is statistically plausible and, hence, less interesting. Similar considerations also apply to the other side. For instance, if a citing article refers to most or all authors under consideration, the resulting cooccurrences carry little or no meaning. We emphasize that these considerations apply to links (e.g., cocitations) and not nodes (e.g., authors). Clearly a highly cited author is important in his/her own right.

The main challenge then is determining Exp. The expected number of cooccurrences for a given node pair can be interpreted as the mean number of cooccurrences over all possible networks that maintain certain structural aspects of G. The set $\mathcal{G}(G)$ of all networks that maintain certain aspects of G is a *network model*. Perhaps the most famous model is the $\mathcal{G}(n, m)$ model by Erdős and Rényi (1959), the set of all networks with n nodes and m links.

Since we are dealing with two-mode networks, we obviously need a network model that only allows links between the t author nodes on the one hand and the b citing article nodes on the other. Even so, there are still many variations one can choose from. Zweig and Kaufmann (2011) provide compelling theoretical arguments that a simple model in which each citing article has an equal probability of citing a given author, is inadequate for non-artificial networks. Instead, they propose using the Fixed Degree Sequences Model (FDSM). FDSM is a network model $\mathcal{G}(T, B)$ for bipartite networks that keeps the degree sequences T and B fixed. In other words, $\mathcal{G}(T, B)$ is the set of all bipartite networks $G' = (A \cup C, E')$ with degree sequences T and B.

There currently exists no closed-form expression of Exp in the FDSM. Moreover, ignoring networks of trivial size, the number of networks in $\mathcal{G}(T, B)$ is too large to consider all of them. Hence, we can only estimate Exp by sampling from $\mathcal{G}(T, B)$. Averaging over all observed cooccurrences leads to an (approximation of) expected cooccurrence. From the same data – i.e., a large number of observed cooccurrences in different networks belonging to $\mathcal{G}(T, B)$ – we can also obtain the standard deviation σ.

We use Markov Chain Monte Carlo sampling to obtain networks from $\mathcal{G}(T, B)$. Starting from the observed network G, we randomly choose two links (a_i, c_x) and (a_j, c_y). We then swap the two links, such that we obtain (a_i, c_y) and (a_j, c_x) (unless either of these already exist). Repeating this swapping procedure enough times results in a bipartite network with the same degree sequences, but independent from G. To ensure independence from the starting network G, we 'chain' the sampled networks. That is, from G we obtain G'; from G' we obtain G''; and so on. The expected number of cooccurrences between two authors then is the mean of the observed number of cooccurrences in all samples.

We only consider those cooccurrences with probability $p < 0.001$, i.e., whose z-score $z > 3.29$. These are retained as links in the resulting network. One can use the z-scores as link weight to further distinguish specifically interesting interactions. Note that we may also have negative correlations ($z < -3.29$).

A small experimental result: author cocitation

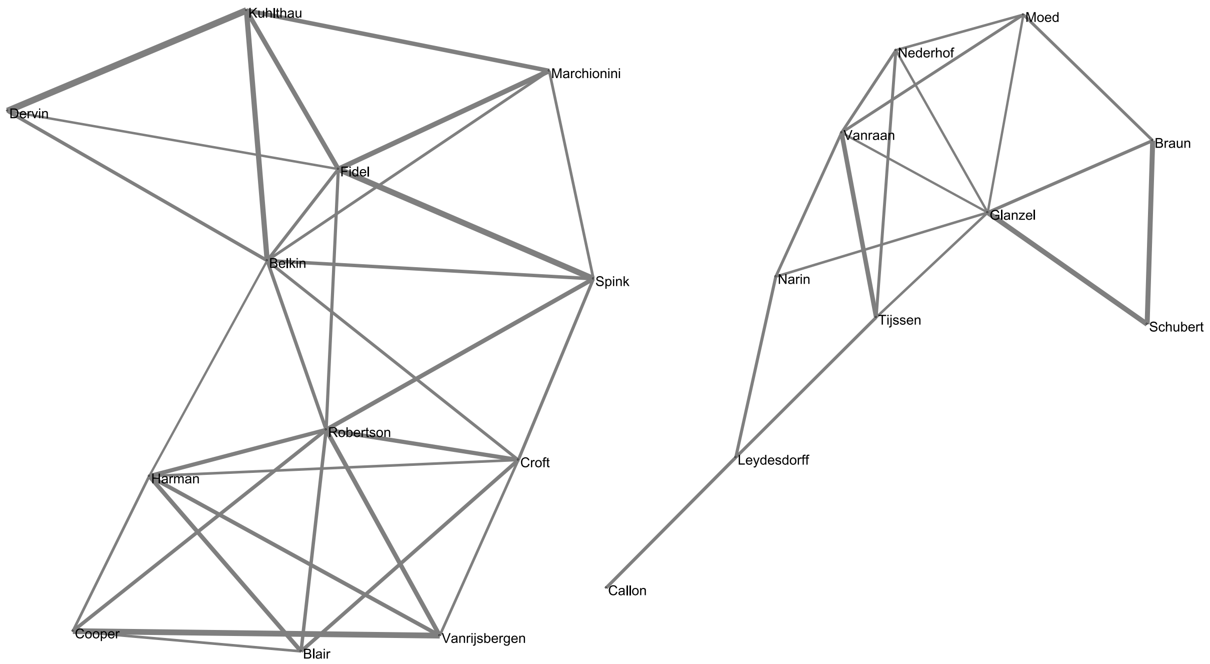
We use a small dataset that has been explored in previous studies (Ahlgren, Jarneving & Rousseau, 2003; Egghe & Leydesdorff, 2009; Leydesdorff & Vaughan, 2006). This dataset consists of all publications in *Scientometrics* and *Journal of the American Society for Information Science* (*JASIS*) in the period 1996–2000. We find 498 publications in *Scientometrics* and 494 in *JASIS*. Due to changes in newer versions of Web of Science, the number of publications in *Scientometrics* is higher than the 469 publications reported by Leydesdorff & Vaughan (2006). Just like the previous studies, we focus on author (co-)citations to 12 authors from information retrieval and 12 authors from bibliometrics (see Table 1). There are 471 articles that cite at least one of these authors.

Table 1. Authors in cocitation analysis

Field	Name
Bibliometrics	Braun T, Callon M, Cronin B, Glanzel W, Leydesdorff L, Moed HF, Narin F, Nederhof AJ, Price DJD, Schubert A, Tijssen RJW, Vanraan AFJ
Information retrieval	Belkin NJ, Blair DC, Cooper WS, Croft WB, Fidel R, Harman DK, Kuhlthau CC, Marchionini G, Robertson SE, Spink A, Vanrijsbergen CJ

Figure 1 shows the cocitation network for these 24 authors. Link weights are normalized using the cosine measure. We can see that all bibliometric authors are placed on top, whereas all IR authors are situated in the bottom half, although the exact boundary is unclear. Likewise, it is unclear which links are most important. Now we compare this with FDSM.

Figure 3. Cocitation map based on FDSM and z-scores (Pajek, Kamada-Kawai)



Using the approach described in the previous section, we randomly sampled 5000 networks from the set of all bipartite networks with the same top and bottom degree sequences. Each new network was obtained by performing 3000 link swaps. For each pair of cocited authors the z-score was determined and only those pairs with $z > 3.29$ (31% of the original number of links) were retained as links. This method yields the network shown in Figure 3. The bibliometric and IR authors are now clearly separated, as two separate components. Moreover, because the number of links is much lower, the structure of each component emerges more clearly. It seems, for instance, that the top half of the IR component contains researchers in information seeking, whereas the bottom half contains researchers in 'hard' system-oriented IR. A possible disadvantage of the FDSM approach is that it can easily cause some nodes to become isolates. This is the case for Cronin and Price (not shown in Figure 3), although these authors appear to occupy a central position in

Figure 1. A possible explanation is that they are regularly cocited with both bibliometric and IR authors and therefore occupy a less clear position with either bibliometricians or IR researchers.

Interestingly, Figure 3 closely resembles a map obtained by Egghe and Leydesdorff (2006) on the basis of the cosine measure (only retaining those links whose corresponding Pearson correlation coefficient cannot be negative). More generally, we find that ranking author pairs by their z-score according to the FDSM procedure has almost the same result as ranking by their cosine similarity (Spearman rank correlation $r = 0.96$).
In summary, we obtain similar results with cosine normalization and FDSM for this particular case.

A larger application: bibcoupling of *JASIST* articles (2009–2010)
Now we consider a larger empirical case. We study bibliographic coupling of all 371 articles published in *JASIST* in the years 2009–2010, in order to obtain a map that visualizes the intellectual structure of the journal. We construct a two-mode network, wherein papers are linked to their references. In total, 12 981 unique references were found. The bibliographic coupling map based on cosine normalization is shown in Figure 5. Map of JASIST articles (2009–2010, bibliographic coupling, FDSM). The corresponding map based on FDSM is shown in Figure 5.
Both maps were created using VOSviewer (Van Eck & Waltman, 2010). Major clusters in each map were manually labelled. While the two maps roughly exhibit a similar structure, the FDSM map is more spread out. Specifically, while the topic of information behaviour forms a fairly coherent cluster in the cosine map, the FDSM map shows that this group actually consists of two separate clusters, pertaining to online communities like Wikipedia on the one hand and to information behaviour on the Web on the other. In our opinion, the FDSM map makes the existence and relative importance of these subtopics much clearer.

Figure 4. Map of JASIST articles (2009–2010, bibliographic coupling, cosine)

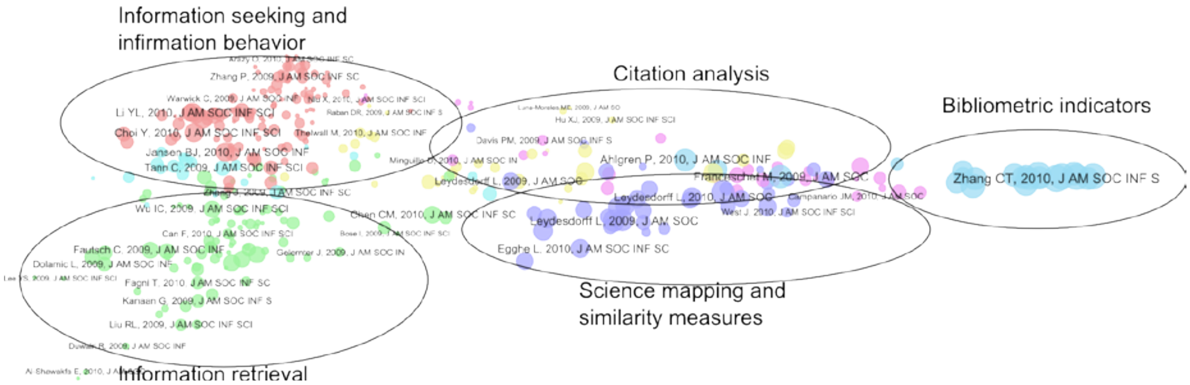
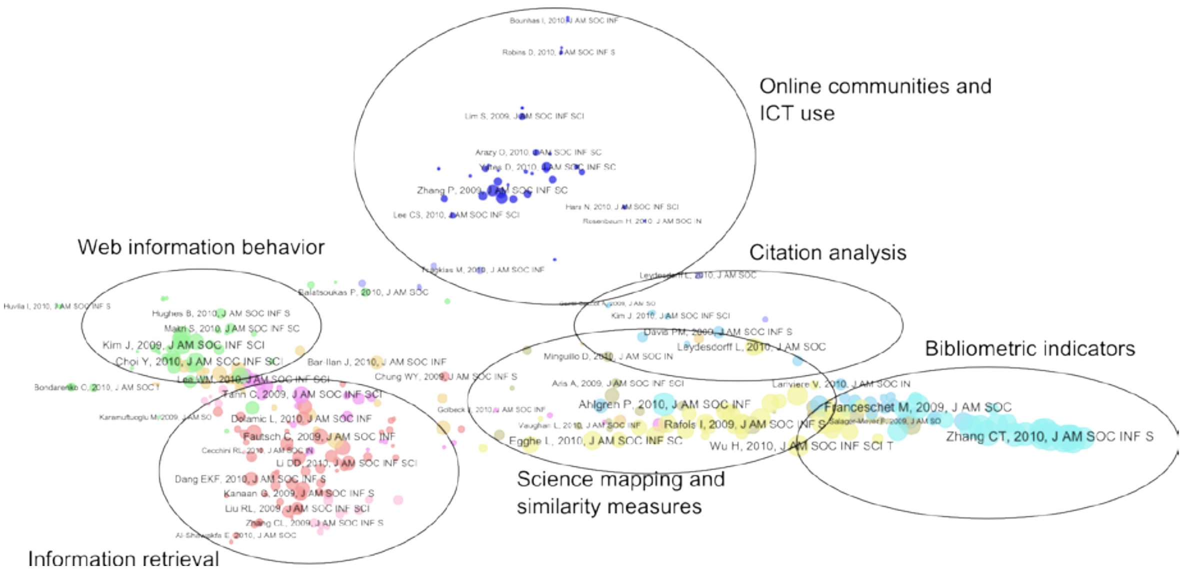


Figure 5. Map of *JASIST* articles (2009–2010, bibliographic coupling, FDSM)



Comparison of FDSM and cosine similarity

The examples in the preceding sections illustrate that FDSM and cosine similarity may result in similar mappings. Does this mean that the two are interchangeable? In our opinion, there are three important differences between FDSM and cosine normalization.

First, contrary to cosine normalization, FDSM reveals positively significant cooccurrences as well as negatively significant ones. Most current mapping techniques only consider positive edges, but in principle the additional information of negative edges could be exploited to obtain better and more accurate maps. Second, because we work with z-scores to determine interestingness, FDSM allows us to establish a threshold that corresponds to a specific *p*-value. Third, the cosine formula ignores size differences between bottom nodes (citing articles in the ACA case). In other words, while cosine does account for large variations in the degree sequence of the top (projected) nodes, it cannot account for similar variations in the degree sequence of bottom nodes. Because variations in our ACA example are fairly small (for both top and bottom nodes), FDSM and cosine yield very similar results.

We now give a hypothetical example to highlight the difference between FDSM and cosine. Assume that we have the following matrix whose six rows denote authors and whose four columns denote citing articles.

$$\begin{array}{ccccc} & c_1 & c_2 & c_3 & c_4 \\ a_1 & 0 & 1 & 1 & 0 \\ a_2 & 1 & 0 & 1 & 0 \\ a_3 & 1 & 1 & 0 & 0 \\ a_4 & 0 & 1 & 1 & 1 \\ a_5 & 0 & 1 & 1 & 1 \\ a_6 & 0 & 1 & 1 & 1 \end{array}$$

We will focus on the author pairs (a_1, a_2) and (a_2, a_3). It is easy to see that both author pairs have the same cosine similarity:

$$\cos(a_1, a_2) = \cos(a_2, a_3) = \frac{1}{\sqrt{2}\sqrt{2}} = \frac{1}{2}$$

Authors a_1 and a_2 are cocited by c_3, whereas a_2 and a_3 are cocited by c_1. The probability of two authors being cocited by these articles is quite different: for c_3 it is 2/3, whereas the probability of being cocited by c_1 is only 1/15. This suggests that the cocitation of a_2 and a_3 is more interesting. Because the cosine formula is based on vectors rather than full matrices, it cannot take this difference into account.

Discussion and conclusions

Cooccurrence networks are frequently used in informetric research, but their density may make them more difficult to use and interpret. We have introduced a new method for pruning cooccurrence networks, which is based on sampling from all two-mode networks with the same degree sequences. This procedure accounts for large variations in the degrees of both top and bottom nodes, which is not the case for other pruning techniques.

The main limitations of the method are the fact that it is computationally intensive and may result in a larger amount of isolate nodes compared to other methods. Future research should try to alleviate these concerns and explore the method's potential in other settings. Finally we note that the method may also yield negative relations; these may augment existingmethods for, for instance, visualization.

References

Ahlgren, P., Jarneving, B., & Rousseau, R. (2003). Requirements for a cocitation similarity measure, with special reference to Pearson's correlation coefficient. *Journal of the American Society for Information Science and Technology*, 54(6), 550–560.

Åström, F. (2007). Changes in the LIS research front: Time-sliced cocitation analyses of LIS journal articles, 1990–2004. *Journal of the American Society for Information Science and Technology*, 58(7), 947–957.

De Nooy, W., Mrvar, A., & Batagelj, V. (2005). *Exploratory social network analysis with Pajek*. Cambridge: University Press.

Egghe, L., & Leydesdorff, L. (2009). The relation between Pearson's correlation coefficient *r* and Salton's cosine measure. *Journal of the American Society for Information Science and Technology*, 60(5), 1027–1036.

Erdős, P., & Rényi, A. (1959). On random graphs, I. *Publicationes Mathematicae (Debrecen)*, 6, 290–297.

González-Pereira, B., Guerrero-Bote, V. P., & Moya-Anegón, F. (2010). A new approach to the metric of journals' scientific prestige: The SJR indicator. Journal of Informetrics, 4 (3), 379-391.

Leydesdorff, L. (2007). Visualization of the citation impact environment of scientific journals: an online mapping exercise. *Journal of the American Society for Information Science and Technology*, 58 (1), 25-38.

Leydesdorff, L., Rafols, I. and Chen, C. (2013). Interactive overlays of journals and the measurement of interdisciplinarity on the basis of aggregated journal-journal citations, *Journal of the American Society of Information Science and Technology*, 64(12), 2573-2586.

Leydesdorff, L., & Vaughan, L. (2006). Co-occurrence matrices and their applications in information science: Extending ACA to the Web environment. *Journal of the American Society for Information Science*, 57(12), 1616–1628.

Milo, R., Shen-Orr, S., Itzkovitz, S., Kashtan, N., Chklovskii, D., and Alon, U. (2002). Network motifs: Simple building blocks of complex networks. *Science*, 298(5594), 824–827.

Persson, O. (2010). Identifying research themes with weighted direct citation links. *Journal of Informetrics*, 4(3), 415–422.

Schvaneveldt, R.W. (1990). *Pathfinder associative networks: Studies in knowledge organization*. Westport: Ablex Publishing.

Van Eck, N.J., & Waltman, L. (2010). Software survey: VOSviewer, a computer program for bibliometric mapping. *Scientometrics*, 84(2), 523–538.

Zweig, K., & Kaufmann, M. (2011). A systematic approach to the one-mode projection of bipartite graphs. *Social Network Analysis and Mining*, 1(3), 187–218.

International scientific collaboration index: an analysis of Brazilian science (1980-2009)

Renata Cristina Gutierres Castanha*, Carla Mara Hilário** and Maria Cláudia Cabrini Grácio***

*regutierres@gmail.com,**carla.hilario@hotmail.com, ***cabrini@marilia.unesp.br
UNESP – Univ Estadual Paulista, 737 Hygino Muzzi Filho Avenue, 17525-900 Marília (Brazil)

Introduction

Rapidly growing since 1988, Brazilian science has consolidated as an important scientific community in the last decades and has led Brazil to stand out in the mainstream science (Leta, Glänzel & Thijs, 2006). In this scenario, the networks of scientific collaborations are highlighted, whether at the individual, institutional or country level, consolidated by the technological developments (Glänzel & Moed, 2002). For Katz & Martin (1997, p.7), scientific collaboration is "[...] the working together of researchers to achieve the common goal of producing new scientific knowledge." It is one kind of social network, an activity that allows more favorable conditions for scientific production as it enables knowledge sharing, "optimizes" resources and enhances the possibilities of approaches and tools to meet the proposed objective (Olmeda Gómez, Perianez-Rodriguez & Ovalle-Perandones, 2008).
According to Glänzel & Moed (2002), the articles published in international co-authorship, in general, have greater impact, because besides the joint efforts of the researchers, the prestige of each of them contributes for these publications to be regarded with greater reliability.
Thereby, the general objective of the study is to diachronically analyze the indicators of scientific collaboration of Brazilian science in the period of three decades (1980-2009).
Specifically, the objective is to analyze the evolution of the international co-authorship index in Brazilian scientific production during the 1980-1989, 1990-1999 and 2000-2009 periods. Moreover, this study aims to identify and group the major collaborating countries to Brazil, evaluate the relative co-authorship index of each of these countries in the context of the Brazilian production and these countries' rank in relation to Brazilian collaborators in order to highlight the contribution of scientific collaboration for the scientific panorama of the country.

Methodological procedures

Articles in Scopus database were retrieved for the 1980-2009 period, using advanced search: AFFILCOUNTRY (BRASIL OR BRAZIL) AND PUBYEAR > 1979 AND PUBYEAR < 1990 AND DOCTYPE (AR), changing the decades in the expression. For each decade, we identified: total number of Brazilian articles, total number of Brazilian articles with international collaboration and the ten major collaborating countries, with the total co-authored article and ranking among the collaborating countries. Then, we calculated the percentage of collaboration of the countries in relation to the total number of Brazilian articles published in each decade. Finally, we performed a hierarchical cluster analysis, using Ward's method, with Euclidean Distance measure, in order to group the countries according to the similarities regarding ranking and percentage of collaboration with Brazil in the three decades.

Presentation and analysis of data

In the 1980-1989 decade, we identified 2,356 (18.9%) Brazilian papers in international scientific collaboration from a total of 12,450 published articles. In the 1990-1999 decade, we identified 16,629 (29.1%) Brazilian articles in international scientific collaboration from a total of 57,094 articles. In the 2000-2009 decade, 52,905 (25.7%) Brazilian papers in international scientific collaboration from a total of 205,877 publications were identified. The data indicate an increase in international research cooperation during the first two decades and then a small decrease from second to third decade, mainly due to the consolidation of graduate studies in Brazil, abroad scholarships were reduced, stimulating decentralization from the national survey that also occurs through collaboration (Faria et al., 2011).

Based on Table 1, we observed that USA holds first place and with collaboration percentage above 7% (mean percentage equal to 9.4%) throughout the period. The three following countries (France, UK and Germany) holds 2nd, 3rd and 4th ranking position throughout the period of 30 years with little variation between UK and Germany in the second decade and percentages above 2% of collaboration, except UK in the first decade.

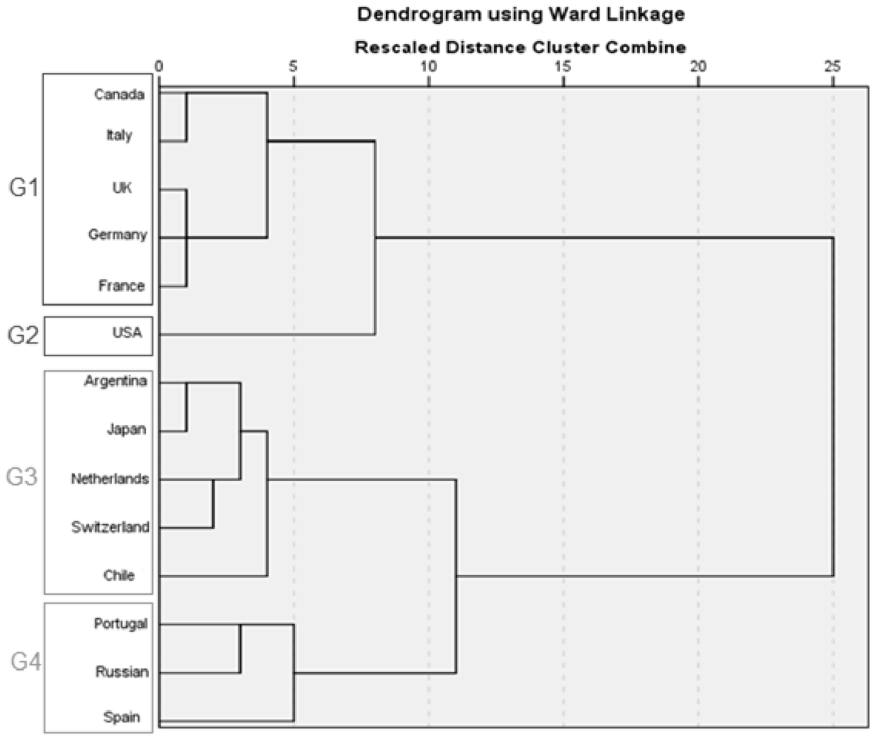
In addition, we observed an increasing scientific collaboration of Spain, Portugal and the Netherlands in the last decade. We also noted a decrease in collaboration from Italy, Argentina, Russia, Japan, Switzerland and Chile, while Canada holds constant position (6th) and percentage between 1.2 and 1.7%.

Table 1. Major Brazilian collaborating countries, rankings and percentage of articles in co-authorship, by decade.

Country	1980-1989		1990-1999		2000-2009	
	Ranking	%	Ranking	%	Ranking	%
USA	1	7.29	1	11.05	1	9.85
France	2	2.43	2	4.21	2	3.21
UK	4	1.65	3	3.72	3	3.05
Germany	3	2.19	4	3.03	4	2.70
Spain	14	0.36	7	1.62	5	1.75
Canada	6	1.22	6	1.74	6	1.71
Italy	5	1.53	5	2.21	7	1.65
Argentina	7	0.86	9	1.33	8	1.55
Portugal	21	0.19	14	0.79	9	1.02
Netherlands	11	0.38	13	0.83	10	0.86
Russian	22	0.16	8	1.42	12	0.78
Japan	8	0.46	10	1.03	11	0.81
Switzerland	9	0.44	11	0.93	15	0.64
Chile	10	0.39	17	0.63	16	0.62

Figure 1 shows four clusters of countries. G1 cluster consists of countries with constant collaboration with Brazil during the full period: between 2nd and 6th position and collaboration percentage between 1% and 5%. G2 cluster consists of USA alone, which presents distinct and prominent behavior in relation to all other collaborating countries: 1st position throughout the period and percentage always above 7%. G3 is composed by countries with less significant positions: positions below 7th and general percentage between 0.3% and 1%. G4 consists of countries with ascending collaboration behavior.

Figure 1. Clusters of Brazilian co-author countries.



Final considerations

International collaboration in Brazil's science grew significantly, especially between the first and second analyzed decades. We observed four clusters of countries according to their similarities in relation to cooperation with Brazilian research. We noticed that major international scientific powers have contributed to the consolidation of the scientific scenario in Brazil, and that these relations have been strengthened since 1980.

References

Faria, L. I. L.; Gregolin, J. A. R.; Hoffmann, W. A. M.; Quoniam, L. (2011). Análise da produção científica a partir de publicações em periódicos especializados. In: Brentani, R. R.; Cruz, C. H. B.; Suzigan, W.; Furtado, J. E. M. P.; Garcia, R. C. (Org.) *Indicadores de Ciência, Tecnologia e Inovação em São Paulo 2010*, 1, 1-71.

Glänzel, W. & Moed, H. F. (2002). Journal impact measures in bibliometric research. *Scientometrics*, 53, 171-193.

Katz, J. S. & Martin, B. R. (1997). What is research collaboration? *Research Policy*, 26, 1-18.

Leta, J., Glänzel, W. & Thjis, B. (2006). Science in Brazil. Part 2: sectorial and institutional research profiles. *Scientometrics*, 67, 87–105.

Olmeda Gómez, C., Perianez-Rodriguez, A. & Ovalle-Perandones, M. A. (2008). Estructura de las redes de colaboración científica entre las universidades españolas. In *Ibersid 2008*: revista de sistemas de información e comunicación, pp.129-140.

A Comparison of Three Prominent Journal Metrics with Expert Judgement of Journal Quality

Peter Haddawy* and Saeed-Ul Hassan**

* *peter.had@mahidol.ac.th*
Faculty of ICT, Mahidol University, Phuttamonthon 4 Road, Nakhon Pathom, 73170 (Thailand)

** *saeed-ul-hassan@itu.edu.pk*
Scientometrics Lab, Information Technology University - Punjab, Ferozepur Road, Lahore, 54770 (Pakistan)

Introduction

In response to the demand for accurate measures of journal impact, quality, and prestige, numerous refinements of the traditional Journal Impact Factor (JIF) have been developed. Two prominent alternatives are the Source Normalized Impact per Paper (SNIP) and the SCImago Journal Rank (SJR) (Colledge et. al., 2010). SNIP is similar to JIF but corrects for differences in topicality between subject fields (Moed, 2010). It is a ratio of a journal's citation impact and the citation potential of its subject field. A journal's subject field is defined as the collection of articles citing the journal. SJR, inspired by Google's PageRank algorithm, is intended as a measure of a journal's prestige. It recursively assigns higher weight to citations from journals that are highly cited (González-Pereira et. al., 2010). Both SNIP and SJR use citation windows of 3 years, while JIF uses a citation window of 2 years.

Arguments for the appropriateness of SNIP and SJR have been made based on the logic underlying their design and studies have been carried out comparing statistical properties of SJR, SNIP, and JIF (Colledge et. al., 2010). But if these metrics are to be used as measures of journal quality, then it is also important to assess the extent to which they agree with human perception of quality. While small scale discipline-specific studies comparing JIF with expert judgement of journal quality have been carried out (Rousseau, 2008), no extensive multi discipline study has yet been carried out comparing alternative journal metrics with expert judgment. Such a study requires a sizable database of journals spanning a broad array of fields, rated by experts in the various fields. Precisely such a rating exercise was carried out by the Australian Research Council as part of its 2010 Excellence in Research for Australia (ERA) initiative. In that exercise journals were assigned to four tiers A*, A, B, C based on the perceived quality of their papers[1]. The process of producing the ranked list of 20,712 journals began in 2007 with a ranking exercise by four Learned Academies and a number of discipline peak bodies and was finalized in the consultation phase in 2010 that involved over 700 expert reviewers[2]. In this paper we study the correlation between the ERA rating and the quantitative journal metrics SJR, SNIP & JIF.

[1] http://www.arc.gov.au/era/tiers_ranking.htm

[2] The use of the ranked journal list was removed from the ERA exercise in 2011 not due to problems with the quality of the exercise but rather because "there is clear and consistent evidence that the rankings were being deployed inappropriately within some quarters of the sector, in ways that could produce harmful outcomes, and based on poor understanding of the actual role of the rankings." Senator Kim Carr, Minister for Innovation, Industry, Science and Research, May 2011.

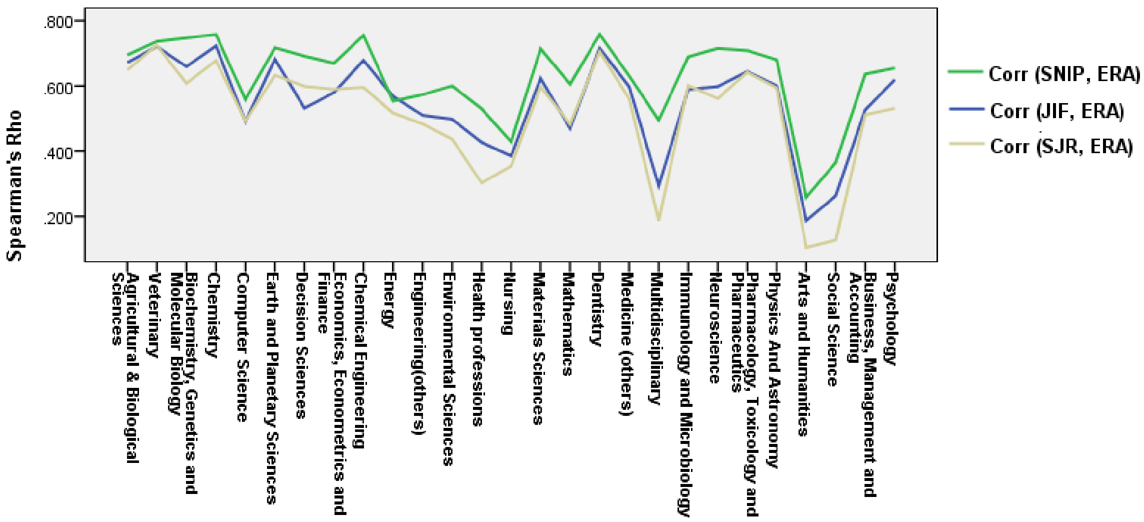
Data Collection & Methods

The 2010 SNIP and SJR metrics were downloaded from www.journalmetrics.com (retrieved 12 January 2013). We computed JIF by applying the definition to the Scopus database. Since SNIP and SJR are defined over Scopus, this controlled for the effect of the database in the comparison of the metrics. JIF for 2010 was computed by taking the ratio of the number of citations in 2010 to citable items in 2008 and 2009 divided by the number of citable items. Citable items are taken to be articles, reviews, proceedings, and notes. We identified those journals in the 2010 ERA list that are indexed in Scopus to produce the list of 11,137 journals for this analysis. We utilized the All Science Journal Classification (ASJC) to group journals for analysis by subject area. We analysed the correlation of JIF, SJR, and SNIP with the ERA rating using the Spearman's coefficient (ρ) overall and in each of 27 subject areas. We used SPSS v. 2.1 to compute the statistics.

Results and Discussion

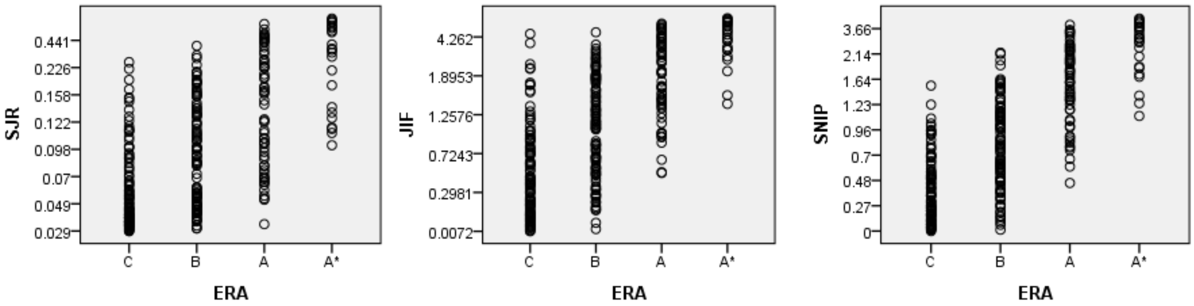
Among the selected metrics, SNIP shows the highest correlation with the ERA rating (ρ = *.537*), followed by JIF (ρ = *.374*) and then SJR (ρ = *.222*). The results are statistically significant at the .000 level, with N=11,137. Figure 1 shows the correlations of the three metrics with the ERA rating broken down by subject area. In every subject area except Energy SNIP has higher correlation than the other two metrics. SNIP has highest correlation in the areas of Dentistry (ρ = *0.758*), Chemistry (ρ = *0.758*), and Chemical Engineering (ρ=0.755). Not surprisingly, the correlation of all three metrics is lowest in the areas of Arts and Humanities, Social Science, and Multidisciplinary.

Figure 1: Correlation by subject area



More insight can be gained by viewing scatter plots of the metrics against the ERA rating. Figure 2 plots the journal metric values of 280 journals indexed under the area Chemical Engineering against the ERA rating. The correlations of the three are relatively high yet differ significantly as well: SNIP (ρ=0.755), JIF (ρ=. 678), SJR (ρ=. 595). All three metrics seem to do a better job at differentiating between A*, A, and B than between B and C while SNIP is the only metric that shows no overlap in values of A* and C journals.

Figure 2: Scatterplots of the three metrics versus the ERA rating in Chemical Engineering



Concluding Remarks

Among the three metrics, SNIP has the highest correlation with the ERA rating, followed by JIF and then SJR. This is despite the fact that one might expect the judgements of the experts to be influenced by their knowledge of the impact factors of journals. The dominance of the correlation with SNIP may have to do with the fact that the ERA rating is focused on journal quality rather than popularity so that journals could be rated highly even if they are in subfields with low citation rates. SNIP is the only one of the three metrics that normalizes for differences in citation potential across fields and subfields.

Acknowledgement

We thank Henk Moed for his guidance in mapping ERA ranked outlets to the ASJC classification.

References

Colledge, L., Moya-Anegón F., Guerrero-Bote, V., López-Illescas, C., Aisati, M. & Moed, H. (2010). SJR and SNIP: two new journal metrics in Elsevier's Scopus. *Serials: The Journal for the Serials Community*, 23(3), 215-221.

González-Pereira, B., Guerrero-Bote,V., Moya-Anegón, F. (2010). A new approach to the metric of journal’s scientific prestige: The SJR indicator. *Journal of Informetrics*, 4(3), 379–391.

Moed, H.F. (2010). Measuring contextual citation impact of scientific journals. *Journal of Informetrics*, 4(3), 265–277.

Rousseau, S. (2008). Journal evaluation by environmental and resource economists: A survey. *Scientometrics*, 77(2), 223–233.

Usage patterns of scientific journals and their relationship with citations

Gali Halevi *, Henk F. Moed **

*g.halevi@elsevier.com
Informetric Research Group, Elsevier, 360 Park Av. South, New York, NY 10011 (USA)

**h.moed@elsevier.com
Informetric Research Group, Elsevier, Radarweg 29, 1043 NX Amsterdam (The Netherlands)

Introduction

In the past decade, many scientific literature publishers have implemented usage monitoring systems based on data including clickstreams, downloads and views of scholarly publications recorded on an article level, that allow them to capture the number of times articles are downloaded in their PDF or HTML formats. This type of data is not only used by publishers as a way to monitor the usage of their journals but also by libraries who wish to monitor and manage the usage of their collections (Duy & Vaughan, 2006). The growing need for this type of monitoring resulted in the launch of COUNTER (Counting Online Usage of Networked Electronic Resources), an international initiative which aimed to set standards and facilitate the recording and reporting of online usage statistics in a consistent, credible and compatible way. Nowadays, COUNTER is an industry standard, used by most publishers and libraries and allows for downloads data to be analyzed and compared more easily by subscribers and publishers alike. This development could be one of the reasons that research in this area has seen such significant growth.

Research on the relationships between citations and downloads has expanded in various studies attempting to understand the relationship between the two as usage phenomenon and as a way to measure research impact. (e.g., Schloegl and Gorrais, 2011; Gorraiz, Gumpenberger & Schloegl, 2014). Kurtz et al. (2005a; 2005b) published two pioneering papers analyzing usage mainly of the NASA Astrophysics Data System (ADS), and comparing the number of electronic accesses – which they term “reads” – of individual articles in astronomy and astrophysics journals with citation counts.

In their review article published in 2010, Michael Kurtz and Johan Bollen describe “Usage Bibliometrics “as the statistical analysis of how researchers access their technical literature, based on the records that electronic libraries keep of every user transaction (Kurtz & Bollen, 2010). They underline that many “classical”, citation-based measures have direct analogs with usage, and that an important approach to validation of usage statistics is to demonstrate the similarities and differences between citation and usage statistics. An important class of usage statistics is based on the number of times articles from publication archives are downloaded in full text format, denoted as “downloads” below. Kurtz and Bollen claim that “….the relation between usage and citation has not been convincingly established”(p. 23) and that “….direct comparisons over the same set of input documents are rare”(p. 23).

The second author of the current paper published in 2005 an analysis of the statistical relationship between citations and full text article downloads for articles in one particular journal: Tetrahedron Letters, published by Elsevier (Moed, 2005). A main objective of the current paper is to expand the analyses presented in the 2005 article in the following ways:

- Analyze a much larger set of journals covering all domains of science and scholarship.
- Analyze in more detail download patters as a function of time;
- Examine the statistical correlation between downloads and citations both at the level of journals and of individual articles;

A full discussion, interpretations of the new findings and their positioning within the framework of the review article by Kurtz & Bollen (2010) will be given in a full article to be published in a later phase. The base assumption underlying this paper is that a sound statistical analysis of relationship between downloads and citations, and a thorough reflection upon its outcomes, contributes to a better understanding of what both download counts and citation counts measure, or more generally, to more insight into information retrieval, reading, and referencing practices in scientific-scholarly research. It is the very combination of the two types of data that enlarges so to speak the horizon, and provides a perspective in which each of the two types can be positioned. In the quantitative study of research activity and performance, downloads and citations provide complementary data sources. In this article the term “usage” is reserved for the use made of electronic publication archives in the broadest sense, and recorded in the archive’s electronic log files. It includes activities such as downloading in pdf, viewing in html format, browsing through abstracts, and also saving, sharing or annotating documents in reference managers.

Data collection

One of the main challenges of analysing downloads and citations figures lays in the availability and completeness of the data collected. The database used to collect the data, whether citations or downloads, might be incomplete. Thus, for example, downloads collected for Scopus™ covered journals, might not be representative of usage in general, because not all literature searches use Scopus™ as their platform of choice. In addition, Scopus™ citations are biased by incomplete source coverage as complete citations are only available from 1996 forward which is a well-documented limitation of the database. Unlike Scopus, ScienceDirect™ is a very specific source of full text articles which is mostly used to either view or download content. Therefore, usage data is fairly complete in ScienceDirect™.

Downloads vs. citations examined in this paper were aggregated in 3 levels: 1) database (e.g., all ScienceDirect™ articles); 2) journal; 3) individual article. The data was collected in two sets of citations and usage data; one at the level of journals and the second at the level of individual articles.

1. Journal Level Data: the first set of data contained all 20,000 peer-reviewed journals covered in Scopus™. For each journal, citations counts for the years 2004-2010 were aggregated per year and per journal. Download counts were aggregated for all 2,500 journals covered both in Scopus™ and ScienceDirect™, Elsevier full text database per year and per journal.
2. Document Level Data: Citations and counts on a per document basis were collected for all individual document published in 63 ScienceDirect™ journals between 2008 and 2012 covering all domains of science and scholarship. Downloads and citations counts on document level are up to September 2013

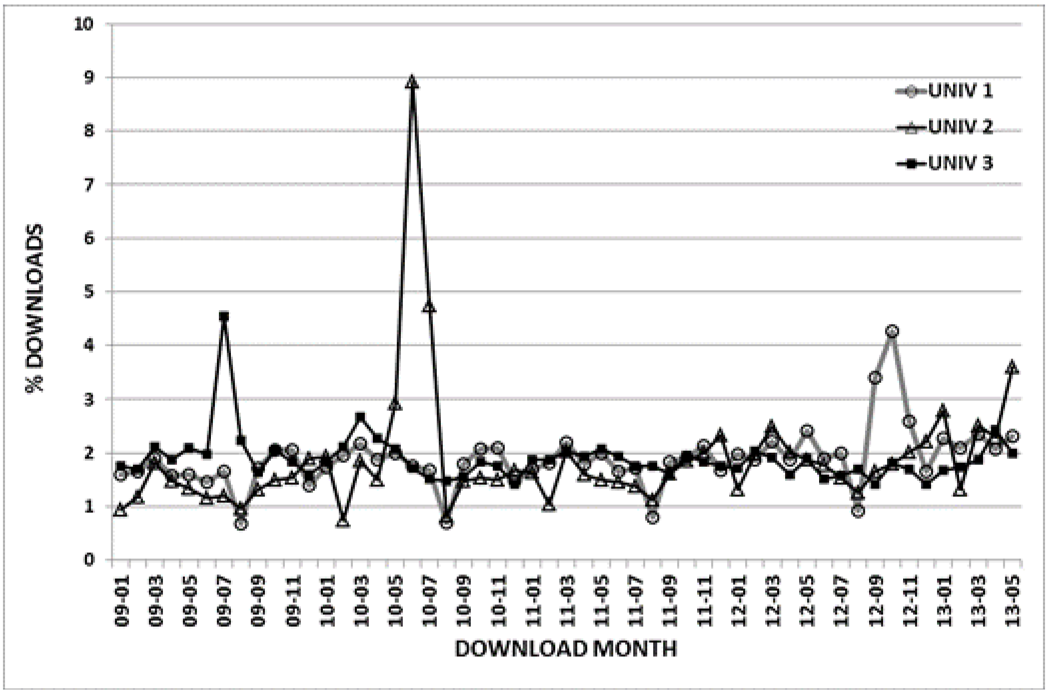
It must be noted that the journals studied are not a random sample from the set of journals in ScienceDirect™. The aim of the selection was to include journals from different disciplines and cover all major disciplines, in order to study differences among disciplines, and also to include journals that were originally sections of one and the same "parent" journal, so that one could even obtain indications of differences within a journal.

Results

Downloads by user institution

Figure 1 presents data on monthly full text downloads from ScienceDirect that users from 3 academic institutions made between January 2008 and May 2013. The data show a clear peaky behaviour. University 1 represented in Figure 1 participated in a national research assessment exercise, in which research staff members could submit full text PDF downloads of their best articles to an evaluation agency for assessment by an expert panel, with a submission deadline in October 2012. For the peaks of Institutions 2 and 3 no explanation is available as of yet. Whether or not these peaks are caused by bulk downloading can be examined by grouping the downloaded articles by user session and by journal volume and issue, and determining the number of downloads per session, journal volume or issue. The three institutions were selected as they provide good illustrations of peaky usage behaviour. In a follow-up study the frequency at which this type of behaviour occurs across all user institutions will be further analyzed.

Figure 1: Longitudinal download counts for three user institutions.

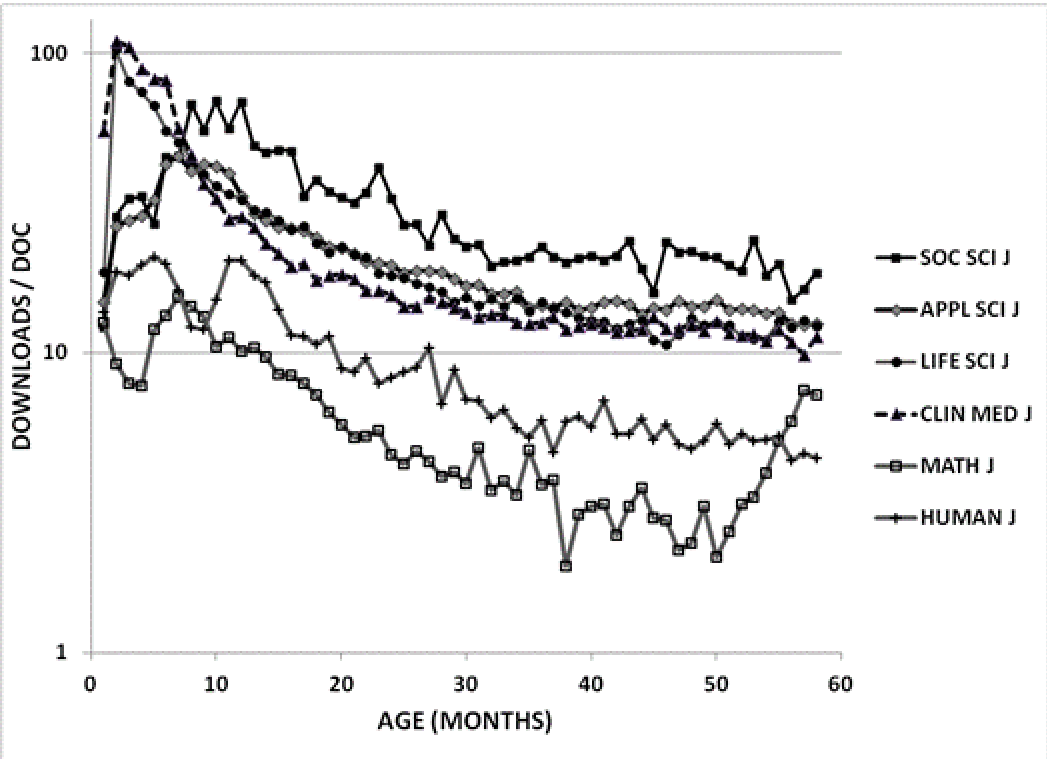


Legend to Figure 1: The vertical axis gives the percentage of downloads in a month, relative to an institution's sum of downloads during the total time period. For University 2 the actual percentage of downloads in July 2010 is 9 %, which is 4.5 times the level one would find if the number of an institution's downloads would be constant over time.

Downloads time series per journal and document type

Figure 2 shows the average number of downloads per full length article for journals in social, applied, life, clinical medicine, mathematics and humanities sciences over time. The overall phenomenon seen in figure 2 is that all journals display peak downloads in the first months following publications, despite the difference in the amount of downloads which varies considerably between journals. Yet, there are differences among the represented journals in the month in which download counts peak. For instance, for the clinical medicine and life sciences journal downloads peak one month after the month in which they were published online, whereas for the applied science and the mathematics journal in the seventh month. Moreover, large differences exist in the decline rates in the various journals. These decline rates themselves tend to decline as the documents grow older. This is consistent with the two-factor models explored by Moed (2005), and the four-factor models explored by Kurtz et al. (2005b).

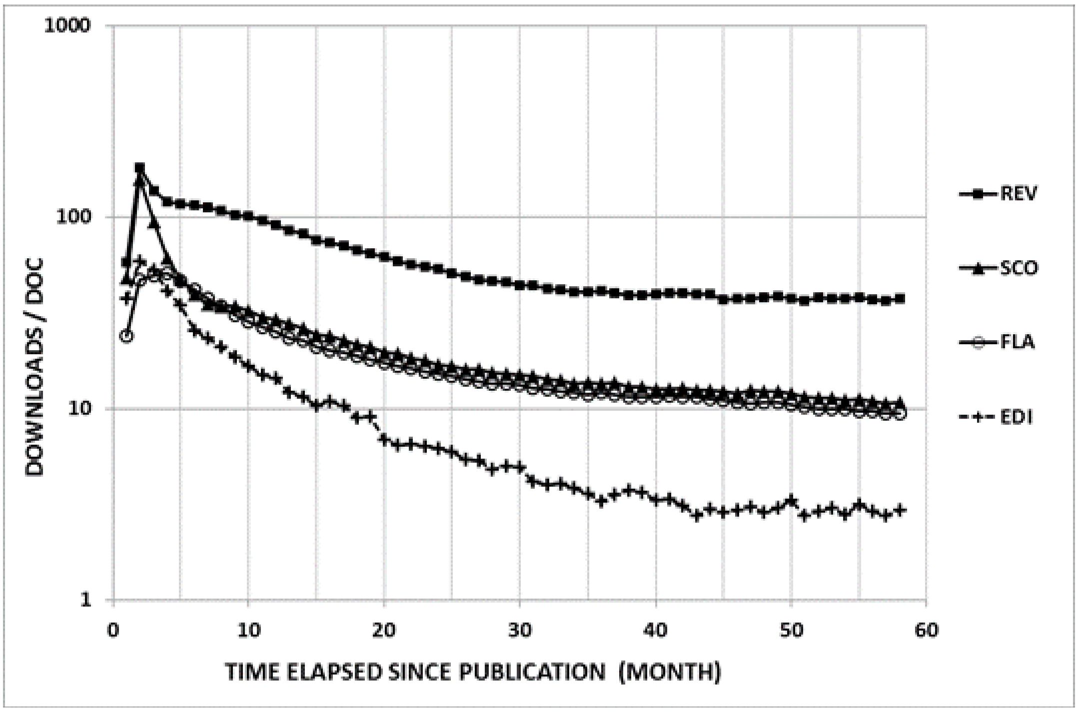
Figure 2: The number of downloads per full length article as a function of the articles' age for 6 journals



Legend to Figure 2: The journals cover the subject fields of Social Sciences (SOC SCI), Applied Sciences (APPL SCI), Life Sciences (LIFE SCI), Clinical Medicine (CLIN MED), Mathematics (MATH) and Humanities (HUMAN), respectively. AGE=1 indicates the months in which the articles were published.

Figure 3 displays the development of downloads over time for four document types in the set of 63 journals: full length articles (FLA, reviews (REV), short communications (SCO) and editorials (EDI). As can be seen in the graph, reviews, short communications and editorials reach their peak downloads in the first month after publication, and full length articles in the third month. Short communications and editorials show the most rapid decline during the first and 24th month after publication. After two years, the decline rates of the four types are similar. The level of downloads is highest for reviews, and lowest for editorials, at least in the set of 63 journals analysed in this section.

Figure 3: The number of downloads per document type as a function of the documents' age.

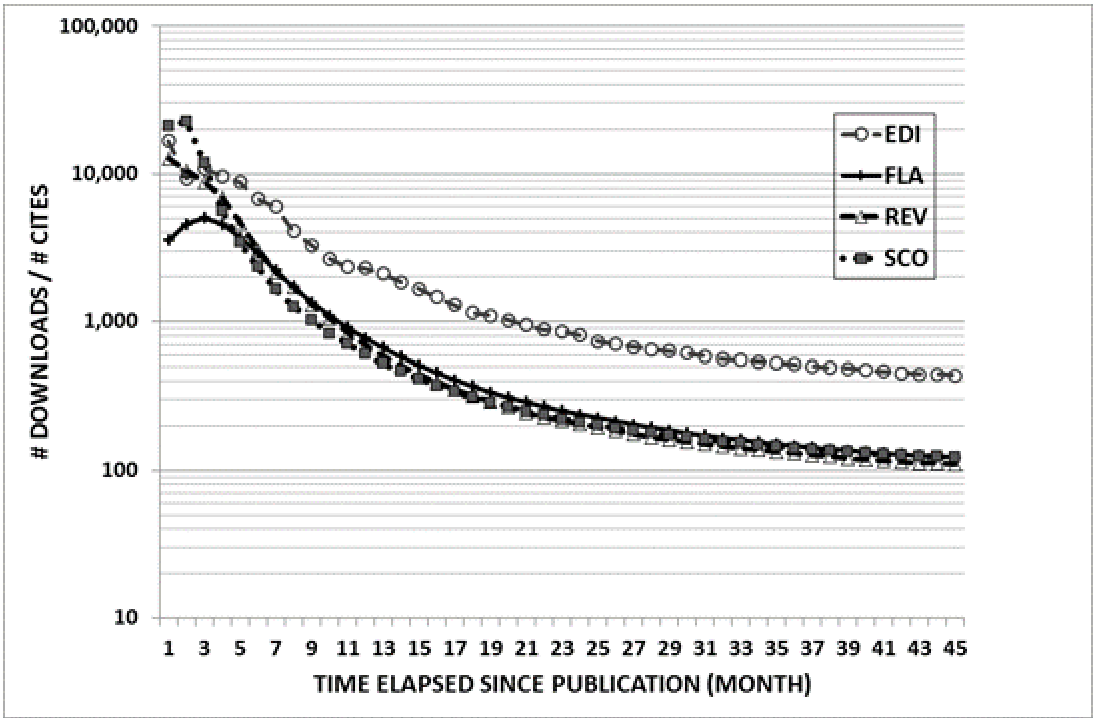


Legend to Figure 3: Data are shown for 4 document types published in the 63 journal set: full length articles (FLA), reviews (REV), short communications (SCO) and editorials (EDI).

Download-versus-citation ratios

Adopting a diachronous approach, Figure 4 presents for documents published during 2008-2009 the ratio of the number of downloads and citations as a function of the documents' age, or, in other words, of the time elapsed since their publication date, expressed in months. In this figure the documents from all journals in the 63 Journal Set are aggregated into one "super" journal. Ratios of downloads and citations are calculated for four types of documents: editorials, full length articles, short communications and reviews. Figure 4 clearly shows that the ratio of downloads and citations very much depends upon the type of document and upon the time elapsed since their publication date. For full length articles, reviews and short communications this ratio reaches a value of about 100 after 45 months.

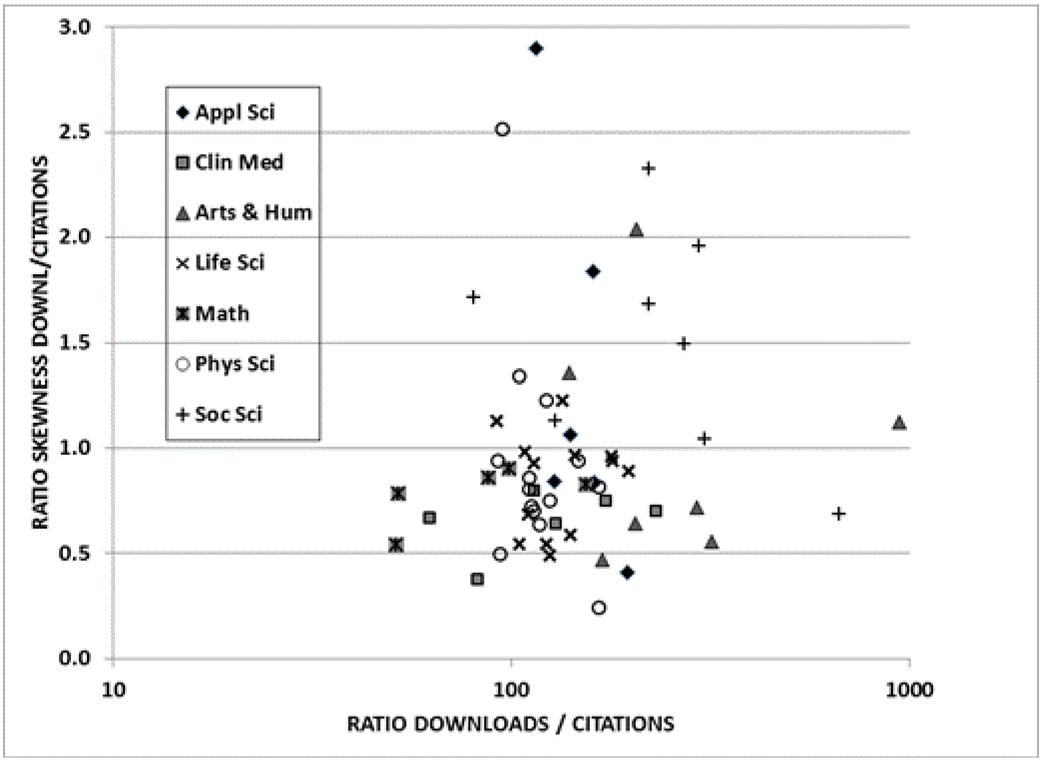
Figure 4: Ratio of downloads and citations of documents as a function of their age (63 Journal Set)



Legend to Figure 4: EDI: Editorials; FLA: Full Length Article; REV: Review; SCO: Short Communications

Figure 5, however, shows large differences in this ratio among the 63 journals. It displays on the vertical axis the ratio of downloads and citations for the aggregate of full length articles published in the 63 Journal Set, and on the vertical axis the ratio of the skewness values of the download and citation article distribution, respectively, further discussed in the next section. Each symbol represents a particular journal. Distinct symbols indicate the main discipline covered by a journal. Figure 5 shows that journals in social sciences and humanities tend to have large downloads ratios versus citations, and several mathematics periodicals relatively low ratios. Clinical medicine journals show large variations.

Figure 5: Ratio of mean and skewedness of the article download and citation distribution for 63 journals set (full length articles (FLA) only)



Statistical correlations between downloads and citations at the journal and article level
Figure 6 presents an analysis at the journal level. It is based on download counts in the year of publication and citations in the third year after publication and shows the Pearson correlations per discipline. Spearman rank coefficients per discipline tend to be somewhat lower than the Pearson values, due to the skewness of the underlying distributions, but the overall picture presented in Figure 6 does not change if the former type is plotted rather than the latter. Analysing the correlation per discipline between a journal's average number of downloads per article against the number of cites per article, Figure 6 shows that in the areas of biochemistry & molecular biology, neuroscience and veterinary sciences downloads and citations are highly correlated followed by chemical engineering, pharmacology and immunology. Disciplines which display the lowest correlation coefficients between downloads and citations are arts & humanities and health professions. The factors responsible for these differences in correlation must be further studied. For instance, the low correlation in Arts & Humanities may be due to the fact that the citation database used does not cover the publication output in this domain sufficiently well, and particularly misses citations in and to books.

Figure 6: Correlation between downloads and citations at the journal level by discipline

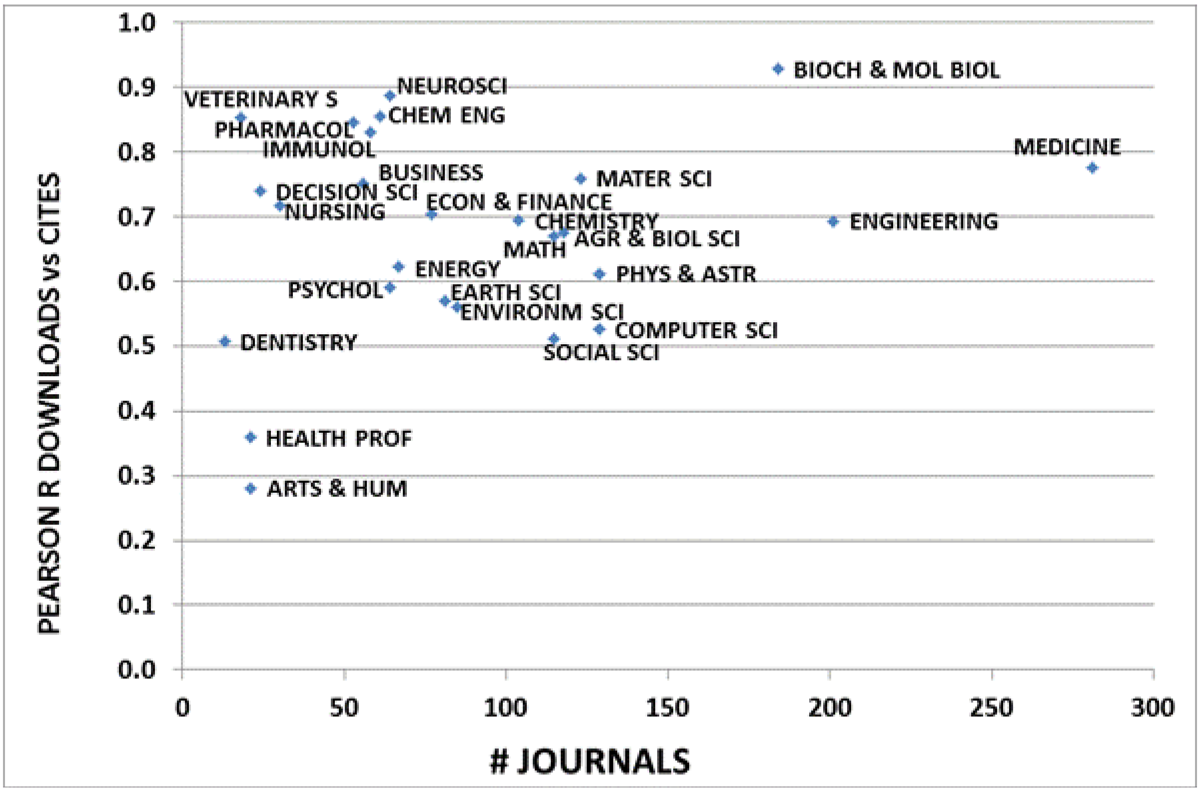
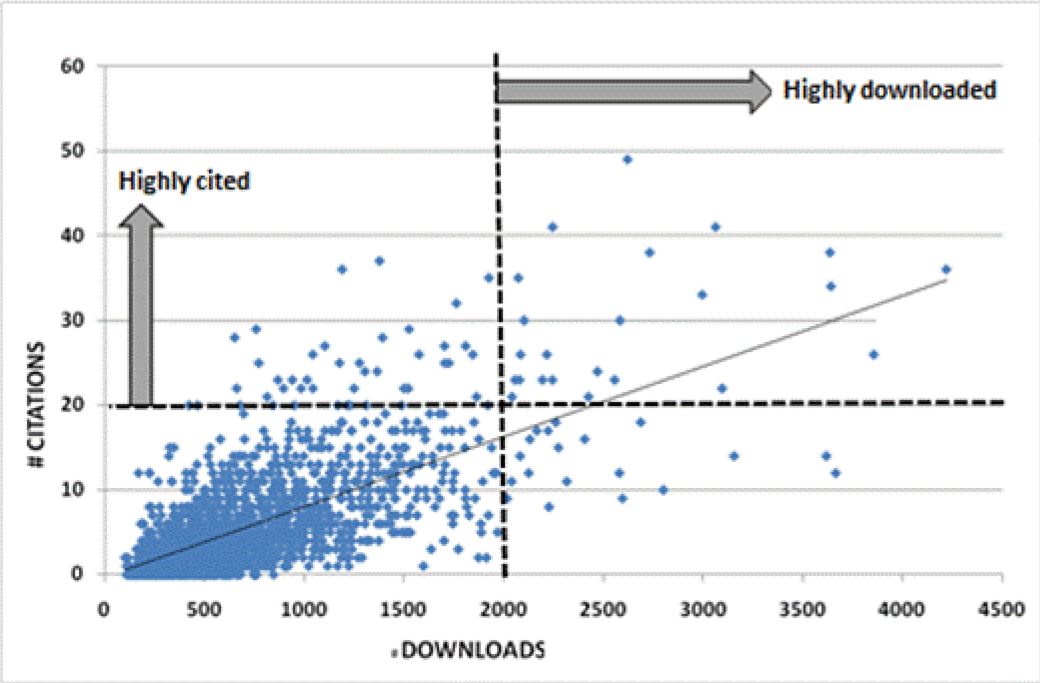


Figure 7 analyses the correlation between downloads and citations at the article level. It presents a scatterplot of downloads versus citation counts of articles in an applied science journal. The diagonal represents the linear regression line. It shows that the articles that are frequently downloaded (tentatively defined as those with more than 2,000 downloads) almost all have a minimum citation count of about 10. In other words, among the articles cited less than 10 times, there are no highly downloaded articles. This is so to speak one side of the correlation coin. But apart from this observation, the citation counts of the highly downloaded articles show a strong scatter. Such a scatter is even more clearly visible among the download counts for articles that are highly cited (tentatively, more than 20 times). But all these highly cited articles have a download rate that exceeds 500.

Figure 7: Downloads versus citation counts for a journal in applied sciences.



Discussion and conclusions

Analyses by user country and institution

The fact that seasonal and academic cycles are reflected in longitudinal download patterns is not surprising. What is of interest is the peaky behavior at the level of user institutions, and the apparent lack in many cases of solid explanations for such behavior. Even if the overall contribution of number of downloads made in peak months across institutions is perhaps only a few per cent of the total number of downloads, more understanding of the cause of outliers is desirable. A combined qualitative-quantitative approach seems the most promising, in which interviews with librarians at institutions is complemented with a more detailed analysis of the underlying usage patterns. Typical questions that should be addressed are: is downloading in peak months a form of bulk downloading, in which large numbers of documents are downloaded issue by issue, journal by journal, in a single user session. Moed (2004) gives typical examples of how bulk sessions can be identified, for instance, an analysis of the average number of downloads per used journal in a session. This parameter tends to obtain extremely high values if complete journal issues or (annual) volumes of a journal are downloaded article-by-article in one single user session.

Downloads time series per journal and document type

Perhaps the main observation of the outcomes presented in this article is that they show such large differences among journals, subject fields, and types of document. It must be underlined again that the journals studied are not a random sample from the total population of journals in ScienceDirect. The aim of the selection was to include journals from different disciplines and cover all major disciplines and include sectionalized journals as well. Our outcomes thus show how large the variability across journals and subject fields can be.

The analyses at the journal level presented in the current paper show that, adopting a diachronous approach, during the first 4 years after online publication date, all journals show a delay in downloads, in the sense that the average number of downloads per month increases after the month of publication and reaches its peak after 2 to 8 months, depending upon the journal. Such a behavior is qualitatively similar to that of citation obsolescence: both processes show a delay.

Moed (2005) used in a diachronous approach a two factor model based on monthly rather than annual usage counts. Although the model showed a reasonable fit when applied in 2005 to Tetrahedron Letters, a journal publishing on a monthly basis short communications with a relatively short life cycle, download obsolescence patterns per journal reveal that a two factor model tends to be inappropriate.

Full length articles, reviews, short communications and editorial have different download obsolescence patterns; their differences are similar to those found for citations. The ratio of the number of the number of downloads per review to that per article is similar to the same ratio for citations. And short communications mature more quickly than full length articles both in terms of downloads and citations.

Download-versus-citation ratios

Findings in this paper illustrate that the actual ratio of downloads and citations strongly depends upon the age of the used articles. It must be noted, however, that the rate of decline decreases over time, and that the value of downloads per citation ratio seems to stabilize somewhat after three years or so to a value of approximately 100. The conclusion is that, after four years following the online publication date, the number of downloads of the articles in a journal is two orders of magnitude higher than the number of citations. This result applies both for full length articles, reviews and short communications. For editorials, however, the ratio is a factor of 2 higher than it is for the other document types.

Statistical correlation between downloads and citations

Large differences in the degree of linear correlation were found among subject fields at the journal level, the Pearson correlation coefficients varied between around 0.3 in the humanities to 0.9 in molecular biology. Intuitively one might conjecture that subject fields in which the correlation is high tend to be very specialized fields, such as molecular biology and biochemistry, in which the main users or readers of publications are the researchers active in that field, in other words, fields in which the author and the reader populations tend to coincide. Fields in which the reader population is probably much wider than the research community – including for instance interested readers from other disciplines of publications made by humanities and social science researchers, or practitioners (engineers or nurses) using technical information from engineering and nursing journals – the correlation is lower. But the analysis did not define or measure more precisely the degree of overlap between author and user population, so that rigorous testing of the hypothesis that the degree of correlation between downloads and citation counts is positively related to this overlap, has not been carried out, due to a lack of information about the user or reader population.

References

Duy, J. & Vaughan, L. (2006). Can electronic journal usage data replace citation data as a measure of journal use? An empirical examination. The Journal of Academic Librarianship 32, 512-517.

Kurtz, M.J., & Bollen, J. (2010). Usage Bibliometrics. Annual Review of Information Science and Technology, 44, 3-64.

Kurtz, M.J., Eichhorn, G., Accomazzi, A., Grant, C., Demleitner, M., & Murray, S. (2005a). Worldwide use and impact of the NASA Astrophysics Data System Digital Library. Journal of the American Society for Information Science and Technology, 56, 36–45. .

Kurtz, M.J., Eichhorn, G., Accomazzi, A., Grant, C., Demleitner, M., Murray, S.S., Martimbeau, N., & Elwell, B. (2005b). The bibliometric properties of article readership information. Journal of the American Society for Information Science and Technology, 56, 111-128.

Moed, H. F. (2005). Statistical relationships between downloads and citations at the level of individual documents within a single journal. Journal of the American Society for Information Science and Technology, 56(10), 1088-1097.

Schloegl, C. & Gorraiz, J. (2011). Global usage versus global citation metrics: The case of pharmacology journals. Journal of the American Society for Information Science and Technology, 62(1), 161–170. doi:10.1002/asi.

Gorraiz J., Gumpenberger, C. & Schlögl, C. (2013). Differences and similarities in usage versus citation behaviours observed for five subject areas. In Proceedings of the 14th ISSI Conference Vol 1, 519-535.

http://www.issi2013.org/Images/ISSI_Proceedings_Volume_I.pdf.

Semantic Analysis of Knowledge Flows using Scientific Literature

Saeed-Ul Hassan* and Peter Haddawy**

*saeed-ul-hassan@itu.edu.pk
Scientometrics Lab, Information Technology University - Punjab, Ferozepur Road, Lahore, 54770 (Pakistan)

** peter.had@mahidol.ac.th
Faculty of ICT, Mahidol University, Phuttamonthon 4 Road, Nakhon Pathom, 73170 (Thailand)

Introduction

As observed by Hagel et al. (2009), "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". Indeed, participation in international knowledge flows has become an accepted measure of the quality of a research environment. But while increasing use is being made of metrics such as student and staff mobility and international co-authorship, metrics that directly measure international impact of research have yet to be widely used.

Citation patterns have been used to measure knowledge flows among scientists (Zhuge 2006), among journals (Zhou and Leydesdorff 2007), and among subject categories (Zhou et al. 2010). In a recent paper (Hassan and Haddawy 2013) we introduced a metric to measure knowledge flow among institutes and countries. The present paper extends that work by introducing an approach to semantically analyze knowledge flows. The approach sheds light on how knowledge produced by researchers in one country is utilized by researchers in other countries.

Methodology

The approach starts by identifying Research Topic (RT) clusters for the publications produced by a given country in a given research area. We then procure the sets of papers (authored by researchers outside the given country) citing the papers in the RTs. Finally, we cluster the citing papers and identify frequent keywords to determine how the knowledge in the papers in the RTs is being used.

In order to select the keyword terms, we use author defined keywords and noun terms extracted from the abstracts and titles from the procured papers. We then identify synonyms of the selected terms and include them as keyword terms as well. The RTs are identified using the author-topic model with distance matrix (Hassan and Ichise, 2009). In order to obtain the optimum number of RTs, we compute inter cluster similarity and average intra cluster similarity. Finally, RT keywords are visualized using Wordle.Net (http://www.wordle.net/).

We present a case study in the subject area Energy. Using All Science Journal Classification (ASJC), we procured 7602 papers from the Scopus database (journal articles, reviews and conference papers) published by researchers from the United States in Energy during 2004-2009 that are cited by researchers from other countries (excluding the co-authored papers with the US researchers) in the same time period.

Results and Discussion

We obtain eleven RTs in the field of Energy cited by researchers from outside the US. Figure 1 shows the five clusters with the largest numbers of papers. Cluster #1 is the largest cluster, containing 33% of the 7602 papers. The cluster covers research topics related to Solar Cells (such as Thin Film Solar Cells, Tandem Solar Cells and Photovoltaics).

To examine the different ways countries use this knowledge, we compare publications of researchers from China and Japan that cite the papers in cluster #1. We procure all the papers (journal articles, reviews and conference papers) authored by researchers from China and Japan that cite papers in cluster #1 and then identify RTs of those papers. Figure 2 shows RTs of the paper by Chinese researchers during 2004-2009. Cluster #1 is the largest cluster, containing 77% of the 1575 papers and mainly covers research topics related to Power Systems. Cluster #2 contains research topics related to Solar Cells (such as Thin Film Solar Cells and Dye-sensitized Solar Cells). Figure 3 shows the research topics for the papers produced by Japanese researchers during 2004-2009. Cluster #1 in Figure 3 shows that the Japanese researchers utilize the same knowledge for rather different research themes than the Chinese researchers. The Japanese researchers focus on topics related to Superconductivity and High Temperature Superconductors in the context of Solar Cells. Note that Superconductors play a vital role in providing low-cost renewable energy. It is also interesting to note that that one cluster accounts for 96% of the citing papers, with the remaining 4% in cluster #2 which contains the topics of efficient use of Photovoltaics, Energy Conversion and Solar Cells.

Concluding Remarks

The method of semantic analysis presented in this paper provides an understanding into 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.

References

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.

Hassan, S., Ichise, R. (2009). Discovering research domains using distance matrix and co-authorship network. *Society for Industrial and Applied Mathematics - 9th SIAM International Conference on Data Mining 2009, Proceedings in Applied Mathematics, 3,* 1252-1257.

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.

Figure 1: Top five Research Topics cited by researchers outside the United States in the Field of Energy during 2004-2009.

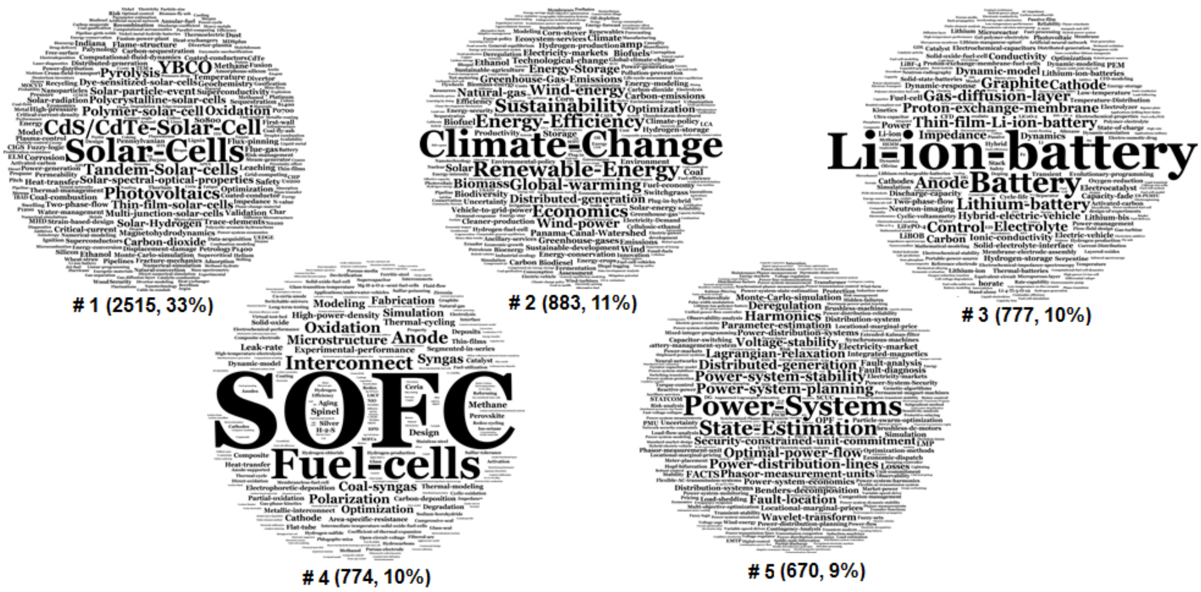


Figure 2: Research Topics of the papers produced by Chinese researchers (during 2004-2009) that cite the Research Topic in Cluster#1 in Figure 1.

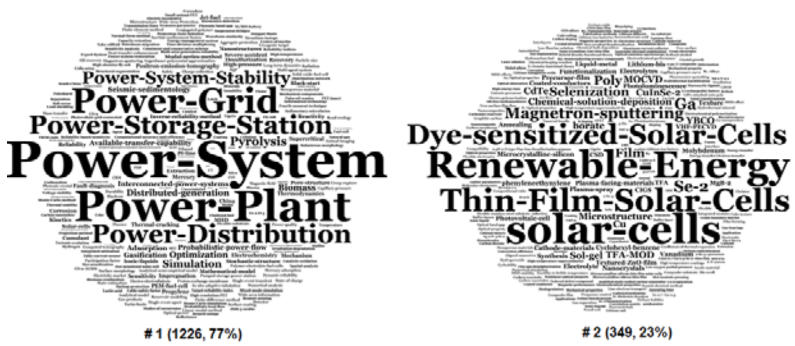
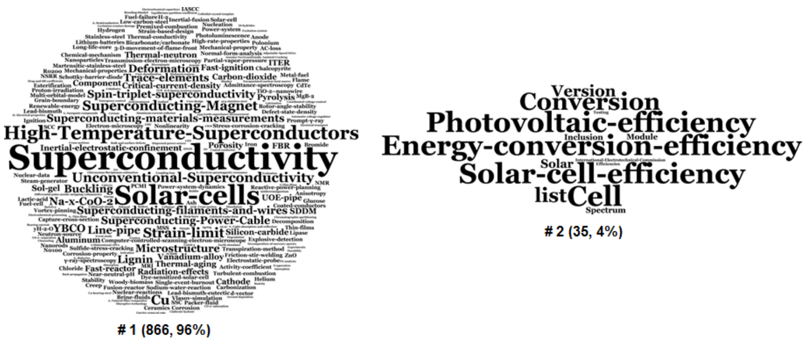


Figure 3: Research Topics of the papers produced by Japanese researchers (during 2004-2009) that cite the Research Topic in Cluster#1 in Figure 1.



Long-distance interdisciplinary researchleads to higher citation impact[1]

Stefanie Haustein*, Vincent Larivière** and Katy Börner***

[*] *stefanie.haustein@umontreal.ca*
École de bibliothéconomie et des sciences de l'information, Université de Montréal, C.P. 6128, Succ. Centre-Ville, Montréal, QC. H3C 3J7, Canada

[**] *vincent.lariviere@umontreal.ca*
École de bibliothéconomie et des sciences de l'information, Université de Montréal, C.P. 6128, Succ. Centre-Ville, Montréal, QC. H3C 3J7, Canada and Observatoire des Sciences et des Technologies (OST), Centre Interuniversitaire de Recherche sur la Science et la Technologie (CIRST), Université du Québec à Montréal, CP 8888, Succ. Centre-Ville, Montréal, QC. H3C 3P8, Canada

[***] *katy@indiana.edu*
Cyberinfrastructure for Network Science Center, School of Informatics and Computing, Indiana University, 10th Street & Jordan Avenue, Wells Library, Bloomington, IN, 47405, USA
Data Archiving and Networked Services (DANS), Royal Netherlands Academy of Arts and Sciences (KNAW), Anna van Saksenlaan 10, 2593 HT The Hague, The Netherlands

Introduction

Interdisciplinary research has long played a central role in research evaluation. By definition, it focuses on solving complex scientific problems by combining methods and concepts from different disciplines and is said to be greater than the sum of its disciplinary parts. Many bibliometric studies have tried to prove the perceived success of interdisciplinary research through higher citation impact, leading to contradicting results depending on the definition of both interdisciplinarity and citation impact (see Wagner et al., 2010 for a review). This study focuses on the identification of the success of interdisciplinary relationships—defined as co-cited subdisciplines—with a particular focus on the distance between the two subdisciplines involved. Specifically, it aims to answer three questions: (1) Which (sub)disciplines benefit most from interdisciplinary research? (2) Which combination of subdisciplines leads to the highest citation impact? (3) How does the distance between two co-cited subdisciplines influence the impact of the citing paper?

Methods

The dataset used in this paper is drawn from Thomson Reuters' Web of Science (WoS) database, including the Science Citation Index Expanded, the Social Sciences Citation Index and the Arts and Humanities Citation Index. It comprises all 11.1 million articles and reviews published between 2000 and 2012 including cited references published during the same period and covered in the same databases as source items. Disciplines and subdisciplines were assigned to paper references using the UCSD classification system and map of science, which comprises 13 disciplines and 554 subdisciplines computed using bibliographic coupling and keywords at the journal level (Börner et al., 2012). The 40 journals that are assigned to more than one subject category were omitted to ensure that a reference is assigned to exactly one

[1] This work was supported by the Canada Research Chairs program, the Social Science and Humanities Research Council of Canada and the Fonds Recherche Québec – Société et Culture and the U.S. National Institutes of Health under Grant No. U01 GM098959.

(sub)discipline. A paper is defined as interdisciplinary if it contains references from more than one subdiscipline. Although not without limitations, this binary definition of interdisciplinarity was chosen to avoid a more arbitrary threshold. While 1.9 million papers (17.2%) were strictly disciplinary, 9.2 million (82.8%) were interdisciplinary. 80,997 pairs were co-cited at least 30 times and are used in this study. To determine the citation impact of interdisciplinary relationships, each co-cited subdiscipline pair was assigned the citing paper's citations as the observed citation. Larivière and Gingras (2010) found that the impact of interdisciplinary papers depends strongly on the citation potential, i.e. expected citations of involved disciplines. Thus, two relative citation rates for each subdiscipline pair s_1-s_2 were computed. The *expected citation* rate of s_1 in year *y* represents the average citation rate of all papers citing s_1 in *y*. The *relative citation rate* of s_1-s_2 for all years relative to s_1 and s_2 respectively represents the average of all *observed* vs. *expected* ratios for each co-citing paper. That is, each co-cited interdisciplinary pair s_1-s_2 obtained two *relative citation rates*—one relative to s_1 and another relative to s_2, resulting in a total of 161,994 relative citation rates. When the observed citations exceeded the expected, i.e., world average citation rate, then the relative citation rate is greater than 1. The success of a subdiscipline pair (i.e., interdisciplinary collaboration) was classified into four categories: *win-win* (both citation rates above 1), *win-lose* & *lose-win* (one rate above, one equal or below 1), and *lose-lose* citation outcomes (both equal or below 1).

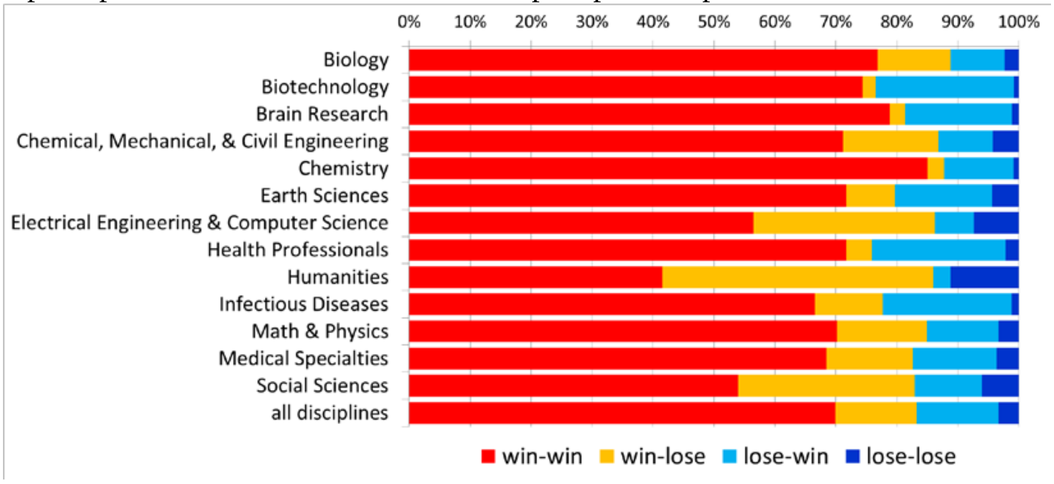
The distance between two subdisciplines seems to affect success. Distance was calculated using the x-y positions of the 554 subdisciplines on the UCSD map, which represents a widely used reference system of the research landscape. Since the position of each of the 554 subdisciplines in the map of science is determined by bibliographic coupling and keyword similarity to each other, the distance between two nodes in the map can be considered as an indicator of topical distance, where close nodes represent closely related subdisciplines and distant nodes are less related. Note that this map wraps around a cylinder with a circumference of 624, i.e., the left most nodes are connected to the nodes on the far right. The distances for the 80,997 subdiscipline pairs ranged from 0.64 to 281.10. Subdiscipline pairs were grouped into 10 categories of distances (A-J, see Table 1) with a comparable number of pairs in each category.

Results

Overall, 69.9% of the subdiscipline pairs were win-win, 26.8% were win-lose or lose-win and only 3.3% were lose-lose. The mean *relative citation rate* of all subdiscipline pairs is 1.54, i.e., papers co-citing publications from two different subdisciplines were on average cited 54% above world average level. As relative citation rates represent a zero-sum game, this shows that co-citing literature from different subdisciplines pays off in terms of citation impact. In fact, the relative citation rate of papers increased with the number of subdisciplines cited and remained 60% below expectations for the 1.9 million papers citing only one subdiscipline.

Aggregated on the level of the 13 UCSD disciplines, all disciplines benefit from interdisciplinarity (Figure 1), especially Biology, Chemistry and Chemical, Mechanical, & Civil Engineering and Electrical Engineering, where more than 86% of co-cited subdiscipline pairs were cited above the average of the particular subdisciplines (win-win and win-lose combined). Chemistry (85.0%), Brain Research (78.8%) and Biology (76.9%) are the disciplines where both subdiscipline sides benefit the most from being co-cited (win-win). While being co-cited with subdisciplines from Biotechnology, Chemistry and Brain Research is most beneficial to the co-cited subdiscipline, as more than 96% of all co-cited subdisciplines are cited above average (win-win and lose-win), combinations with subdisciplines from the Humanities (55.7%), Electrical Engineering & Computer Science (37.1%) and Social Science (35.0%) are more disadvantageous because over one third of co-cited subdisciplines do not exceed their world average citation rates (lose-lose and win-lose). Interdisciplinary combinations with the Humanities and Social Sciences are least beneficial to either side, as 11.2% and 6.0% of all subdiscipline pairs do not meet expected citation in either of the co-cited subdisciplines (lose-lose).

Figure 1: Percentage of win-win, win-lose, lose-win and lose-lose relationships of co-cited subdiscipline pairs based on relative citation impact per discipline.



As shown in Table 1, the mean relative citation rate of co-cited pairs increases with the distance between the two subdisciplines. Except for category A which contains co-cited pairs closest to each other in the UCSD map, i.e., the most similar subdisciplines—71.4% of the subdiscipline pairs in this category were assigned to the same discipline—the percentages of win-win relationships is around 71% for all distance categories. This suggests that the increase in citation impact is not due to the increase in the number of win-win pairs but rather by an actual growth of relative citation impact with increasing distance. The highest *relative citation impact* was obtained by papers co-citing Child Abuse (Social Sciences) and Leukemia (Brain Research) with a relative citation rate of 27.5 (relative to all papers citing Child Abuse; distance category E), Thoracic Surgery (Brain Research) and Air Quality (Earth Sciences) cited 27.2 (I) and Child Abuse (Social Sciences) and Clinical Chemistry (Brain Research) cited 26.3 (E) on average.

Table 1. Mean *relative citation rates* and percentage of win-win, win-lose & lose-win, and lose-lose relationships of subdiscipline pairs per distance category.

Distance category	Number of subdisc. pairs	Relative citation rate			Percentage of win vs. lose relationships		
		mean	std. dev.	median	win-win	win-lose & lose/-win	lose-lose
A: 0<28	15,790	1.26	0.45	1.20	62.4%	31.5%	15.8%
B: 28<48	16,584	1.38	0.52	1.30	70.3%	25.9%	13.0%
C: 48<67	16,404	1.45	0.64	1.34	70.5%	25.9%	13.0%
D: 67<86	15,854	1.49	0.79	1.36	69.0%	28.0%	14.0%
E: 86<107	15,882	1.51	0.80	1.38	69.7%	27.7%	13.9%
F: 107<130	16,284	1.58	0.75	1.44	71.4%	26.0%	13.0%
G: 130<152	16,160	1.64	0.79	1.48	71.2%	26.4%	13.2%
H: 152<172	16,308	1.66	0.82	1.49	70.2%	27.1%	13.6%
I: 172<181	16,358	1.71	0.96	1.50	71.0%	26.1%	13.1%
J: 181<282	16,370	1.72	0.97	1.50	73.0%	23.0%	11.5%
All distance	**161,994**	**1.54**	**0.78**	**1.38**	**69.9%**	**26.8%**	**13.4%**

The findings support the assumption that interdisciplinary research is more successful and leads to results greater than the sum of its disciplinary parts. Papers citing references which are positioned far away from each other in the conceptual space of the UCSD map of science manage to attract the highest relative citation counts on average. From a research policy perspective this suggests that interdisciplinary connections should be especially supported where it is most challenging: between distant areas of research. Future research will involve visualizing the most beneficial win-win relationships on the map of science and analyzing the relationship between citation impact and distances in depth.

References

Börner, K., Klavans, R., Patek, M., Zoss, A.M., Biberstine, J.R., Light, R.P., Larivière, V. & Boyack, K.W. (2010). Design and update of a classification system: the UCSD map of science, *PLOS ONE*, e39464.

Larivière, V., Gingras, Y. (2010). On the relationship between interdisciplinarity and scientific impact. *Journal of the American Society for Information Science and Technology*, 61(1), 126-131.

Wagner, C.S., Roessner, J.D., Bobb, K., Thompson Klein, J., Boyack, K.W., Keyton, J., Rafols, I. & Börner, K. (2010). Approaches to understanding and measuring interdisciplinary scientific research (IDR): A review of the literature. *Journal of Informetrics*, 165, 14-26.

Patterns of Globalisation: Geographical, Sectoral and Technological Developments in Corporate Invention.

Gaston Heimeriks[*], Floortje Alkemade[*], Antoine Schoen[**], Lionel Villard[**], Patricia Laurens[***]

G.J.Heimeriks@uu.nl
[*]Innovation Studies, Copernicus Institute, Utrecht University
Heidelberglaan 2, 3584 CS Utrecht, The Netherlands.)

[**, ***]Université Paris-Est, [**]ESIEE/[***]CNRS– LATTS - IFRIS - FRANCE
2, bd Blaise Pascal 93160 NOISY LE GRAND

Introduction

The continuing globalisation of the R&D activities of firms is a subject of considerable interest to policymakers, as innovation is recognised as a main driver of productivity and growth for countries, as well as a vital resource in addressing societal challenges. We are witnessing a surge in knowledge-intensive activities by firms in catch-up economies, altering the geography of ideas and their commercialisation across the globe. The expansion of the knowledge base is reflected in an increasing diversity of locations, applications and fields in knowledge productions. New ideas, methods and tools are continuously introduced in research practices, the landscape of knowledge production is continuously in flux and new applications are being developed at any given moment (Heimeriks, 2012). This clearly has implications in terms of international competition and growth strategies.

As there is a strong link between corporate R&D and innovation, policy concerns focus on the potential loss of jobs and economic benefits as well as on the potential impoverishment of the local knowledge base due to the internationalisation of R&D. Especially the increasing attraction of Asian countries as R&D location (Heimeriks & Boschma, 2014) leads to a growing concern among policy makers for hollowing out the national innovation system (Narula & Zanfei, 2005).

Yet, patterns of global corporate invention remain poorly understood. Whether we look at it from the point of view of geography, economic activities and technological invention in different fields, we observe that advances in technological invention are unevenly distributed. In this paper, we explore these patterns of corporate invention among countries, sectors and technological fields. What is the nature of the process and how can it be managed? The central research question of this paper is thus:

Can we specify the national, sectoral and technological characteristics of the globalisation of corporate knowledge production over time? Furthermore, how do these different dynamics interact?

Clearly, there are three analytical dimension relevant for understanding patterns of corporate knowledge production (Leydesdorff, 2010). Geographically positioned units of analysis (e.g., firms), economic exchange relations, and (technological) novelty production cannot be reduced to one another. However, these independent dimensions can be expected to interact to varying extents.

First, from a geographical perspective, we witness globalisation processes that involve geographically disparate firms and their subsidiaries, whose technologies are disseminated over vast distances. Yet, corporate knowledge production is extremely unevenly distributed across space (Florida, 2005). Studies in economic geography attribute this to the fact that regions (and their aggregates in countries) tend to expand into activities that are closely related to their existing capabilities. Corporate knowledge production results from locally available skills, tacit knowledge, institutions and infrastructures that both enable and constrain the evolution of knowledge (Boschma, 2005).

Second, from an economic perspective, different firms (and their aggregates in sectors) rely on knowledge to a different extent and are able to produce and apply knowledge to different degrees. Innovation scholars have argued that organisational routines of knowledge producing organisations respond to satisfying the knowledge needs of those entities outside the organisation (governments, customers, users and investors) that provide the resources for organisations to survive. As a consequence, organisations will be successful if they produce knowledge that translates into solutions, goods, services and profits that those external entities require. The availability of resources thus enables and constrains the production of knowledge (Pfeffer & Salancik, 2003).

Third, from a cognitive perspective, codified knowledge developments are unevenly distributed among topics (and their aggregates in fields). Researchers in Science and Technology Studies (STS) and information science argue that the evolution of codified knowledge is characterised by a path-dependent process of branching; new knowledge is developed from recombinations of existing knowledge. The existing body of codified knowledge thus enables and constrains the production of new knowledge (Arthur, 2007).

Data and Methods

In order to address the need for more systematic analysis of patterns of corporate invention, we use a unique database, the Corporate Invention Board (CIB). The CIB includes 2289 multinational corporations (MNC) that have at least one transnational patent application between 1993 and 2005 and for which information on both inventor and applicant location is available. Of the 2289 MNC's, 730 have their corporate headquarters in Asia, 1002 in Europe and 538 in the Northern America (1 in Africa, 7 in Latin America and the Caribean and 11 in Oceania). The CIB complements the industrial R&D Investment Scoreboard[1] which analyses the performances of companies with the highest annual R&D investments.

The CIB combines this scoreboard data with data on the patents of these companies taken from the PATSTAT[2] patent database published by the European Patent Office. CIB covers a very significant share of private R&D investments: the industrial corporations account for 80% of world total private R&D. Through patents' statistics, we focus on the outputs of these R&D investments providing information on sectors (through the Industry Classification Benchmark), technologies and on geographical location of these investments.

While patent classification systems provide a starting point for identifying patents that belong to a specific technological domain, they do not constitute a classification of technological

[1] http://iri.jrc.ec.europa.eu/scoreboard.html
[2] http://www.epo.org/

fields (OECD 2009). In order to overcome this problem, we developed an original classification of technology that distributes all inventions in 389 non-overlapping classes. This classification is based on the well-established WIPO hierarchical classification that distinguishes, at its finest aggregation level, 35 technological fields, these 35 fields, being grouped in 5 technological domains (WIPO 2008). The global technology map depicts how these technological fields are connected.

Over the 20 year period 1986-2005, the corporation included in the CIB have applied for 5.667.253 priority patents, of which 1.019.989 are transnational priority patents, i.e. the protection for the invention has been asked for in more than one country.

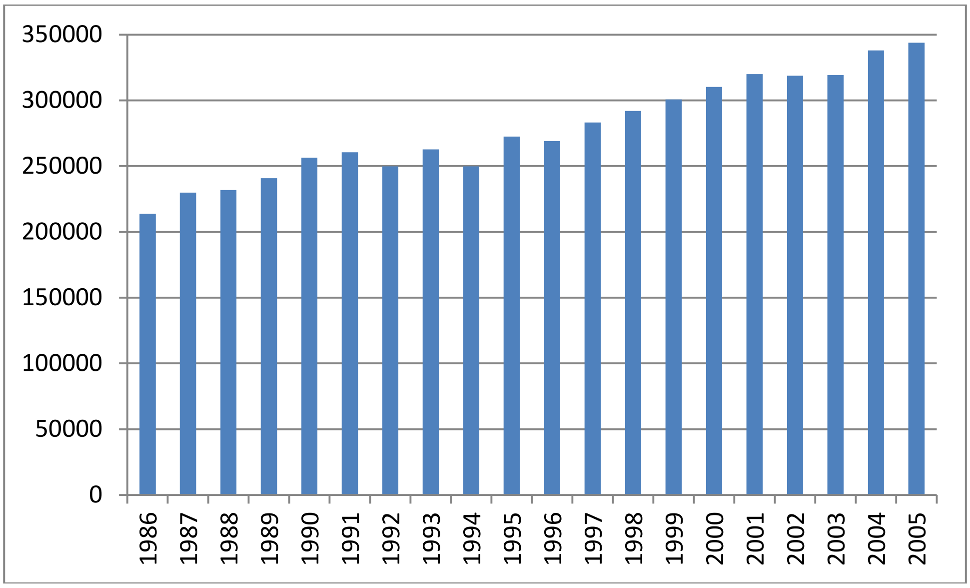


Figure 1. The growth of corporate invention between 1986 and 2005.

Using patent occurrences along the dimensions of countries, technologies and sectors, entropy analysis can be used to quantify the relationship among these dimensions (Theil, 1972). We focus on pairs of distributions; countries and technologies, countries and sectors and sectors and technologies. The entropy value of each two-dimensional matrix is given by;

$$H\ (A) = -\sum_{a=A} \left(pa \cdot \log_2 pa \right)$$

The entropy is zero when all distributions are equal since then there is no uncertainty, and is positive otherwise. The larger the entropy value, the larger the variety within a distribution of technologies.

The expected mutual information is a measure of dependence between two dimensions, i.e., to what extent events tend to co-occur in particular combinations. Mutual information is given by:

$$J(X,Y) = \sum_{i=1}^{m} \sum_{j=1}^{n} p_{ij} \log_2 \left(\frac{p_{ij}}{p_{i.} \cdot p_{j}} \right)$$

The mutual information value equals zero when there exists no coupling/dependence between two dimensions, and the higher the mutual information value the higher the degree of coupling.

Results

From a geographical point of view, a multi-polar world is emerging with an increasing number of public and private research hubs spreading across North and South. In general, corporate knowledge production in seems to shift away in relative terms from the US towards Asian regions. Especially, Korea and China have established themselves at the top of the corporate invention rankings. The sectoral distribution of the growth in patenting activity seems even more unevenly distributed. Almost all sectors show an increase in patenting activities in the period under study. However, the largest growth takes place in a limited number of sectors, most notably related to Electronics, Automotive, Chemicals and ICT. Likewise, the growth of technological knowledge production is unevenly distributed over technological fields and can be attributed to a limited number of fields, mostly related to ICT.

Most countries expanded their technological capabilities, as indicated by the number of technological fields. The largest increase in number of technologies occurs in the EU countries, Japan and the US as well as in emerging economies such as Korea, China, Taiwan, India and Russia. The emergence of a multi-polar world, is not only associated by an increase in the number of patents outside the traditional Triadic countries, but also with more diverse technological capabilities in different locations around the world (Figure 2).

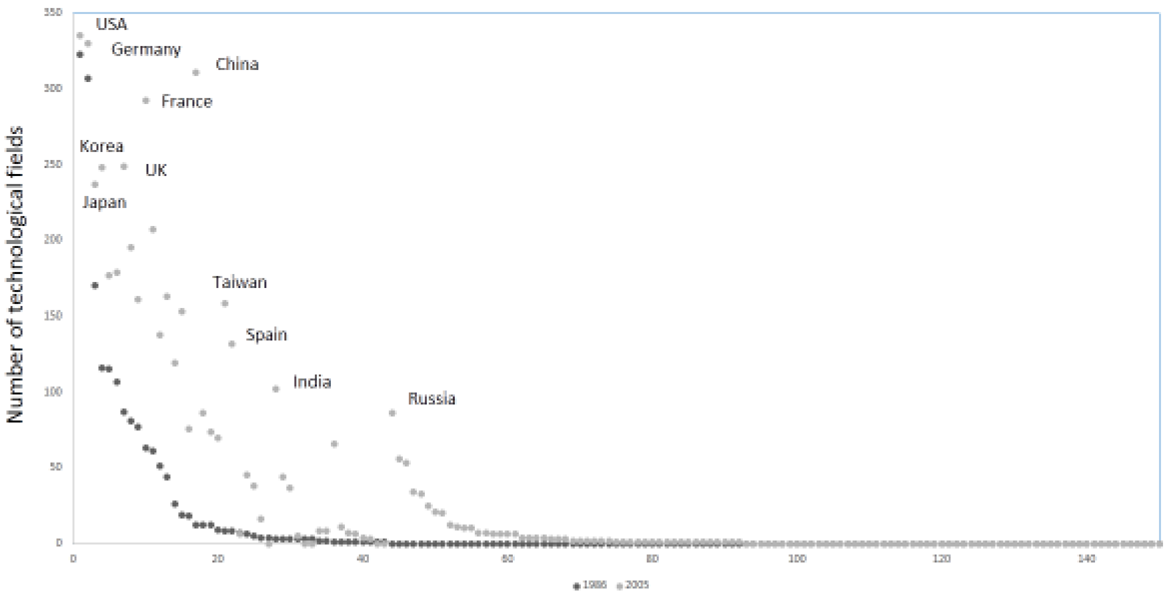


Figure 2. The number of technological fields per country in 1986 and 2005.

Reversely, the distribution of countries over technological fields informs us about the globalisation of technologies (Figure 3).

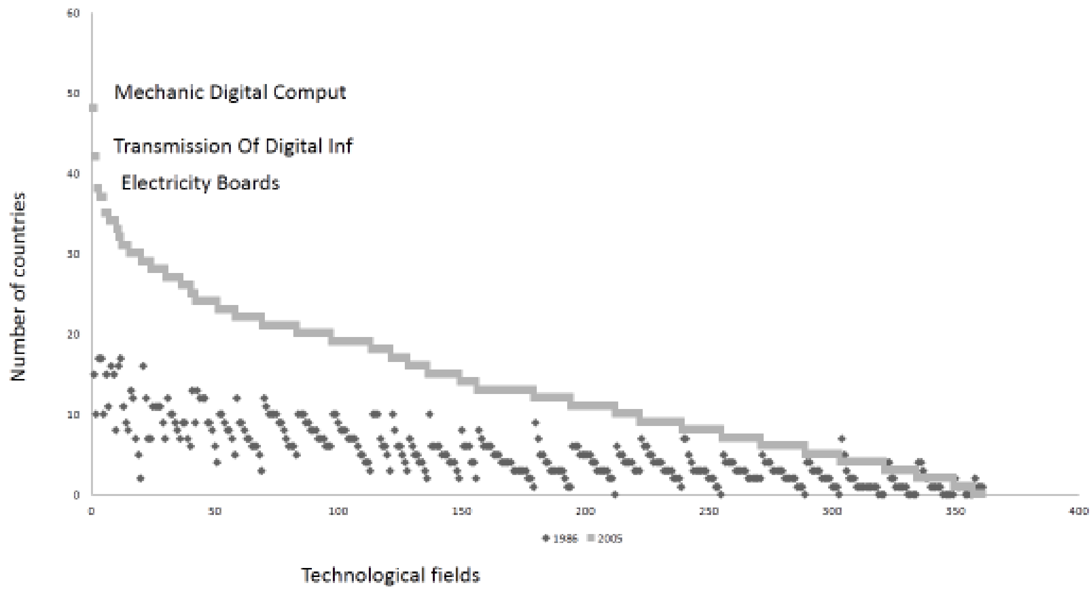


Figure 3. The number of countries per technological field in 1986 and 2005.

A relatively flat distribution is visible the number of countries contributing to technological fields. Only two ICT related fields ("Digital Computing" and "Transmission of Digital Info") occur in more than 40 countries in 2005. However, virtually all technologies are truly globalised. Knowledge production in the large majority of technological fields occur in more than 20 countries.

The question arising is whether the pattern of increased knowledge intensity and technological diversification is associated with more diversified economic activities, as indicated by the diversity of sectors. Again, we observe an increase in in the diversity. However, the diversification of sectoral activities is much less pronounced than the diversification in technologies (Figure 4).

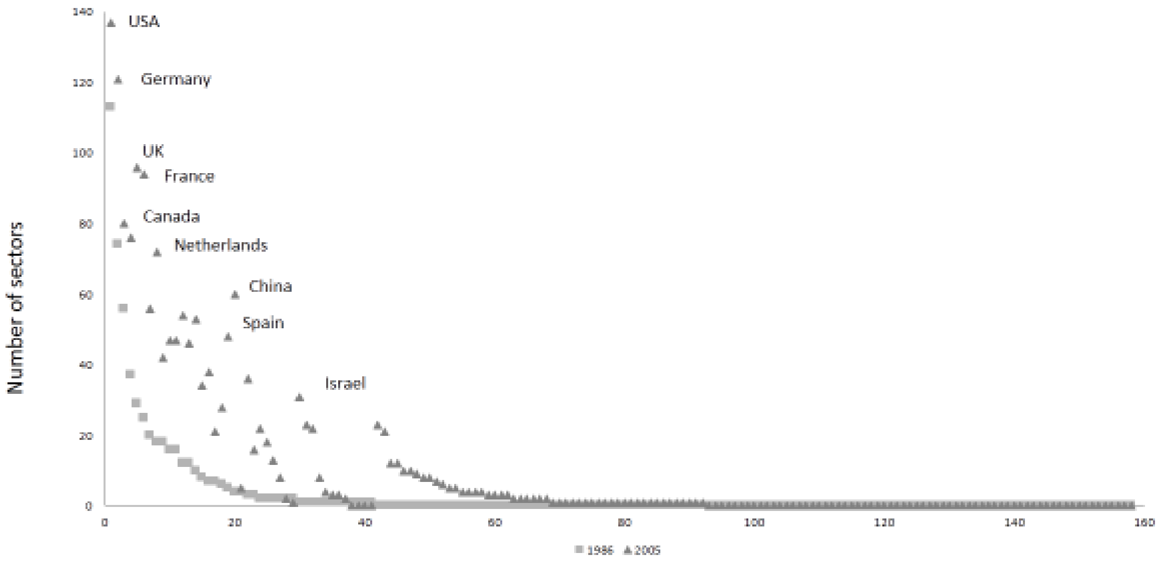


Figure 4. The number of sectors per country in 1986 and 2005.

The increase in sectoral activities is much more concentrated than the increase in technological activities discussed above. The largest increase in number of sectors can be found in France (+69), UK (+67), China (+56), The Netherlands (+54) and Germany (+47). Moreover, only a small number of countries (e.g. China) manage to move towards the core of the map, where the diversified countries are located. Most countries are very stable in their sectoral composition. Countries diversify into related sectoral activities that are gradual in comparison to the more diverse technological development.
To study the globalisation on sectoral activities in more detail, we turn to the number of countries involved in different sectoral activities (Figure 5).

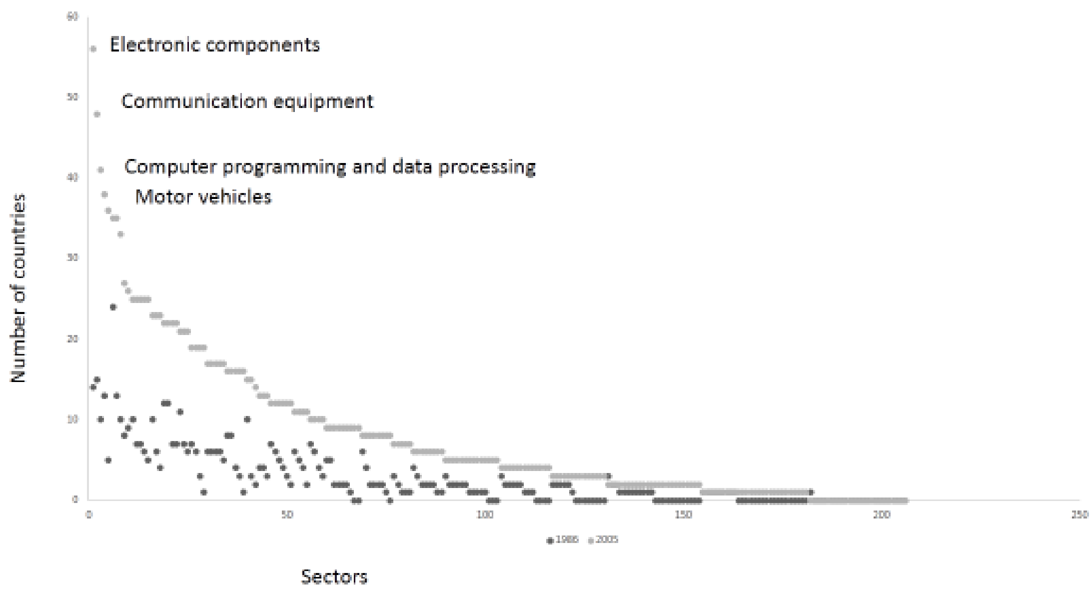


Figure 5. The number of sectors per country in 1986 and 2005.

The most globalised sectors are all related to the broader ICT industry. The sectors “Electronic components”, “Communication equipment” and “Computer programming and data processing” are most globalised sectors and showed the largest increase (in number of countries) between 1986 and 2005. Also the automotive industry (“Motor vehicles”) is among the most globalised. However, the knowledge intensity as indicated by the number of technological fields shows only a modest increase in most sectors (Figure 6).

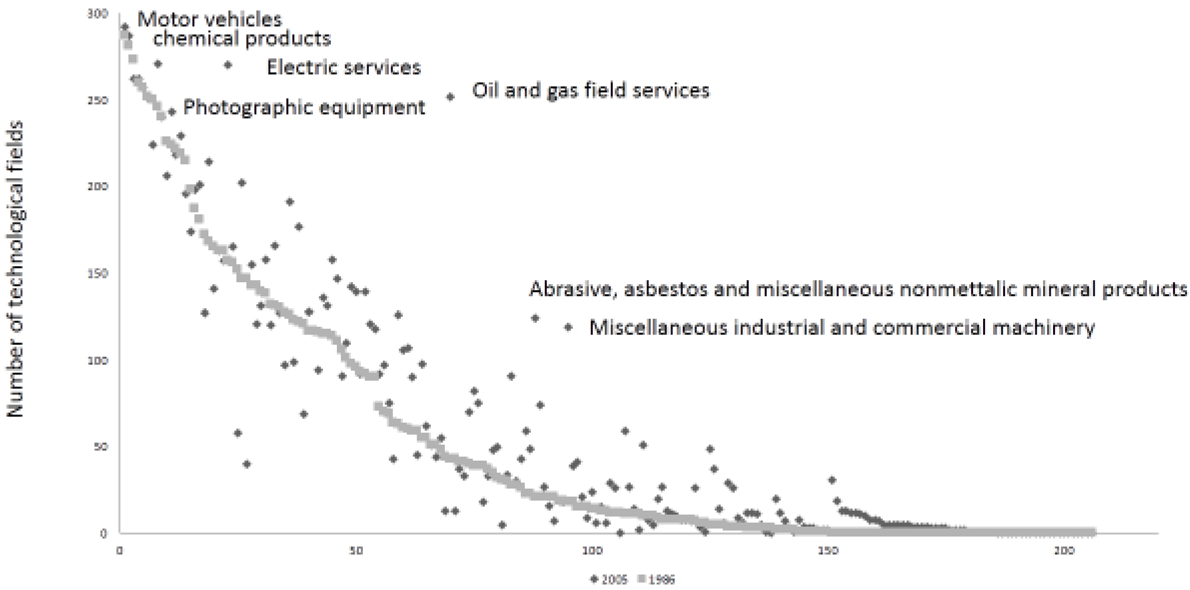


Figure 6. The number of technological fields per sector in 1986 and 2005.

The most knowledge intensive sectors, as indicated by the *variety* of technological fields involved are “Motor vehicles and motor vehicle equipment “, “Miscellaneous chemical products” and “Electrical industrial apparatus”. The largest increase in number of technological fields can be found in the sectors of “Oil and gas field services” (+209), “Electric services” (+113).

While the technological variety increases rapidly, only a limited number of sectors draw knowledge from a wider variety of technological fields. Thus, technological diversification is greater than sectoral diversification. Sectors rely on wide range of technologies in order to develop and produce products and services. Thus, most sectors could be labelled multi-technology, even if they are specialised in just one line of business (Granstrand, 1998).

The number of sectors associated with technological fields shows a relatively flat distribution (Figure 7).

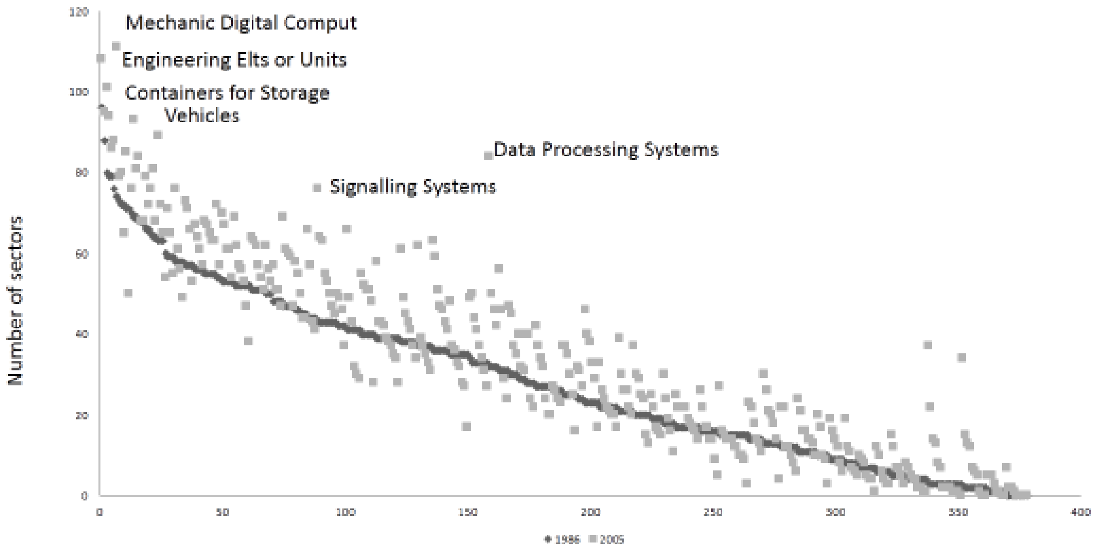


Figure 7. The number of sectors per technological field in 1986 and 2005.

On average, there is a modest increase visible in the number of sectors associated with each technological field between 1986 and 2005. In general, many ICT related fields have become more general purpose with respect to the number of sectors involved in knowledge production activities.

Mutual information between different dimensions

The previous section highlighted the uneven and 'spiky' distribution of technological knowledge production among countries, technologies and sectors. The entropy value of the distribution of patents can be calculated from the distribution of occurrences along each dimension, and any combination of dimensions. Figure 8 shows the Entropy values (H) of distributions over countries, sectors and technologies between 1986 and 2005.

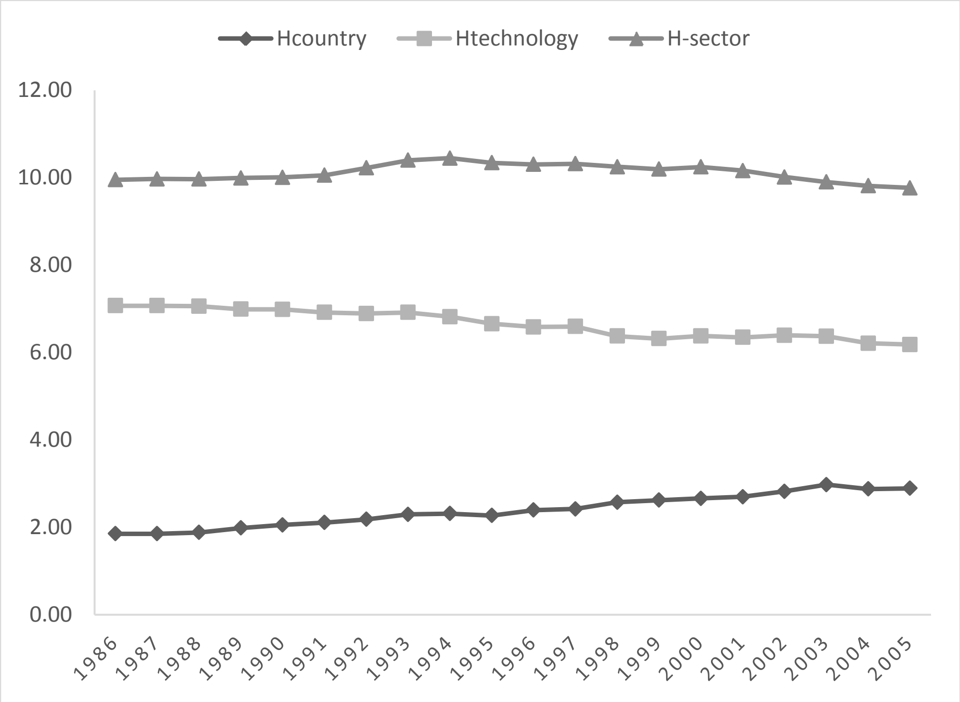


Figure 8. Entropy values (H) of distributions over countries, sectors and technologies between 1986 and 2005.

The entropy analyses show that corporate invention is increasingly diverse from a geographical dimension. However, diversity in the technological dimension is decreasing. In line with the results discussed above, this suggests that corporate invention is relatively increasingly concentrated The sectoral distributions show a stable pattern in the period under study.

In the next step, we quantify the mutual information between the dimensions through calculation of the Transmission values.

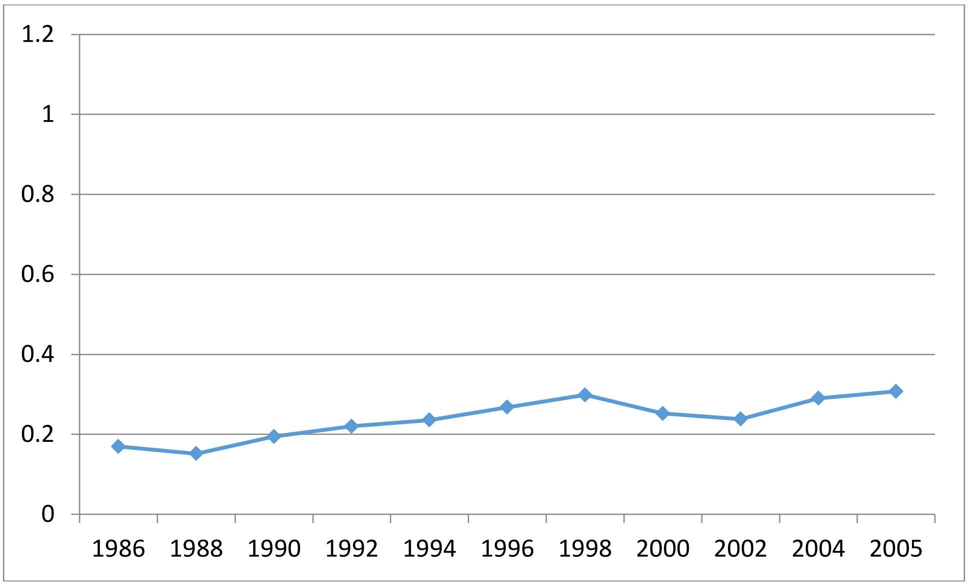


Figure 9. Transmission between technologies and countries between 1986 and 2005.

The mutual information content between technologies and countries slows a slight increase in the period under study. However, the overall values remain low, indicating a low degree of coupling between geography and field of technological invention.

However, individual countries exhibit different patterns of technological specialisation (figure 10). The USA shows a stable pattern of a high level of specialisation, while Germany shows a slow decline. Korea joins the USA as most technologically specialised in later years.

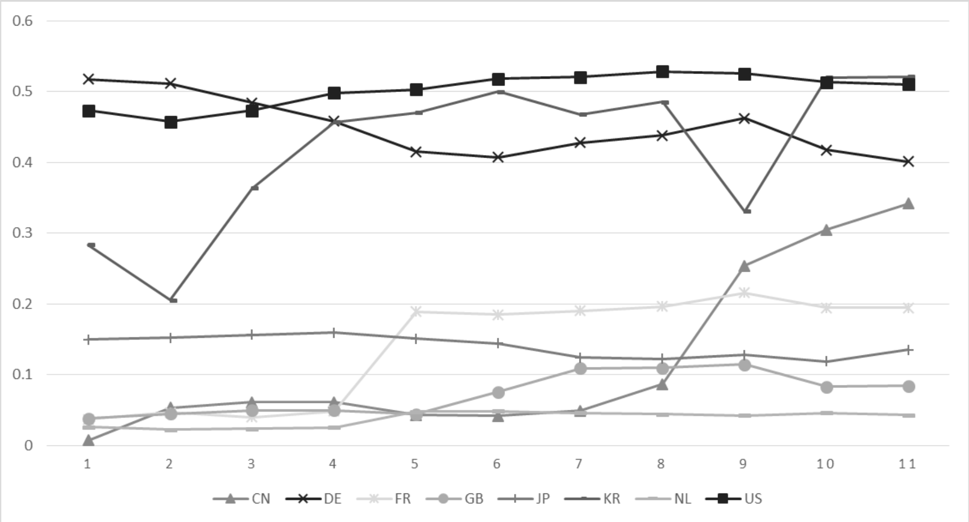


Figure 10. Transmission between technologies and selected countries between 1986 and 2005.

The mutual information between sectors and countries is considerably higher than between countries and technologies, and is slowly rising. This rise suggests that countries increasingly specialise along socio-economic dimensions rather than technological dimensions.

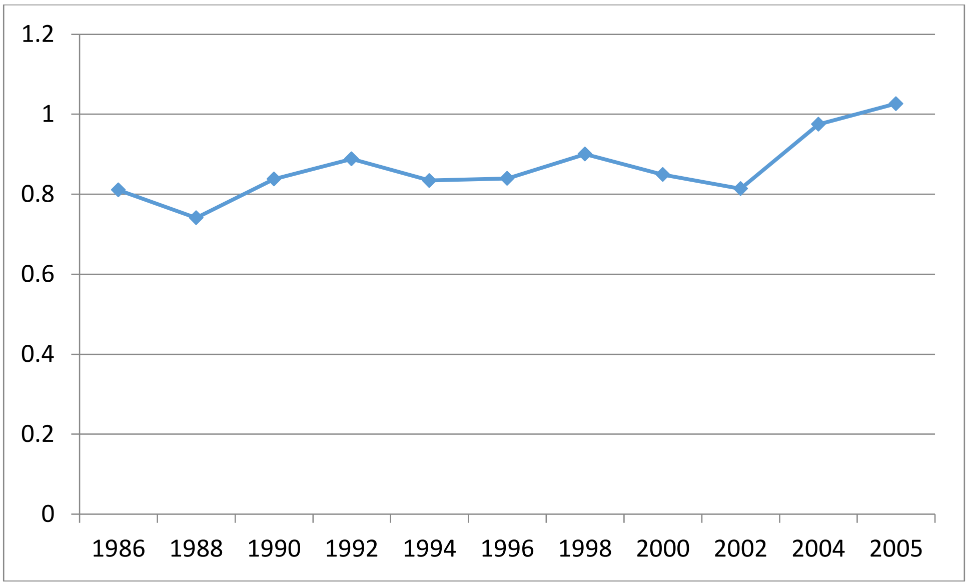


Figure 11. Transmission between sectors and countries between 1986 and 2005.

Again, individual countries show different patterns of sectoral specialisation. Germany exhibits a very high level of sectoral specialisation, which slowly declines between 1986 and 2005 (Figure 12). China shows a strong increase in sectoral specificity. France shows a strong increase in sectoral specialisation between 1994 and 996, in line with the observed increase in technological specialisation in the same period.

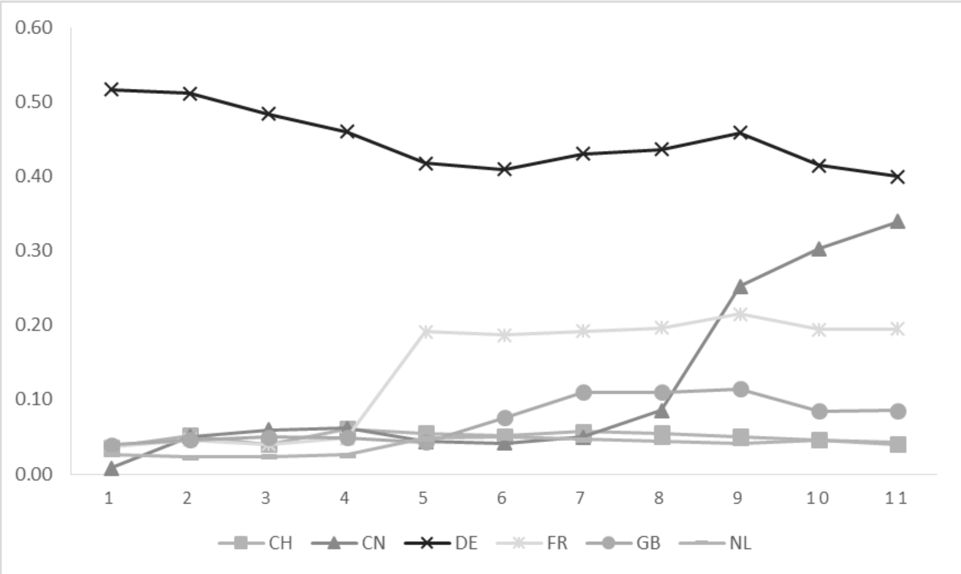


Figure 12. Transmission between sectors and selected countries between 1986 and 2005.

The mutual information between sectors and technologies (Figure 13) is considerably higher than between countries and technologies, but lower than between countries and sectors. The stable pattern suggests that the growth of corporate knowledge production remains equally distributed over (growing) sectors. As shown before, the profile of technological diversification of sectors is rather stable. It changes slowly over time as a consequence of the inertia of specialisation, incremental changes in knowledge production and modifications in firms' competencies (Cantwell, 1999).

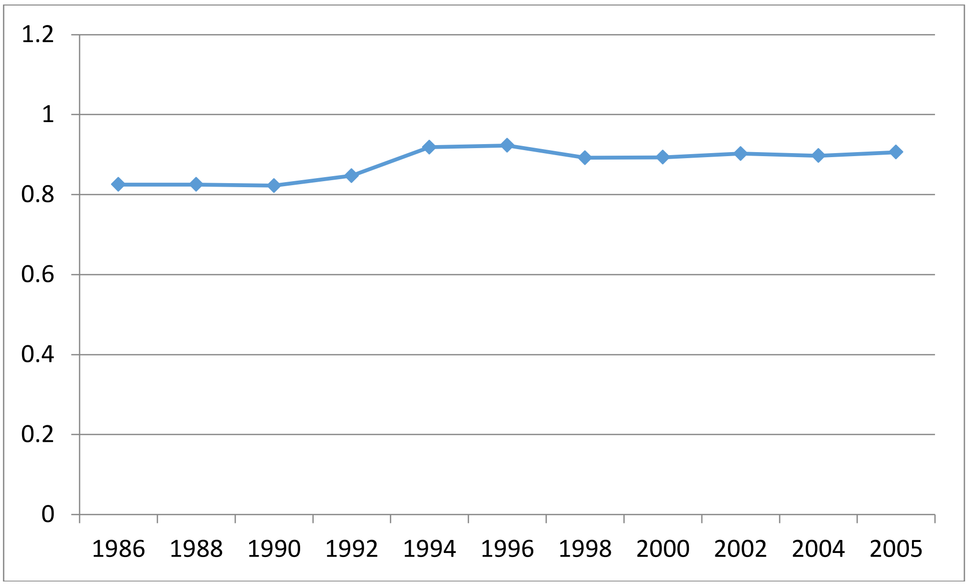


Figure 13. Transmission between technologies and sectors between 1986 and 2005.

Conclusion

Geographically, the past decades have seen a remaking of the global map of world corporate invention. Especially, Korea and China have established themselves at the top of the corporate invention rankings. Almost all sectors show an increase in patenting activities in the period under study. However, the largest growth takes place in a limited number of sectors, most notably in ICT related sectors. Likewise, the growth of technological knowledge production is unevenly distributed over technological fields and can be attributed to a limited number of fields, mostly related to ICT.

Most countries expanded their technological capabilities, as indicated by the number of technological fields. The largest increase in number of technologies occurs in the EU countries, Japan and the US as well as in emerging economies such as Korea, China, Taiwan, India and Russia. Technologies are truly globalised. A relatively flat distribution is visible the number of countries contributing to technological fields. Knowledge production in the large majority of technological fields occur in more than 20 countries. The increase in sectoral activities is much more concentrated in the traditional Triadic countries than the increase in technological activities, with the exception of China. The most globalised sectors are all related to the broader ICT industry, and the automotive industry.

On average, there is a modest increase visible in the number of sectors associated with each technological field between 1986 and 2005. In general, many ICT related fields have become more general purpose with respect to the number of sectors involved in knowledge production activities. The increasingly many-to-many correspondence between products and technologies results in the emergence of multi-product (generic, general purpose) technologies and multi-technology products that require closer association among sectors and technological fields. These development are largely limited to ICT.

The increase in sectoral activities was shown to be much more concentrated than the increase in technological activities. The mutual information between sectors and countries is considerably higher than between countries and technologies, and is slowly rising. This rise suggests that countries increasingly specialise along socio-economic (‘sectoral’) dimensions rather than technological dimensions.

The mutual information between sectors and technologies is considerably higher than between countries and technologies, but lower than between countries and sectors. The stable pattern suggests that the growth of corporate knowledge production remains equally distributed over (growing) sectors.

References

Arthur, W. B. (2007). *The Nature of Technology: What it is and How it Evolves*. New York: The Free Press.

Boschma, R. A. (2005). Proximity and innovation. A critical assessment. *Regional Studies*, *39*(1), 61–74.

Cantwell, J. (1999). Technological globalisation and innovative centres: the role of corporate technological leadership and locational hierarchy1. *Research Policy*, *28*(2-3), 119–144.

Florida, R. (2005). The World Is Spiky. *THE ATLANTIC MONTHLY*.

Granstrand, O. (1998). Towards a theory of the technology-based firm. *Research Policy*, *27*(5), 465–489.

Heimeriks, G. (2012). Interdisciplinarity in biotechnology, genomics and nanotechnology. *Science and Public Policy*, *40*(1), 97–112.

Heimeriks, G., & Boschma, R. (2014). The path- and place-dependent nature of scientific knowledge production in biotech 1986-2008. *Journal of Economic Geography*, *14*(2)

Leydesdorff, L. (2010). The Knowledge-Based Economy and the Triple Helix Model. *Annual Review of Information Science and Technology*, *44*, 367–417.

Narula, R., & Zanfei, A. (2005). The Oxford Handbook of Innovation. In J. Fagerberg, D. C. Mowery, & R. R. Nelson (Eds.), *The Oxford Handbook of Innovation* (pp. 318–345). Oxford: Oxford University Press.

Pfeffer, J., & Salancik, G. R. (2003). *The external control of organizations: A resource dependence perspective* (2nd Edn.). Stanford, California: Stanford University Press.

Theil, H. (1972). *Statistical* Decomposition *Analysis* (p. Amsterdam: North–Holland).

WIPO (2008). World Patent Report – A Statistical Review. Available at http://www.wipo.int/export/sites/www/freepublications/en/patents/931/wipo_pub_931_2008.pdf

Is the publication behavior of Danish researchers affected by the national Danish publication indicator? A preliminary analysis

Dorte Henriksen* and Jesper Wiborg Schneider*

*dh@ps.au.dk; jws@ps.au.dk
Danish Centre for Studies in Research & Research Policy, Department of Political Science & Government, Aarhus University, Bartholins Allé 7, Aarhus C, DK-8000

Introduction

The aim of this study is to investigate what are the effects of the national Danish publication indicator on the Danish researchers' publication behavior? Research evaluations have become a regular phenomenon at universities and research institutions and governments are increasingly using bibliometric indicators to allocate funds and increase research performance (Hicks, 2012), while ignoring the possibility of unintended effects on the research community and the scientific communication system (Weingart, 2005). Studies have shown that indicators can have a negative steering effect on the publication behavior of researchers (Butler, 2003). The Danish government introduced in 2009 the national Danish publication indicator (NDPI). The goal of the indicator is to measure and assess the Danish research productivity, besides motivate the researchers to only published in prestige and acknowledge publication channels (FIVU, 2013). The indicator assign points according to an authority list of journals, publishers and conference series and 25% of the universities basic funding (approx. 2.05B DKR) is allocated via the indicator. The further allocation of funds to the departments is decided by the university.

This poster presents a preliminary analysis of the effects of the NDPI on the Danish researchers' publication behavior. The purpose of the study is to illuminate the nuances of the process of producing a publication, the different choices made in the process, and how external factors, such as bibliometric indicators, may influence the publication of research results. The study is part of an ongoing Danish research project Governance, Funding and Performance of Universities[1].

Method

In the period October 2013-february 2014 we conducted 43 interviews with Danish researchers from six universities in Denmark; Aalborg University (AAU), Aarhus University (AU), University of Copenhagen (KU), University of Southern Denmark (SDU), Technical University of Denmark (DTU) and Copenhagen Business School (CBS). We interviewed 8-12 researchers from each of the four main academic research areas; Humanities (HUM), Social Sciences (SOC), Health & Life sciences (HL) and the Science & Technologies (ST). All the researchers who participated in the interviews had completed a questionnaire about the NDPI in 2011 and had published a point-receiving publication in the period 2009-2011. The participants consisted of 1 PhD student, 9 post docs, 24 associate professors and 9 professors.

---

1 www.sdu.dk/en/om_sdu/institutter_centre/i_statskundskab/forskning/forskningsprojekter/university+governance

The interviews were semi-structured and lasted an hour on average. The interviews had one publication of the researchers as a point of departure. These specific publications were chosen using the following criteria: the publication had to be recent, it had to be published in a journal or a publisher on the NDPI publishing authority list, and if possible the researcher had to be first or last author.

The interviews were semi-structured and lasted an hour on average. The interviews had one publication of the researchers as a point of departure. These specific publications were chosen using the following criteria: the publication had to be recent, it had to be published in a journal or at a publisher on the NDPI publishing authority list, and if possible the researcher had to be first or last author.

The interviews were structured in four parts with questions about:

1. Publishing Process
    a. Their own publication
    b. General questions about collaboration, authorship and publishing
2. Publication practice and culture
3. Publication pressure
4. NDPI

The researchers were not directly ask about the NDPI during the interviews, the hypothesis were that if researchers do not mention the indicator, they are probably not affected by the indicator.

Results

The Danish institutions in the study have very different approaches to the NDPI. AU & KU do not officially use the NDPI and it's up to the departments if they chose to pay attention to the indicator. CBS have created awareness of the NDPI, but prefer international publication ranking lists. The boards of directors at AAU & SDU have decided that some of the allocation of the basic funding to the faculties and departments depend on the points obtain in the NDPI. A summary of the researchers' knowledge of NPDI can be found in table 1.

Table 1. Researchers' knowledge of NPDI.

University	**Research Area**	**Participants**	**Awareness of NDPI**	**Checked the authority list**	**Could recall the NDPI after introduction**
AU	HL	4	0	0	4
AU	HUM	4	4	1	*
AU	SOC	4	1	1	3
CBS	SOC	4	4	4	*
DTU	ST	4	1	0	1
KU	HL	4	0	0	2
KU	HUM	4	4	3	*
SDU	SOC	4	2	2	2
SDU	ST	4	4	4	*
AAU	HUM	3	3	3	*
AAU	ST	4	4	4	*
Total		43	27	22	13

All the researchers at the humanity faculties (AAU, AU, KU) were aware of the indicator and the majority had at least once checked the authority list before choosing publication channel and describe an increasing focus on publishing articles. Some of the researchers thought that the value of books and book chapters in the indicator were too low, and it were suggested that the number of pages should influence the number of points. The researchers from AAU were extremely aware of the indicator, because they could lose research time by not obtaining enough points and therefore adjusted their publication behavior to the indicator.

9 out of 12 researchers from the science and technology faculties knew the indicator. The eight researchers from SDU and AAU had checked the authority list of journals, before publishing because it was important for the allocation of funds to their department. 2 of the researchers from DTU had never heard of the indicator, while one of the other researchers could recall some mention of it a couple of years ago.

The 4 social science researchers at CBS knew the indicator and checked the authority list before choosing a publication channel. Only 3 of the other 8 social science researchers (AU, SDU) mention the indicator before asked about the indicator. The 3 researchers had all selected publication channels based on the authority list.

The 8 researchers in the health and life sciences (KU, AU) did not mention the indicator during the interviews, though half of them had heard about it. The indicator did not affect their choice of publication channel and they did not perceive it as being of any importance. Some considered it to be another administrative hassle.

Conclusion

The impact of the indicator on the researchers' publication behavior depends on the research area and how the universities managed the indicator. The researchers at universities, that use the indicator to allocate funds, where more focused on publishing accordingly to the lists and obtaining points.

References

Butler, L. (2003). Modifying publication practices in response to funding formulas. *Research Evaluation, 12*(1), 39-46.

FIVU. (2013, 25-01.2013). Spørgsmål og svar om den bibliometriske forskningsindikator. Retrieved 24-01, 2013, from http://fivu.dk/forskning-og-innovation/statistik-og-analyser/den-bibliometriske-forskningsindikator/sporgsmal-og-svar

Hicks, D. (2012). Performance-based university research funding systems. *Research Policy, 41*(2), 251-261.

Weingart, P. (2005). Impact of bibliometrics upon the science system: Inadvertent consequences? *Scientometrics, 62*(1), 117-131.

Using Patent Indicators to Evaluate the Strategic Priorities of Public Research Institutions: An exploratory study

Yuen-Ping Ho* and Poh-Kam Wong**

* yuenping@nus.edu.sg
Entrepreneurship Centre, National University of Singapore, 21 Heng Mui Keng Terrace, Singapore 119613

** pohkam@nus.edu.sg
Entrepreneurship Centre, National University of Singapore, 21 Heng Mui Keng Terrace, Singapore 119613

Introduction

As a key actor in the Triple Helix of University-Industry-Government interactions in an economy (Etzkowitz and Leydesdorff, 1999), Public Research Institutes (PRIs) play an important role in the national innovation system (NIS). As discovered by Mazzoleni and Nelson (2007), PRIs have contributed significantly in the S&T catch up process, especially in the East Asian NIEs. The importance of PRIs lies in the basic rationale for their existence: PRIs perform essential R&D functions in the NIS that cannot be efficiently performed by enterprises and universities, whether due to resource constraints or strategic reasons. In this regard, an important role of PRIs is to bridge academia and industry through applied and translational research. Examples of other functions of PRIs include industry or technology-specific research, contract research aligned to national industrial development strategies, public-interest research and hosting critical large scale infrastructures.

The functions that a PRI is expected to perform is dictated by the policy emphasis of decision-makers responsible for governance of the public research agenda. The work of Sanz-Menendez and colleagues documents how policy changes and government intervention has shaped the strategies and management practices of PRIs, and evolved new forms of PRIs (Cruz-Castro and Sanz-Menendez, 2007; Cruz-Castro, Sanz-Menendez and Martinez, 2011; Sanz-Menendez and Cruz-Castro, 2003). The OECD (2011) identified four "ideal types" of public research organizations, each with a different main focus.

To fulfil the expected roles premised on its main focus, each PRI develops organizational strategies aligned to its resources and missions. The literature has documented several proposed approaches for priority-setting in public research at the national or innovation system level (Stewart, 1995; Klerkx and Leeuwis, 2009). There is no equivalent focus on the strategy-formulation process of PRIs at the organizational level, or the tools for PRIs to benchmark their adopted strategic positions. This is in part due to the lack of easily-available information and consistently-measured data on PRI strategies. In this paper, we propose a framework which uses patent indicators to evaluate the strategic priorities of PRIs, allowing for comparisons across different organizations and time periods.

Patents as Indicator of Strategy

A patent represents a significant advancement made by inventors – and by extension, their affiliated organizations – in a technology field. As such, patents data provide a window to understanding patterns of technology development and accumulation. Scientometrics indicators based on patent data have been extensively used to measure innovation in organizations in many different contexts.

An organization's patent portfolio is a treasure trove of information about its research and innovation activities. Porter and Newman (2004) demonstrate that patent analysis plays an important role in Competitive Technological Intelligence, in which firms attempt to discern the technological trajectory and future strategic moves of competitors. While corporate-style competition is less of a salient consideration in the context of PRIs, the indicators that can be developed from analysing patent portfolios are useful to all types of organizations, whether private or public sector, or profit or non-profit oriented. Ernst (2003) presents a wide range of patents-based indicators that inform an organization's strategic management of technology. Adopting a different approach, Debackere and Luwel (2004) show that an organization's stock of patents can be used to benchmark Science and Technology (S&T) portfolios, following the portfolio management models developed in the wake of studies proposing the concept of technological S-curves (Martino, 1983).

Despite the wealth of information to be mined from patent portfolio analysis, research on PRIs has not fully exploited this data source. In studies on PRIs, patents and publication statistics typically form the basis of performance evaluation, used as output indicators to assess the efficiency or productivity of PRIs (Coccia, 2004; Matsumoto et al, 2010). In this paper, we posit that beyond benchmarking performance, a PRI's patent portfolio reflects its R&D and innovation strategies. We develop a framework comprising a suite of indicators to analyse the patent portfolios of four PRIs. The analysis aims to evaluate PRI strategic priorities and how they differ among the PRIs, and to detect changes or shifts in R&D focus over time. This framework is useful for researchers and PRI stakeholders to understand whether stated strategies are aligned to R&D outcomes as reflected in patent portfolios.

Method

Framework for dimensions of PRI strategic priorities

We develop a framework which incorporates six dimensions of PRI strategic priorities, as listed below. These dimensions represent key decisions made by PRIs when fulfilling their function in the NIS. Due to the unique role played by PRIs, these decisions often require resolution of tensions arising from their position in the Triple Helix of University-Industry-Government. While each dimension is framed as a choice between two contrasting strategic orientations, it is noted that organizations may adopt middle-ground positions that straddle both ends of the strategy spectrum.

i) Industry-pull vs Technology-push
ii) High vs Low Science-based intensity
iii) Quantitative growth vs Quality improvement
iv) Specialization vs Diversification
v) Indigenous capabilities vs International collaboration
vi) Autonomous control vs Joint ownership

Scorecard of Patents-based Indicators

We identify relevant patents-based indicators, as summarized in Figure 1, to form an institutional patent scorecard. Figure 2 illustrates how the patent indicators are mapped onto the dimensions of strategic priorities in our framework.

The selected indicators are largely drawn from established measures in the literature, including several reported in The Patent Scorecard ™ published by the Patent Board, formerly published in MIT's Technology Review. Indicators drawn from The Patent Scorecard ™ are flagged with an asterisk below.

a) Number of patents is a simple measure of patent counts
b) Growth of portfolio is measured as average annual growth rate in a PRI's patent stock
c) Average number of forward citations measures the quality of a patent by the number of times it is cited as prior art by subsequent patents. To account for the issue of truncation, this indicator is computed within 5 years of the referenced patent's date of grant.
d) Technology Impact Index (TII) is the share of a PRI's patents in the pool of highly cited patents relative to its share in total patents. This draws conceptually on King's (2004) measure of publication quality. A highly cited patent is one which is among the top 5% most frequently cited in its cohort. A cohort is defined by year of grant and technology class.
e) Technology Cycle Time * is an indicator of a PRI's speed in turning leading edge technology into IP. It is defined as the median age of patents cited as prior art by the reference patent.
f) Current Impact Index (CII) * is measured by examining how often a PRI's patents from the previous five years are cited as prior art in the current year's global batch of patents. CII is a relative measure with 1 representing the global average.
g) Bibliographic Citations Ratio (BCR) is the ratio of Non-Patent References to total backward citations reflecting a patent's prior art. This is a proxy measure for scientific content in a patent and is further discussed below.
h) Share of Science-based Patents is another measure of scientific content and is the share of patents with BCR value higher than 50%.
i) Technology Specialization is measured using the Herfindahl Index which quantifies the degree to which a portfolio is specialized or concentrated in a small number of technology areas versus being distributed across a range of technologies.
j) Share of Complex Multi-Technology Patents is an original indicator which we developed to assess the extent of technological complexity in a patents portfolio. A "complex" patent is one which is classified in multiple technological areas in its technology classification field. The derivation of this indicator is described more fully below.
k) Co-patenting quantifies the degree to which a PRI engages in external collaborations resulting in joint creation and ownership of IP.
l) International Co-invention quantifies the degree to which a PRI's inventors engage in collaborative R&D with inventors outside the home economy.

The Bibliographic Citations Ratio (BCR), and by extension the Share of Science-based Patents, use non-patent references (NPRs) in patents to represent scientific content. Narin, Hamilton & Olivastro (1997) proposed that NPRs directly signal the influence of science on technology and can therefore be used to measure the scientific intensity in patents. However, subsequent research shows that the relationship between NPR's and the patented technology is not as straightforward. Meyer (2001) and Tjissen (2001) conclude that NPRs should not be seen as an indicator of the direct link between science and technology. From a sample of approximately 5000 NPRs extracted from the USPTO and EPO, Callaert et al (2006) found that NPRs comprise a mix of both scientific knowledge and technological information. As such, we view the two NPR-based measures as indicators of science-relatedness or science-related content of patents, rather than direct indicators of scientific intensity.

The indicator Share of Complex Multi-Technology Patents is based on the notion of technological complexity (von Graevenitz et al., 2008). The complexity of technologies is often framed in terms of to the industry or product in which technologies are applied. Cohen et al (2001) suggested a breakdown between discrete and complex industries/products which has been adopted in studies of patenting behaviour by firms (Hall, 2005; von Graevenitz et al., 2007). The concept of "complexity" is founded on the idea that technologies can be multi-faceted in nature, whether it is in their application or content. There is however no measure of technological complexity that directly encapsulates this idea of multiple facets. These existing approaches in the literature address technological complexity at the level of the firm or industry. We propose a measure of the *technological complexity of patents* based on the technological nature of patents themselves, rather than the technologies applied in patent-owning firms or industries.

In a patent document, the technologies germane to the invention are summarized in the technology class field. We propose that a complex patent is one which is classified in multiple (more than one) technological areas. In this paper, a technological area is defined at the one-digit level of the NBER patent classification, which aggregates the detailed US Patent Classification (USPC) schematic (Hall et al, 2001). At the one-digit level, the NBER classification identifies six technological areas: (i) Chemical, (ii) Computers & Communications, (iii) Drugs & Medical, (iv) Electrical & Electronics, (v) Mechanical, and (vi) Others. A complex patent is identified if it has technology classes spanning at least two of these 6 areas. The Share of Complex Multi-Technology Patents indicates the proportion of the PRI's portfolio which comprises complex patents. Higher share of complex patents indicates that the PRI is producing inventions with greater opportunities for generating cross-sector economic activity. As found by Cohen et al (2001) and Ziedonis and Hall (2001), patents in complex industries are important for cross-licensing and trading purposes.

Figure 1: Indicators in Institutional Patent Scorecard

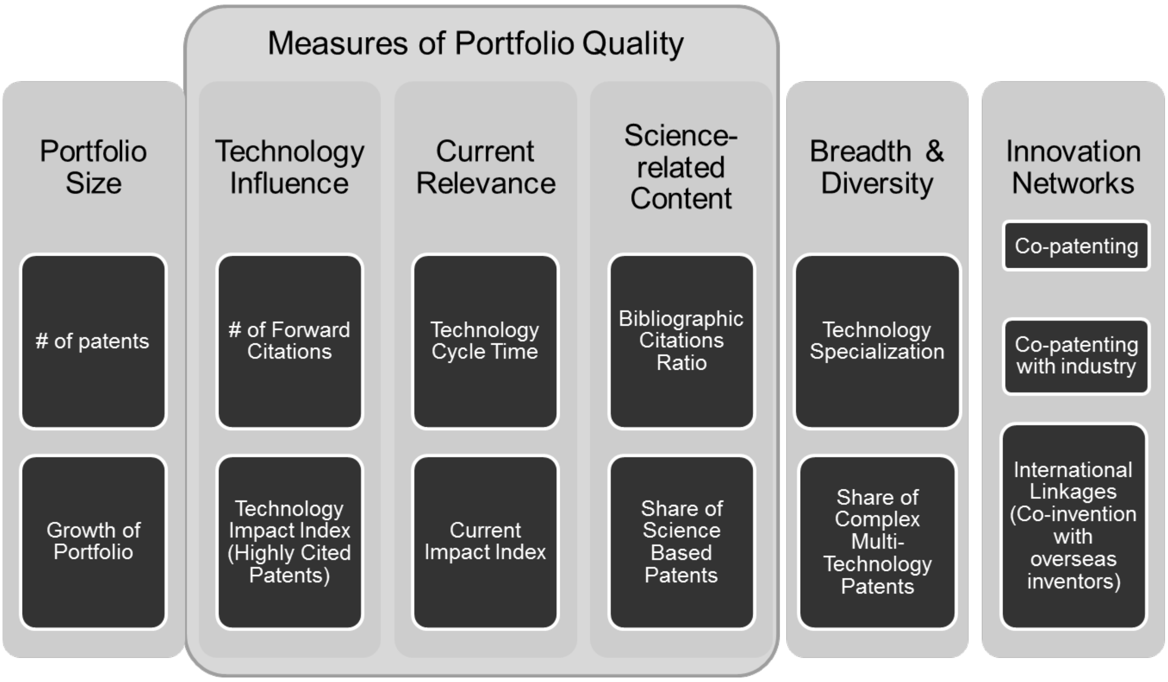
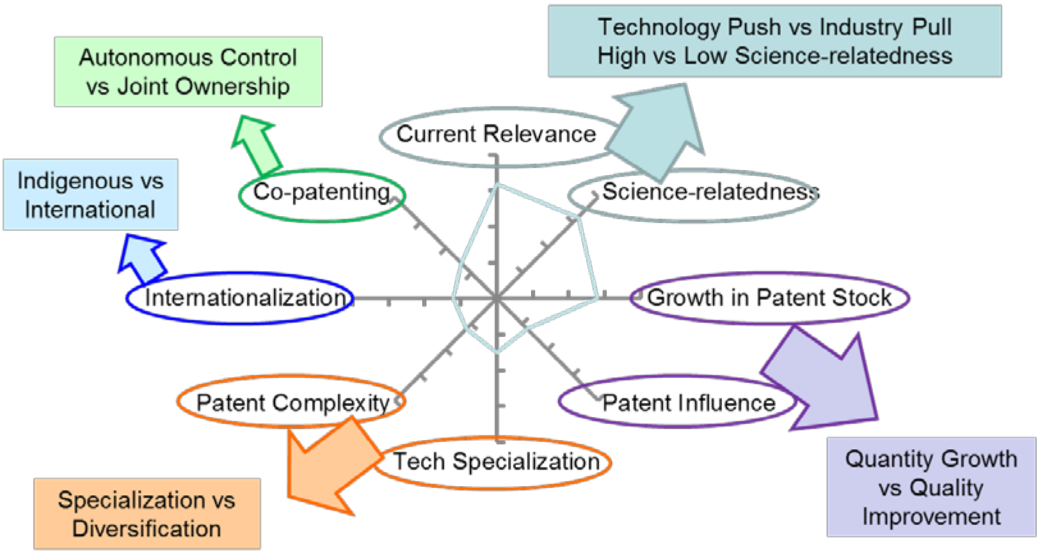


Figure 2: Mapping Patent Indicators to Dimensions of Strategic Priorities



Construction of PRIs' Patent Portfolios
We compiled patent data for 4 PRIs, three of which are from East Asian NIEs and 1 from a developed economy. The 4 PRIs were selected to cover a range of organizations in terms of age and size, as shown in Table 1.

- ASTAR (Singapore): Agency for Science, Technology and Research
- ITRI (Taiwan): Industrial Technology Research Institute
- KIST (South Korea): Korea Institute of Science and Technology
- CSIRO (Australia): Commonwealth Scientific and Industrial Research Organization

Table 1: Profile of Public Research Institutes

PRI	Year Formed	USPTO Patents, 1978-2011	Scientific Publications, 1978-2011 (SSCI & SCI-E)	R&D Spending, 2011/12
ASTAR (Singapore)	1991 (as NSTB)	497	3,832	SGD 976 m
ITRI (Taiwan)	1973	4,484	5,521	Na
KIST (South Korea)	1966	730	14,476	Na
CSIRO (Australia)	1926	750	51,142	AUD 1130.4 m

For each PRI, we extracted patents of invention granted by the United States Patents & Trademark Office (USPTO) from January 1978 to December 2011. The portfolio for each PRI is constructed by identifying all patents with the PRI as a named assignee. This would include all patents co-assigned to other organizations, but excludes patents granted to individual researchers where the PRI is not named as an assignee. The final numbers of patents extracted are reported in Table 1.

Institution-Level Results

We computed the patent indicators for each of the four PRIs and compiled the figures into a scorecard as presented in Table 2. From the first few rows of Table 2, we note that almost all of ASTAR's patents were granted in the last 10 years, while 60% of CSIRO's patents were granted prior to 2000. The pattern of patent production is more consistent in ITRI and KIST throughout the last 3 decades.

There are several notable differences between the PRIs as seen in Table 1. ASTAR has the highest average citations and is the only organization with Current Impact Index above the global average of 1. ITRI has the shortest Technology Cycle Time. CSIRO has the most science-oriented patents, while ITRI and KIST have patents with relatively low science content. To examine these inter-organizational differences in greater detail, we further break down selected indicators by technology field, as reported in the next section.

When plotted graphically in the form of a strategic priority map (Figure 3), these indicators collectively reveal the strategic orientation of the PRI along the dimensions of our framework. To construct the maps in Figure 3, we firstly standardized the figures such that each indicator has a mean of 0 and standard deviation of 1 across the sample of 4 PRIs. The dotted line in the map represents the group average of 0. By comparing a PRI's map (depicted by the solid line) with the group mean, we can benchmark the PRI's strategic priorities against those of its counterparts. Where two sets of indicators represent two contrasting positions in a strategy dimension (for example, patenting growth vs quality improvement), the relative values of the two indicators will show the PRI's priority.

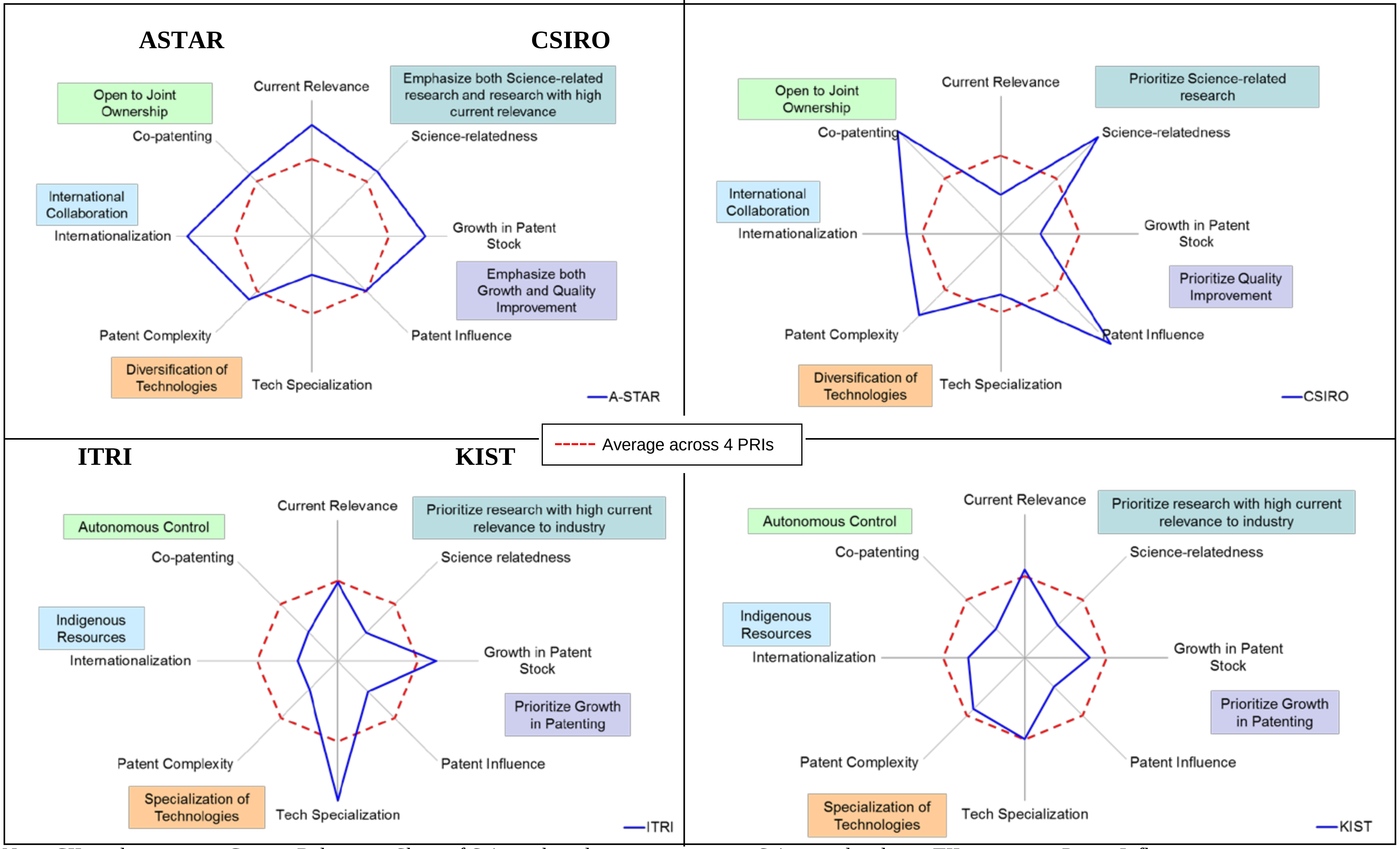
Figure 3 summarizes the strategic priorities of all 4 PRIs in the dimensions of our framework, using portfolios of patents granted in 2006-2011. ITRI and CSIRO present an interesting study in contrasts. ITRI prioritizes R&D with high industry relevance and lower scientific content, suggesting an industry-pull strategy. Patenting growth is prioritized over improving the quality of patents, and the portfolio is specialized rather than diversified. The ITRI approach also emphasizes internal resources and asserting autonomous control on ownership. On the other hand, CSIRO boasts high science-intensity in its patents and relatively low current industry relevance, suggesting a technology-push approach. There is emphasis on producing high quality patents over growth in the portfolio. CSIRO patents tend to be diversified across multiple fields, are invented with international collaborators and jointly-owned. The map for ASTAR indicates an approach which attempts to balance multiple priorities. ASTAR patents have relatively high current relevance as well as scientific content. While there is high patenting growth, the quality of ASTAR patents is also above average.

Figure 4 shows how the framework can be used to trace shifts in priorities over time. The PRI's priority map for 2001-2005 is depicted by the thinner line and for the later period 2006-2011 by the thicker line. The difference between the two lines represents changes in strategic priorities. To illustrate, ASTAR's maps indicate an increased emphasis on science-based research in the later period, and a shift towards greater diversification in the portfolio as seen by changes along the Tech Specialization and Patent Complexity axes. CSIRO's maps show a significant shift towards quality improvement as a priority, while ITRI's emphasis shifted towards growth. ITRI and KIST also oriented towards greater specialization in their patents portfolios in the last 5 years over the previous period.

Table 2: Patent Scorecard for selected PRIs

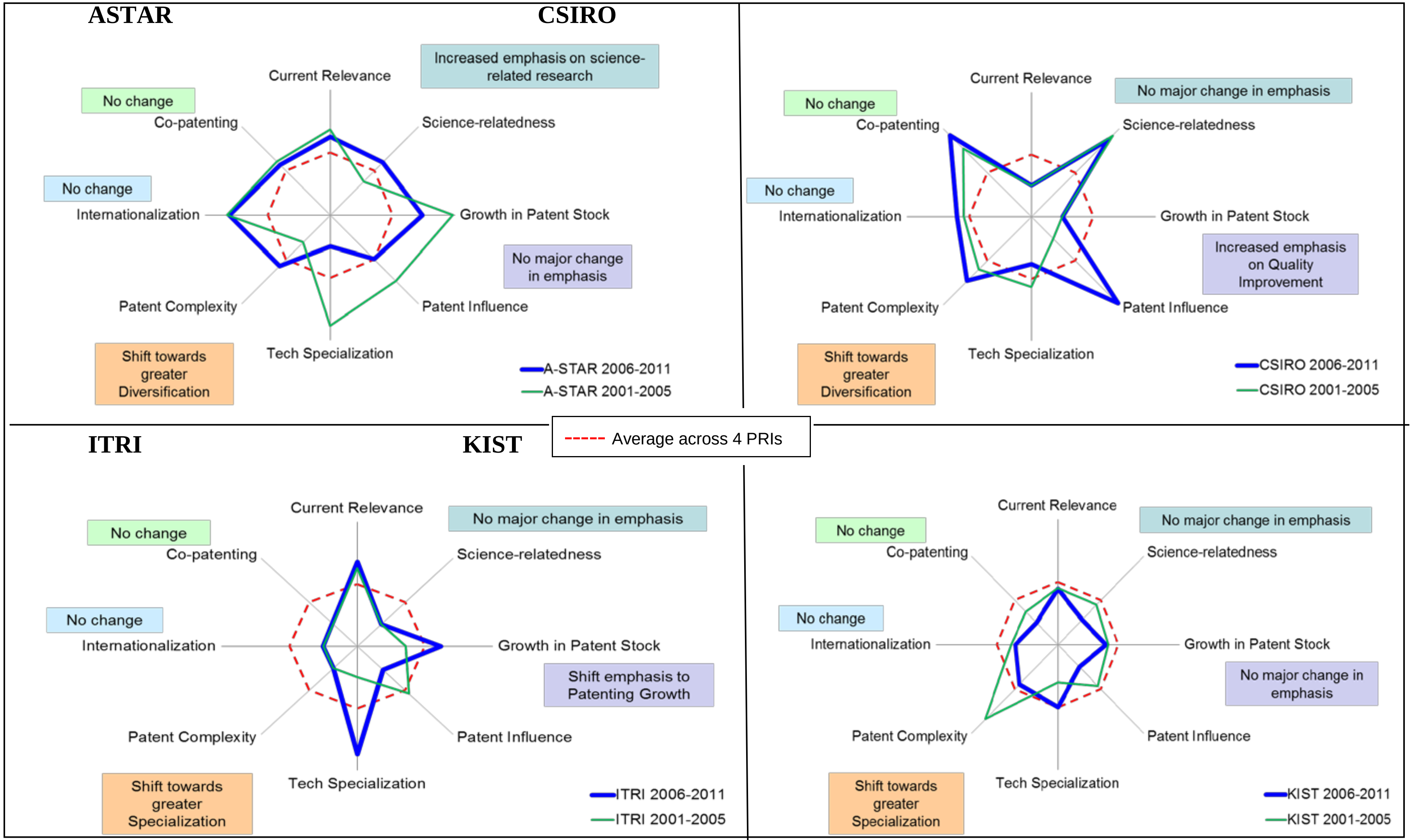
	ASTAR	CSIRO	ITRI	KIST
Cumulative USPTO Invention Patent Stock, 1978-11	497	750	4,484	730
1978-2000	18	450	1,427	272
2001-2005	205	160	1,006	222
2006-2011	274	140	2,051	236
Growth in USPTO Invention Patent Stock (% p.a.)				
2001-2005	36.5	5.7	10.3	12.5
2006-2011	13.5	3.0	10.9	6.2
Average Forward Citations within 5 years				
2002-2006	5.61	1.42	3.91	2.98
Technology Impact Index – Highly Cited Patents				
2001-2005	1.36	0.38	0.96	0.81
2006-2011	1.02	2.00	0.56	0.53
Technology Cycle Time (in years)				
2001 to 2005	6.1	12.2	7.1	9.4
2006 to 2011	9.0	13.5	8.4	11.2
Current Impact Index 2011	1.24	0.15	0.71	0.83
Bibliographic Citations Index (%)				
2001-2005	16.15	37.94	5.15	18.43
2006-2011	35.51	48.95	9.35	13.50
Share of Science-based Patents (%)				
2001-2005	11.11	41.40	4.00	15.02
2006-2011	31.97	53.24	5.50	8.04
Herfindahl Index of Technology Specialization				
2001-2005	0.35	0.29	0.23	0.24
2006-2011	0.23	0.24	0.28	0.25
Share of Complex Multi-Technology Patents (%)				
2001-2005	23.56	30.22	22.21	35.75
2006-2011	27.01	33.13	14.54	22.03
Co-invention with overseas inventors (% of patents)				
2001-2005	31.4	16.9	1.7	8.6
2006-2011	25.3	16.4	1.6	5.7
Co-assignment with other organizations (% of patents)				
2001-2005	22.51	31.88	3.38	9.91
2006-2011	12.64	22.14	4.44	4.24
Co-assignment with industry partners (% of patents)				
2001-2005	13.61	21.88	3.08	8.11
2006-2011	5.78	17.14	2.97	2.12

Figure 3: Strategic Priorities Maps for Selected PRIs (2006-2011)



Note: CII used to measure Current Relevance, Share of Science-based patents to measure Science-relatedness, TII to measure Patent Influence

Figure 4: Change in Strategic Priorities for Selected PRIs



Note: Tech Cycle Time (reverse coded) used to measure Current Relevance, Share of Science-based patents to measure Science-relatedness, TII to measure Patent Influence

Comparisons across Technology Classes

Table 3 shows the technological composition of patent portfolios in the four PRIs and how it has changed over time. For each patent, the technology field is derived from the primary USPC code of the patent, mapped onto one of the six technology categories of the NBER classification scheme (Hall et al, 2002). There are clear differences in technology focus among the PRIs. Over 60% of CSIRO's patents are in the Drugs & Medical and Chemical areas, with the concentration in Chemical having increased in the most recent period. A-Star's portfolio has shifted away from a strong concentration in Electronics towards more equal representation of other classes across the board. Comparably, the composition of ITRI and KIST's portfolios has remained largely unchanged.

Table 3: Technological Composition of PRIs' Patent Portfolios

% distribution by sector	A-STAR		CSIRO		ITRI		KIST	
	2001-05 (N=205)	2006-11 (N=274)	2001-05 (N=160)	2006-11 (N=140)	2001-05 (N=1006)	2006-11 (N=2051)	2001-05 (N=222)	2006-11 (N=236)
Chemical	9	16.5	32.9	45.6	12.5	19.9	32.3	38.6
Computers & Communications (ICT)	22.2	31.9	7.1	4.4	29.6	21.6	7.5	5.5
Drugs & Medical	7.4	15	30.7	24.4	3.6	1.5	13.3	11.4
Electrical & Electronic	52.9	27.1	15	10	39.9	34.3	33.2	20.5
Mechanical	4.2	7.1	9.3	8.1	10.3	15.3	9.3	16.4
Others	4.2	2.3	5	7.5	4	7.4	4.4	7.7

We next examine three of the strategic dimensions in our framework – current relevance, science relatedness and patent influence - disaggregated by technology classes. This allows us to understand whether certain strategic orientations, such as CSIRO's emphasis on science-based patents, are driven by technology class-level effects.

Class-level details reveal that different PRIs achieved fast cycle times in different technology sector. As seen in Table 4, ITRI has shortest cycle time among the four PRIs in the two areas where it has the most patents – Electronics and ICT. This suggests that specialization in these two fields has led to accumulation of capabilities to respond to current technology trends. On the other end, CSIRO's cycle times are consistently slower than the other PRIs', with the exception of Drugs & Medical. In this field, CSIRO has the fastest cycle time in the group.

A-STAR's strong overall Current Impact Index is due to the high current relevance of its patents in 3 technology fields: ICT, Electronics and Mechanical. ITRI has relatively strong CII values in its major focus technologies of ICT and Electronics. Similarly, KIST's patents in its two main areas of Chemical and Electronics also have relatively strong current impact. As with the findings on cycle time, the CII results show that PRIs like KIST and ITRI concentrate R&D resources in current technologies. This is in line with their role as essentially Research Technology Organizations that develop and transfer public S&T to industry (OECD, 2011). In contrast, CSIRO's patents across all technologies have relatively low current relevance, as may be expected from a mission-oriented PRI which performs research in specific sectors in support of policy making (OECD, 2011).

Table 4: Current Relevance by Technology Class

	A-STAR	CSIRO	ITRI	KIST
Technology Cycle Time 2011 (in years)	9	13.5	8.4	11.2
By technology class				
Chemical	9.7	16.3	10.7	11.5
ICT	8.4	11.6	6.8	6.8
Drugs & Medical	9.5	8.8	11.2	10.8
Electrical & Electronic	8.7	12.1	7.7	10.1
Mechanical	12.1	18.6	10.8	15.7
Others	6.3	20.4	11.5	16.2
Current Impact Index (CII) 2011	1.24	0.15	0.71	0.83
By technology class				
Chemical	0.22	0.21	0.44	0.94
ICT	1.73	0.24	0.88	0.7
Drugs & Medical	0.3	0.22	0.3	0.47
Electrical & Electronic	1.84	0	0.73	1.17
Mechanical	1.25	0	0.77	0.23
Others	0	0	0.5	0

The difference in science-relatedness of PRIs' patents is partially due to the composition of their portfolios. As seen in Table 5, patents in Drugs & Medical tend to have higher shares of NPR citations. CSIRO, with almost one quarter of patents in this field, has a high overall share of science-based patents. Notwithstanding the composition effect, CSIRO and A-STAR have the highest proportion of science-related patents across all technology classes. Correspondingly, patents by ITRI and KIST have relatively lower proportion of science-based patents in all technologies, confirming the technology-oriented strategic priority of these PRIs.

Table 5: Science-Relatedness by Technology Class

	A-STAR	CSIRO	ITRI	KIST
Share of Science-Based Patents 2006-11 (%)	31.97	53.24	5.5	8.04
By technology class				
Chemical	69.0	55.6	7.3	4.2
Computers & Communications (ICT)	18.8	10.0	6.9	5.9
Drugs & Medical	73.7	93.0	30.9	20.7
Electrical & Electronic	11.3	19.0	3.1	8
Mechanical	10.5	23.1	0.5	9.5
Others	16.7	14.3	1.2	0

The strong overall performance of CSIRO in patent quality is due to disproportionately large shares of highly-cited patents in Chemical and Drugs, the two focus areas of CSIRO. This is in stark contrast with ITRI and KIST, where the patent influence in the fields of concentration is relatively low. ITRI and KIST prioritize current relevance and growth in these areas, rather than quality of patents. A different pattern in observed in A-STAR's portfolio, where patent influence is strong in Drugs & Medical, Mechanical and Electronics. Two of these fields - Mechanical and Drugs - are among the faster-growing in A-STAR's portfolio, while growth in Electronics has slowed (Table 3). This suggests that A-STAR's strategic orientation in the growth versus quality dimension is differentiated for various technologies. In Drugs and Electronics, quality and growth are jointly emphasized, while quality improvement in prioritized over quantitative growth in Electronics.

Table 6: Patent Influence by Technology Class

	A-STAR	**CSIRO**	**ITRI**	**KIST**
Technology Impact Index 2006-11 (TII)	1.02	2.0	0.56	0.53
By technology class				
Chemical	0.45	2.17	0.56	0.82
Computers & Communications (ICT)	0.94	0	0.34	0
Drugs & Medical	1	3.72	0	0
Electrical & Electronic	1.67	0.95	0.58	0.8
Mechanical	1.05	0	1.45	0
Others	0	0	0.49	0

Conclusion

The overall similarities and differences between the strategic priorities of the 4 PRIs are summarized by the difference indices reported in Table 7. The difference index is calculated as the aggregate of squared differences between PRI pairs in the patent scorecard vectors. Values for 2006-2011 are used to compute the difference indices in Table 7. ITRI and KIST are seen to have the lowest pairwise difference index, indicating similarities in their indicators and strategy profiles. On the other hand, A-STAR and CSIRO have high difference index values when paired with all other PRIs.

Table 7: Pairwise Difference Indices

	A-STAR	**CSIRO**	**ITRI**	**KIST**
A-STAR	0			
CSIRO	18.7	0		
ITRI	20.3	42.2	0	
KIST	13.2	30.5	4.7	0

The patent-output based scorecard described in this paper provides useful insights into the strategic priorities and R&D focus of organizations. Based on our scorecard for the 4 PRIs, we identify 3 different strategic profiles. Firstly, CSIRO prioritizes basic and science-based research in a diversified portfolio, with a focus on high quality and influential patents. The strategy at CSIRO is open to international collaborations and co-patenting. Secondly, there is the more industry-oriented approach shared by KIST and ITRI, who have similar profiles. This profile has an industry-pull dynamic, is growth-oriented and concentrated on a narrow range of specialized fields. There is relatively little internationalization of invention activities and the PRIs assert autonomous control patent ownership. The third strategic profile is ASTAR's, which is one that attempts to balance multiple priorities. This strategy emphasises both basic and industry-focused research, as well as both growth and quality improvement in a diversified portfolio. The ASTAR model is open to international collaborations and co-patenting.

In addition to expanding the coverage of PRIs for benchmarking purposes, future research would involve validating our analysis and findings by examining PRI strategy documents and where possible, conducting interviews with PRI management. This would allow us to determine if the priorities revealed by the patent portfolio are aligned to the PRI's stated objectives and mission, and are thus intentional. In this regard, our framework and patent-based strategy map could be a tool for monitoring the execution of R&D strategies and uncovering potential unintentional prioritization, where outcomes suggest strategic orientations which were not planned.

The scorecard can be refined with additional patent indicators - such as the Generality and Originality Indices (Hall, Jaffe & Tratjenberg, 2001), and measures of exploratory and exploitative propensity; and by including bibliometric indicators – such as co-publication intensity and field-specific publication-patent ratios. These refinements would enhance the relevance of the scorecard as a tool for evaluating strategic priorities in different types of organizations, allowing this analysis to be extended to universities and private sector firms.

References

Callaert, J. Van Looy, B. Verbeek, A., Debackere, K. and Thijs, B. (2006). Traces of Prior Art: An analysis of non-patent references found in patent documents. *Scientometrics*, 69 (1), 3-20.

Coccia, M. (2004). New models for measuring the R&D performance and identifying the productivity of public research institutes. *R&D Management*, 34 (3), 267–280.

Cohen, W.M., Nelson, R.R. and Walsh, J.P. (2000). “Protecting their intellectual assets: Appropriability conditions and why US manufacturing firms patent (or not)”. *NBER Working Paper* No. 7522. Cambridge, MA: National Bureau of Economic Research.

Debackere, K. & Luwel, M. (2004) Patent data for monitoring S&T Portfolios. . In H.F. Moed, W. Glanzel & U. Schmoch (Eds.), *Handbook of Quantitative Science and Technology Research* (pp. 569-586). Netherlands: Kluwer Academic Publishers.

Ernst, H. (2003). Patent information for strategic technology management. *World Patent Information* 25, 233–242.

Etzkowitz, H. & Leydesdorff, L. (1995). Emergence of a Triple Helix of university-industry-government relations. *Science and Public Policy*, 23, 279-286.

Hall, B.H. (2005). Exploring the patent explosion. *Journal of Technology Transfer*, **30**, 35-48.

Hall, B.H., Jaffe, A.B. & Trajtenberg, M. (2002). The NBER Patent Citation Data File: Lessons, Insights and Methodological Tools. In A. Jaffe & M. Trajtenberg (Eds.), *Patents, Citations and Innovations* (pp. 403-460). Cambridge, MA: The MIT Press.

King, D.A. (2004). The Scientific Impact of Nations. *Nature,* 430, 311-316.

Klerkx, L., & Leeuwis, C. (2009). Operationalizing demand-driven agricultural research: institutional influences in a public and private system of research planning in The Netherlands. *Journal of Agricultural Education and Extension*, **15** (2), 161-175.

Martino, J.P. (1983). *Technological forecasting for decision making*. Amsterdam: North Holland, Amsterdam.

Matsumoto, M., Yokota, S., Naito, K. & Itoh, J. (2010). Development of a model to estimate the economic impacts of R&D output of public research institutes. *R&D Management*, 40 (1), 91-100.

Mazzoleni, R. & Nelson, R.R. (2007). Public research institutions and economic catch-up. *Research Policy* 36 (10), 1512-1528.

Meyer, M (2001). "Patent citation analysis in a novel field of technology: An exploration of nano-science and nano-technology", Scientometrics, 51 (1), 163–183.

Narin, F., Hamilton, K.S. & Olivastro, D. (1997). The Increasing Linkage between U.S. Technology and Public Science. *Research Policy*, 26 (3), 317-330.

OECD (2011). Actor brief on Public Research Organizations (PROs). OECD Innovation Platform. Downloaded from: www.oecd.org/dataoecd/52/62/48136051.pdf

Porter, A.L. & Newman, N.C. (2004). Patent profiling for competitive advantage. In H.F. Moed, W. Glanzel & U. Schmoch (Eds.), *Handbook of Quantitative Science and Technology Research* (pp. 587-612). Netherlands: Kluwer Academic Publishers.

Stewart, J. (1995) Models of Priority-Setting for public sector research. *Research Policy* 24 (1), 115–126.

Tijssen, R.J.W. (2001). Global and domestic utilization of industrial relevant science: Patent citation analysis of science-technology interactions and knowledge flows. *Research Policy* **30**, 35-54.

von Graevenitz, G., Wagner, S. and Harhoff, D. (2008). Incidence and growth of patent thickets – the impact of technological opportunities and complexity. *CEPR Discussion Paper* no. 6900. London: Centre for Economic Policy Research.

Exploring the effects of the motivation of a research project on the research team composition, management, and outputs[1]

Masatsura IGAMI* and Sadao NAGAOKA**

* igami@nistep.go.jp
National Institute of Science and Technology Policy, 3-2-2 Kasumigaseki, Chiyoda-ku, Tokyo, 100-0013 (Japan)

** *nagaoka@iir.hit-u.ac.jp*
Institute of Innovation Research, Hitotsubashi University, 2-1 Naka, Kunitachi, Tokyo, 186-8603 (Japan)

Introduction

Scientists' motivation to conduct research can be broadly classified as either external motivation, such as fame or financial gain, or internal motivation (Deci and Flaste, 1996). Various cases have demonstrated the importance of intellectual curiosity as an internal motivation (Misu, Horoiwa & Tsunoda, 2008; Stephan, 2012). At the same time, scientific research is a competitive process that seeks to establish priority (Merton, 1973). Consequently, the external motivation of achieving fame and recognition via the establishment of priority in research is also important.

The quadrant model introduced by Donald Stokes provides a method for classifying research motivation by content (Stokes, 1997). Stokes applied this concept to overcome the classification of research as one-dimensional (i.e., either basic or applied) and categorized research motivation into "pursuit of fundamental principles/understanding" and "solving specific issues in real life." In this model, the Pasteur's quadrant covers such "use-inspired basic research" exemplified by the research by Pasteur, while the Bohr's quadrant covers pure basic research and the Edison's quadrant covers pure applied research.

Analyses of recent scientific papers reveal an increase in the number of authors per paper over time (Adams, Black, Clemmons, and Stephan, 2005; Wuchty, Jones, and Uzzi, 2007). This trend indicates a shift in the unit of scientific research from the individual to a team of scientists and implies that the composition and management of such teams have become increasingly important in scientific research.

On the basis of the previous studies, the purpose of this work in progress paper is to analyse the following questions: Is there relationship between motivation of a research project and (i) the composition of the research team; (ii) research management; and (iii) research outputs?

In team-based research, the processes from motivation to outputs can be considered as follows. First, the project leader or leaders forms the concept for a research project in the wake of internal and external motivations. Then, depending upon motivation for the project,

[1] This research has been being conducted in collaboration between Institute of Innovation Research (IIR) of Hitotsubashi University and National Institute of Science and Technology Policy (NISTEP) as a part of "Industry-university-government joint research on the innovation process" with support of Special Funds for Education and Research of The Ministry of Education, Culture, Sports, Science and Technology. This work has also been supported by JSPS KAKENHI 21243020 .

research is managed by gathering a research team/environment within the scope of the resources available. New knowledge is created through research activities and outputs in various forms are generated. An understanding of these processes is crucial for the interpretation of the knowledge creation process in science. To our knowledge, the topic has not been studied sufficiently and therefore our study gives evidence how the context of a research, the motivation of a research, impacts on the research team composition, management, and outputs.

Overview of scientists survey

We conducted large scale survey to scientists both in Japan and the United States, the Hitotsubashi-NISTEP-Georgia Tech scientist survey (scientist survey) (Nagaoka *et al.*, 2010; Nagaoka *et al.*, 2011). The scientist survey identified the focal papers, top 1% highly cited papers and normal papers among the papers published in 2001–2006, and asked the corresponding authors or equivalents about the research projects from which the focal paper was yielded. The response rate of the survey was 27.2% (2,081 respondents) in Japan and 26.3% (2,329 respondents) in the US.

Among the collected responses, this paper focuses on research projects in natural science conducted in universities. The number of complete answers from university respondents is 2,264[2]. The scientist survey gathered comprehensive data on research projects such as project motivation; structure of the research team; research management; amount of research funding used in the project; and outputs of the project.

Motivation of research project and the composition of the research team

Adopting the framework of Pasteur's quadrant, we asked each respondent to evaluate the importance of the following two basic motivations for initiating the research project that yielded the focal paper and the other closely related papers: (1) pursuit of fundamental principles/understandings and (2) solving specific issues in real life. "Pursuit of fundamental principles/understandings" is defined to be gaining a new knowledge of the principles, underlying natural phenomenon and observed facts, through experiments and/or theoretical analyses and "solving specific issues in real life" is defined to be solving practical and specific problems such as for industrial applications, following *Frascati Manual* of OECD.

Each motivation was graded from 1 to 5, based on the subjective evaluation by respondents, where 1 indicating that the motivation was "not important at all" and 5 indicating that it was "very important." This study considers these degrees of importance as a proxy to measure the strength of the motivation.

Motivation of research project is thought to affect to the composition of the research team. When a research project is trying to "solve specific issues in real life," there is a particularly real need to combine the knowledge resources required to solve the issue, and to achieve this goal, the research team may incorporate multiple fields of expertise and skills.

For the composition of the research team, we considered four dependent variables: (1) diversity of fields of expertise, (2) diversity of skills, (3) university-industry collaboration,

[2] The survey asked a scientist to identify the sector of the organization with which he/she was affiliated when the focal paper was submitted. This sector was used for analysis. The five-sector classification shown the following is used; (1) higher education institutions; (2) public research institutions; (3) private firms; (4) private non-profit organisations, including hospitals; (5) others.

and (4) international co-authorship. These dependent variables are dummies that can take a value of 1 or 0. "Diversity of fields of expertise" or "diversity of skills" was assigned a value of 1 if the team consisted of scientists from multiple fields of expertise, or with multiple skill sets. "University-industry collaboration" is assigned a value of 1 if one of the authors of the focal paper is from the industrial sector. Similarly, "international co-authorship" is assigned a value of 1 if the research paper is internationally co-authored.

The main independent variables are as follows. "Pursuit of fundamental principles /understanding" and "solving specific issues in real life" indicate the respective importance of both as direct motivation to initiate the research project. We also included the logarithmic values of the number of authors as a variable to control the size of the projects. We assumed that the diversity of the scientists constituting the research team would increase with the increase in the number of authors of the focal papers. In addition, we considered the amount of research funding to control the size of the projects.

Motivation of research projects strongly depends on the field of science; therefore we included a dummy variable related to the field of science[3]. For field dummies, we summarized the 21 ESI journal fields (excluding the multi-disciplinary field) into an 8 or a 3 field classification. In our estimation, we pooled the normal papers and top 1% highly cited papers. We also introduced a dummy variable to control for the type of papers.

Table 1 shows results of a logistic regression. For "diversity of fields of expertise," "solving specific issues in real life" has statistically significant and positive coefficients in both Japan and the US (significant at the 1% level in Japan and 5% in the US). In addition, the "pursuit of fundamental principles/understanding" shows a statistically significant and negative coefficient at the 1% level in the US.

Focusing on "diversity of skills," "solving specific issues in real life" shows a statistically significant and positive coefficient at the 1% level in Japan," whereas the "pursuit of fundamental principles/understanding" shows a statistically significant and negative coefficient at the 5% level in the US.

There were considerable differences between Japan and the US in trends for "university-industry collaboration." This variable has a positive coefficient of 1% statistical significance for "solving specific issues in real life" in Japan, whereas it has a positive coefficient of 5% significance for "pursuit of fundamental principles/understanding" in the US."

For "international co-authorship," both Japan and the US show statistically significant and negative coefficients at 1% level for "solving specific issues in real life." In addition, Japan shows a positive coefficient of 1% statistical significance for the "pursuit of fundamental principles/understanding."

[3] Using the survey outcomes, we applied Stokes' Quadrant model to projects both in Japan and the United States. We found the strong linkage between fields of science and quadrant model. The balance of Pasteur's quadrant is relatively high in clinical medicine and psychiatry/psychology (both Japan and the United States), agricultural sciences (the United States), and materials science (Japan). Bohr's quadrant is dominant in space science, physics, mathematics, and molecular biology & genetics (both Japan and the United States).

Table 1. Results of regression

	Model J-1	Model U-1	Model J-2	Model U-2	Model J-3	Model U-3	Model J-4	Model U-4
	Diversity in field of expertize		Diversity in skills		University-industry collaboration		International co-authorship	
	Logit		Logit		Logit		Logit	
	Marg eff.		Marg eff.		Marg eff.		Marg eff.	
	JPN	USA	JPN	USA	JPN	USA	JPN	USA
Pursuit of basic principles/understanding(1-5)	0.004	-0.058***	-0.003	-0.041**	-0.011	0.022**	0.055***	0.017
	[0.018]	[0.021]	[0.017]	[0.020]	[0.007]	[0.011]	[0.019]	[0.020]
Solving specific issues in real life(1-5)	0.039***	0.034**	0.046***	0.016	0.033***	0.001	-0.047***	-0.037***
	[0.011]	[0.014]	[0.011]	[0.013]	[0.006]	[0.005]	[0.011]	[0.011]
Number of authors(Log)	0.153***	0.248***	0.133***	0.157***	0.039***	0.038***	0.199***	0.307***
	[0.025]	[0.035]	[0.023]	[0.032]	[0.010]	[0.010]	[0.027]	[0.031]
Research funding (categorized value)	0.008	0.011	-0.012	-0.001	0.001	-0.003	0.003	-0.027**
	[0.011]	[0.014]	[0.010]	[0.013]	[0.005]	[0.005]	[0.011]	[0.012]
Paper type (0: Normal papers, 1: top 1% highly cited papers)	-0.023	0.022	0	-0.006	-0.006	-0.024**	0.117***	-0.041
	[0.032]	[0.039]	[0.032]	[0.038]	[0.015]	[0.012]	[0.034]	[0.031]
Field dummy	YES	YES	YES	YES	YES	YES	YES	YES
Chi-squared	91.835***	139.096***	142.680***	110.078***	54.739***	27.683**	203.960***	150.604***
Log-likelihood	-780.031	-575.249	-726.53	-577.753	-350.643	-184.976	-702.694	-479.606
Psuedo-R2	0.056	0.108	0.089	0.087	0.072	0.07	0.127	0.136
Observations	1313	951	1313	951	1313	951	1313	951

* p<0.1, ** p<0.05, *** p<0.01

Outline of the analytical work to be completed

In this work in progress paper, we analysed the effects of the motivation of research project on the composition of the research team. Combination of expertise of different field of science and combination of skill sets are more common in research teams strongly motivated to "solve specific issues in real life." Our results also indicate that research projects strongly motivated to "solve specific issues in real life" involved a lower proportion of international co-authorship.

We should note that we found considerably different results for "university-industry collaboration" in Japan and the US. In Japan, there was a positive, statistically significant correlation between "solving specific issues in real life" and "university-industry collaboration." However, in the US, there was a positive, statistically significant correlation with "pursuit of fundamental principles/understanding." Further study is needed in this regard.

Focusing on the motivation of research projects, the paper will analyse how the motivation of a science research project affect to the management, e.g., ambitious goal settings, and outputs, e.g., incidence of a patent application and a start-up, of the research project.

References

Misu, T., Horoiwa, A. & Tsunoda, H. (2008). Investigative Survey on the Research Activities and Awareness of Postdoctoral Fellows: Analysis on Career Choices of Postdoctoral Scholars. *Research Material-161*, National Institute of Science and Technology Policy.

Adams, J.D., Black, G.C., Clemmons, J.R. & Stephan, P.E. (2005). Scientific teams and institutional collaborations: Evidence from U.S. universities, 1981–1999. *Research Policy*, 34, 259-285.

Deci, E.L. and Flaste, R. (1996). *Why we do what we do: understanding of self-motivation*. Penguin Press.

Merton, R.K. (1973). *The Sociology of Science: Theoretical and Empirical Investigations*. University of Chicago Press, Chicago, II.

Stephan, P.E. (2012). *How Economics Shapes Science*. Harvard University Press.

Stokes, D.E. (1997). *Pasteur's Quadrant: Basic Science and Technological Innovation*. Brookings Institution Press.

Wuchty, S., Jones, B. & Uzzi, B. (2007). The increasing dominance of teams in the production of knowledge. *Science*, 316, 1030–1036.

Nagaoka, S., Igami, M., Eto, M. & Ijichi, T. (2010). Knowledge Creation Process in Science: Basic findings from a large-scale survey of researchers in Japan. *IIR Working Paper,* WP#10-08

Nagaoka, S., Igami, M., Walsh, J.P. & Ijichi, T. (2011). Knowledge Creation Process in Science: Key Comparative Findings from the Hitotsubashi-NISTEP-Georgia Tech Scientists' Survey in Japan and the US. *IIR Working Paper*, WP#11-09

Assessment of expertise overlap between an expert panel and research groups[1]

A.I.M. Jakaria Rahman *, Raf Guns **, Ronald Rousseau*** and Tim C.E. Engels ****

*jakaria.rahman@uantwerpen.be
Centre for R&D Monitoring (ECOOM), Faculty of Political and Social Sciences, University of Antwerp, Middelheimlaan 1, B-2020 Antwerp (Belgium)

** raf.guns@uantwerpen.be
Institute for Education and Information Sciences, University of Antwerp, Venusstraat 35, B-2000 Antwerp (Belgium)

***ronald.rousseau@uantwerpen.be
ronald.rousseau@kuleuven.be
Institute for Education and Information Sciences, University of Antwerp, Venusstraat 35, B-2000 Antwerp and KU Leuven, Dept. of Mathematics, B-3000 Leuven (Belgium)

**** tim.engels@uantwerpen.be
Centre for R&D Monitoring (ECOOM), Faculty of Political and Social Sciences, University of Antwerp, Middelheimlaan 1, B-2020 Antwerp, and
Antwerp Maritime Academy, Noordkasteel Oost 6, B-2030 Antwerp (Belgium)

Abstract
Discipline-specific research evaluation exercises are typically carried out by committees of peers, expert panels. Currently, there are no available methods that can measure overlap in expertise between a panel and the units of assessment. This research in progress paper explores a bibliometric approach to determining the overlap of expertise, using the 2010 research evaluation of nine physics research groups of the University of Antwerp as a test case. Overlay maps were applied to visualize to what extent the groups and panel members publish in different Web of Science subject categories. There seems to be a moderate disparity between the panel's and the groups' expertise. The panel was not as diverse as the groups that needed to be assessed. Future research will focus on journal level overlay maps, similarity testing, and a comparison with other disciplines.

Keywords: Research assessment, Expert panel, Research group

Introduction

Discipline-specific research evaluations are a common practice at many universities worldwide. These evaluations are carried out by committees of peers. As is the case with research proposals submitted to research funding organizations, expert panel review is considered the standard for determining research quality of individuals and groups (Nedeva, et al, 1996; Butler & McAllister, 2011; Lawrenz, Thao, & Johnson, 2012). The principal objective of such evaluations is to improve the quality of scientific research. The University of Antwerp, Belgium, implemented evaluative site visits by expert panels in 2007. Using data collected in the frame of one of these evaluations, this papers explores the expertise overlap

[1] This research has been made possible by, among others, the financial support of the Flemish Government to the ECOOM. The opinions in the paper are the authors' and not necessarily those of the government.

between the expert panel and physics research groups involved in the evaluation. To the best of our knowledge, no methods have been established to measure and quantify overlap in expertise between panels and the units of assessment. However, in research evaluation the extent to which the expertise of the panel members charged with research assessment is congruent with the research of the units, is crucial to the trustworthiness of the assessment (Engels et al., 2013). Only panel members that are credible experts in the field can deliver an assessment that can contribute to the improvement of the quality of the research. Moreover, Langfeldt (2004) explored expert panel evaluation and decision making processes, and concluded that overlap of expertise between experts is highly needed in order to foster cooperation among panel members. For the evaluation of research groups, it is expected that the research of each group is well covered by the expertise of the panel members.

The goal of this research in progress is to inform the process of expert panel composition. In this paper, we present a bibliometric analysis of the overlap of expertise between the physics expert panel and the (whole of the) units of assessments in the Department of Physics of the University of Antwerp. Hence, the research questions are:

1) To what extent is there overlap between the panel's expertise and the whole of the research to be assessed?
2) To what extent is the individual research group expertise covered by the panel's expertise?

Data and Methodology

As a test case we present an analysis of the 2010 assessment of the Department of Physics' nine research groups of the University of Antwerp. The reference period is a time interval of eight years preceding the evaluation. The citable items from the Science Citation Index Expanded of the Web of Science (WoS) published by the research groups in the period 2002 to 2009 have been taken into account.

The panel was composed of six members including the chair. All the publications of the panel members since their respective first scientific publication to the year 2009 have been taken into account. Potential panel members had no prior involvement with the research groups that were evaluated (i.e. no prior affiliations, no co-publications, no common projects). In total, the six panel members have 1,104 publications, none of which are co-authored with another panel member. The number of publications per panel member ranges from 117 to 282. In total, these publications were published in 204 journals.

Table 1: Publication profile of the physics research groups

Group code	Number of Publications	Number of WoS categories	Number of Journals
Physics group A	125	44	53
Physics group B	486	25	66
Physics group C	525	46	147
Physics group D	269	7	17
Physics group E	159	28	55

Group code	Number of Publications	Number of WoS categories	Number of Journals
Physics group F	42	13	23
Physics group G	43	12	26
Physics group H	132	12	31
Physics group I	115	49	63
Total	**1732**	**102**	**353**

Table 1 summarizes the number of publications for the nine research groups. A total of 164 publications was co-authored by members of two or more groups.

The VOSviewer computer program is used to visualize the overlap of groups and panel publications based on a global map of science incorporating the new WoS subject categories (Leydesdorff, Carley, & Rafols, 2013) Overlay maps were created for the panel, the separate research groups, and the nine research groups taken together. The Spearman's rank correlation coefficient is calculated between the panel's and groups' publications based on WoS subject categories.

Analysis and Results

a) Panel profile versus Groups profile

The overlay maps for the panel and the groups as a whole (figure 1 and 2) visually show that the groups taken together publish more widely than the panel members. The panel members publications are strong (58.54%) in the categories of 'Physics condensed matter', 'Physics multidisciplinary', 'Chemistry physical', 'Physics applied' whereas, the groups' publications are mostly (44.92%) concentrated in the 'Physics condensed matter', 'Physics multidisciplinary', 'Physics applied', and 'Materials science multidisciplinary' subject categories.

Figure 1: Panel members publications overlay map

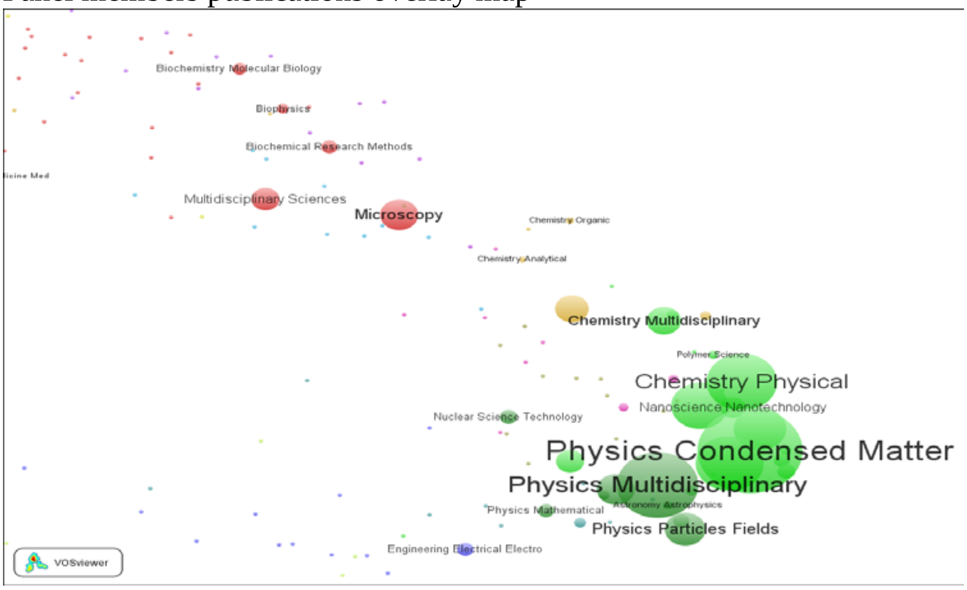
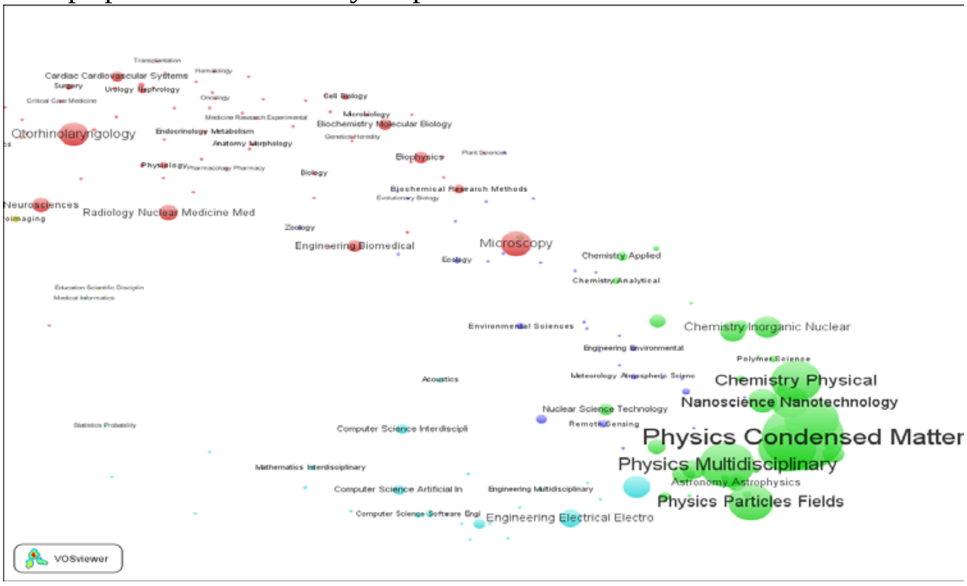


Figure 2: Groups publications overlay map



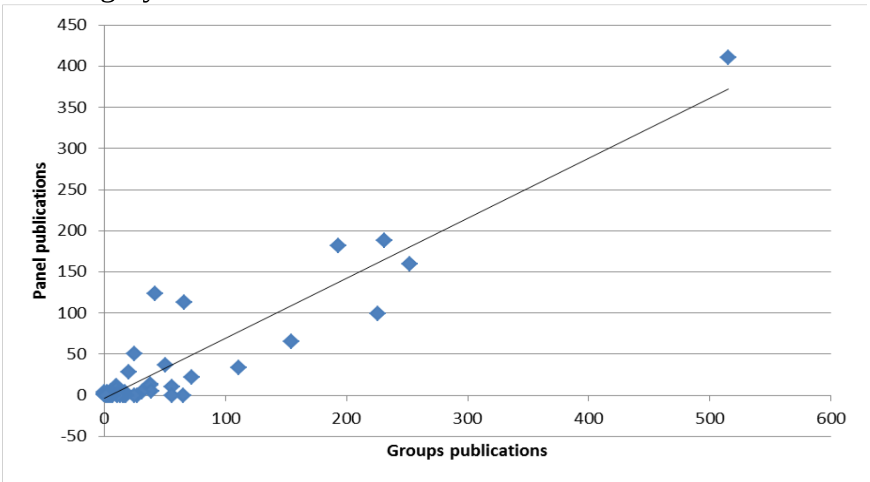
Panel publications fall in 39 WoS subject categories whereas the groups cover 102 WoS subject categories. Table 2 shows that the panel (23.58%) and the groups (18.9%) have the majority of their publications in 'Physics condensed matter', followed by 'Physics multidisciplinary' (panel 14.28%, groups 8.48%)', 'Chemistry physical' (panel 10.65%, groups 7%)' and 'Physics applied' (panel 10.03%, groups 9.25%).

Table 2: Top ten WoS subject categories

Panel publications			**Groups publications**		
Web of Science Categories	**Number of records**	**%**	**Web of Science Categories**	**Number of records**	**%**
Physics condensed matter	416	23.58	Physics condensed matter	515	18.90
Physics multidisciplinary	252	14.28	Physics applied	252	9.25
Chemistry physical	188	10.65	Physics multidisciplinary	231	8.48
Physics applied	177	10.03	Materials science multidisciplinary	226	8.29
Physics atomic molecular chemical	125	7.08	Chemistry physical	193	7.0
Materials science multidisciplinary	104	5.89	Physics particles fields	154	5.6
Physics particles fields	65	3.68	Nanoscience nanotechnology	111	4.09
Microscopy	56	3.17	Microscopy	72	2.64
Optics	56	3.17	Physics atomic molecular chemical	66	2.42
Chemistry multidisciplinary	45	2.55	Otorhinolaryngology	65	2.3

The Spearman's rank correlation coefficient is 0.524. This indicates a positive yet moderate correlation between the panel's and the groups' publications occurrence in the WoS subject categories.

Figure 3: Scatter plot of the panel's and the groups' publication numbers per WoS subject category

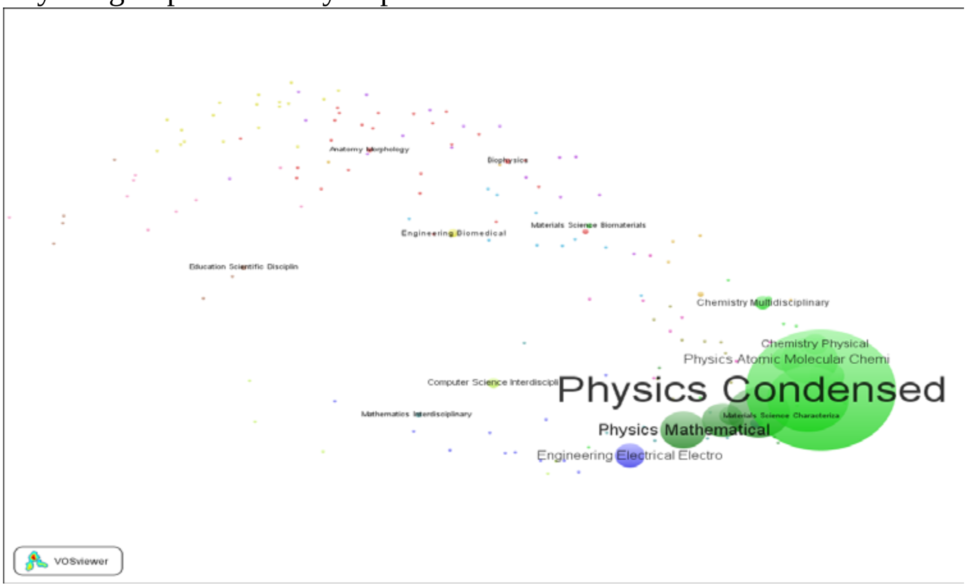


From the above discussion, it appears that there is visible disparity between panel and group publications according to WoS subject categories. The groups publish more diversely than the panel, which might be due to the interdisciplinary orientation of some of the groups.

b) Panel versus Individual groups

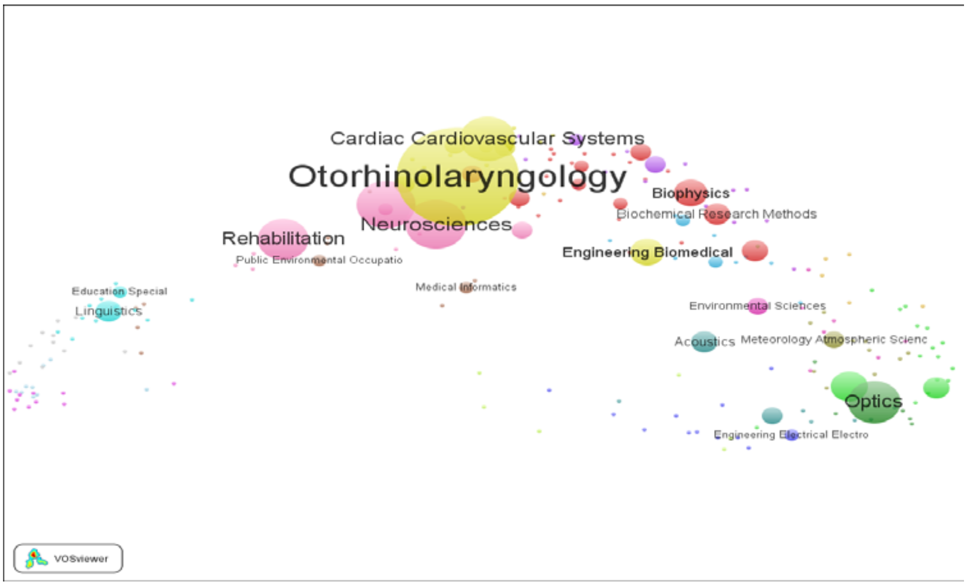
We have created overlay maps of individual group publications in the WoS subject categories, and compare them with the panel overlay map (Figure 2). Group 'B' focuses on 'Physics condensed matter' (45.24%), and 'Physics applied' (14.66%) subject categories (Figure 4). Similarly, group 'C' focuses on 'Materials science multidisciplinary' (19.04%), 'Chemistry physical' (15.99%), and 'Physics condensed matter' (13.54%); group 'E' focuses on 'Physics multidisciplinary' (14.39%), 'Physics particles fields' (14.03%), and 'Physics condensed matter' (11.87%); group 'F' focuses on 'Physics Multidisciplinary' (37.88%); and group 'H' focuses on 'Physics condensed matter' (47.06%). Physics groups 'B', 'C', 'E', 'F', and 'H' are well covered by the panel's expertise, as the panel's publications mostly fall into these subject categories.

Figure 4: Physics group 'B' overlay of publications



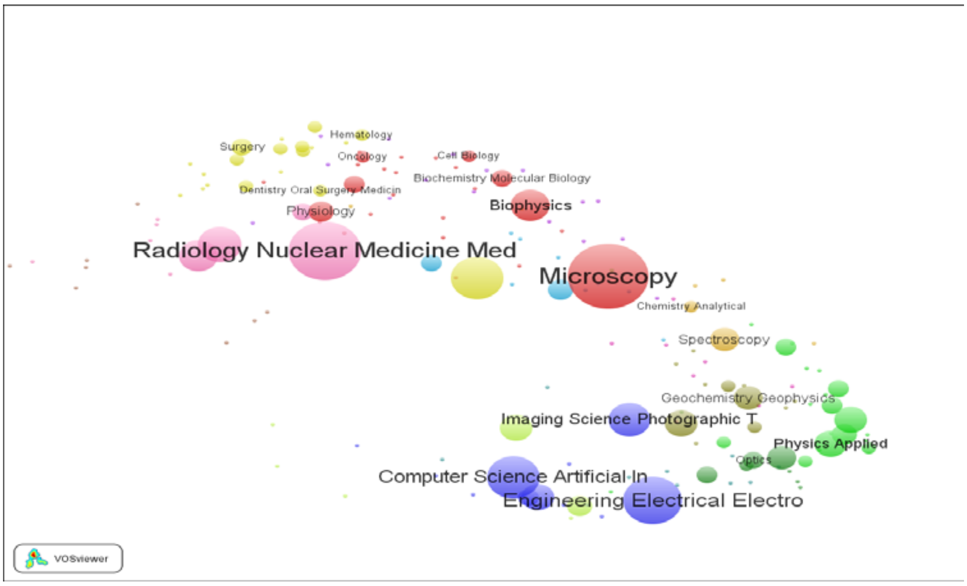
The publications of group ‘A’ fall in 42 subject categories with a focus on ‘Otorhinolaryngology’ (29.23%; Figure 5). Physics group ‘D’ publications fall in only seven subject categories, and focus on ‘Physics particles fields’ (47.96%) and ‘Physics multidisciplinary’ (34.48%) subject categories. The panel has few publications in these subject categories, therefore groups ‘A’ and ‘D’ are partially covered by the panel expertise.

Figure 5: Physics group ‘A’ overlay of publications



Physics group ‘G’ publications are concentrated in 12 WoS subject categories; this group focuses on ‘Physics atomic molecular chemical’ (22.06%) and ‘Chemistry physical’ (20.59%). Physics group ‘I’ publications belong to 49 subject categories; this group focuses on ‘Microscopy’ (13.95%) and ‘Radiology nuclear medicine medical imaging’ (11.16%), as shown in Figure 6. However, the panel has no overlap with the categories where group ‘G’ and ‘I’ have a largest share of their publications.

Figure 6: Physics group ‘I’ overlay of publications



Conclusion

The results indicate that there is some disparity between the panel's and the groups' publications according to WoS subject categories, and the visual map supported by the Spearman's rank correlation coefficient indicates a moderate correlation. In future research, we may explore other correlation coefficients, since the large number of zeroes may influence Spearman's rho. The panel was not as diverse as the groups that needed to be assessed. This could be expected, as the panel members have been selected primarily because of their expertise and not necessarily because of the match thereof with the research in the groups. In subsequent analysis we will look at overlay maps on the journal level (Leydesdorff, Rafols, & Chen, 2013), and will quantify the similarity between groups and panel at this level. The results will be compared with at least one other discipline to identify what overlap leads to the best standard for evaluation, as well as to find a suitable method for the expert panel composition.

References

Butler, L., & McAllister, I. (2011). Evaluating University research performance using metrics. *European Political Science*, *10*(1), 44–58.

Engels, T. C. E., Goos, P., Dexters, N., & Spruyt, E. H. J. (2013). Group size, h-index, and efficiency in publishing in top journals explain expert panel assessments of research group quality and productivity. *Research Evaluation*, *22*(4), 224–236.

Langfeldt, L. (2004). Expert panels evaluating research: decision-making and sources of bias. *Research Evaluation*, *13*(1), 51–62.

Lawrenz, F., Thao, M., & Johnson, K. (2012). Expert panel reviews of research centers: The site visit process. *Evaluation and Program Planning*, *35*(3), 390–397.

Leydesdorff, L., Carley, S., & Rafols, I. (2013). Global maps of science based on the new Web-of-Science categories. *Scientometrics*, *94*(2), 589–593.

Leydesdorff, L., Rafols, I., & Chen, C. (2013). Interactive overlays of journals and the measurement of interdisciplinarity on the basis of aggregated journal–journal citations. *Journal of the American Society for Information Science and Technology*, *64*(12), 2573–2586.

Nedeva, M., Georghiou, L., Loveridge, D., & Cameron, H. (1996). The use of co-nomination to identify expert participants for Technology Foresight. *R&D Management*, *26*(2), 155–168.

Constructing Composite Indicators of Basic Research: Conceptual Issues

Naeyang Jeong*, Jun Young Lee** and Dae Nyoung Heo*

* *nyjeong@ibs.re.kr; nyoung@ibs.re.kr*
Policy Research Team, Institute for Basic Science,
70 Yuseong-dearo 1689-gil Yuseong-gu, Daejeon, 305-811 (Korea)

** *jylee@ibs.re.kr*
Policy Research Team, Institute for Basic Science and
Science & Technology Policy, University of Science & Technology(UST)
70 Yuseong-dearo 1689-gil Yuseong-gu, Daejeon, 305-811 (Korea)

Introduction

This research intends to develop composite indicators that can inclusively express the achievements or progress of basic research[1] activities at the national level. Composite indicators are recognized to be useful tools in analysing policies or promoting them toward the people. However, attention should be paid as composite indicators may mislead facts or cause incorrect policy conclusions when they are unduly designed or interpreted. It should be a crucial procedure for us to select specific detail indicators that comprise the composite indicators based on a solid theoretical framework. Therefore, this study focused on the construction of a structure of composite indicators through a diverse literature survey or collection of expert opinions.

Development of a framework of the Basic Research Index

The concepts to be measured may be made concrete or specific in the process of systematically structuring basic research activities. To that end, first, we referred to the examples of composite indicators for measurement of scientific and technological activities, which had already been developed. Second, we also analysed major keywords of government policies related to the basic research. Third, we derived factors that significantly affected Nobel prize-winning by treating it as a major achievement of the basic research. Finally, we complemented a draft framework of the Basic Research Index by obtaining advice by experts in various areas after preparing the draft framework.

Table 1. A framework of the Basic Research Index

Main Type	Dimensions	Indicators	Source
Basic research inputs	Human resources	Total researchers in higher education sector	OECD
		New doctorate graduates in science	OECD
	Financial resources	Basic research expenditure by government, higher education, and private non-profit	OECD
		Basic research expenditure by business enterprise	OECD
	Intellectual resources	Number of scientific papers published in the past 10 years	Thomson ISI
		Number of top 1% cited scientific papers	Thomson ISI

[1] This research followed the definition of basic research under the OECD Frascati Manual.

		published in the past 10 years	
Basic research environment	Education environment	Number of world leading universities in science	QS
		PISA scores	OECD
		Degree to which the importance of science is emphasized in education*	IMD
	Institutional environment	Ratio of basic research budget among government R&D budget	OECD
		Degree to which researchers and scientists are attracted by country or government*	IMD
		Degree to which the law encourages innovation in scientific research*	IMD
	Cultural environment	Ratio of internationally co-authored papers out of total papers	OECD STI
		Degree of flexibility or accommodation people have against new challenges*	IMD
Basic research outputs	Generation of outputs	Number of scientific papers published in the most recent year	Thomson NSI
		Publications in the top-quartile journals (per GDP)	OECD
		Degree of superiority of private or public science research to international standards*	IMD
		Number of Nobel science prize winners	Nobel foundation
	Diffusion of outputs	Ratio of patents citing scientific papers out of total patents	KIPI
		Number of citations of papers	Thomson ISI
		Degree of active diffusion of knowledge between businesses and universities*	IMD
		Patents filed by universities and public labs (per GDP)	OECD

* Questionnaire indicators

Discussion of conceptual issues

This section summarizes conceptual issues that were discussed while the Basic Research Index framework was derived.

(i) Is the framework MECE?

Whether the lower level concepts defined for basic research activities under the indicator structure were Mutually Exclusive and Collectively Exhaustive (MECE) were examined. No significant objections were raised for the dimensions of indicators, but issues were commonly raised as to the fact that some specific indicators comprising some dimensions have strong characteristics of proxy indicators. In particular, it was inevitable to most appropriate indicates among the available indicators as proxy indicators in order to secure comparability among countries. By following this reasoning, we could define the criteria for selecting proxy indicators. When there is no indicators that can directly measure basic research activities, we should select indicators that can measure achievements in science out of the options between science and engineering or science and technology.

(ii) Should the country's size be controlled?
In selecting specific indicators, we needed to adjust scales adequately if an indicator depended on the factors related to a country's GDP or population. On the other hand, the size of phenomena to be measured can be an important factor in itself. For example, it was unnecessary to adjust the scale if the composite indicator was designed to measure investment itself when investment and investment efficiency are compared. Based on our summing up of comments by experts, we concluded it would desirable to discuss whether to adjust the scales later, after performing simulations for all of the methods after the data was collected.

(iii) Are patents basic research achievements?
We started such discussion because our analysis of Nobel science prize-winning countries revealed a quantitative co-relationship between the patent-related indicators and the number of Nobel prize winners. To summarize the discussion, we concluded it would be desirable to include patents in wider-range basic research outputs and in the specific indicator of 'diffusion of outputs' for the basic research indicators. It may not be desirable to require basic research performers to produce patents. However, the problem may be resolved when we separate basic researchers from patent experts as in the case of WIS and Yeda in Israel. We concluded it would be desirable to resolve the problem by further exploring specific indicators that can measure patents derived from basic research.

Conclusion
This study derived a composite indicator structure that can cover basic research activities at a government or state level. We employed proxy indicators that can indirectly measure the inclusive phenomena rather than conceptual categorizations as our efforts for discovering proper indicators had to be limited for measuring the basic research environment in a country.

In our future research, more stakeholders and experts should participate in the development of composite indicators. So far, only a limited number of experts participated in the research – only policy researchers participated. It is necessary to derive an agreement to a certain level on issues that cannot be discussed by expanding the participation of experts from diverse layers of research, including policy developers. If the Basic Research Index is completed and data are accumulated through follow-up research, the outcome could be utilized for developing key policies concerning basic research.

References

EU, Innovation Union Scoreboard, each year

IMD, The World Competitiveness Yearbook, each year

OECD(2008), Handbook on Constructing Composite Indicators

OECD(2011), Government at a Glance 2011

OECD(2011), National Accounts of OECD Countries: General Government Accounts

OECD(2011), OECD Science, Technology and Industry Scoreboard 2011

OECD(2012), Research and Development Statistics

OECD(2013), Main Science and Technology Indicators 2012-2

OECD(2013), Main Science and Technology Indicators 2013-1

Genetic patents in plant biotechnology

Koen Jonkers*1, Catalina Martinez**

koen.jonkers@ec.europa.eu
European Commission, Joint Research Centre, Institute for Prospective Technological Studies (IPTS), C/ Inca Garcilaso No. 3, E-41092 Seville, Spain

** *catalina.martinez@csic.es*
Institute of Public Goods and Policies, CSIC-IPP, C/ Albasanz 26-28, Madrid, 28037,Spain

Introduction

Genetic patenting is an area of heated debate in scientific, legal and economic fora, especially as regards the human genome, but increasingly also with respect to plant biotechnology given the large economic stakes involved (Louwaars al, 2009; Baillie & Connett Porceddu, 2012). In 2008, the global seed industry had an annual turnover of 2.7 billion Euro. These seeds do not only form the basis for a total product market with an annual turnover of 250 billion Euro, but are of central importance for the global food supply (Louwaars et al, 2009). Intellectual property rights on plants are governed through two different regimes: plant breeders' rights and patents. This paper focuses on the latter.

Our aim is to assess whether policy relevant questions in the field of plant biotechnology, such as ownership of genetic patents, science-industry links and the relation between patents and follow-on research, can be addressed with publicly available data in patent databases linked to DNA sequence repositories. In doing so, we first present a review of the literature and background on the relation between plant biotechnology and genetic patenting, and then we examine the distribution of patent applications claiming plant gene sequences by major patent office and type of organism, number of sequences per patent, type and country of origin of applicants as well as other patent features. We pay special attention to the distinction between patents including genes of the model organism *Arabidopsis thaliana* – as a model organism it is a research tool and therefore a potential indicator for science relatedness – and patents including genes of commercial crops, as well as to the dynamics of public/private nature of ownership and business concentration in the plant seed sector, where five companies hold 64% of all plant sequence EPO applications and over 90% of all the plant gene sequences listed in them. We conclude with a more in-depth exploratory quantitative and qualitative analysis of the value and scope of protection of EPO filings listing gene sequences of different kinds of plant organisms.

Database construction and preliminary findings

In order to build our database, we first downloaded all the patent sequences files (GBPAT) from NCBI GENBANK flat file release 183.0 (February 15 2011 and April 15 2011)[2] which were used to create a relational database and linked those files to patent information from

[1] The information and views set out in this publication do not necessarily reflect the opinion of the European Commission. The EC does not guarantee the accuracy of the data included in this study. Neither the EC nor any person acting on its behalf may be held responsible for the use which may be made of the information contained herein.
[2] ftp://ftp.ncbi.nih.gov

PATSTAT (October 2010), using patent numbers. GENBANK has 19,189,921 different gene sequence entries and 424,238 patent publication numbers from different patent offices. For other studies using GENBANK see e.g.: Arnaud-Haoud *et al*. 2011 on marine species. The data contained in NCBI GENBANK allows us to come to an alternative, more specific, delimitation of life science patents than would be possible on the basis of IPC classes. Only patents that contain genetic sequences are studied in this paper.

A first noticeable finding is that GENBANK does not have information on USPTO gene sequences broken down by type of organism.[3] We therefore focus our analysis on plant sequence patents filed at EPO and through the PCT route worldwide (table 1). In EPO patent applications, plant sequences comprise around 4.1% of the total number of patented sequences, in PCT filings this share is around 3.3 %. This makes plantae one of the largest source groupings after bacteria, mammalia, and synthetic constructs.

Table 1. Patents containing gene sequences and number of sequences by type of filing and source organism.

	EPO filing			**PCT filing**	
	number of sequences	**number of patents**		**number of sequences**	**number of patents**
Unclassified	47664	1823		81905	5540
Synthetic	570022	6969		2074033	21481
Mammalia	2231844	3316		2186310	13469
Bacteria	227932	1352		146193	3929
Plantae	133995	581		175420	2376
Virus	9849	562		71685	1484
Fungi	18344	385		431636	1457
Arthropoda	10201	126		43399	420
Archae	4590	94		1886	258
Other	8530	532		12868	1349
Total	3262971	15740		5225335	51763

Table 2 lists the top ten source plants, accounting for around 81% of all plant sequences in EPO patent applications and 77% in PCT patent applications. In terms of gene sequences, Arabidopsis is the most important source organism in PCT filings with plant gene sequences, with 20% of all, and the second most important at EPO, with 18%, after Maize with 28%. In terms of patent applications, Arabidopsis ranks first, as 35% of all EPO applications and 39% of all PCT filings include at least one sequence of Arabidopsis.

[3] We explored the use of BLAST searches to classify a subsample of the unidentified sequences by type of organism, as a pilot exercise. This proved to be quite a labour intensive effort and will be described in detail in the paper.

Table 2. Top 10 plant gene sequences in patent filings, by organism.

	EPO sequences		EPO patents		PCT sequences		PCT patents	
	Count	% Total Plantae	Count	% Total Plantae	Count	% Total Plantae	Count	% Total Plantae
Maize	37293	27.8%	123	21.2%	28056	16.0%	551	23.2%
Arabidopsis	23845	17.8%	204	35.1%	34613	19.7%	917	38.6%
Soybean	19644	14.7%	76	13.1%	12863	7.3%	331	13.9%
Rice	19614	14.6%	126	21.7%	30793	17.6%	510	21.5%
Wheat	2338	1.7%	83	14.3%	5470	3.1%	263	11.1%
Tomato	1869	1.4%	60	10.3%	2569	1.5%	220	9.3%
Tobacco	1748	1.3%	66	11.4%	6625	3.8%	242	10.2%
Potato	1108	0.8%	58	10.0%	1471	0.8%	210	8.8%
Barley	968	0.7%	51	8.8%	12687	7.2%	176	7.4%
Pea	494	0.4%	23	4.0%	498	0.3%	72	3.0%
Top 10 / Total Plantae		**81%**				**77%**		
Total Plantae	**133995**	**100%**	**581**	**100%**	**175420**	**100.0%**	**2376**	**100%**

Taking the 581 EPO filings including plant gene sequences shown in Table 2, we perform a more in-depth quantitative and qualitative analysis of different patent features. Several preliminary findings are worth mentioning:

- First, patent applications filed by public research organisation (PRO), alone or jointly with firms, have grown in recent years, although companies are the main applicants of plant genetic sequence patents in EPO.
- Second, small firms and PROs are significantly more likely to take the PCT route than large firms, which reflects the role of the PCT route as a way to gain time and look for funding to proceed with the patenting process and commercialisation of the protected invention, which is especially relevant for financially constrained small firms and PROs.
- Third, a handful of companies own the majority of the patents, following an intense process of concentration in which large agrochemical and pharmaceutical firms have bought up many smaller seed companies and biotech firms. We trace this process and find that at present the top five companies (Monsanto, Bayer, Basf, Syngenta and Dupont Pioneer) and their subsidiaries (including firms that have been bought by these five firms or are under shared ownerships) control 64% of all plant sequene patents in EPO corresponding to over 90% of the sequences for which a patent application has been made at EPO.
- Fourth, preliminary results of econometric estimations on the relation between patent characteristics and number of forward citations received show that patents filed by Bayer and Syngenta are more likely to receive a higher number of forward citations than those filed by BASF, conditional on other patent characteristics. Based on a 2009 report by Louwaars et al, we believe these firms have different business models: Syngenta and Bayer Cropscience develop seeds whereas BASF is into patenting traits and licensing these patents to seed companies. These different business models are also visible in simple correlation analysis, where we observe that BASF patent filings also tend to have a larger number of gene sequences than those of the other top companies.
- Fifth, Arabidopsis gene sequences, which are included in 35% of all the patents in our sample, seem to have a negative effect on the number of forward patent citations received

by the patents in our sample. This may be due to its use as research tool in a broad range of applications or the lack of direct commercial relevance of this organism.

Further quantitative and qualitative analyses, including an assessment of the specific subject matter and scope of the EPO filings included in the sample, are currently being undertaken to better interpret these preliminary results and the forces at play in the field of plant biotechnology in relation to genetic patenting.

Acknowledgements We thank Stéphane Maraut for for providing IT help and advice throughout the project. Laura Manzano Baena and Maria Eugenia Espinosa Calvo helped with the codification of USPTO patents. We also wish to acknowledge funding from CSO2011-2350, CSO2009-10845 and CSO2012-32844.

References

Arnaud-Haond, S., Arrieta, JM, Duarte, C.M., 2011, Global Genetic Resources - Marine Biodiversity and Gene Patents, *Science*, Vol. 331 no. 6024 pp. 1521-1522

Baillie, B., Connett Porceddu, M., 2012, Patent Landscape of the Arabidopsis genome and related plant species, http://www.patentlens.net/daisy/ArabidopsisGenome/1662.html

Louwaars, N., Dons, H., van Overwalle, G., Raven, H., Arundel, A., Eaton, D., Nelis, A., Breeding business, the future of plant breeding in the light of developments in patent rights and plant breeder's rights, CGN Report 2009-14, Centre for Genetic Resources, the Netherlands.

NCBI, 2012, BLAST - Basic Local Alignment Search Tool, available at: http://blast.ncbi.nlm.nih.gov/

How has International Collaboration been beneficial to Young Universities? A Case Study

K. A. KHOR* and L.-G. YU*

*mkakhor@ntu.edu.sg; mlgyu@ntu.edu.sg
Research Support Office and Bibliometrics Analysis, Nanyang Technological University, #B4-01, Block N2.1, 76, Nanyang Drive, Singapore 637331, Singapore

Introduction

It is widely assumed that research collaboration, especially international collaboration, has benefits for both the researchers and the organisations involved, and enhances the quality of research (Van den Besselaar et al., 2012). However, research also suggests that the effects of international collaboration may vary across disciplines and the authors' countries (Moed, 2005). Scholars in developing nations especially favour international collaboration, as their internationally collaborated papers will be more visible and more frequently cited in prestigious journals than their traditional papers without international collaboration (Cronin & Shaw, 1999).

In this study, the effect of international collaboration on the impact of publication of selected young universities (listed in the Times Higher Education 100 under 50 Universities) and old renowned universities (> 150 years old) was investigated. The 5-year citations per paper data, the percentages of overall publications and collaborated publications fall in the top 1% and 10% of global highly cited publications are used as the impact indications.

Method and Data

The Thomson Reuters (T-R) Web of Science (WoS) database is used to extract the publication data for collaboration analysis, and the international collaboration rate was calculated from the number of international collaborating publications therein. The 5-year citation per paper data in the T-R Essential Science Indicator (ESI) is used as one of impact indicator. The comparison of the percentage of publications that fall in the ESI global top 1% and 10% highly cited publications for all journal publications and for international collaboration publications is another indicator of impact. The analysis is based on papers published from 2003 to 2013. Only publications in the WoS Core Collection: Citation Index Expanded (SCI-EXPANDED), Social Sciences Citation Index (SSCI) and Arts & Humanities Citation Index (A&HCI) are used for publication search.

The international collaboration rates for the selected institutions are obtained by analysing yearly publication of the institutions and grouped by 5-year intervals. The collaborating papers among the selected young universities in the last 5 years are downloaded from WoS for international collaboration mapping. The mapping is carried out using the VOSViewer developed by CWTS in Leiden University.

Results and Discussion

Correlation between International Collaboration rate and Citations per Paper in 5-year interval

Figure 1 shows the 5-year Citation per Paper (CPP) Trends as a function of 5-Year International Collaborations Rate Trends for Selected Young and Old Universities. It can be seen that there is quite strong correlation between CPP increase and international collaboration increase for both young universities and old universities.

Figure 1: 5-Year Citation per Paper Trends vs. 5-Year International Collaborations Rate Trends for Selected Young and Old Universities.

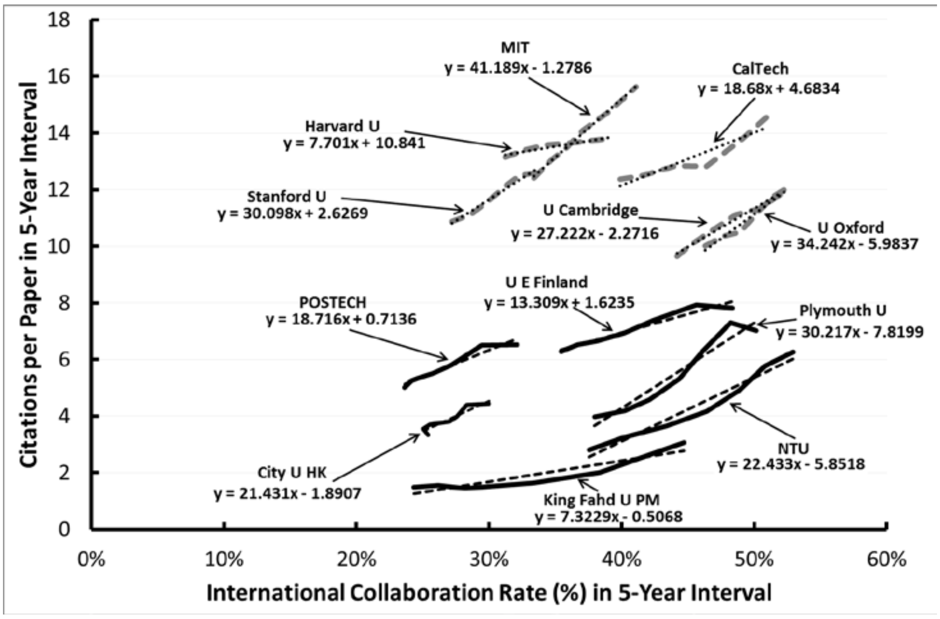


Figure 2 shows the relative increment of citations per paper of selected young and old institutions for internationally collaborated publications. It can be seen that although international collaboration benefits young universities as well as old universities, young universities has a higher relative citation per paper increment.

Figure 2. Relative Citations per Paper Increment over last 5-Year period

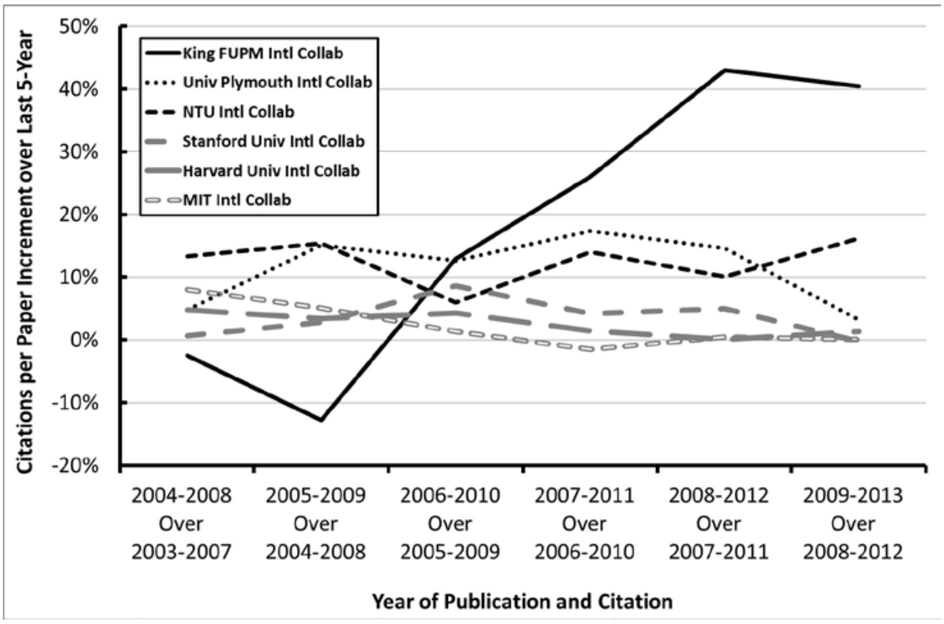
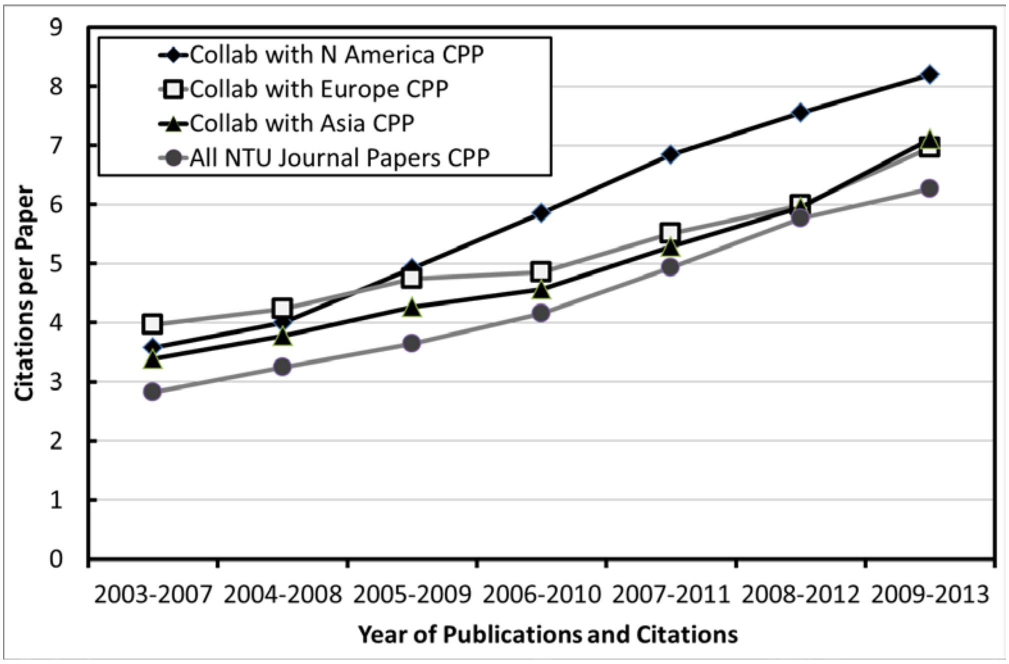


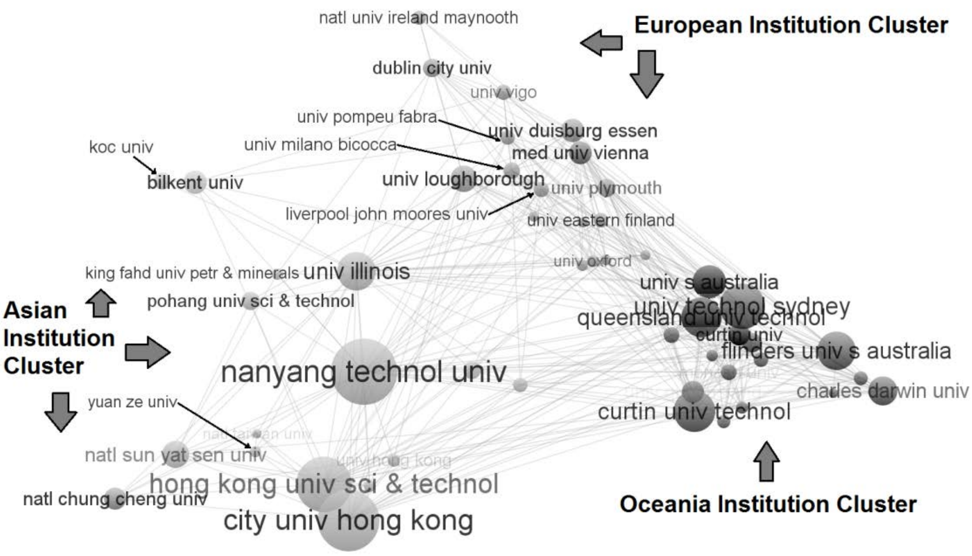
Figure 3 shows the citations per paper for publications of Nanyang Technological University (NTU) collaborating with institutions in North America, Europe and Asia. It can be seen that the collaboration with North American institutions gives highest citations per paper among the international collaborations. This trend applies to most of the selected young universities studied in this paper.

Figure 3. Citations per paper for publications of NTU collaborating with institutions in different regions of the world.



The mapping of the collaboration among the selected young institutions based on their international collaboration publications shown in Figure 4 indicates that, the regional collaboration is still stronger than long distance international collaborations. This may because the fact that international collaboration is still more costly and inconvenience.

Figure 4: Mapping Collaborations among Selected Young Universities.



The investigation of the percentage of publications of the selected institution fall in the ESI global top 1% and 10% highly cited publications for all journal publications and for international collaborated publications shows that the international publications have a higher rate of high citation publications compared to that of the overall publications of the institution.

Conclusions

In this study, a positive trend of international collaboration on the impact of the publication of the selected young and old institutions was found. Yet, the benefit of international collaboration on the impact of institutional publications varies from one to another institution. For example, for MTI, it is 4.12 CPP increment per 10% intl collab increment; for NTU, it is 2.24 CPP per 10% Intl Collab increment, and that for Plymouth Univ is 3.02 CPP per 10% Intl Collab increment, and 0.73 CPP per 10% Intl Collab increment for King Fahd Univ of Petr and Min. Although international collaboration benefits young universities as well as old universities, young universities has a higher relative CPP increment for one 5-year period over the previous 5-year period. The contributions of collaboration of young institutions with different institutions from different region of the world are also different; with the collaborations with institutions from North America have the highest citation per papers.

Although international collaboration is beneficial for young institutions, regional collaboration is the dominant style of collaboration between the institutions.

The percentage of publications fall in the ESI global top 1% and 10% highly cited publications for international collaborated publications is generally higher than that for all journal publications of the same institution.

References

Cronin, B. & Shaw, D. (1999). Citation, funding acknowledgement and author nationality relationships in four information science journals. *Journal of Documentation* 55(4), 402-408.

Moed, H.F. (2005). Citation analysis in research evaluation. Dordrecht, Netherlands: Springer.

THE World University Ranking, (2013/2014). From: http://www.timeshighereducation.co.uk/world-university-rankings/2013-14/world-ranking/methodology

Van den Besselaar, P., Sven Hemlin, S. & van der Weijden, I. (2012). Collaboration and Competition in Research. *Higher Education Policy*, 25, 263–266.

Indicators of Innovative Research

Richard Klavans*, Kevin W. Boyack**, Henry Small***, Aaron A. Sorensen****, John P. A. Ioannidis*****

* *rklavans@mapofscience.com*
SciTech Strategies, Inc., Berwyn, PA, 19312 (USA)

** *kboyack@mapofscience.com*
SciTech Strategies, Inc., Albuquerque, NM, 87122 (USA)

*** *hsmall@mapofscience.com*
SciTech Strategies, Inc., Bala Cynwyd, PA, 19004 (USA)

**** *tuf21130@temple.edu*
Temple University School of Medicine, Philadelphia, PA 19140 (USA)

***** *jioannid@stanford.edu*
Stanford University School of Medicine, Stanford, CA 94305 (USA)

Abstract
Disruptive innovation is considered to be essential for major progress in science, but it has been difficult to identify indicators that would predict that a given highly-cited paper is also highly innovative. We invited 400 highly prolific authors in biomedical science to complete a survey in which they rated their high citation impact papers along axes related to innovation, continuous progress and synthesis. These authors rated only one-fifth of their 1,233 highly cited papers as innovative. Using the survey data we tested a number of indicators designed to separate innovative high impact papers from other high impact papers. Traditional indicators, such as the number of references and whether the paper is coded as a review paper, have predictive value. Two new indicators, based on atypical knowledge relationships and cited journal influence, were also found to have predictive value. Implications are discussed.

Introduction
The development of an article-level indicator of innovativeness is critical for institutions and nations intending to pursue an innovation strategy. Such indicators are needed both for planning (assisting in the selection of innovative proposals) and evaluation (determining if the institution has taken appropriate risks or is primarily funding safe, i.e., conforming, research). This is a central question for science policy as evidenced by a recent commentary in *Nature* which claims that the NIH has a “conform and be funded” profile (Nicholson & Ioannidis, 2012). Every funding institution is sensitive to this issue, and tries to implement policies that will attract innovative research proposals.

The specific motivation behind this study was to test and validate bibliometric indicators of innovativeness recently proposed by Klavans & Boyack (2013) and Uzzi et al. (2013) (hereafter referred to as K&B and UMSJ, respectively). Given the lack of definitive data on which papers are truly innovative, we conducted a structured survey to gather these data. The data were then correlated with a number of indicators, most of which are based on work in K&B and UMSJ.

In the first section we describe how papers are classified as innovative or not for the purpose of this analysis. We then describe the indicators used in this study. Our primary findings focus on the four indicators that were able to statistically discern between innovative and non-innovative high impact research. While our results are significant, there is far more work to be done before we can claim that we have developed a reliable document-level indicator of innovativeness. The paper concludes with a discussion of possible directions for future research.

Classification of High-Impact Papers

As described elsewhere in detail, we invited 400 highly influential biomedical researchers to participate in a web-based survey. Using Scopus data, authors were identified using a combination of total citation counts (min=25,142) and h-index (min=76) (Boyack, Klavans, Sorensen, & Ioannidis, 2013). We identified the 10 most highly cited papers (published between 2005-2008) for each author based on citation counts as of end-2011 normalized by year, document type, and discipline. Authors were asked to rate each of their 10 papers on a scale of 0-100 along six different potential dimensions of impact. Authors were also allowed to add an 11th paper if their most important paper (published during this time period) had not been identified. Only 20 additional papers were added in this fashion.

The six dimensions of impact investigated were disruptive innovation, continuous progress, synthesis of existing literature, broader interest, obtaining surprising results, and difficulty in getting a paper through the review process. An innovative paper was defined in the survey as one that "introduces a radically new solution or framing of a problem that, if successful, changes the status quo." We received 123 responses, with corresponding evaluations of 1,233 documents. Detailed descriptive results about the survey are currently under review (Ioannidis, Boyack, Small, Sorensen, & Klavans, 2014).

For the purposes of the current analysis, each paper was categorized into one or more of the six types of impact using its maximum scores. For example, Table 1 lists the ratings for the ten articles by one of the authors. The first six papers can be unambiguously assigned – the highest score is only associated with one type. Papers G, H and I have two types receiving the same high score. These papers were each assigned to these two types with 50% confidence. There's a 50% chance that paper G should be assigned to synthesis and a 50% chance that the paper should be assigned to interest. Note that paper J is placed in a category called unassigned. Papers were considered unassigned if all six of the questionnaire responses were below the median or mean for the author. We found 58 papers matching this profile. 1,175 papers were correspondingly assigned to the six impact categories with different levels of confidence.

Table 1. Assigning papers to impact categories.

Paper	Innov	Prog	Synth	Inter	Surp	Diff	Assignment
A	50	60	***100***	90	10	10	Synthesis
B	***90***	40	80	70	80	40	Innovation
C	20	***90***	10	20	30	50	Progress
D	***80***	60	50	70	70	70	Innovation
E	30	70	80	***90***	20	30	Interest
F	10	***80***	70	70	10	10	Progress
G	50	50	***80***	***80***	30	10	Synthesis + Interest
H	***80***	40	40	40	***80***	50	Innovation + Surprise
I	***70***	40	***70***	60	20	10	Innovation + Synthesis
J	0	0	0	0	0	0	Unassigned

The results from recoding the data in this fashion are presented in Table 2. The most prominent type of high impact paper reports on normal progress. The second most popular type was synthesis. This type is correlated with progress (Pearson correlation of 0.334). The third most popular type was broad interest (correlated with synthesis at the 0.348 level).

Table 2. Types of high impact papers (n=1,233).

Type	# Papers	% Papers
Progress	355.4	28.8%
Synthesis	262.9	21.3%
Broad Interest	220.0	17.8%
Innovation	195.9	15.9%
Surprise	99.3	8.0%
Difficulty	41.6	3.4%
Unassigned	58.0	4.7%

Only 15.9% of the high impact papers are considered to be primarily innovative by their authors. This is probably a conservative number since (a) there are many papers that were assigned to more than one category, and (b) two of the categories (surprise and difficulty) were explicitly designed to capture other aspects of innovativeness (Pearson correlations between Innovation and Surprise, Surprise and Difficulty, and Innovation and Difficulty are 0.551, 0.388, and 0.276, respectively). Specifically, there were 255 papers that were considered innovative, surprising or difficult, but did not have a split vote with progress, synthesis or interest. The percentage of innovative papers is therefore estimated to be about one in five for this sample.

Bibliometric Indicators

The following seven indicators were used in this study:

<u>Distance (K&B Model):</u> At the 2013 STI conference, Klavans & Boyack (2013) proposed a new paper-level indicator of innovativeness based on science mapping. Millions of citing and cited papers are located on a two dimensional surface using a combination of direct citation clustering and textual analysis for layout (Boyack & Klavans, 2014b). Proximity of two objects on the map equates to non-innovativeness in how those two objects relate to each

other. We posited that non-innovative papers would be citing references that were in the same knowledge area (distance equals zero). Highly innovative papers would be citing references that were distant. We presented strong empirical support that distance affects citation rates, but no empirical support to claim that the indicator identifies innovative (vs. non-innovative) high-impact papers.

Atypical and Typical Knowledge Relationships (UMSJ Model): Uzzi et. al. (2013) recently proposed two paper-level indicators based on the assumption that highly innovative papers (1) focus on a specific area of knowledge in which to introduce new concepts and (2) make connections between areas of knowledge that are not typically made. The intent of their model is similar to ours, but implemented in a completely novel way. The authors used co-cited journal-journal relationships to determine whether any pair of cited references is typical or atypical. Using cited references from nearly 18 million articles, they calculated actual and expected counts for each co-cited journal pair and converted those counts into Z-scores. Negative Z-scores indicate that actual counts are less than expected, and reflect atypical knowledge relationships. Positive Z-scores indicate typical knowledge relationships. The authors show that articles that have higher than average typical relationships (using the median Z-score) combined with a high level of atypical relationships (using the left 10th percentile Z-scores) are twice as likely to be highly cited as the average article. We replicated this procedure and core results using Scopus data and the K50 metric (Klavans & Boyack, 2006) instead of Z-scores.

Co-cited Journal Influence: We are exploring the possibility that the number of times a journal is co-cited might be the underlying driver behind the UMSJ indicator of novelty. This is based on the finding that highly influential multidisciplinary journals (such as *Science* and *Nature*) are twice as likely to be associated with atypical knowledge relationships (Boyack & Klavans, 2014a). We therefore used the total number of journal co-citations as our indicator of co-cited journal impact, and calculated the mean score for all co-citation pairs in an article.

Review paper: Scopus tags each record with a document type (e.g., article, review, note, letter, etc.) We expect that synthesis papers and papers with a wider interest are more likely to be classified as reviews, and that innovative papers are less likely to be classified as reviews. We also expect papers coded as surprising, or that were difficult to get through the review process, to be less likely to be review papers. We noted that the UMSJ model did not include review papers (they are commonly considered less innovative). This study will therefore allow us to determine if the UMSJ indicators are able to discern between high impact innovative papers and high impact review papers.

Number of References: We hypothesize that synthesis and interest papers will have much larger reference lists than innovative papers. Innovative papers are targeted – the authors are introducing a new idea into a specific area. Synthesis and interest papers are broader – they are reviewing a wide landscape and should correspondingly have more references.

Number of Authors: We hypothesize that teams are larger when one is providing overviews (synthesis and wider interest) and when one is reporting on normal progress. We correspondingly expect the other three types (innovative, surprising and difficult) will have fewer authors. We note that the average paper in our sample has an exceptionally large number of authors (over 15). The authors in our sample seem to be embedded in large team efforts.

Findings

The results for all indicators by impact type are listed in Table 3. The indicator of atypical knowledge relationships was the only one of the three new K&B and UMSJ indicators that yielded significant results. The statistic for the papers coded as Innovative (.636) means that 63.6% of the papers in this group had a 10th percentile score that was negative. As expected, papers associated with synthesis, wide interest and normal progress did not build extensively on atypical co-citation relationships. Papers that were innovative, had difficulty during the review process or had surprising results had higher levels of novelty (the 10th percentile score was more likely to be negative). The unassigned papers had the highest level of atypical knowledge relationships.

The co-cited journal influence statistic also has a statistically significant ability to discern between types of high impact papers. This statistic tells us the average impact of the co-cited journals. Highly innovative papers are mostly citing papers that are appearing in highly co-cited journals (e.g., *Science*, *Nature*, *PNAS*, *Journal of Biological Chemistry*, *New England Journal of Medicine*). The less innovative papers are citing papers with lower average cited-journal co-citation rates.

Table 3. An Evaluation of Seven Potential Indicators (defined in the previous section) of Innovative High Impact Papers

Type	**Distance [K&B]**	**Atypical [UMSJ]**	**Typical [UMSJ]**	**Jnl CC x10^6**	**%Rev**	**#Ref**	**#Auth**
Synthesis	2.01	.506	.480	23.3	26.1	71.0	16.2
Interest	2.13	.574	.475	25.1	17.3	59.8	15.5
Progress	2.05	.491	.511	22.7	12.5	51.7	17.8
Unassigned	2.04	.672	.500	30.8	12.1	55.0	14.9
Difficulty	1.98	.620	.451	28.6	12.7	51.1	9.8
Innovation	2.05	.636	.467	24.9	6.3	42.9	12.3
Surprise	2.04	.608	.500	27.1	4.5	41.7	11.4
F-stat	0.42	**4.61**	0.41	**5.52**	**13.8**	**9.63**	**1.90**
Prob	.867	**.0001**	.875	**.0000**	**.0000**	**.0000**	**.0773**

By far, the most predictive indicators are whether the paper is a review article and has a large number of references. These two indicators are highly correlated (Pearson correlation of 0.439), and also generate the same rough ordering of types. At one extreme are papers doing synthesis and having broad interest; these are more likely to be a review paper and have many references. Papers that are unassigned, reporting progress or were difficult to get through the review process are least likely to be review papers and will have fewer references. The more innovative papers and those presenting surprising results are the least likely to be review papers and have the fewest references. The number of authors result in a similar ordering of types, but with lower statistical significance.

Discussion

Our initial analysis of the survey data in the summer of 2013 was disappointing in that the K&B distance indicator did not make meaningful distinctions between types of high impact papers. The UMSJ study (published in the fall of 2013) proposed a set of indicators we had not been considered and revitalized our interest in the subject. Our initial findings are encouraging; their indicator of atypical knowledge relationships is able to discern between different types of high impact papers.

These are tentative findings, and we are very concerned that the underlying driver is journal co-citation influence rather than atypical knowledge relationships. There are two reasons for this concern. First, we have seen both discipline-level and journal-level effects in the UMSJ results (Boyack & Klavans, 2014a). Second, we note the inability of the indicator for atypical knowledge relationships to identify review papers, even though it does a credible job in identifying papers that do a synthesis of the literature. If review papers tend not to be innovative, then an indicator of innovativeness should be able to make this discernment.

Implications

We posit that it is inappropriate to assume that high impact is equivalent to highly innovative. This is an assumption that underlies many bibliometrics studies about innovative research. An innovative paper may be highly cited, but the odds that a highly cited paper is innovative are only one in five (based on the sample in this study) and only one in three if we had excluded review papers.

Future studies also need to identify indicators for other types of significant contributions, such as synthesis and replication. While it is generally reasonable to assume that the most highly cited papers are making a contribution to science, we do not have effective means for discerning between different types of contributions. Survey methods are ideal for developing and testing bibliometrics indicators, and are being scheduled for other disciplinary areas. Future studies could also identify indicators of insignificant contributions. We did not explicitly ask this question in the survey, but are planning to in the future. It may be that the unassigned category, because of their low overall scores, is mostly composed of highly cited papers that have made relatively insignificant contributions to science. An exploration of this possibility is warranted.

Future studies might also take advantage of the fact that high impact review papers are less likely to be innovative. One approach worth exploring is a matched pair analysis - selecting a highly cited review paper and a highly cited research paper for specific knowledge areas. This would be an exploratory study to identify candidate indicators that can effectively discern between these two types. We may find that indicators of innovative research are field specific.

Finally, the role of journal influence, in both citing and cited papers, deserves far more attention in the future. Why would an innovative high impact paper tend to have more references to papers in high impact journals? Since it appears that innovative high impact papers have fewer overall references, does that mean that the references to the low influence journals were excluded? Or, to put it differently, if one is going to successfully review a field or report on normal progress, is there a tendency to include more references to the less influential journals? The use of cited journal influence as an indicator of an innovative high impact paper is an anomalous finding worth investigating. This finding may challenge our preconceived views of how to identify high impact innovative research.

References

Boyack, K. W., & Klavans, R. (2014a). *Atypical combinations are confounded by disciplinary effects*. Paper presented at the 19th International Conference on Science and Technology Indicators.

Boyack, K. W., & Klavans, R. (2014b). Creation of a highly detailed, dynamic, global model and map of science. *Journal of the Association for Information Science and Technology, DOI: 10.1002/asi.22990*.

Boyack, K. W., Klavans, R., Sorensen, A. A., & Ioannidis, J. P. A. (2013). A list of highly influential biomedical researchers, 1996-2011. *European Journal of Clinical Investigation, 43*(12), 1339-1365.

Ioannidis, J. P. A., Boyack, K. W., Small, H., Sorensen, A. A., & Klavans, R. (2014). What do the most-cited authors think about their most-cited papers? *under review at Science*.

Klavans, R., & Boyack, K. W. (2006). Identifying a better measure of relatedness for mapping science. *Journal of the American Society for Information Science and Technology, 57*(2), 251-263.

Klavans, R., & Boyack, K. W. (2013). *Towards the development of an article-level indicator of conformity, innovation and deviation*. Paper presented at the 18th International Conference on Science and Technology Indicators.

Nicholson, J. M., & Ioannidis, J. P. A. (2012). Conform and be funded. *Nature, 492*, 34-36.

Uzzi, B., Mukherjee, S., Stringer, M., & Jones, B. (2013). Atypical combinations and scientific impact. *Science, 342*, 468-472.

How accurately does output reflect the nature and design of transdisciplinary research programmes?[1]

Elizabeth Koier and Edwin Horlings*

*e.koier@rathenau.nl; e.horlings@rathenau.nl
Science System Assessment Department, Rathenau Instituut, Anna van Saksenlaan 51, The Hague, 2593 HW (the Netherlands)

Introduction

Governments around the world are using funding programmes and coordinating instruments to guide research towards themes and grand challenges that are relevant to society (Lyall & Fletcher, 2013). Grand challenges are wicked problems that affect the interests of a multitude of stakeholders. Developing solutions for wicked problems requires a new, transdisciplinary approach to the organisation of science (Rosenfield, 1992; Funtowicz & Ravetz, 1993; Spangenberg, 2011). A major challenge is that transdisciplinary science requires different criteria for quality assessment.

There is a growing body of literature on the evaluation of interdisciplinary and transdisciplinary science (Klein, 2008; Wagner et al., 2011). This literature shows that though definitions vary the objective of the evaluation of interdisciplinary and transdisciplinary research is always to assess to what extent the integration of knowledge was achieved. However, no single metric exists with which to assess the integration of knowledge (Wagner, et al., 2011). A multi-method approach is required (Stokols et al., 2003).

One of the methods which is consistently used in multi-method approaches to evaluation is bibliometric analysis. In this paper we examine the accuracy of output indicators for two climate programmes, investigate the usability of download statistics, a new and upcoming indicator, and assess the accuracy of links between data on participants, projects, and outputs.

Data

To assess the accuracy of programme-related data, we collected data on the participants and projects as well as on the outputs and downloads of two Dutch climate programmes. These data came from the programme's own databases as well as from the Web of Science. Within the programmes' databases, links between programmes, projects, participants, and outputs were embedded. Links between participants and authors were established by means of names, affiliations and – in ambiguous cases – using Google results and LinkedIn profiles.

The two climate programmes have a specific governance structure (foundation) that sets them apart from universities and research councils. Climate changes Spatial Planning (2004-2011; €80m) was set up to equip government and industry in the Netherlands with an operational knowledge infrastructure tailored to the relationship between (anthropogenic and natural)

[1] This work was supported by the Dutch national research programme Knowledge for Climate (www.knowledgeforclimate.org) [project SSA01]. This work was part of a meta-project to monitor the organisation, effects, and impacts of the programme called *Comparative Monitoring of Knowledge for Climate*. The Knowledge for Climate research programme is co-financed by the Ministry of Infrastructure and the Environment.

climate change, climate variability, and land use. It included 155 universities, public research organisations, (lower) governments, environmental NGOs and engineering consultancies. Most projects were carried out by more than one institution and often in collaboration with academic and non-academic organizations. Knowledge for Climate (2008-2014; €92m) was designed to develop the knowledge and tools with which it is possible to determine if the spatial and infrastructural investments of the next twenty years are climate-proof and, if necessary, to adapt them. Government (national and local) and firms actively participated in agenda setting.

Testing accuracy

The accuracy of bibliometric indicators needs to be tested before they can be used. Also, additional sources of information are needed to give a more reliable indication of the output and impact of a programme. The two climate programmes provide incentives to collaborate with stakeholders and encourage the production of a wider range of outputs, including tools, models, and policy studies. This means that we need to take into account both the scientific output and the non-scientific output of the programmes as well as the interactions between the different types of output.

We assess accuracy using the following questions:

1. Does scientific output accurately reflect the activities, nature and design of the programmes?
2. Do the programme and publication data accurately reflect the contribution of programme members to the performed research?
3. Are there alternative reproducible metrics that provide a more accurate picture of the output of transdisciplinary science?

Scientific output

The first step in our accuracy test concerns scientific output, the key performance indicator in the evaluation of research programmes. We compare two sources: (1) the programme's own list of peer-reviewed scientific articles reported by project leaders, and (2) Web of Science (WoS) metadata for those articles in the list that were published in journals indexed by the WoS.

It is a well-known fact that the WoS does not provide full coverage of scientific output. Some disciplines (e.g. medicine, physics) are represented more completely than other disciplines (e.g. the humanities). In many disciplines incomplete coverage is not considered an obstacle to bibliometric evaluation as is shown by its abundant use. The main argument is that the WoS is considered an index of high-quality journals that publish knowledge of global importance. In many disciplines, evaluators feel they can afford to forego articles not covered by the WoS. This is definitely not true for the social sciences and humanities, but it is less clear how this effects the evaluation of transdisciplinary research programmes.
We compare the list of peer-reviewed scientific articles reported by project leaders with a list of articles in the programme database that were published in journals indexed by the WoS What percentage of total output is indexed? And is coverage random or biased in relation to

the programmes' nature and design? Table 2. Peer-reviewed publications in the programme output database according to Web of Science coverage

	Number of publications	Percentage of the total number of publications
Total publications	480	100%
Found in the WoS	397	83%
Not found in the WoS	83	17%
of which:		
published in an international journal not indexed by the WoS	30	6%
published in an national (Dutch) journal not indexed by the WoS	19	4%
other publications not covered by the WoS (e.g. books)	17	4%
too recent to be found in the WoS (2012 or more recent)	2	0%
unclear [a]	15	3%

[a] The publication source is unclear in the programme database (n=2) or the publication source is a journal indexed by the WoS but the article was not found in the WoS (n=13).

We were able to find about 83% of the peer-reviewed articles in the output database in the WoS, using both automatic retrieval and manual corrections (Table 1). This is roughly comparable to the results of Ingwersen and Larsen (2007), who find that the WoS covers 79% of the reported output of the Danish Strategic Environmental Research Program SMP. A large part of the output that could not be found was published in Dutch journals or in non-indexed international journals. The latter often have a more practical orientation. This implies that we underrate performance with respect to one of the primary goals of the programmes: to produce practical solutions for local problems.

In many cases, a list of articles produced by the programme is not available. Since August 2008, the WoS provides a second method of entry by extracting funder information from the acknowledgements in publications. As a result, funder information is increasingly used to assess programme outcomes (Costas & van Leeuwen, 2012; Rigby, 2011, 2013). We have tested to what extent the programmes' publications can be retrieved from the WoS using funder information.

Table 2 presents the results. First we see that researchers do not always acknowledge their funders. We found that about 53% of the programmes' output after 2008 contains an acknowledgement to one of the programmes. The second observation is that there are inaccuracies in the retrieval of funder information where researchers do acknowledge their funders. The WoS derives its funding information by first recognising and retrieving the acknowledgement, and then extracting individual funders from the acknowledgement text. We found that both steps contain inaccuracies. The first step – recognition of the acknowledgement – is responsible for a loss of information on 24% of the articles. The second step – extraction of individual funders where the acknowledgement was correctly retrieved – involves a loss of information on another 12% of the articles. Of the 117 articles that do acknowledge funding by one of the two Dutch climate programmes (57+26+34), 60 articles (26+34, 51%) were not recognizable as such from the WoS funder list.

Table 3. Retrieval of acknowledgements and funders by the Web of Science

	Number of publications	Percentage of the total number of publications
Total publications	221	
Acknowledgement was correctly retrieved	***142***	***64%***
–the programmes' funders were correctly identified	57	26%
–the programmes' funders were not correctly identified	26	12%
–funders do not include one of the programmes	59	27%
Acknowledgement was mistakenly not retrieved	***54***	***24%***
–funders include one of the programmes	34	15%
–funders do not include one of the programmes	20	9%
There is an acknowledgement but it does not contain funder information	***7***	***3%***
There is no acknowledgement	***18***	***8%***

We can conclude that funder information in the WoS is as yet not sufficiently accurate to reconstruct a programme's output and researchers' acknowledgement practises are not completely reliable. An alternative explanation for the latter problem is that some outputs in the programme database are not the result of the programmes. We will examine this particular phenomenon in closer detail in the next subsection.

The contribution of programme members to output

For the purposes of evaluation, accurate attribution of output is of vital importance. The first, general question on attribution is whether the articles in the output database were written by programme members participating in the projects to which the articles were assigned. The second question is if the output of a programme reflects its nature and design, specifically with respect to stakeholder involvement.

Are project members co-authors of the papers assigned to the projects in which they participate? We find that in most articles (83%) at least one author is indeed a member of the project to which the article was assigned. Of the remaining 17% of articles, 5% involves authors who were members of other projects in one of the programmes than the project to which the article was assigned, and 12% involves articles of authors who were not project members themselves but who were colleagues of project members.

Co-authorship by non-academic organisations has been used to evaluate patterns and trends in public-private research collaboration (Abramo, D'Angelo, & Solazzi, 2010; Tijssen, 2011, 2012). Tijssen (2012) does acknowledge that public-private co-authorships probably underestimate the true amount of collaboration between academics and firms.

We have examined to what extent stakeholder involvement in the programmes and their projects is reflected in programme output. Since the Knowledge for Climate programme is still in progress, the analyses only relate to the Climate changes Spatial Planning programme. In 17 of the 33 CcSP projects with scientific output (52%), all consortium partners are represented in the affiliations of authors of scientific output in the WoS. In the remaining

48%, the division of the missing consortium partners is as follows: industry (33%), PRO's (23%), NGO's and universities (both 13%) and governments (10%).

The left side of table 3 shows the number of articles of which at least one author is affiliated to a type of organization. If we look at the same data from a project perspective (right side of the table), most of the projects (11 of 13) involving governments, firms and NGO's could not be recognised through author affiliations. This clearly shows the bias described by Tijssen (2012).

Table 4. Affiliations in project output

	Articles co-authored by at least one author affiliated to types of organisation N=304		Projects of which the consortium partners are recognisable from author affiliations in project output N=33		
type of organisation	Dutch and foreign affiliations	Dutch affiliations only	total number of projects involving organisations	projects not recognisable from author affiliations	percentage not recognisable
University	251	202	23	4	17%
PRO	193	155	26	8	31%
Government	6	5	3	2	66%
Firm	27	17	7	6	86%
NGO	6	0	3	3	100%

Alternative reproducible metrics

Transdisciplinary research programmes produce more than (peer-reviewed) scientific articles and aim to generate an impact outside science. Programmes collect information that provide insight into the diversity and usage of their outputs. Can such information provide an alternative reproducible metric for the output of a transdisciplinary research programme? We examine two types of programme data: non-scientific output and the number of downloads of programme output.

The non-scientific output reported by projects and uploaded into the programme database consists of a wide range of items including factsheets, presentations, reports, and newsletters. Table 3 shows the number of items per output type uploaded by projects into the programme's output database. Figure 1 shows that non-scientific items dominate total output throughout the programme's duration, especially from 2008 onwards.[2]

There are two problems associated with the programme's non-scientific output. First, the database contents depend mostly on self-reporting by project leaders who have different reporting habits. These differences affect the programme's aggregate output as well as the comparability of individual projects. Some projects seem to have added all the (scientific) presentations that were held at conferences, whereas others only report published scientific papers.

The second problem is that there is no common definition of the output of transdisciplinary research. Since the individual types of output are not comparable, it is also impossible to produce an aggregate measurement of non-scientific or societal output. In pure science programmes, there is at least agreement on the unit of output, namely articles, preferably in

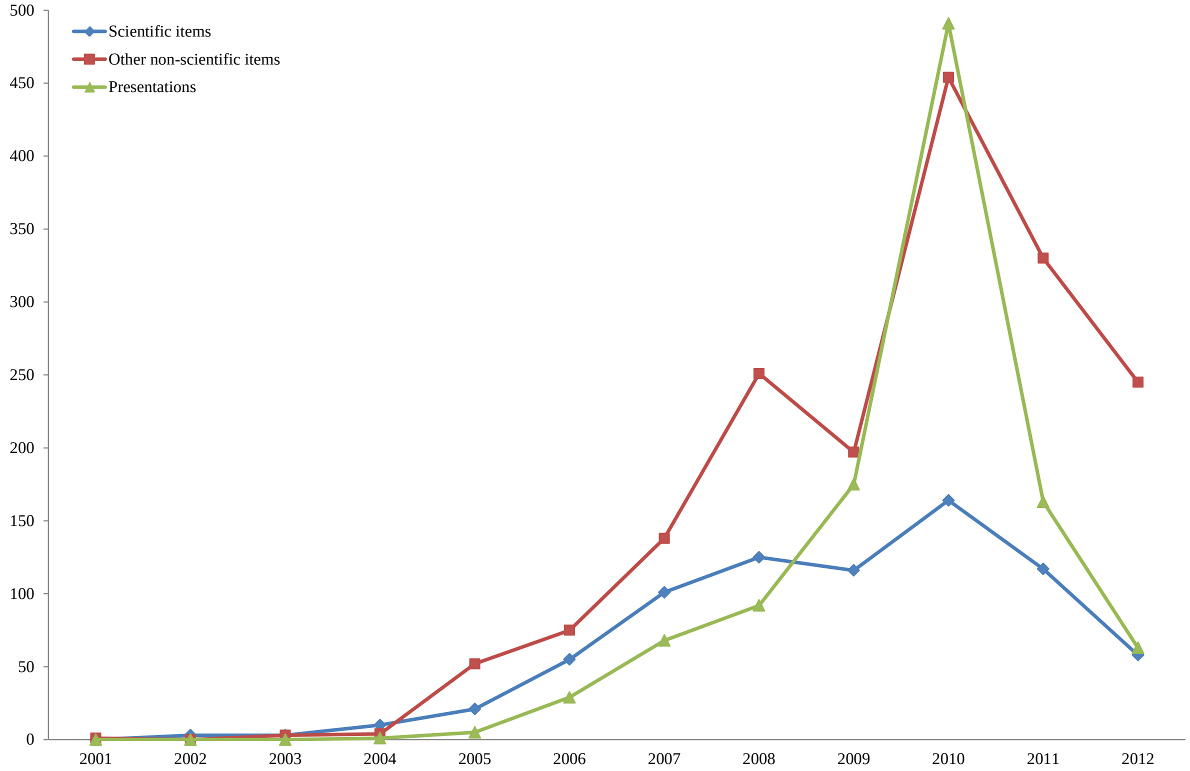
[2] The peak in 2010 may be due to the overlap and merger of the two programmes as well as the fact that KVR was in its final year.

high-impact journals, papers in the proceedings of high-status conferences, and books published with prestigious publishers.

Table 5. Programme output by type according to the programme database, 2001-2013

Type of output	Count
Scientific	
Scientific papers	106
Peer-reviewed scientific papers	542
Proceedings	115
PhD theses	11
Non-scientific	
Audio	2
Books	43
Brochures	68
Final project reports	171
Media	322
Press releases	26
Popular articles about science	117
Posters	159
Presentations	1087
Project factsheets	247
Project newsletters	81
Reports	515

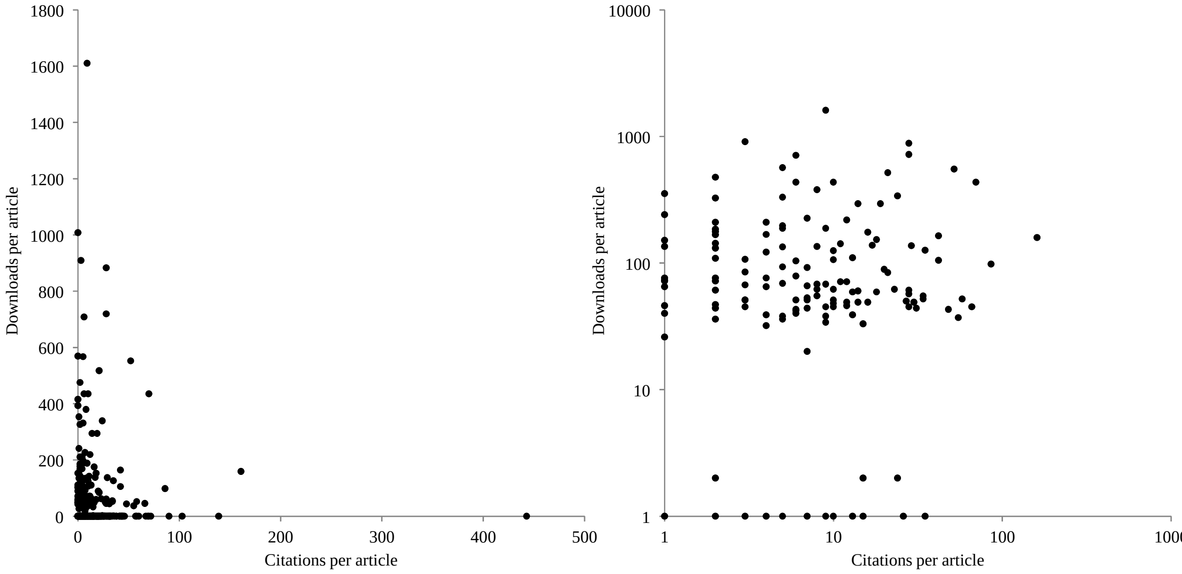
Figure 6. Development of the number of scientific and non-scientific items produced by the programmes according to the programme database, 2001-2012



One possible solution to the incomparability of different items of output is to use the statistics on annual downloads per item. If we consider downloads as the equivalent of citations, they may indicate the value of an item relative to other items. The use of download statistics and other similar indicators is part of a movement to produce alternative impact measures or altmetrics (Priem, Taraborelli, Groth, and Neylon (2010). The general idea is that alternative metrics show how scientific results are actually used by different audiences in different contexts, other than by science itself.

Figure 2 reveals that there is no relationship – either linear or logarithmic – between the total number of citations received and the number of times an item was downloaded from the programme database. There are two possible explanations, either the data are flawed or citations and downloads reflect different processes.

Figure 7. Linear (left) and logarithmic (right) relationship between total citation impact and downloads in 2009-2012 of peer-reviewed scientific articles in the programme database



The download statistics of the programmes are inherently flawed for two reasons. First, the download data of the Dutch climate programmes do not distinguish individuals or IP addresses on an item level. Consequently, it is possible for one individual to have been responsible for several downloads, for instance because it is an easy way to find the final version of one's own paper. Should download statistics become a formal metric, they can easily be manipulated. Therefore, they are not suited for comparative evaluation.

The second reason is that items in the database can be downloaded from multiple sources. Citation counts are generally derived from a single source (e.g. the Web of Science or Scopus). These sources may not cover all of science, but it is clear which journals are included and each citation is measured only once. The coverage of download data is less clear since publications, presentations, and other outputs may be available on the internet at several websites. Some items may not even be downloadable, such as links to websites. Download statistics are consequently incomplete and have an unknown bias.

Even if we were to overcome these flaws, downloads and citations may measure different aspects of usage and impact. For example, non-academic actors may use the programme outputs they have downloaded but are much less likely to cite them in a scientific paper.

Conclusions and discussion

In this paper we have examined the accuracy of one particular method from the mix of methods required for evaluation of transdisciplinary science, that of bibliometrics. Using all available information on the output of and participation in two large climate adaptation research programmes in the Netherlands, we have assessed the accuracy of the information on output and its relation to the programmes' organisation.

We can draw three conclusions. First, the Web of Science (WoS) covers a high percentage of the scientific activities of the programmes: 83% of scientific publications could be retrieved from the Web of Science. However, the recently added information on funding agencies is not yet sufficiently accurate to reconstruct a programme's output through the WoS. For now it is advisable to rely on the programme's own databases.

Second, scientific output in the WoS does not accurately reflect the nature and design of the programmes in that the WoS appears to contain less locally oriented and practically oriented research. Non-academic actors rarely co-author scientific publications and the contributions of non-academic organisations to projects could not accurately be recognised from author affiliations.

Third, alternative reproducible metrics, such as downloads, social media references, blog references, and hyperlinks, are seen as a promising expansion on current citation-based impact indicators. Our exploration of two such metrics – download statistics and the number of non-scientific output items – shows that it is too early to introduce such metrics into evaluation practices. There is no common understanding of the relative importance of different non-scientific outputs nor a standard for what needs to be reported. Download statistics are not accurate enough for comparative evaluation. They are incomplete, most likely biased, and can be manipulated. Successful implementation of altmetrics for the evaluation of transdisciplinary science will require a degree of regulation and standardisation across the sciences.

References

Abramo, G., D'Angelo, C. A., & Solazzi, M. (2010). Assessing public-private research collaboration: is it possible to compare university performance? *Scientometrics*, 84(1), 173-197.

Costas, R., & van Leeuwen, T. N. (2012). Approaching the "reward triangle": General analysis of the presence of funding acknowledgments and "peer interactive communication" in scientific publications. *Journal of the American Society for Information Science and Technology*, 63(8), 1647-1661.

Funtowicz, S. O., & Ravetz, J. R. (1993). Science for the post-normal age. *Futures*, 25(7), 739-755.

Ingwersen, P., & Larsen, B. (2007). Evaluation of strategic research programs: the case of Danish environmental research 1993-2002. Research *Evaluation*, 16(1), 47-57.

Klein, J. T. (2008). Evaluation of interdisciplinary and transdisciplinary research - A literature review. *American Journal of Preventive Medicine*, 35(2), S116-S123.

Lyall, C., & Fletcher, I. (2013). Experiments in interdisciplinary capacity-building: The successes and challenges of large-scale interdisciplinary investments. *Science and Public Policy*, 40(1), 1-7.

Priem, J., Taraborelli, D., Groth, P., & Neylon, C. (2010). *Alt-metrics: a manifesto*.

Rigby, J. (2011). Systematic grant and funding body acknowledgement data for publications: new dimensions and new controversies for research policy and evaluation. *Research Evaluation*, 20(5), 365-375.

Rigby, J. (2013). Looking for the impact of peer review: does count of funding acknowledgements really predict research impact? *Scientometrics*, 94(1), 57-73.

Rosenfield, P. L. (1992). The potential of transdisciplinary research for sustaining and extending linkages between the health and social-sciences. *Social Science & Medicine*, 35(11), 1343-1357.

Spangenberg, J. H. (2011). Sustainability science: a review, an analysis and some empirical lessons. *Environmental Conservation*, 38(3), 275-287.

Stokols, D., Fuqua, J., Gress, J., Harvey, R., Phillips, K., Baezconde-Garbanati, L., et al. (2003). Evaluating transdisciplinary science. *Nicotine & Tobacco Research*, 5(Supplement 1), S21-S39.

Tijssen, R. J. W. (2011). Joint Research Publications: A Performance Indicator of University-industry Collaboration. *Evaluation in Higher Education*, 12, 19-40.

Tijssen, R. J. W. (2012). Co-authored research publications and strategic analysis of public-private collaboration. *Research Evaluation*, 21(3), 204-215.

Wagner, C. S., Roessner, J. D., Bobb, K., Klein, J. T., Boyack, K. W., Keyton, J., et al. (2011). Approaches to understanding and measuring interdisciplinary scientific research (IDR): A review of the literature. *Journal of Informetrics*, 5(1), 14-26.

Altmetrics-based Visualizations Depicting the Evolution of a Knowledge Domain[1]

Peter Kraker*, Philipp Weißensteiner** and Peter Brusilovsky***

*pkraker@know-center.at
Marshall Plan Scholar, Know-Center, Inffeldgasse 13/VI, Graz, 8010 (Austria)

** philipp.weissensteiner@student.tugraz.at
Graz University of Technology, Rechbauerstraße 12, Graz, 8010 (Austria)

*** peterb@pitt.edu
Personalized Adaptive Web Systems Lab, University of Pittsburgh, 135 North Bellefield Avenue, Pittsburgh, 15260 (USA)

Introduction

Altmetrics have recently received a lot of attention in the scientometric community and beyond. A lot of effort is going into assessing the potential of altmetrics for evaluative purposes. Yet another strand of research has emerged, focusing on altmetrics for establishing relations between articles, journals and research fields. One problem that is considered within the latter strand of altmetrics research is knowledge domain visualization. So far, clicks (Bollen and van de Sompel, 2006) and readership (Kraker et al., 2013) have been successfully employed to map a scientific domain. These efforts, however, represent the state of a domain, but not the evolution of a domain.

The evolution of scientific domains has been addressed in the past primarily using citation-based analysis; see e.g. Garfield (1966) and Small (1999). There is, however, a significant problem with citations: it takes around two to six years after an article is published before the citation count peaks (Amin and Mabe, 2000). Therefore, citation-based visualizations, and indeed all analyses that are based on incoming citations, have to deal with this time lag. Altmetrics have thus emerged as a potential alternative to citation data. Compared to citation data, altmetrics have the advantage of being available earlier, many of them shortly after the paper has been published. In many instances, usage statistics are also easier to obtain and collect (Haustein and Siebenlist, 2011).

In this paper, we present work towards showing the evolution of scientific domain using data from a scientific conference scheduling software. Conference Navigator 3 (Parra et al., 2012) allows conference attendees to create a personal conference schedule by bookmarking talks from the program that they intend to follow. This scheduling data represents an altmetrics source which – to the best knowledge of the authors – has not been studied before.

Data and Method

All data is sourced from Conference Navigator 3. As a use case, we have chosen to analyse the 19th and 20th iteration of the Conference on User Modelling, Adaptation and

[1] This contribution is partly funded by the Know-Center, which is funded within the Austrian COMET program - Competence Centers for Excellent Technologies - under the auspices of the Austrian Federal Ministry of Transport, Innovation and Technology, the Austrian Federal Ministry of Economy, Family and Youth, and the State of Styria. COMET is managed by the Austrian Research Promotion Agency FFG.

Personalization (UMAP), representing the conference years of 2011 and 2012. The data sets relevant to this analysis were article metadata of the two UMAP conferences and the scheduling information associated with these articles.

The procedure for creating knowledge domain visualizations follows the approach used in the knowledge domain visualization Head Start which employs Mendeley readership data to map a research field (Kraker et al. 2013). It is adapted from the knowledge domain visualization process described in Börner et al. (2003).

At first, individual knowledge domain visualizations were created for each year. Therefore, the article metadata and the scheduling data have been extracted from the system ranked by the number of bookmarks received. A threshold of 4 (3) bookmarks was introduced for 2011 (2012) to ensure that enough scheduling data for the subsequent processing steps was present. Then, a co-bookmarking matrix for the remaining articles was created. On top of this matrix, we performed multi-dimensional scaling (MDS) and hierarchical agglomerative clustering (HAC) to create the visualization. We used HAC with Ward's method to establish the research areas. We then employed non-metric MDS to place the articles on the map. To unclutter the articles in the map, a force-directed layout with a collision detection algorithm was used. Finally, article titles and abstracts were sent to the text mining services Zemanta and Open Calais to find appropriate labels for the areas. Both services return a number of concepts that describe the content.

As far as time series visualization goes, there are many types of visualizations, most prominently index charts and stacked graphs. In the case of knowledge domain visualizations, simple visualizations are unfortunately not able to convey all necessary dimensions of the data (in terms of ordination, size of research areas and closeness). One possibility would be to use animation to show changes in a domain over time. Psychological studies have shown, however, that people are bad at recognizing change in an object or a scene. This phenomenon is called change blindness (Simons and Rensink, 2005). Therefore, the approach of small multiples (Tufte, 1990) was chosen. In small multiples, a graph is drawn for each of the steps in a time series. Then the graphs are positioned next to each other. This approach thus allows for direct visual comparison between different representations.

Results

The result of the visualization procedure for a single year detailed above can be seen in Figure 1. The blue bubbles represent research areas. The centre point of each circle was calculated as the means of the coordinates of the articles based on the MDS result. The size of the areas is determined by number of bookmarks that the papers have received. Spatial closeness implies topical similarities. As can be seen, “User modeling” is the area with most papers and most bookmarks. It is closely connected to several other larger areas, including “Recommender system”. A second cluster of areas can be found on the right hand side of the visualization, involving “Intelligent tutoring system”, “Adaptive system”, and “Problem solving”.

Figure 2 shows the evolution of the topical overview using small multiples. To aid the user in detecting changes between the representations, we introduced two visual helpers. First, a grid is drawn to help with comparing size and position of the research areas. Second, whenever users hover over an area, the corresponding area is highlighted in the other representation, and a line is drawn between the two entities. There are three areas that are present in both years: “User modelling”, “Recommender system” and “Intelligent tutoring system”. While the relative position of the areas to each other has not changed much, the area with the most papers and bookmarks is now “Recommender system”.

Figure 1: Topical visualization of UMAP 2011.

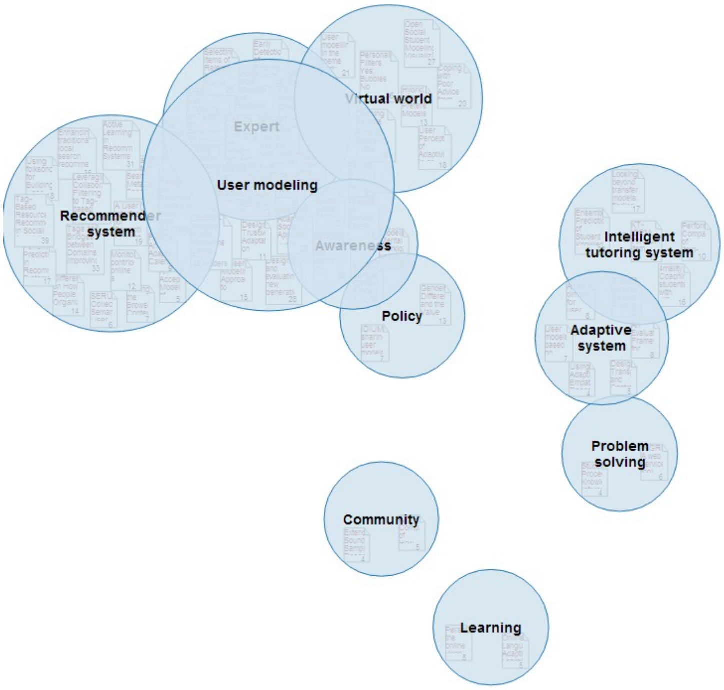
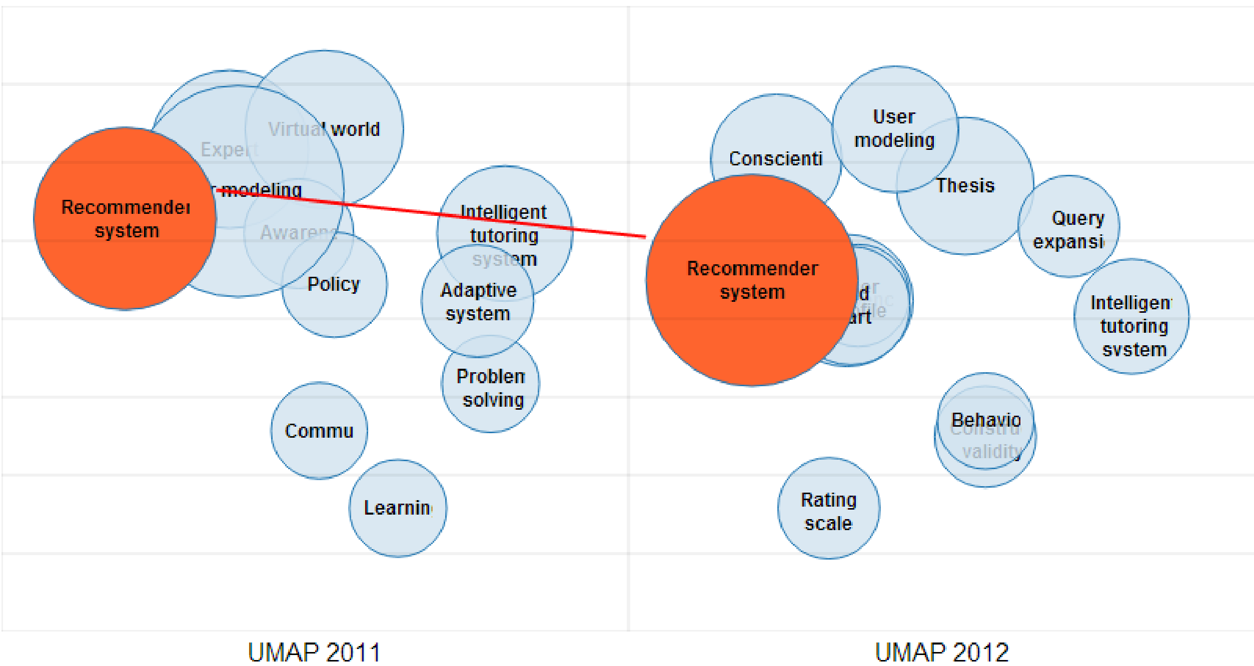


Figure 2: Evolution of the domain of UMAP from 2011 to 2012, with the area “Recommender system” highlighted



Discussion and Future Work

The first results from this type of analysis are encouraging. Using small multiples allows for a comparison of knowledge domain visualizations over various years. The work, however, also revealed several weaknesses of the current approach. First, the topology needs to be improved. As many of the areas are overlapping, it becomes harder and harder to disambiguate between the different clusters. One way to overcome this problem might be to employ force-directed placement.

Second, scheduling data in CN3 is sparser than readership data in Mendeley due to the fact that the audience of CN3 is restricted to conference participants. This means that the results of the clustering vary more when choosing different threshold values. Therefore, we want to explore supplementing bookmarks with content-based measures.

Third, the continuity between the two years is very low. There are only three areas that are present in both iterations. Therefore, it might be worthwhile to explore moving time windows of two years to show how the different papers move when associated with earlier or later years.

Finally, it will be important to evaluate the method and the interface. One way to go about this would be to ask experts to critically review the visualizations, and to give them the ability to manipulate size and location of papers and areas. This would allow for comparing the automatically created visualizations to the experts' perception of the field. Another possibility would be to contrast the visualization based on conference papers with the evolution of the field based on other types of literature (e.g. journal articles).

References

Bollen, J., & van de Sompel, H. (2006). Mapping the structure of science through usage. *Scientometrics*, *69*(2), 227–258. doi:10.1007/s11192-006-0151-8

Börner, K., Chen, C., & Boyack, K. W. (2003). Visualizing knowledge domains. *Annual Review of Information Science and Technology*, *37*(1), 179–255. doi:10.1002/aris.1440370106

Garfield, E., Sher, I., & Torpie, R. (1964). *The use of citation data in writing the history of science* (p. 75).

Haustein, S., & Siebenlist, T. (2011). Applying social bookmarking data to evaluate journal usage. *Journal of Informetrics*, *5*(3), 446–457.

Kraker, P., Jack, K., Schlögl, C., Trattner, C., & Lindstaedt, S. (2013). Head Start: Improving Academic Literature Search with Overview Visualizations based on Readership Statistics. In *Web Science 2013*.

Kraker, P., Körner, C., Jack, K., & Granitzer, M. (2012). Harnessing User Library Statistics for Research Evaluation and Knowledge Domain Visualization. In *Proceedings of the 21st International Conference Companion on World Wide Web* (pp. 1017–1024). Lyon: ACM.

Parra, D., Jeng, W., Brusilovsky, P., López, C., & Sahebi, S. (2012). Conference Navigator 3: An online social conference support system. In *UMAP Workshops* (pp. 1–4).

Simons, D. J., & Rensink, R. A. (2005). Change blindness: past, present, and future. *Trends in Cognitive Sciences*, *9*(1), 16–20. doi:10.1016/j.tics.2004.11.006

Tufte, E. R. (1990). *Envisioning* Information (p. 127). Graphics Press.

Performance Analysis of Major Basic Science Research Institutes By Bibliometric Methods: The status of basic science research in Korea and establishment of IBS

Jun Young Lee, Naeyang Jeong, Kyoung Hun Kim, Dae Nyoung Heo and Choong Han Song

jylee@ibs.re.kr; nyjeong@ibs.re.kr; khkim@ibs.re.kr; nyoung@ibs.re.kr; chsong@ibs.re.kr
Institute for Basic Science (IBS), 70, Yuseong-daero 1689-gil, Yuseong-gu, Daejeon 305-811(Republic of Korea)

Introduction

Until now, the science and technology policies of Korea have only been focusing on short term application and development research. Subsequently the investment in the basic (fundamental) science that requires the long term commitment has been lacking and the economic growth through the catch-up R&D centered on applied and development research has reached the limit. Furthermore, there has been a greater emphasis on the importance of challenging and creative research recently. In response the Korean government has increased the investment in basic research with the intention of using it as the driving force to boost the creative economy, and it established the Institute of Basic Science (IBS) as a representative institute that will lead such a move. This study aims to identify the current position of basic science of Korea and propose the future direction through the bibliometric analysis of research performances of major basic science research institutes.

Methods

To identify the status of basic science in Korea, publications from major basic science research institutes were comparatively analysed. Research papers and researchers within the Top 1%, referred as a highly cited paper and highly cited researchers, of the citation in the research fields related to the basic science between 2003 and 2012 were extracted by the bibliometric analysis using Elsevier's Scopus DB. In order to conduct this analysis, basic science is sorted using their ‘All Science Journal Classification (ASJC)’ codes to achieve an effective classification. Moreover, the ratio of the Top 1% researchers from institutes was compared. IBS and world-class Institute research groups were selected among representative basic science institutes, as two groups in similar research field & scale, and a comparison of research performances between those groups was also performed.

Results

Table 1. Comparison of the performance of representative basic science institutes in Korea(K) and overseas.

Institutes	Total number of papers	Total number of citations	Institutes	Average number of citations per paper
Institute A	86,679	1,872,044	Institute C	31.22
Institute B (K)	45,714	505,319	Institute F	24.02
Institute C	22,015	687,399	Institute A	21.60
Institute D	21,191	403,762	Institute D	19.05
Institute E (K)	18,599	187,481	Institute B (K)	11.05
Institute F	11,540	277,200	Institute E (K)	10.08
Institute G (K)	7,383	74,040	Institute G (K)	10.03

According to the result obtained by comparing the performance of Korea's representative national research institute and an outstanding university with the performance of basic science institutes in overseas over the period from 2003 until 2012, the total number of research papers and paper citations of Korean institutes except Institute B is considered to be very low. The rank of Korean research institutes was also lower in terms of the average number of citations per paper compared to basic science institutes in overseas, eventually demonstrating that the performance level of Korean research for basic science was the lowest. These comparisons clearly showed that the performance of basic science in Korea has been evaluated in a quantitative way, and the qualitative performance is only in the beginning stage and needs to be improved. For revitalizing such a low performance in basic science research and enhancing its quality, outstanding researchers in the basic science field from Korea and overseas have gathered together and established IBS.

Table 2. Comparison of the ratio of Top 1% researchers of IBS with the ratio of Top 1% researchers of representative research institutes in overseas.

Institutes	Num# of researchers (A)	Num# of Top 1% researchers (C)	Percentage of Top 1% researchers (C/A, %)
Institute A	5,470	318	5.8
Institute C	1,685	48	2.8
Institute F	1,000	57	5.7
IBS	176	23	13.1

When the ratio of Top 1% researchers of IBS is compared to that of representative research institutes in overseas, Institute C with 48 people of Top 1% researchers showed the lowest value as 2.8%, while IBS had 23 out of 176 researchers as Top 1%. The IBS' proportion of Top 1% researcher was 13.1% showing the highest level among the representative basic science institutes in overseas, and it is expected that IBS will play a significant role in raising the Korea's basic science research level.

Table 3. Comparison of research performance between IBS and Institute A research group.

	Variable	IBS	Institute A
Scale comparison of research center	Director	1	1
	Group leader	4	6
	Research fellows	11	36
	Trainees (Students)	82	42
	Total	98	83
Comparison of research performance	Total number of papers	972	1,106
	Total number of citations	29,357	22,352
	Total number of citations ('08-'12)	15,407	12,041
	Average number of papers per person	9.9	13.3
	Average number of citations per person	299.6	269.3

The result from comparing the level of the Institute A research groups, which were selected from representative basic science institutes as in a similar research field and scale, obviously showed that IBS research group was equivalent or even at a higher level in quality than world-class Institute A research groups. Currently, the number of IBS' researchers is relatively small in scale, but IBS is expected to have a continuous growth with the aim of world's best research institute.

Conclusions

From the result of comparing paper performance of major basic science research institutes, it has shown that the research performance of Korean basic science institutes was inferior to that of overseas basic science institutes. However, it was found that IBS was the highest in the ratio of Top 1% researcher among representative basic science institutes in overseas, and IBS was at an equal level of quality when compared to world-class Institute A. In order to establish a climate for Korea's basic science research to make more adventurous and long-term researches and produce performance of creative basic researches that can lead the global academic world, the autonomy and excellence of IBS' researches will be secured, and a new advanced system that allows "honorable failure" in a research process for an immersion in creative research will be driven. Such an effort is expected to demonstrate the positive influence of IBS' creative and outstanding research performance on the whole basic researches and IBS' important role in upgrading the level of Korean basic science.

An in-depth analysis of papers retracted in the Web of Science

T.N. Van Leeuwen and M. Luwel

{Leeuwen,Luwel }@ cwts.leidenuniv.nl
Center for Science and Technology Studies (CWTS),
Leiden University, PO Box905, 2300 AX Leiden (Netherlands)

Introduction

By publishing their results in the open literature researchers claim the ownership and make them accessible for scrutiny by other researchers. Before a journal with an international standing accepts a manuscript it is reviewed by one or more reviewers and many are rejected for various reasons.

However this peer review process can be flawed. If results presented in the scientific literature turn out to be not trustworthy, the publications are (partially) retracted. In principle these papers are removed from the common stock of knowledge. There are many reasons for retracting a paper, scientific misconduct being only one of them. However these cases attract a considerable amount of attention and undermine the public trust in science.

Over the last couple of years the retraction of publications became a popular subject in science studies (Grieneisen & Zhang, 2012 and references therein). Authors studied mainly retractions of papers processed for MEDLINE or subsets of retracted papers published in journals processed for Thomson Reuters Web of Science (WoS). There is a general impression that the number of retracted papers is increasing over the years and scientific misconduct is becoming more important (Steen, Casadevall & Fang, 2013 and references therein).

The study’s aim is a comprehensive analysis of all retracted papers published in journals processed for WoS. A detailed classification is produced of the motives for the retraction as stated in the retraction notices and of the indexing of these notices in the table of content of the online version of the journals, two subjects paid little attention to in previous work. We report on the preliminary results of ongoing work.

Methods and data

In the WoS the suffix “Retracted article” is added to the title of articles that are officially retracted (Chen et al., 2012). Using the advanced search mode on the online version of the WoS on January 7, 2014, in total 2.479 publications were identified with this suffix (search string: all years, all languages and all document types). The full bibliographic record of these retracted publications was downloaded.

The database created at CWTS from WoS data was used to make a bibliometric analysis of the retracted papers. We present the results of the distribution of the countries mentioned in the papers’ address by-line. A full counting scheme at country level is applied.

The journals processed for the WoS are assigned at least to one subject category. The retracted papers are assigned to the subject category or categories of their journals. The distribution over the subject categories of the retracted papers is compared with the distribution of all papers processed for the WoS. Again a full counting scheme is used: a paper published in a journal assigned to several subject categories, is counted in full to each category.

In the title of the retracted article, the suffix “Retracted article” is followed by the bibliographic data of the retraction notice (volume, page, year). We manually retrieve from the Leiden electronic library the pdf-version of the retraction notices or if not available, we search for the hard copy. One notice may contain information on more than one retracted article.

From the retraction notices the information on the party responsible for the retraction, the reason of retraction (further called retraction type), the journal section under which the retraction notice is classified and the Digital Object Identifier (DOI) is retrieved. A classification scheme is developed for the first two items. This very labour intensive work is still in progress and in this paper we report on preliminary results based on a subset of 200 couples ‘retracted paper – retraction notice’.

It is well known that retracted papers are cited often years after the retraction date (see e.g. Chen et al., 2012). Although interesting the number of the citations gives little information of the continuing impact of these publications. For a few highly publicised fraud cases we plan to carry out a sentiment analysis of citations to retracted papers before and after retraction (Li et al., 2013).

Preliminary results

Table 1 gives the countries most often mentioned in the address by-line of the retracted publications. Papers (co)signed by authors from the USA, PR China, and especially Japan and India are retracted more frequently compared to these countries’ share in the total number of publications processed for the WoS. Overall for each country the fraction of retracted papers in its total number of WoS papers is very small, but it varies between countries by a factor of 4.

Table 1. Country of affiliation of the authors of retracted papers for countries with more than 10 retracted papers: the number of retracted papers (# retracted), the percentage share in the total number of retracted papers (% retracted),the country’s percentage share in the total number of papers processed for the WoS in the period 2009-2013 (% WoS) and the per mille share of the retracted papers in the country’s total number of papers processed for the WoS in this period.

Country	# retracted	% retracted	% WoS	‰ retracted in WoS
USA	624	26,70	21,48	0,11
PEOPLES R CHINA	316	13,52	9,72	0,14
JAPAN	217	9,29	4,57	0,18
GERMANY	163	6,97	5,61	0,13
INDIA	124	5,31	2,75	0,21
ENGLAND	97	4,15	6,09	0,08
SOUTH KOREA	79	3,38	2,63	0,11
CANADA	58	2,48	3,48	0,07
NETHERLANDS	55	2,35	2,03	0,12
ITALY	53	2,27	3,38	0,06
FRANCE	49	2,10	3,99	0,05
AUSTRALIA	48	2,05	2,70	0,08

About 40% of the retracted papers are published in journals covering medicine and life science, three times higher than the share of these categories in total number of publications processed for the WoS. The same ratio is found for the multidisciplinary journals such as Nature and Science. 'Crystallography' is the only category from the physical sciences in the subject categories with more than 50 retracted papers. The share of retracted papers in the total number of papers assigned to the 14 subject categories varies by a factor of 10: the category 'Anaesthesiology' has the highest (3,81‰) and 'Pharmacology & Pharmacy' the lowest fraction (0,34‰).

Table 2. Subject categories of the journals publishing the retracted papers with more than 50 papers: the number of retracted papers (# retracted), the percentage share in the total number of retracted papers (% retracted), the category's percentage share in the total number of papers processed for the WoS in the period 2009-2013 (% WoS) and the per mille share of the retracted papers in the category's total number of papers in this period.

Journal Subject Category	# retracted	% retracted	% WoS	‰ retracted in WoS
BIOCHEMISTRY & MOLECULAR BIOLOGY	216	7,54	2,36	0,83
CELL BIOLOGY	135	4,71	1,12	1,11
MULTIDISCIPLINARY SCIENCES	127	4,43	1,37	0,84
CRYSTALLOGRAPHY	118	4,12	0,42	2,61
ONCOLOGY	112	3,91	1,50	0,69
ANESTHESIOLOGY	97	3,39	0,23	3,81
IMMUNOLOGY	95	3,32	0,99	0,88
SURGERY	80	2,79	1,62	0,45
MEDICINE, RESEARCH & EXPERIMENTAL	66	2,30	0,76	0,82
CARDIAC & CARDIOVASCULAR SYSTEMS	60	2,09	0,93	0,60
PHARMACOLOGY & PHARMACY	60	2,09	1,62	0,34
HEMATOLOGY	59	2,06	0,55	1,00
NEUROSCIENCES	59	2,06	1,59	0,35
MEDICINE, GENERAL & INTERNAL	53	1,85	1,16	0,42

Based on the sample already analysed a classification of the parties responsible for the retraction (retracting party) is made. Table 3 gives the different parties and the corresponding percentages of retracted papers. A retracted paper has been assigned to only one type. With nearly 70% the 'Authors' are the dominant retracting party. However, within this group it has to be noticed that within 10% of the retractions, one of the authors disagreed ('Authors except one'), and a few papers were retracted by one single author. The editors, and in some cases together with the publisher retracted 30% of the papers. It should also be notice that in few cases the editor and the authors together retracted the paper.

Table 3. The distribution of the retracted papers over the retracting parties

Retracting party	%
All authors	51,9
Authors except one	10,5
Editor (& publisher)	29,3
Editor and authors	3,8
One author	2,8
Ambiguous	1,9

Only for a few retractions the retraction notice mentions the reason. However as already remarked in other studies (Fang, Steen & Casadevall, 2012), the retraction notices can be uninformative opaque, hiding the underlying arguments, and secondary sources are necessary to clarify the real reasons. This is illustrated by the notice published to retract a number of papers published in Science in the well-known ‘Schön case’ (Reich, 2009):

We are writing as co-authors on the following manuscripts published in Science, which were, in part, the subject of an independent investigation conducted at the behest of Bell Laboratories, Lucent Technologies. The independent committee reviewed concerns related to the validity of data associated with the device measurements described in the papers.
….
As a result of the committee’s findings, we feel obligated to the scientific community to issue a retraction of the above articles. We note that although these papers may contain some legitimate ideas and contributions, we think it best to make a complete retraction.’
(Boa et al, 2002)

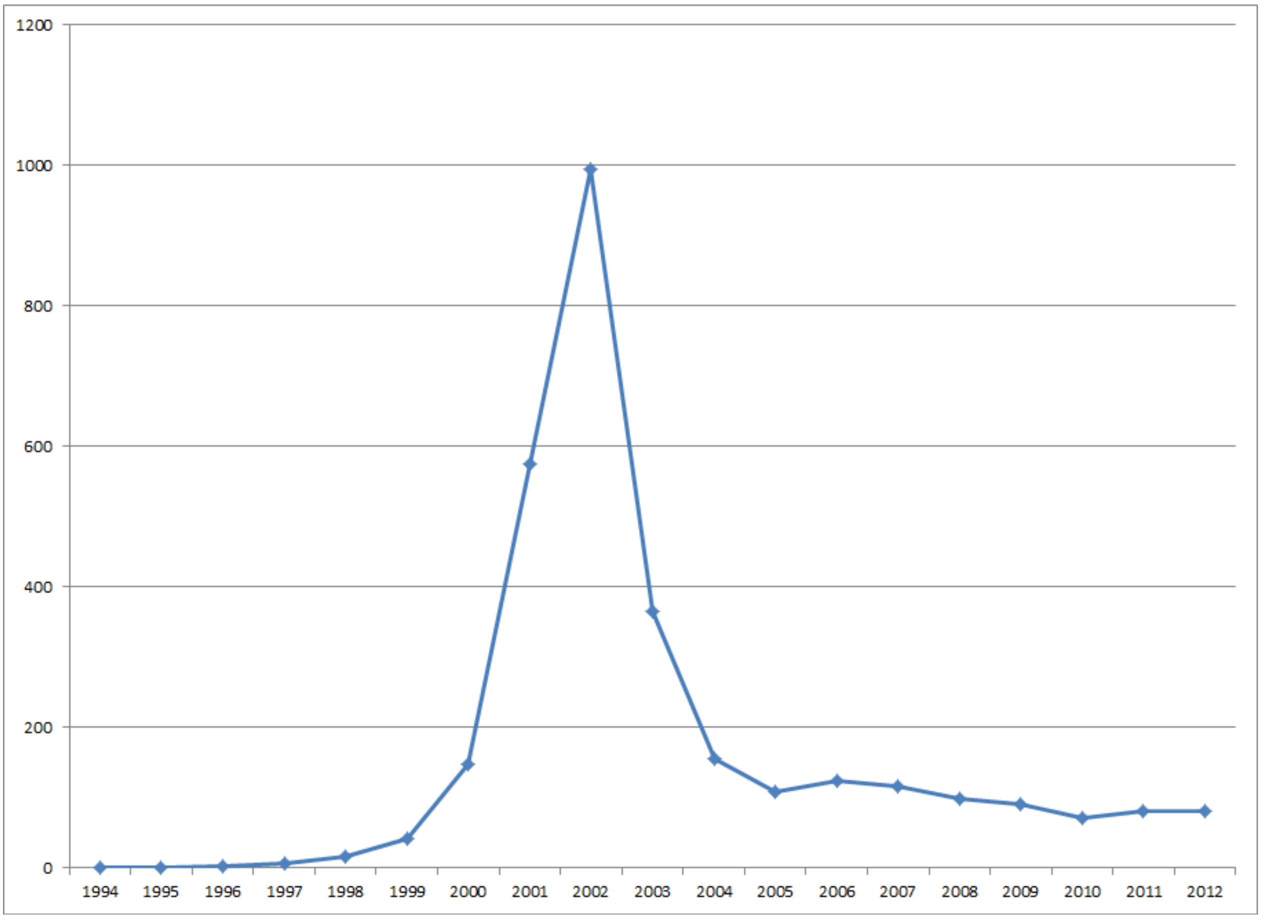
Based on the limited sample already analysed Table 4 gives a fine grained classification of the reasons for retracting the papers. ‘Fraud’ and ‘Fraud by one author’ together represent one third of the retractions. The category ‘Fraud’ is a mixed bag, only in 10% of the notices the word fraud is used, instead of a rather vague description as ‘No experimental data / Not reproducible / No notebook / Data not supporting findings/…’. Around 20% of the selected publications are retracted because of duplicate publishing of the same results or plagiarism. A small but not negligible number of papers is retracted because the researchers did not dispose of the appropriated authorisation to carry out the experiments. A small fraction of the retracted publications consists of classification errors by either the journal or the WoS; not the full paper but only a figure or one of the conclusions is retracted (partial retraction) or a critical comment is classified as a retraction notice.

Table 4. Reasons given in the retraction notice for retracting a paper

Reasons for retraction	**%**
Fraud	22,1
Errors	21,2
Fraud by 1 author	12,4
Duplicated / concurrent publishing	11,5
Plagiarizing	8,0
No motivation given for retraction	6,2
No approval by competent authority for experiments	5,3
Classification errors in journal or WoS	4,4
Independent review	4,4
Incomplete consultation between authors / listed a author without consent	2,7
Errors by editors	1,8

As already discussed in other studies often retracted papers continue to be cited after being retracted for a period longer than could be expected due to publication delays. This is illustrated in Figure 1 by the relative high number of citations to the papers retracted after 2002 in the ‘Schön case’. We started to analyse the distribution of the citations in the body of the citing papers and to make a sentiment analysis.

Figure 1: Number of citations to the retracted publications with J.H. Schön as co-author for the period 1994-2012. The retraction notices were published in 2002



Discussion

From these preliminary results the first conclusion is the lack of homogeneity of the material and the absence of guidelines for managing retractions.

Searching in the WoS the retracted articles turn out to be more complicated than discussed in the methodological section. An analysis learned that the search based on the suffix 'Retracted article' does not results in the complete set of retracted papers published in journals processed for the WoS. We identified retracted papers without this suffix but with 'Retraction' added as starting word in the title of the retraction notice. For example the pair: Sharma, 2011 (retracted paper) and Sharma, 2013 (retraction notice). Ongoing work has to identify the size of this set.

For a number of papers with the suffix 'Retracted article' the bibliographic data of the retraction notice were either erroneous or referred to a correction, a comment or a partial retraction. It seems the WoS has no classification of partial retractions.

Most journals do not have a separate section for retractions, but classify them under 'Corrections, Comments., Erratum, …'. Journals such as '*Journal of Biological Chemistry*' do not give the reason of the retraction.

The analysis of parties responsible for the retraction and of the reasons for the retractions shed some light on the management of research groups and dynamics of research collaboration. In a non-negligible fraction of retracted papers one of the authors refuses to co-sign the retraction notice and contest its validity. What information are collaborating researchers exchanging before submitting a manuscript as there are no data available to support a result or no proper archiving is done of results and methodological steps? In the ongoing work special attention will be paid to the group of retracted papers due to the lack of authorisation of the ethical committee to carry out the experiment as the patient health could be in jeopardy.

The ambiguous and evasive phrasing of some retraction notices or the outright lack of motivation for the retraction could be explained by the embarrassment of the authors or the editors. However it is not very helpful neither for the public standing of science or for the research community itself. Moreover, most journals have no separate section to publish retractions and/or do not (very visibly) mark in the electronic version that articles are retracted.

These factors could explain why authors continue to cite retracted papers. In the follow up work the sentiment analysis of these citations will shed some light on the 'citation culture' in different disciplines and journals and on the editorial process, in which both editors and reviewers clearly overlook previous retractions, causing retracted scientific output to still play a role in science.

In a next phase of our research, a bibliometric analysis will be made of the retracted papers for the different retraction types, to test for differences between disciplines and the influence of the journals' standing using the JFIS, which stands for Journal-to-Field Impact Score (van Leeuwen & Moed, 2002). The time to retraction, the time lag between retraction date and the publication date, will be analysed for each retraction type.

References

Boa Z., Batlogg B., Berg S., Dodabalapur A., Haddon R.C., Hwang H., Kloc C. Meng H. & Schön J.H. (2002). Retraction. *Science*, 298, 961.

Chen C., Hu Z., Milbank J. & Schultz T. (2013). A visual analytic study of retracted articles in scientific literature. JASIST, 64, 234-253.

Fang F.C., Steen R.G. & Casadell A. (2012). Misconduct accounts for a majority of retracted scientific publications. Proceedings of the National Academy of Sciences of the United States of America, 109, 17028-17033.

Grieneisen M.L. & Zhang M. (2012). A comprehensive survey of retracted articles from scholarly literature. PLOS ONE, 7, e44118.

Li X., He Y., Meyers A. & Grishman R. (2013).Towards fine-grained citation function classification. Proceedings of recent advances in natural language processing (pp. 402-407). Bulgaria.

Reich, E.S. (2009). *Plastic* Fantastic*: How the biggest fraud in physics shook the scientific world*. New York: Palgrave Macmillan.

Sharma S.K., Agawal S., Damodaran D., Sceenivas V., Kadhiravan T, Lakshmy R., Jagia P. & Kumar A. (2011). CPAP for the metabolic syndrome in patients with obstructive sleep apnea. New England Journal of Medicine, 365, 2277-2286.

Sharma S.K., Agawal S., Damodaran D., Sceenivas V., Kadhiravan T, Lakshmy R., Jagia P. & Kumar A. (2011).(2013). RETRACTION: CPAP for the metabolic syndrome in patients with obstructive sleep apnea. New England Journal of Medicine, 369, 1770.

Steen R.G., Casadell A. & Fang F.C. (2013). Why has the number of scientific retractions increased? PLOS ONE, 8, e68397.

Van Leeuwen, T.N., and Moed, H.F. (2002) Development and application of journal impact measures in the Dutch science system, Scientometrics, 53, 249-266

Understanding factors influencing participation in European programs of Higher Education Institutions

Benedetto Lepori*, Barbara Heller-Schuh**, Thomas, Scherngell**, Michael Barber**

* *blepori@usi.ch*
Centre for Organisational Research (CORe), Università della Svizzera Italiana, Via Lambertenghi, 10, Lugano, (Switzerland)

** *barbara.heller-schuh@ait.ac.at; thomas.scherngell@ait.ac.at, Michael.barber@ait.ac.at*
AIT Austrian Institute of Technology GmbH, Donau-City-Strasse 1, Vienna, 1220, Austria

Abstract
This paper aims to investigate determinants of participation of European Higher Education Institutions (HEIs) in European Framework Programs (EU-FPs). It is based on a comprehensive sample of 1376 HEIs in 28 European countries, which constitutes the full population of HEIs, which could potentially participate in the EU-FPs, since they are active in research. Data are derived from the EUMIDA dataset on Higher Education Institutions, as well as from the EUPRO database on participation of the EU-FPs.
The analysis confirms and extends previous findings of a very strong concentration of participations to EU-FPs in a very small numbers of European research universities; furthermore, they demonstrate that number of participations per HEIs can be predicted with high precision from organizational characteristics and, especially, size and international reputation. We show that the chance of participating in EU-FPs programs is quite limited below an academic staff size of around 500 FTEs. Further, the number of participations tends to grow proportionally to organizational size, but it is strongly influenced by international reputation, and to a lesser extent, by research intensity and specialization in sciences and engineering. Finally, our results imply that countries factors have become less important in determining participation to EU-FPs.

Keywords: higher education institutions, European framework programs, organizational size, international reputation

Introduction
This paper aims to investigate determinants of participation of European Higher Education Institutions (HEIs) in European Framework Programs (EU-FPs). It is based on a comprehensive sample of 1376 HEIs in 28 European countries, which constitutes the full population of HEIs, which could potentially participate in the EU-FPs, since they are active in research.
By matching the EUPRO database – which provides data on participation in the EU-FPs – with the EUMIDA database, and gathering data on organizational size, international reputation, research intensity, and subject specialization, we are able to test the impact of HEI characteristics suggested by the literature concerning their number of participations and to investigate the relative importance of country and organizational effects.
Our results confirm and extend those of previous studies on smaller samples (Geuna 1998a; Henriques, Schoen and Pontikakis 2009). First, participation in the EU-FPs remains highly concentrated in the most reputed and research-intensive universities, with about 70 HEIs in

the whole sample accounting for half of the total number of participations in 2008. Second, despite the emergence of second-tier HEIs with a research mission, doctorate-awarding universities account for almost all EU-FPs participations, despite comprising only about 60% of our sample. Third, characteristics of individual HEIs – in particular their size and international reputation – predict with high precision whether they participate and the number of participations in the EU-FPs; country effects are comparatively less important than HEI characteristics in driving their participation.
These results confirm the strong cumulative effects of research funding hinted at by the literature (Geuna 2001), as well as that EU-FPs is successful in targeting the most reputed European universities; they also suggest that access opportunities to EU-FPs are now reasonably equal across European countries, and by large, only depend on the quality of the concerned HEI.

Background

The funding and structure of the European higher education system is currently undergoing a profound modification. In recent years most OECD countries have extended their competitive research funding, while at the same time proportionately reducing institutional core funding (Geuna 2001, Vincent-Lancrin 2006). Higher education institutions (HEI) increasingly rely on external funding sources and the availability of supra-national funding instruments plays an extremely important role. The EU Framework Programme (EU-FPs) for Research and Technological Development represents one of several alternative funding opportunities that universities depend on more and more, since in virtually all European countries, total government university funds (mainly general university funds) have decreased. The increasing relevance of this funding source – the available research funding has grown from 3.75 billion Euro/ECU in FP1 to nearly 80 billion Euro in Horizon 2020 (Roediger-Schluga and Barber 2008) – makes it crucial to understand the patterns and determinants for HEI participations in the EU-FPs. Thus, investigating HEI participation patterns in the FP and identifying determinants affecting HEI engagement in the EU-FPs, has become an important question in a European policy context, in particular with respect to the goals of supporting excellence and integration in the European Research Area (ERA). Some empirical studies in this direction have been conducted in the 1990s (Geuna 1998a; Geuna 1996); since then, only scarce empirical evidence is provided in the literature (see Nokkala, Heller-Schuh and Paier 2011).
The EU-FPs has been conceived as one of the main instruments of European research policy to foster economic competitiveness and to stimulate knowledge diffusion across European countries. By funding basic and applied transnational collaborative research with industrial and societal relevance it facilitates research and networking between various knowledge producers, including higher education institutions, research organizations, and industry and governmental institutions. Since their inception in 1984, seven EU-FPss have been launched. Horizon 2020, the current Framework Programme for Research and Technological Development, is built upon the results and successes of former EU-FPss, the Competitiveness and Innovation Framework Programme (CIP), and the European Institute of Innovation and Technology (EIT).
HEIs have long been recognized as the principal beneficiaries of EU-FPs funding and major players in the development of the ERA. HEIs represent about 25% of all participants since FP1 and constitute about 40% of the top 100 organizations with the highest number of participations and the highest number of partners in the EU-FPs (Heller-Schuh, Barber, Henriques, et al 2011). Henriques, Schoen and Pontikakis 2009 investigated the role and place

of 171 top European research universities[1] in EU-FPs-funded research and address concerns expressed in several evaluation reports, particularly that top research universities are reluctant to participate in the EU-FPs due to 'cumbersome' administrative procedures, 'low content of basic research,' and availability of other, 'more attractive' sources of (national) funding. The study demonstrates that these top research universities represent the core of higher education participation in FP6, measured by the number of participations as well as by the amount of funding awarded. In FP6 HEI received – compared to research and business organizations – the highest percentage of funding (37%) and also held the highest percentage of project participation (36%) (Henriques, Schoen and Pontikakis 2009).

The FP7 interim evaluation report confirms that FP7 attracts the top EU researchers from universities and research organizations (Annerberg, Begg, Acherson, et al 2010). The expert group states that the list of organizations that have obtained the largest amounts of funding can be read as the Who's Who of European research quality. The distribution of funding in the FPs is rather skewed and concentrated on a relatively small number of organizations. The top 50 funding recipients, mainly universities (28 organizations) and RTOs, acquired a quarter of FP funding, while three quarters of the funding is spread among the 14,000 other recipients.

Research questions

In this paper we empirically address, on a very large sample of European HEIs, the following questions.

How are participations in EU-FPss distributed? Is there strong concentration in a small number of HEIs? To which extent has the expansion of the higher education system led to a broadening of the circle of HEI participations?

While previous studies agree that there is a strong concentration of participations in a rather small group of HEIs, the extension of our sample allows us to provide a more comprehensive analysis and to examine the distributions of participation also beyond the core of the most reputed universities. Recent structural changes in European higher education systems, with the extension of the research mission beyond the traditional doctorate-awarding universities (Kyvik and Lepori 2010, might also suggest that these organizations are becoming a relevant actor in the EU-FPs.

To which extent are participations in the EU-FPs predicted by organizational factors like size, reputation, research intensity, and subject specialization? Are there scale effects on EU-FPs participations?

The available literature agrees that international reputation of HEIs is a chief determinant of participation in the EU-FPs. We also expect that participations increase with research orientation, as well as with the specialization in science and technology, given the focus of the EU-FPs on these areas. The relative importance of these effects has hardly been investigated. Finally, whether there are scaling effects in participations in the EU-FPs, i.e. the number increases more rapidly than size, as suggested by some economic literature, is still unclear.

Are there differences in levels of participation by country (once we take into account the importance of HEI-level factors)?

Previous studies displayed that HEIs in less developed countries in the EU, like Ireland and Greece, participated more in EU-FPs, following the convergence objective of European research policy (Geuna 1998b). However, since the year 2000, the context of European research policy deeply changed with the establishment of the European Research Area, a

[1] defined as those universities having published more than 5,000 publications in peer reviewed journals in the period 2000 to 2006

stronger focus on competitiveness and research excellence, and finally, with the enlargement of the EU to Central and Eastern Europe.

Data and methods

Sample and sources

Our sample is composed by the so-called research-active HEIs included in the European MIcroDAta (EUMIDA) dataset (Lepori and Bonaccorsi 2013). EUMIDA includes 2,457 HEIs in 28 European HE systems (European Union members plus Norway and Switzerland; France is not included since it did not deliver any data to EUMIDA), comprising more than 90% of tertiary education students in the considered countries. Research-active HEIs have been identified by EUMIDA as those HEIs that have an institutionalized research activity, meaning they have an official research mandate and identifiable research groups. This definition turns out to be quite extensive: among the 1,378 research-active HEIs in EUMIDA, there are 850 doctorate-awarding HEIs, but most are sizeable second-tier HEIs in countries with binary systems. We can thus consider that this sample includes almost all HEIs that could potentially participate in the EU-FPs. EU-FPs.

The data mostly refers to the year 2008, and includes: identifying information (name, category, foundation year, highest degree granted), expenditure and income, academic and non-academic staff, bachelor and doctoral students per total number and by field, patents and spin-off, and degrees awarded by national/international origin and by field.

For the purpose of this study, the list of HEIs provided by EUMIDA was matched with data on number of European participations from the EUPRO database for the years 2008, 2009, and 2010. The EUPRO database currently comprises information on more than 60,000 research projects funded by EU-FPss (complete for FP1-FP6) and all participating organizations. It contains systematic information on the participating organizations including the full name, the full address, the type of the organization, and when appropriate and possible, the organizational sub-entity involved in the project. EUPRO raw data is based on publicly accessible CORDIS data, with substantial effort to significantly improve quality and the level of standardisation of the data — e.g. correction of heterogeneous spellings of organisation names, different languages, inconsistency of organization types and organisational levels — and to retrieve and add missing data (for a full description of data processing see Roediger-Schluga and Barber 2008).

In this way, 905 HEIs could be identified, of which 717 had at least one EU-FPs project in 2008 (the remaining were included in EUPRO since they had projects in previous years). We set the number of participations of the remaining HEIs in EUMIDA to 0. There is the possibility that in a few cases these HEIs have not been identified in EUPRO, but we consider that this phenomenon is marginal.

We finally drop from the sample two Italian graduate schools which only offer postgraduate education, since they have a completely different structures than other HEIs. Our final sample is thus composed of 1,376 HEIs in 28 countries (EU-27 countries less France plus Norway and Switzerland). The countries with the largest number of observations are Germany (306), UK (148), Poland (91), and Italy (79). The effective sample for the regressions is slightly smaller since data on size and disciplinary composition are not available for all HEIs in the sample. A few countries – Belgium, Cyprus, Czech Republic, Greece, and Portugal – are not included in the regressions since the corresponding data on HEIs were missing in the EUMIDA database.

Variables

Dependent variable

In all analyses, the dependent variable is the *count of participations in EU Framework Programs in the year 2008*, derived from the EUPRO database.

Multiple participations by the same university are considered, if different entities (institutes, departments) of one university take part in the same project and are indicated as a distinct partner in the CORDIS project database (repeated participations account for about 1% of university participations). The projects in the EU FP last between ten months up to nine years. We include all projects with university participation that were ongoing in the year 2008.

We remark that the number of participations are quite stable across years, as counts for the years 2009 and 2010 are correlated to .99*** to those for 2008.

Independent variables

We use following independent variables.

Size is the number of academic staff (Full Time Equivalents).

Reputation. We use the product between normalized impact factor and total number of publications of the concerned HEIs ("brute force" indicator; van Raan 2007), normalized with the number of academic staff. This indicator builds on the insight that the international visibility of an HEI is related both to quality and the number of publications. Data is derived from the SCIMAGO institutional rankings for the year 2011 (http://www.scimagoir.com), which is based on data from the period 2005-2009. We hold data for 482 HEIs in the dataset – the other HEIs are not covered since they had less than 100 publications in Scopus in the reference period. Despite normalization by size, this indicator remains correlated with output (as a result of scaling properties of research output; van Raan 2007); accordingly, this indicator converges to 0 when the number of publications decreases and thus it can be set to 0 for the remaining HEIs in the sample.

PhD awarding is a dummy variable with a value 1 if the concerned HEI has the legal right to award a doctorate and 0 if not. This distinction is highly relevant in distinguishing traditional universities from new second-tier HEIs (Kyvik and Lepori 2010).

Research intensity. The share of PhD students over undergraduate students is a widely used indicator of research intensity (Bonaccorsi, Daraio, Lepori and Slipersaeter 2007).

Disciplinary characteristics. We introduce two variables to characterize HEIs with a stronger orientation towards humanities on the one hand, to natural sciences and technology on the other hand. Since we do not have data on the breakdown of staff by scientific fields, we resort to data on the distribution of undergraduate students by eight fields of education (FOE-1997 classification; UOE 2006). Accordingly, we construct two dummy variables for HEIs which have more than half of their undergraduate students in arts and humanities, respectively more than half of the students in natural and technical sciences.

Methods

Descriptive analysis. We first perform detailed descriptive analysis of data on participation in European programs, as well as on relationships with independent variables. ANOVA is used to provide preliminary evidence of the importance of country factors.

Impact of HEI characteristics. The choice of the regression method must take into account that our dependent variable is highly skewed and is null for half of the ~~whole~~ sample. An approach would be to use a negative binomial regression with a hurdle model, which specifies a logistic regression model in order to predict whether the case is null and passes the case to a negative binomial model to predict a non-zero count; this would be the approach for highly skewed count data (Cameron and Trivedi 1998).

However, in the case at hand, once the null values are excluded, the data is not extremely skewed. Therefore, we prefer running two distinct models as follows (the same approach was adopted by Geuna 1998a): the first model is a binary logistic regression run on a dichotomized dependent variable (no projects = 0; at least one project = 1) and thus tests whether a HEI participates in the EU-FPs.
The second model computes the predicted number of participation through a truncated linear regression applied on the non-zero cases; as a dependent variable, we employ ln(participations) in order to reduce the skewedness of the dependent variable. Truncated regression models take into account in the estimators the fact that values below some threshold (0 in our case) have been removed from the sample (Long 1997).
We run two types of models, one only introducing HEI-level variables and one including country dummies in order to test whether country effects are relevant and significantly account for differences in participations.
In order to avoid problems with samples that are too small, we introduce country dummies only when the number of cases in a country exceeds 10. All other countries are grouped together.

Results

Descriptive analysis
Table 6 provides full descriptive information concerning the variables considered in the empirical analysis.

Table 6. Descriptive statistics

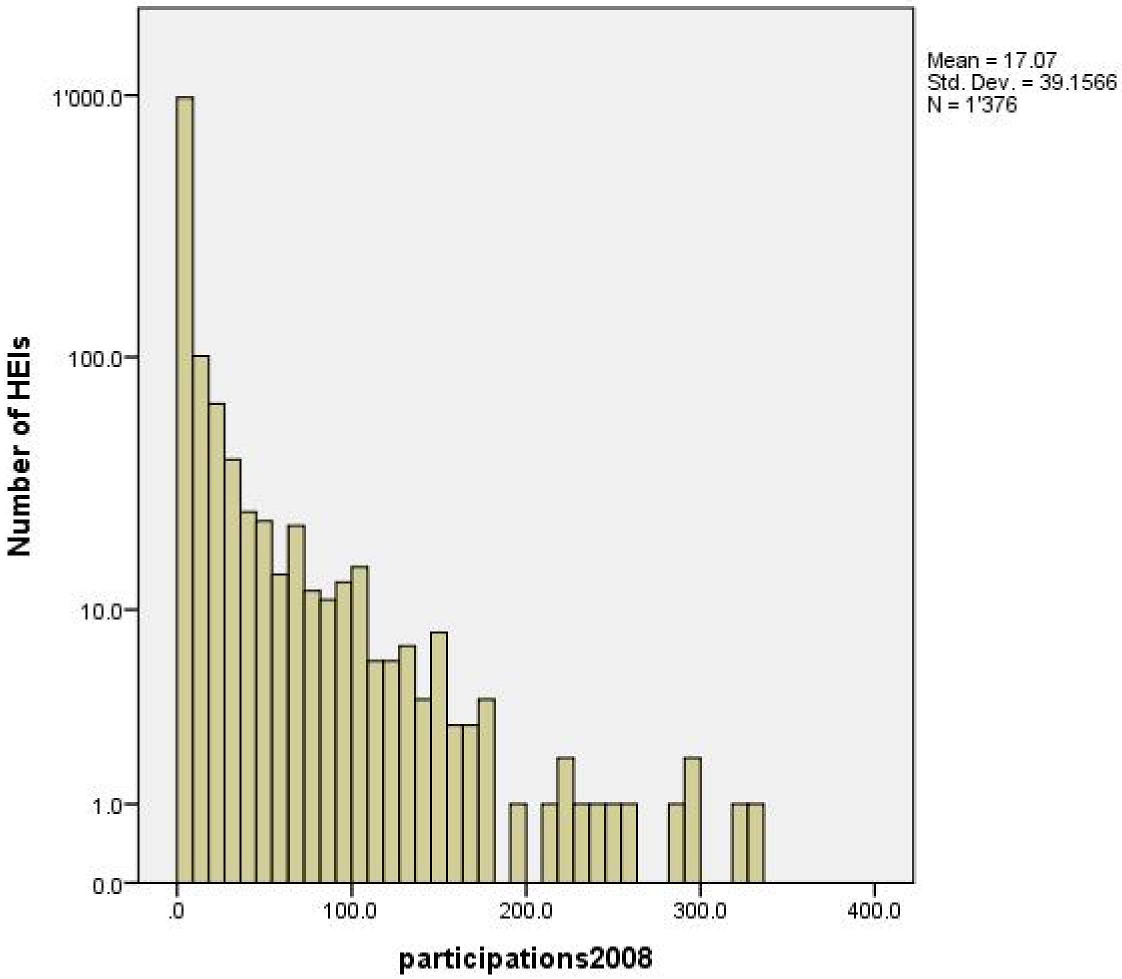
	Valid N	Mean	STDEV	Min	1Q	Median	3Q	Max
participations	1376	17.1	39.2	0.0	0.0	1.0	14.0	329.0
AcademicStaff	1288	685.7	882.5	0.0	111.0	324.0	906.6	6571.0
Reputation	1376	1.5	3.2	0.0	0.0	0.0	1.7	40.6
research_intensity	1376	0.0	0.1	0.0	0.0	0.0	0.0	1.7
		All cases			Non-zero cases			
		1.00	.00		1.00	.00		
engineering_binary	1277	150	1127		99	571		
arts_binary	1279	170	1109		14	658		
phd awarding	1376	856	520		613	102		

All variables are highly skewed, the distribution of academic staff is nearly lognormal, as would be foreseen by Gibrat's law (Ijiri and Simon 1964), whereas reputation and research intensity are even more skewed and have a large number of nulls. The two dummy variables for subject specialization clearly identify small groups of very specialized HEIs (about 10% of the whole sample for each dummy). As expected, most HEIs specialized in natural and technical sciences participate in the EU-FPs, whereas most of those specialized in arts and humanities do not.
The correlation between academic staff and reputation is fairly large (.551***), despite the fact that the latter is normalized by size, owing to the scaling effects of scientific reputation. The correlation with research intensity is fairly small (.158** for academic staff and .167** for reputation). As for the dummy variables, the phd awarding variable is strongly correlated with research intensity (.785**) and to a lesser extent to academic staff (.572**) and reputation (.520**). Correlations between the three dummy variables are quite small.

Finally, there are strong correlations between the number of participations and independent variables: the largest scores are for academic staff (.787**) and reputation (.775**), thus suggesting that these HEI characteristics have a deep impact on participations in the EU-FPs.

Participations to framework programs
As shown by Figure 8, the distribution of participation counts in our sample is highly skewed. About 50% of the HEIs in the sample have no participations, while only 397 HEIs have more than 10 participations and only 72 more than 100 participations.

Figure 8. Participation of EU-FPs, year 2008



Even among the participating HEIs, the concentration is rather high: the maximum count of participations is 329 projects, while 152 HEIs with more than 50 participations account for 70% of total projects.
The top ten list of EU participations reflects very well the top universities in Europe, in terms of university rankings, but also includes a large representation of the technical schools (Table 7)

Table 7. Characteristics of the HEIs with the top-10 counts of participations to EU-FPs

University	**N**	**Country**
The University of Cambridge	329	UK
Technical University of Denmark	319	DK
Imperial College of Science, Technology and Medicine	297	UK
Katholieke Universiteit Leuven	293	BE
The University of Oxford	287	UK
University College London	259	UK
Federal Institute of Technology Zurich	247	CH
Lund University	243	SE
National Technical University of Athens	228	GR
University of Copenhagen	223	DK

The distinction between participating and non-participating HEIs largely matches the one between HEI awarding or not the PhD: 71% of the doctorate-awarding HEIs participated in the FP in 2008, whereas this share was only 20% for the non-doctorate awarding. In fact, non-doctorate-awarding HEIs account for only about 1% of the participations (while constituting 40% of our sample). No non-doctorate awarding HEI had more than 15 participations in EU-FPss in 2008, and just a handful, mostly Swiss and German Universities of Applied Sciences Fachhochschulen, had more than 5 participations.

As expected, there are large differences between countries both in total number of participations and in participations per HEI. The largest number of participations are in the UK (5,107 participations) and Germany (3,529 participations), while the average number of participations per HEI varies between 64 for Denmark and 2.7 for Romania. These averages might however be influenced by more or less extensive definitions of the EUMIDA perimeter by country, and do not take into account the fact that HEI-level characteristics (like size and reputation) systematically vary by country.

As a matter of fact, an ANOVA variance test shows that while the variance between countries is statistically significant, it accounts for only 6.5% of the total variance. This provides preliminary evidence that HEI-level characteristics are highly important, and even within countries with large levels of participations, there are large differences between individual HEIs.

Determinants of the probability to participate to European programs

Table 8 presents the results of a binary logistics regression to predict whether HEIs in the sample had at least one participation in the EU-FPs in the year 2008. The model including the HEI variables performs quite well as it classifies correctly 83.2% of the cases (against 51% of the null model). The model including country dummies improves performance only slightly (to 85.3%) and few of the country dummies are statistically significant, confirming the results of the ANOVA that most of the variation takes place at the HEI level (at least concerning cases where there is at least one participation).

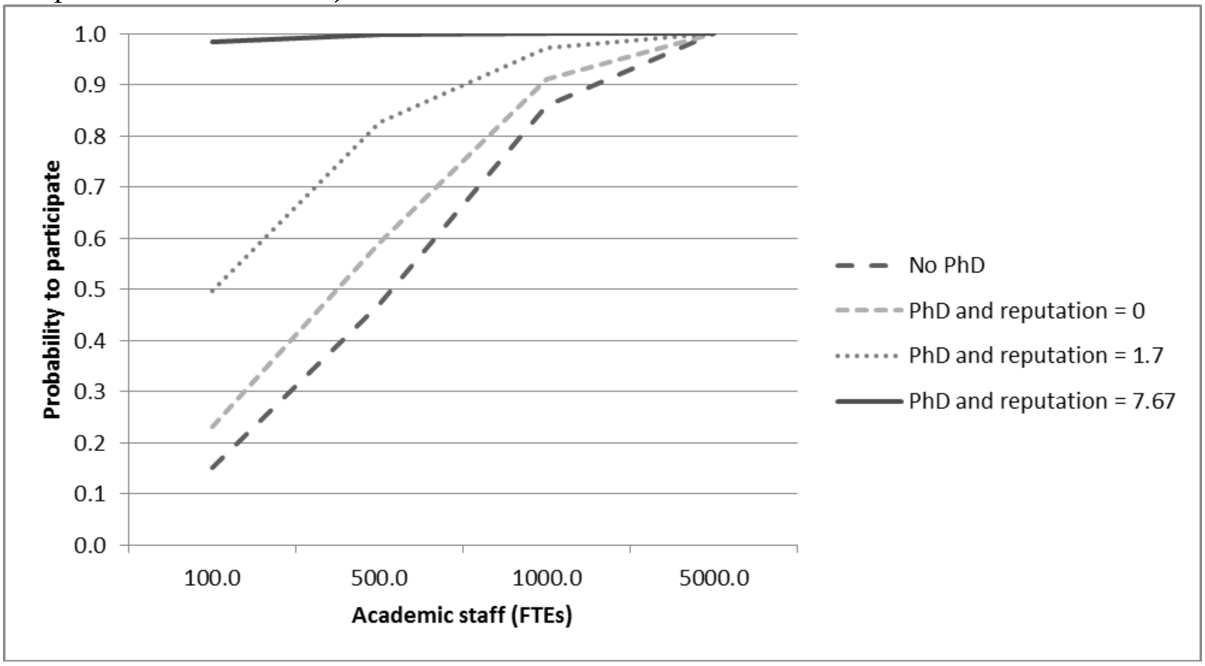
The performance of the model is particularly remarkable for two reasons: first, our sample is equally divided between participating and non-participating HEIs, unlike the one used by Geuna (1998). Second, the distinction between participation vs. non participation for a single year is rather weak, as it might happen that one HEI gets a project for reasons which cannot be controlled by our independent variables.

Table 8. Logistic regression for predicting participations
Dependent variable: 1 = at least one participation in 2008, 0 = no participations in 2008. Country dummy = 00 for Cyprus, Estonia, Luxembourg, Malta and Slovenia. Reference category: UK. Belgium, Czech republic, Greece and Portugal are missing because some independent variables are missing.

	Staff model			Full model			Country model		
				B	S.E.	Sig.	B	S.E.	Sig.
Constant	-1.854	.117	.000	-1.982	.161	0.1377	-2.810	.418	0
AcademicStaff (thousands FTEs)	4.464	.285	.000	2.536	.339	.000	3.930	.454	.000
phd				.842	.197	.000	1.202	.291	.000
Reputation				.788	.171	.000	.705	.183	.000
research_intensity				4.144	1.768	.019	3.315	1.973	.093
natural_engineering				.698	.255	.006	.576	.274	.036
arts				-1.475	.317	.000	-1.417	.342	.000
									.000
00							1.554	.773	.045
AT							1.012	.500	.043
BG							.764	.522	.144
CH							.093	.711	.896
DE							.857	.408	.036
DK							1.385	1.027	.178
ES							-.212	1.150	.854
FI							-.476	.605	.432
HU							-.201	.545	.712
IE							-.819	.726	.260
IT							-.884	.566	.118
LT							.840	.732	.251
LV							.532	.614	.387
NL							-2.042	.782	.009
NO							.357	.546	.513
PL							-1.444	.445	.001
RO							-.383	.510	.453
SE							2.360	.583	.000
SK							.140	.561	.803
-2loglikelihood	1102.44			864.159			781.744		
n	1288			1216			1216		
Correctly classified	81.20%			83.20%			85.30%		

The figure below summarizes the combined impact of size and reputation to the probability of participating in EU-FPs.

Figure 9. Probability to participate to European programs
Reputation level 1.7 corresponds to 75% percentile, 7.97 to 95% percentile (only 5% of the sample exceeds this value).



Accordingly, second-tier HEIs, which don't award the doctorate and don't have international reputation, have a chance to participate in EU-FPss only if they are reasonably large, beyond 500 FTE of academic staff. For universities, this threshold tends to be become lower and highly reputed schools have a high chance to participate even if their size is extremely small. With the exception of very highly reputed HEIs, these results point to the fact that there is a clear size threshold to participate in EU-FPss, around 500 FTEs of academic staff.
Finally, we notice that most country dummies are not statistically significant; the few significant cases rather point to the fact that HEIs in more developed countries have higher chances to participate in EU-FPss. Given the rather small number of cases in each country, this result should however not be taken too seriously.

Number of participations
Table 9 displays the results of a truncated linear regression using as a dependent variable the ln of the number of participations (limited to the HEIs with at least one project in 2008). We also applied a log transformation to the number of academic staff and the reputation (adding 1), while the other variables are not transformed since they are bounded between 0 and 1. We dropped the arts variable since a very limited number of HEIs specialized in the arts are included in the subsample.

Table 9. Truncated linear regression for predicting counts
Standardized coefficients have been calculated for a change in the independent variable equal to one standard deviation for continuous variables and for a change of 1 in the dummies. Reference country category: UK. Country dummy = 00 for Cyprus, Denmark, Latvia, Estonia, Luxembourg, Malta and Slovenia.

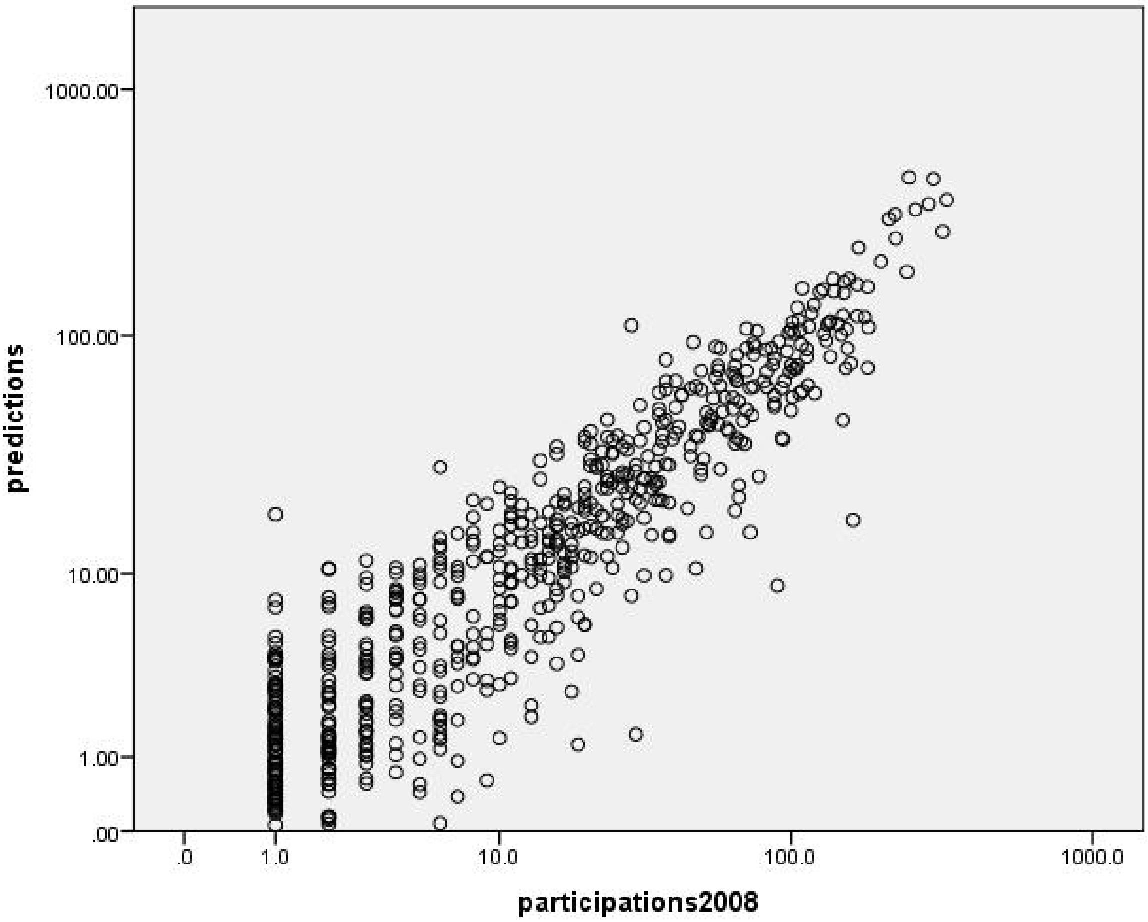
	Staff model			Full model			Country model			
	Coef.	Std. Err.	P>|z|	Coef.	Std. Err.	P>|z|	Coef.	Std. Err.	P>|z|	St. Coefficients
_cons	-8.980	.476	.000	-4.955	.358	.000	-5.546	.349	.000	
ln_academic_staff	1.680	.067	.000	.858	.053	.000	.969	.052	.000	1.008
phd				.508	.173	.003	.661	.167	.000	.661
ln-reputation				.929	.053	.000	.781	.056	.000	.703
research_intensity				2.239	.500	.000	1.910	.493	.000	.229
natural_engineering				.262	.103	.011	.368	.099	.000	.368
0							.607	.175	.001	.607
AT							.138	.187	.459	.138
BG							-.109	.239	.649	-.109
CH							.090	.215	.675	.090
DE							-.097	.113	.392	-.097
ES							-.472	.136	.001	-.472
FI							-.189	.225	.403	-.189
HU							.396	.216	.067	.396
IE							.445	.243	.067	.445
IT							-.378	.130	.003	-.378
LV							.006	.275	.982	.006
NL							.120	.207	.564	.120
NO							-.387	.260	.136	-.387
PL							-.692	.154	.000	-.692
RO							-.895	.206	.000	-.895
SE							.513	.171	.003	.513
SK							.278	.249	.264	.278
loglikelihood	-861.463			-639.026			-590.079			
N	668.000			629.000			629.000			
Wald chi2	623.610			1285.020			1647.490			
Prob > chi2	.000			.000			.000			

The model with all HEI-level variables provides a significantly better fit than the one including only size, and the country dummies model is superior to both; however, coefficients of HEI-level variables are not significantly affected by the introduction of country dummies, thus supporting the robustness of this effect.

When comparing standardized coefficients, it turns out that size and reputation are by far the most important factors and display a similar strength. The other HEI variables are less relevant. The coefficient of size is not statistically different from 1, implying that the number of participations is proportional to academic staff and thus that there are no direct scale effects. We however remark that there are scale effects in reputation, which remains correlated with size despite being normalized; accordingly, large HEIs tend to have more participations in EU-FPs programs in respect to staff as an outcome of their larger reputation.

Few country dummies are statistically significant, showing that country effects have become relatively less important than HEI characteristics in driving participations, and point consistently to the fact that HEIs in less developed and reputed countries tend to have less participations in EU-FPs programs – this applies clearly to Poland and Romania, but also to some extent to Italy and Spain.

Figure 10. Predicted and observed counts of EU-FPs participation
Only cases with at least one participation in 2008, N=629



Importantly, the model presents an excellent fit in terms of predicting correctly the observed counts of participations. The correlation between observed and predicted values is as high as .885*** on the log-log scale and .912*** on the original scale. The quality of prediction remains quite good on the whole range from counts below ten to the highest counts observed and thus predicts very accurately both cases with a small number of participations and high values (Figure 10).

Discussion and conclusion

In terms of our research questions, we can summarize our results as follows.
a) First, there is indeed a strong concentration of EU-FPs participation in a small group of relatively large and highly reputed international HEIs. About 150 HEIs in the EU countries (plus Switzerland and Norway) accounted in the year 2008 for 70% of the total participation in European projects. This compares with our sample of more than 1,300 HEIs performing some research, which would be potential participants to EU-FPs.
Accordingly, while the European higher education system has been characterized by strong expansion and differentiation of types of HEIs, there are strong cumulative effects at work, which concentrate participations in EU-FPss to a small number of core HEIs.

b) Second, HEI characteristics have a strong impact both on whether an HEI participates in the EU-FPs and to the observed number of participations. Except for very highly reputed HEIs with strong research orientation, we could show that the chance of participating in EU-FPs programs is quite limited below an academic staff size of around 500 FTEs. Below this size HEIs lack the critical mass in terms of research to be integrated into European networks.
Further, we showed that the number of participations tends to grow proportionally to organizational size, but it is strongly influenced by international reputation, and to a lesser extent, by research intensity and specialization in sciences and engineering. The excellent fit of the model implies that other factors, like the quality of support services or an explicit strategic orientation towards European funding, do not play a major role anymore. We interpret this as a sign that awareness and support of participation in EU-FPss has spread to a larger extent across the whole ERA and thus there are no longer any significant differences in this respect between HEIs.
c) Third, country effects are much less important on the aggregate than HEI effect in determining participations, with the possible exception of a few individual countries like Poland and Romania The remaining country effects also point to a reversal of the convergence policy observed in the 90s, where less developed European countries had more participations than expected in the EU-FPs; on the contrary, there are signs that less developed countries tend to have less participations, even if this is not a systematic effect across all countries. On the one hand, the move towards ERA has promoted greater integration of all EU countries in EU-FPss, diminishing the importance of country effects; on the other hand, the stronger emphasis on research excellence seems to have replaced the previous cohesion goals in European research funding policies.

References

Annerberg, R., Begg, I., Acherson, H., Borrás, S., Hallen, A., Maimets, T., Mustonen, R., Raffler, H., Swings, J. & Ylihonko, K. (2010). *Interim Evaluation of the Seventh Framework Programme: Report of the Expert Group* European Commission.

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.

Cameron, A. C. & Trivedi, P. K. (1998). *Regression analysis of count data.*Econometric Society Monograph, Cambridge University Press.

Geuna, A. (1998a). Determinants of university participation in EU-funded R & D cooperative projects. *Research Policy, 26*, 677-687.

Geuna, A. (1998b). Determinants of university participation in EU-funded R&D cooperative projects. *Research Policy, 26(6)*, 677-687.

Geuna, A. (1996). The participation of higher education institutions in European Union Framework Programmes. *Science and Public Policy, 23(5)*, 287-296.

Geuna, A. (2001). The changing rationale for European university research funding: are there negative unintended consequences? *Journal of Economic Issues, XXXV(3)*, 607-632.

Heller-Schuh, B., Barber, M., Henriques, L., Paier, M., Pontikakis, D., Scherngell, T., Veltri, G. A. & Weber, M. (2011). *Analysis of Networks* in *European Framework Programmes (1984-2006)* Luxembourg: Publications Office of the European Union.

Henriques, L., Schoen, A. & Pontikakis, D. (2009). Europe's top research universities in FP6: scope and drivers of participation. *JRC Technical Notes, 53681*.

Ijiri, Y. & Simon, H. A. (1964). Business firm growth and size. *The American Economic Review, (54(2))*, 77-89.

Kyvik, S. & Lepori (2010). The Research Mission of Higher Education Institutions Outside the University Sector. Striving for Differentiation.

Lepori, B. & Bonaccorsi, A. (2013). Towards an European census of higher education institutions. Design, methodological and comparability issues. *Minerva, 51(3)*, 271-293.

Long, J. S. (1997). *Regression Models for Categorical and Limited Dependent Variables* Thousand Oaks: Sage.

Nokkala, T., Heller-Schuh, B. & Paier, M. (2011). Ranking lists and european framework programmes. In Anonymous *Public Vices, Private Virtues?* (pp. 111-139). Springer.

Roediger-Schluga, T. & Barber, M. J. (2008). R&D collaboration networks in the European Framework Programmes: Data processing, network construction and selected results. *International Journal of Foresight and Innovation Policy, 4(3)*, 321-347.

UOE (2006). *UOE data collection on education* systems. *Manual. Concepts, definitions, classifications* Montreal, Paris, Luxembourg: UNESCO, OECD, Eurostat.

van Raan, A. F. J. (2007). Bibliometric statistical properties of the 100 largest European universities: prevalent scaling rules in the science system. *Eprint arXiv:0704.0889*, .

Vincent-Lancrin, S. (2006). What is changing in academic research? Trends and futures scenarios. *European Journal of Education, 41(2)*, 169-202.

Investigating the Impact of the Award of the Nobel Prize on Tweets

Jonathan Levitt* and Mike Thelwall*

J.M.Levitt@wlv.ac.uk, m.thelwall@wlv.ac.uk
Statistical Cybermetrics Research Group, University of Wolverhampton, Wulfruna Street, Wolverhampton, WV1 1LY (UK)

Introduction

Social media provide the potential for a new understanding of the social impact of research and how researchers engage with the public. This study is a pilot investigation of how the response of social media to major research events can contribute to understanding of these topics. The award of the Nobel Prize is a particularly suitable research event to investigate, as it is a particularly high profile event and there are several prizes announced in a short space of time, allowing comparisons between similar events. Surprisingly, there seem to be only a few studies about public reactions to Nobel Prizes (e.g., James, 1966; Tötösy de Zepetnek, 2005), perhaps because it has been difficult to get direct evidence of public interest, and so press coverage has been the main source of data.

This pilot study examines the focus and content of tweets posted after the award of the 2013 Nobel Prize that contain the word 'Nobel' (termed 'Nobel tweets'). Twitter was chosen, as it is particularly suited for large-scale investigation: the Twitter API (Applications Programming Interface) enables the automatic downloading of all Nobel tweets.

Related research

Most academic Twitter studies have related to the evaluation of research (e.g., Piwowar, 2013). Significant associations have been found between higher citation rates and higher scores for Twitter and other social media, at least for articles published at a similar point in time (Thelwall et al., 2013). Previous studies had also found a connection between the number of tweets relating to research articles on Twitter and the number of citations those articles later receive for one medical informatics journal (Eysenbach, 2011) and for arXiv.org preprints (Shuai, Pepe & Bollen, 2012).

Some studies have also investigated how the public react to science-related issues on Twitter. One used tweets to analyse the dynamics of climate change protests (Segerberg & Bennett, 2011) and another identified the most tweeted species to gauge levels of public interest (Roberge, 2014).

Despite the above findings, no investigation has investigated how tweets have responded to major research events, other than protests. This pilot investigation studies responses in Twitter to Nobel Prize awards. It addresses two questions:

1. Does the nature of Nobel tweets depend on the subject area of the prize?
2. In general, what is the focus of Nobel tweets?

Methods and data

The main data for this study is an analysis of 716,000 tweets containing the word or hashtag Nobel posted on Twitter between 1pm October 9th and 2pm October 22nd, 2013. The data

does not start with the first Nobel Prize because, for personal reasons, the author's interest started with the Chemistry prize. For future full-scale studies, the data collection period will cover all of the prizes.
The tweets were automatically downloaded using the Twitter API via the free Twitter time series analysis software Mozdeh.

The Nobel Prizes were announced in October, medicine on the 7th, physics on the 8th, chemistry on the 9th, literature on the 10th, peace on the 11th, and economics on the 14th. Figure 1 from the website Topsy.com reveals considerable public or academic interest in the topic and Figure 2 gives an hourly breakdown of the tweets collected by Mozdeh and analysed in this paper, showing that three prizes and part of a fourth are covered.

Figure 1: Tweets per day containing 'Nobel', as reported by Topsy.com.

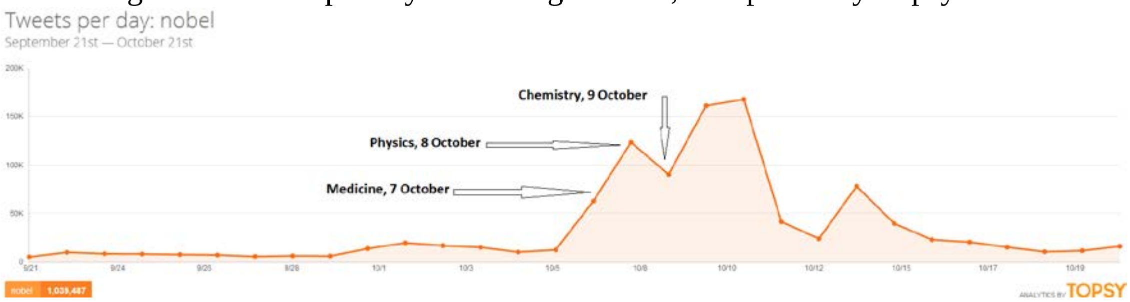
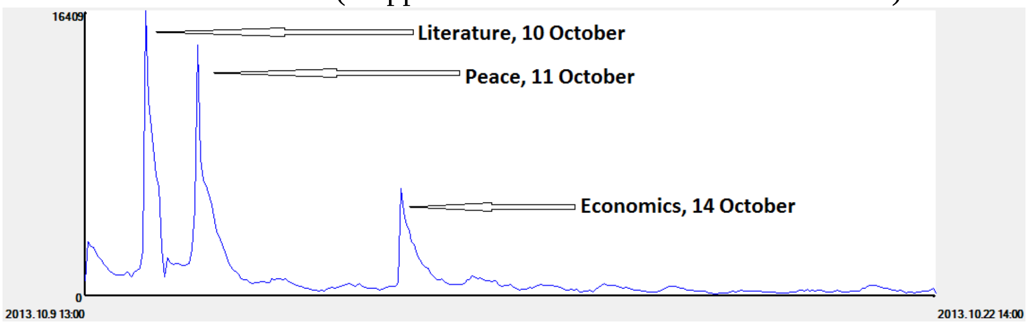


Figure 2: Tweets per hour containing 'Nobel' gathered by Mozdeh. Note the later starting time for the data (cropped and annotated Mozdeh screenshot).



For this pilot study a new word frequency method was developed to compare the prize topics and added to Mozdeh, as follows.

First, a complete list of all words in all 716,000 Nobel tweets was constructed (de-pluralising words, when relevant). Second, for each discipline and each word, the relative frequency of the word in discipline-relevant tweets (e.g., tweets containing Nobel and Chemistry) was compared to the relative frequency of the word in other tweets (e.g., tweets containing Nobel but not containing Chemistry) and a chi-squared test used to assess the significance in the differences between categories. Third, a ranked list of terms was compiled for each discipline based upon chi-square score. This list reveals words used disproportionately often for one Prize compared to the others.

Findings

The table below gives the results of an analysis of the 100 terms with the highest chi squared associated with each prize in the data set.

- Names of winners: Prominent in all cases, presumably to convey information about who won the prizes.
- Geographic location of winners: Extensively mentioned - particularly by the countries or institutions claiming an association with them. The exception is for the international organisation winning the peace prize.
- Prize topic area: Also extensively discussed, with the partial exceptions of the literature and peace prizes. Topic discussions are particularly evident for the chemistry prize, perhaps because the achievements of the three winners are not simple to explain.
- Jokes and political sarcasm: In all cases they relate to the prize topic.
- Alternative winners: Discussed in the non-academic categories.
- Gender: Discussed for the female winner. 1781 out of 19162 literature tweets (9%) mentioned her gender. Although she was the only female winner, in all tweets "woman" and "female" were twice as common as "man" and "male".
- Sentiment: Mainly expressed for the literature prize, presumably because readers of books may feel a personal engagement with the author or her books, or an ability to judge the quality of her works. Altogether, 42% of literature tweets contained at least a mildly positive sentiment.

Limitations and Conclusion

Because of Twitter API restrictions, it was not possible to automatically download the tweets substantially prior to the start of the study. The results are also limited by linguistic issues associated with the word frequency method and may miss topics that were discussed equally for the different Prizes but with a varied vocabulary so that no individual word reflected the topic.

This study found differences between subject areas in the nature of tweets containing ‘Nobel’. This confirms that the word frequency method can identify important themes through a large scale, mainly automatic analysis of tweets.

References

Eysenbach, G. (2011) Can Tweets predict citations? Metrics of social impact based on Twitter and correlation with traditional metrics of scientific impact. *Journal of Medical Internet Research*, 13(4).

James, C. (1966). Soviet press reception of Sholokhov's designation as 1965 Nobel Prize winner. *Modern Language Journal*, 50(1), 18-20.

Piwowar H. (2013). Value all research products. *Nature*, 493(7431), 159.

Roberge, J. M. (2014). Using data from online social networks in conservation science: which species engage people the most on Twitter? *Biodiversity and Conservation*, 23(3), 715-726.

Segerberg, A., & Bennett, W. L. (2011). Social media and the organization of collective action: Using Twitter to explore the ecologies of two climate change protests. *The Communication Review*, 14(3), 197-215.

Shuai, X., Pepe, A. & Bollen, J. (2012). How the scientific community reacts to newly submitted preprints: article downloads, Twitter mentions, and citations. *PLoS ONE*, 7(11).

Thelwall, M., Haustein, S., Larivière, V. & Sugimoto, C.R. (2013). Do altmetrics work? Twitter and ten other candidates. *PLoS ONE*, 8(5).

Tötösy de Zepetnek, S. (2005). Imre Kertész's Nobel Prize, public discourse, and the media. *CLCWeb: Comparative Literature and Culture*, 7(4), 9.

Table 1. Relatively high frequency words by discipline and their rankings. Numbers in brackets represent the total number of terms with similar meanings (e.g., alice, munro, #alicemunro, munroe are grouped together) in the top 100 list.

Theme	**Chemistry**	**Literature**	**Peace**	**Economics**
Names of winners	5: levitt (2), warshel(2), karkplus(2)	2: munro (4)	2: chemical(7) [winner: Organization for the Prohibition of Chemical Weapons]	3: fama(2), Hansen(3), shiller(2),
Geography of winners	19: israeli (5), usc (3), stanford(2), us-based (2), harvard, emigrated	4: canadian(10),	22: global	2: american, chicago(5), yale(2), utah, tuft(3)
Topic	2: computer (2), cyberspace(2), computational, model(4), complex, quantum, multiscale, experiment(2), tube, classical, Newtonian, mixing, simulation, geek, programming, reaction, system, molecular, development, messy, processes, drug, structural, dynamic	8: author (3), short (2), story (2), book(2), eloquent, read, secret, art, chekhov	5: watchdog, monitor(2), destruction, war, anti-chemical	13: asset, market, financial(2), analysis, share, price(2), empirical, trend-spotting(3), adjustment, interpreting, handicapping, forecasting, critique, theory, insight, reality-based, grounding, method, contradictory, unanswered, inequality, inefficient
Competitors	-	74: alexievich, haruki	7: malala(7), putin	-
Gender	-	11: woman(3)	-	-
Sentiment	-	10: congratulations (2), master, wonderful(4), hooray(2), deserved(2), proud, thrilled, happy(3)	-	-
Joke or political sarcasm	57: letterman (2), #breakingbad (3), kasparov(2), taft, cognitive, [E.g., Letterman: "Nobel Prize for Lack of Chemistry went to John Boehner and Barack Obama."]	80: ashbery [e.g., The Onion: Fucking Pathetic John Ashbery Actually Thinks He Has Shot At Nobel Prize In Literature This Year.]	41: joke, obama, bashar, drone [e.g., Obama has killed thousands w drones: can Nobel committee have Peace Prize back?]	78: debt(2) [e.g., Three Americans win the Nobel Prize in Economics. Meanwhile Congress shuts down government and hurtles towards debt default.]

Patents as Instruments for Exploring Innovation Dynamics: Different Perspectives on "Photovoltaic Cells"

Loet Leydesdorff*, Floortje Alkemade**, and Gaston Heimeriks***

*loet@leydesdorff.net
Amsterdam School of Communication Research (ASCoR), University of Amsterdam, Kloveniersburgwal 48, NL-1012 CX Amsterdam (The Netherlands)

** F.Alkemade@uu.nl; ***gheimeriks@gmail.com
Department of Innovation Studies, Faculty of Geosciences, Utrecht University, Heidelberglaan 2, NL-3584 CS Utrecht (The Netherlands)

Introduction

Patents are framed in different contexts: in addition to being among the outputs of the production system of knowledge, patents can also serve as input to the economic process of innovation. Furthermore, intellectual property in patents is legally regulated, for example, in national patent offices (e.g., Granstrand, 1999). Thus, different selection environments are relevant to patenting: the context of technological knowledge production, the economic context, and the legal framework of the state. Patents reflect these different contexts in terms of attributes: names and addresses of inventors and assignees provide information about the locations of inventions, patent classifications and claims within the patents can be used to map technological developments, citations provide measures of impact and value, etc. (Porter & Cunningham, 2005). Can patent analysis and patent maps provide us with an analytical lens for studying the complex dynamics of technological innovations? (e.g., Jaffe & Trajtenberg, 2002; Balconi *et al*., 2004; Feldman & Audretsch, 1999; Mowery *et al*., 2001).

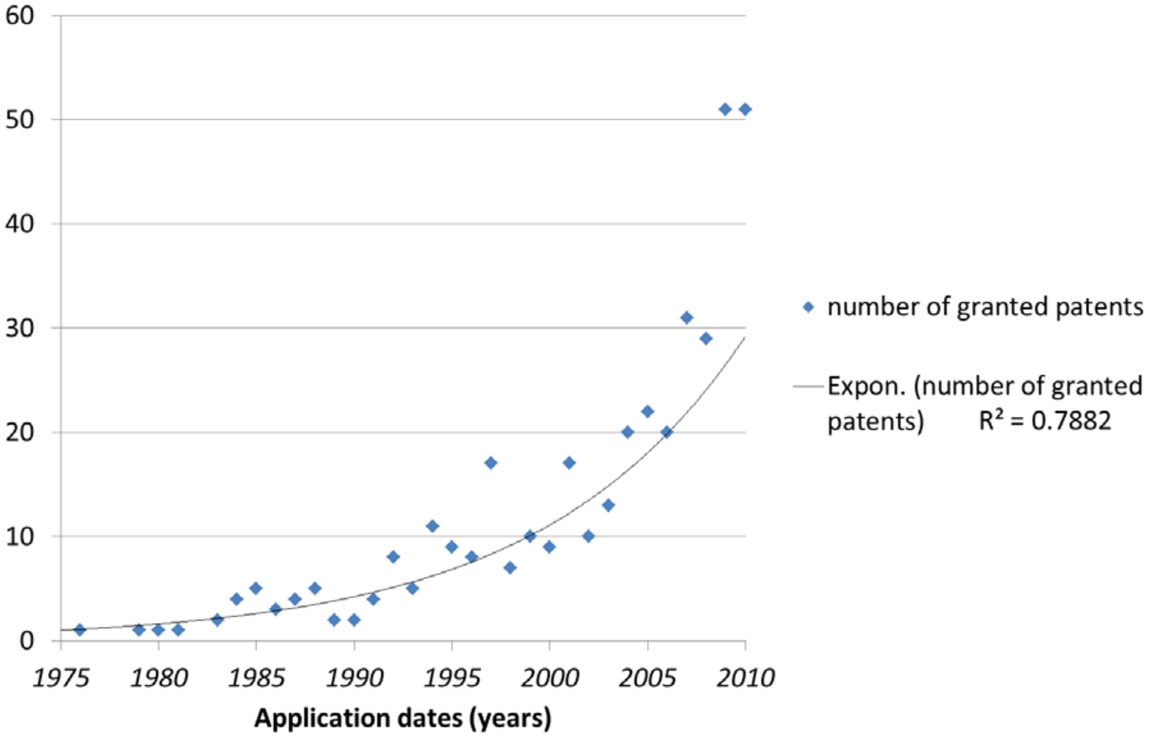
In this study, we argue that a further development of methodologies is required more than of theories (Griliches, 1984) when one understands technologies as complex adaptive systems. The diffusion of a new technology in different dimensions can be simultaneous, but also delayed or changing direction. Thus, one is challenged to combine the different perspectives heuristically and yet analytically. We explore comprehensive base maps in different dimensions (cognitive, geographical, etc.) that (*i*) can be overlaid with information about specifically selected samples, and (*ii*) show the evolution of the technologies over time. Whereas several teams have generated patent maps and overlays for patent classes (Kay *et al*., in press; Schoen *et al*., 2012), our main objective is to make these overlays *dynamic* and *interactive* so that one can use them as versatile instruments across samples gathered for different purposes.

Data

Recently, USPTO and EPO introduced a new system of so-called Cooperative Patent Classifications (CPC) that unlike patent classifications such as International Patent Classifications IPC, and its American or European equivalents, is also indexed with a focus on emerging technologies using specific tags in the new Y-class (Scheu *et al*., 2006; Veefkind *et al*., 2012). Whereas the previous classification systems have grown historically with the institutions, and combine patents that cover product and process innovations at different scales, the classification in terms of CPC provides the opportunity to take a reflexive turn since technological classes are added under the category "Y" from the perspective of hindsight. The new classifications have been backtracked into the existing databases for indexing.[1]

A new CPC tag for emerging technologies was developed as Y02: "Climate Change Mitigating Technologies." This latter tag and its subclasses are now operational in both USPTO and EPO data. The new class follows up on the "Pilot Program for Green Technologies Including Greenhouse Gas Reduction" that USPTO launched in 2009. In the meantime, more than 150,000 patents are tagged with Y02 in USPTO, among which 5,021 US patents with the search string cpc/y02e10/54$ for *material technologies* in photovoltaic cells (PV). We focus on developing the relevant instruments using the first subclass Y02E10/541 that covers "$CuInSe_2$ material PV Cells." In a next study (Heimeriks, Alkemade, & Leydesdorff, 2014), we upscale to comparisons among the nine material technologies, and including PatStat as another database of patent statistics.

Figure 1: 419 Patents granted in USPTO under the CPC tag Y02E10/541 for "CuInSe2 material PV cells", 1975-2010; September 5, 2013



$CuInSe_2$ was first synthesized in 1953 (Hahn *et al.*, 1953), and proposed as a photovoltaic material in 1974 (Shafarman & Stolt, 2003: 567f.). Thin-film technology for solar cells is still considered as the commercially most promising candidate for generating energy at competitive prices although monocrystalline silicon cell technology currently dominates the market at a cost of about 0.5 Euro/Watt-peak.[2] Although this technology has only a small share of the market, it continues to attract most of the funding for R&D among the material technologies for photovoltaic cells (刘壮 and 卢兰兰, 2011, p. 12).[3] We retrieved 419 granted

[1] The USPTO envisages replacing the US Patent Classification System (USPC) with CPC during a period of transition to 2015; at EPO, however, the European classification ECLA has already been replaced with CPC.

[2] One Watts-peak (Wp) is defined as the maximum power output of a one square meter solar panel at 25 degrees centigrade.

[3] The transcripts of these names in the Latin alphabet are: Liu, Zhuang and Lu, Lanlan, respectively.

patents at USPTO with the CPC "Y02E10/541", and brought these records under the control of a relational database management system. Figure 1 shows the trend.

Methods

Existing routines for overlaying patent data to Google Maps (Leydesdorff & Bornmann, 2012) and a map based on aggregated citations among IPC (Leydesdorff, Kushnir, & Rafols, 2012) were further developed for the purpose of *dynamic* mapping. The resulting routines are available at http://www.leydesdorff.net/software/patentmaps/dynamic.

Geographic maps

As specified in Leydesdorff & Bornmann (2012), the proportion of top-cited patents in a sample of USPTO data can be (z-)tested for each location against the expectation, but only in the case of more than five patents at a city-location. Using colors similar to those of traffic lights, cities with patent portfolios significantly below expectation in terms of citedness are colored dark-red and cities with portfolios significantly above expectation dark-green. Lighter colors (lime-green and red-orange) are used for cities with expected values smaller than five patents (which should not statistically be tested) and for non-significant scores above or below expectation (light-green and orange). The precise values are provided in the descriptors which can be accessed by clicking on the respective nodes. See at http://www.leydesdorff.net/photovoltaic/cuinse2/cuinse2_inventors.htm for the aggregated set.

Classification maps

We use the base map of aggregated citation relations among IPC in the USPTO data 1975-2011 provided by Leydesdorff, Kushnir, and Rafols (2012). These maps are available at http://www.leydesdorff.net/ipcmaps for both three and four digits of the current IPC version 8.[4] The initial step for the construction of the time-series is again the construction of the overall map for the aggregated set. Subsequently, the time series are generated by setting filters for consecutive years to this aggregate.

The routine ipcyr.exe (available at http://www.leydesdorff.net/software/patentmaps/dynamic) generates input information for consecutive years in the format of VOSviewer for the mapping (http://vosviewer.com). Two time series of files are generated as input for the mapping for three and four digits of IPC, respectively. The routine additionally writes a file "rao.dbf" which contains Rao-Stirling diversity for both three and four-digit IPC-based maps. Rao-Stirling diversity is defined as follows (Rao, 1982; Stirling, 2007):

$$\Delta = \sum_{ij} p_i p_j d_{ij} \tag{1}$$

where d_{ij} is a disparity measure between two IPC classes i and j at the respective level of specificity; p_i is the proportion of elements assigned to each class i. As the disparity measure, we use (1 – *cosine*) since the cosine values of the citation relations among the aggregated IPC was used for constructing the base map of three and four digits (Jaffe, 1986).

[4] The first four digits of CPC will be identical to IPC 8.

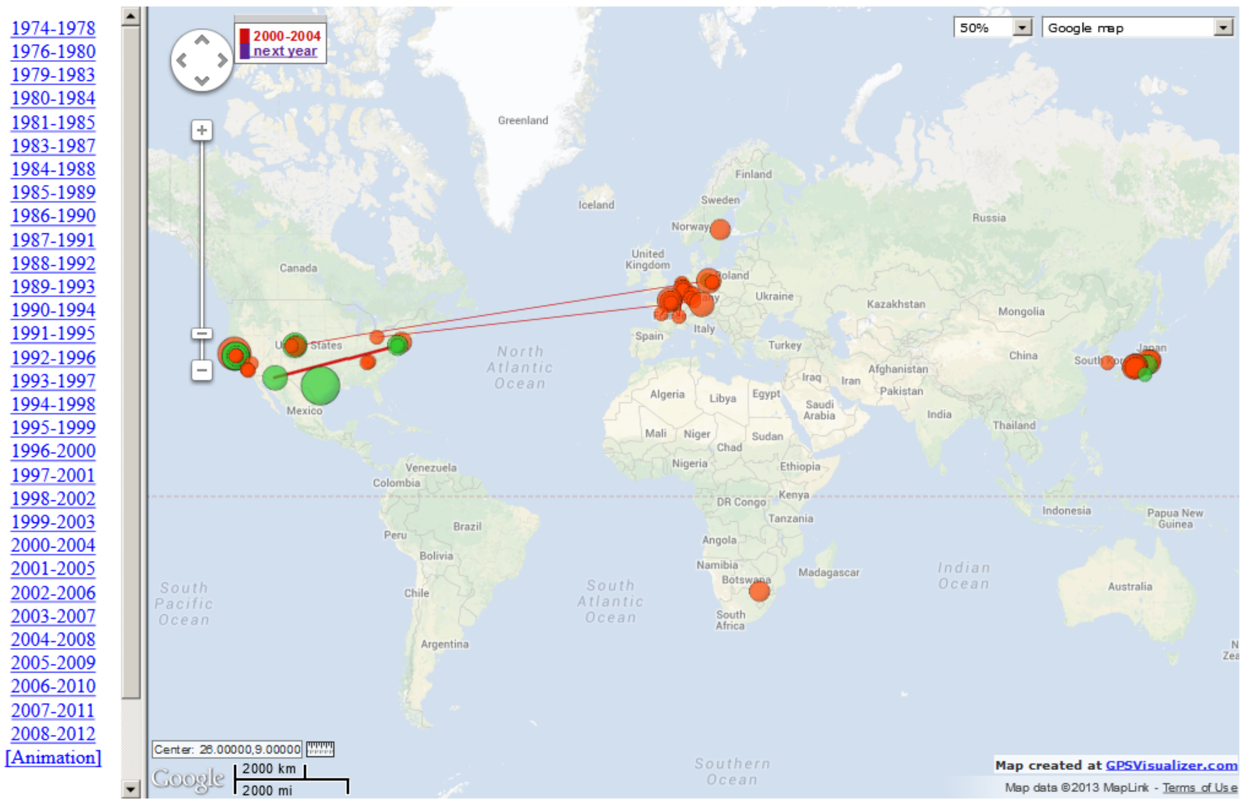
Results

Geographical diffusion

One obtains a series of maps in which the node sizes are proportionate to the logarithm of the number of patents. [We use log(n+1) in order to prevent cities with single patents from disappearing because log(1) = 0.] As noted, the node colors correspond to the quality of the patents in terms of their citedness. One can click on each node to find statistical details. (This statistical data is also stored in the file "geo.dbf" that is generated and overwritten in each run.)

Figure 2 provides the Google map for the five-year period 2000-2004. The numbers of patents are often too small for significance testing, but one can see at a glance that the US is dominant with green-coloured nodes in this (USPTO!) set in terms of both numbers and quality. In addition to the US, Japan and Europe have developed their own networks. (One can zoom in on the map at http://www.leydesdorff.net/photovoltaic/cuinse2/index.html.) During this period, international co-inventorship between the three world regions was limited to transatlantic collaborations.

Figure 2: Patent configuration during 2000-2004 for CuInSe2 material in PV Cells (Y02E-10/541) in USPTO data; an interactive version of this map is available at http://www.leydesdorff.net/photovoltaic/cuinse2/index.html .



Inspection of the animation informs us that patenting began in isolated centers in the USA, then spread first within the U.S. and thereafter also to some centers in Europe (e.g., 1983-1987). During the second half of the 1980s, Japanese and also isolated inventors in Europe began to patent in the USA. In 1990-1994, co-inventorship is found only in the local environments of Munich (Germany) and within Colorado. The latter network reflects that the

National Renewable Energy Laborarory (NERL) of the US Department of Defense is based in Golden, Colorado.

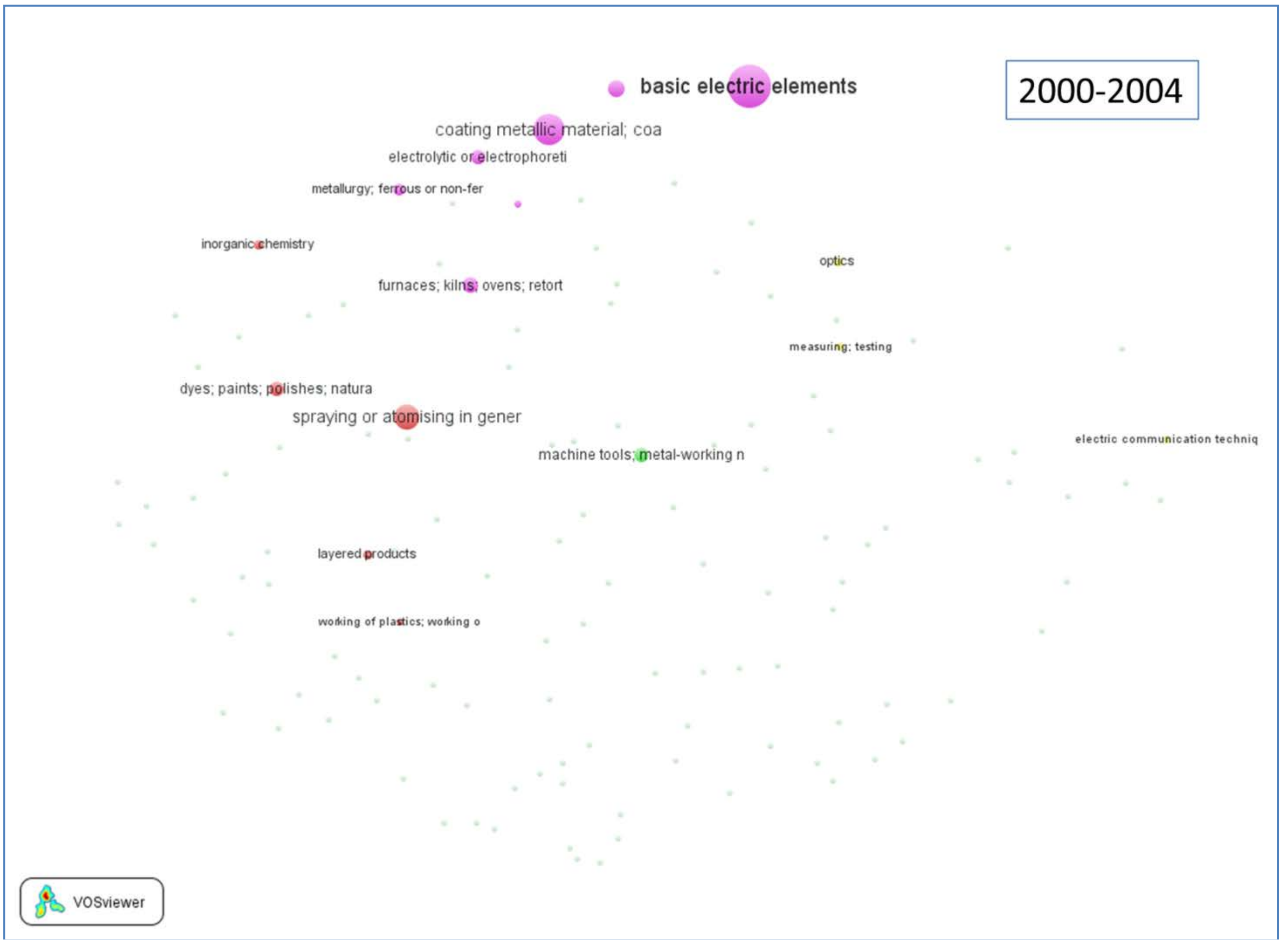
In the second half of the 1990s, there is also more co-invention in the USA and Japan, but within national boundaries. The technology increasingly becomes commercially viable during this period. The number of cities in Europe and Japan with USPTO patents increases, and transatlantic collaboration is resumed towards the end of the 1990s. Since 2003—the commercial phase—one sees co-invention between Japan and the USA, and within Europe. In the European context, France plays a role in addition to a recurrent collaboration between Germany and Spain. An address in the UK (Stirling in Scotland) joins the US networks in the final periods (2007-2011, 2008-2012). During 2008-2012, Europe is otherwise no longer represented in USPTO data.

In summary, these are sparse networks. The majority of the inventors do not collaborate beyond local environments; collaborations within nations are more important than international collaborations.

IPC classes

How can the map in terms of IPC-classes add to our understanding of these geographical dynamics? Figure 3 shows the IPC-based map (three digits) for the same set of patents as used in Figure 2 (2000-2004). The technology originated during the 1970s in the category of “basic electric elements” and remained there during the next 15 years, but has spread during the 1990s into other domains of technology such as “spraying and atomizing” and machine techniques for making thin films in photovoltaic cells. This diffusion increases further during the 2000s.

Figure 3: Map of USPTO patents in terms of IPC at the three-digit level for the period 2000-2004. A dynamic version of this map is available at http://www.leydesdorff.net/photovoltaic/cuinse2/cuinse2.ppsx.



Figures 2 and 3 can be combined using frames in the html for the splitting of the screens (at http://www.leydesdorff.net/photovoltaic/cuinse2/dualmix.html; not shown here. One can animate this figure as Figure 2, but this animation taught us that dynamic changes in two different (split) screens are difficult to handle for an analyst.

A user needs control over the time steps when focusing on the *differences* between two dynamics. Therefore, we propose another solution: by clicking on another year, one opens a new window in the browser with the same figures for this different year. A user is then able to compare among years using, for example, different time intervals (such as five or ten years) by going back and forth between windows, at one's own pace.

Rao-Stirling diversity as a measure of technological change

Figure 4: The development of Rao-Stirling diversity in IPC (three and four digits) among 419 USPTO-patents with CPC Y02E10/541 (“$CuInSe_2$ material PV cells”) during the period 1975-2012.

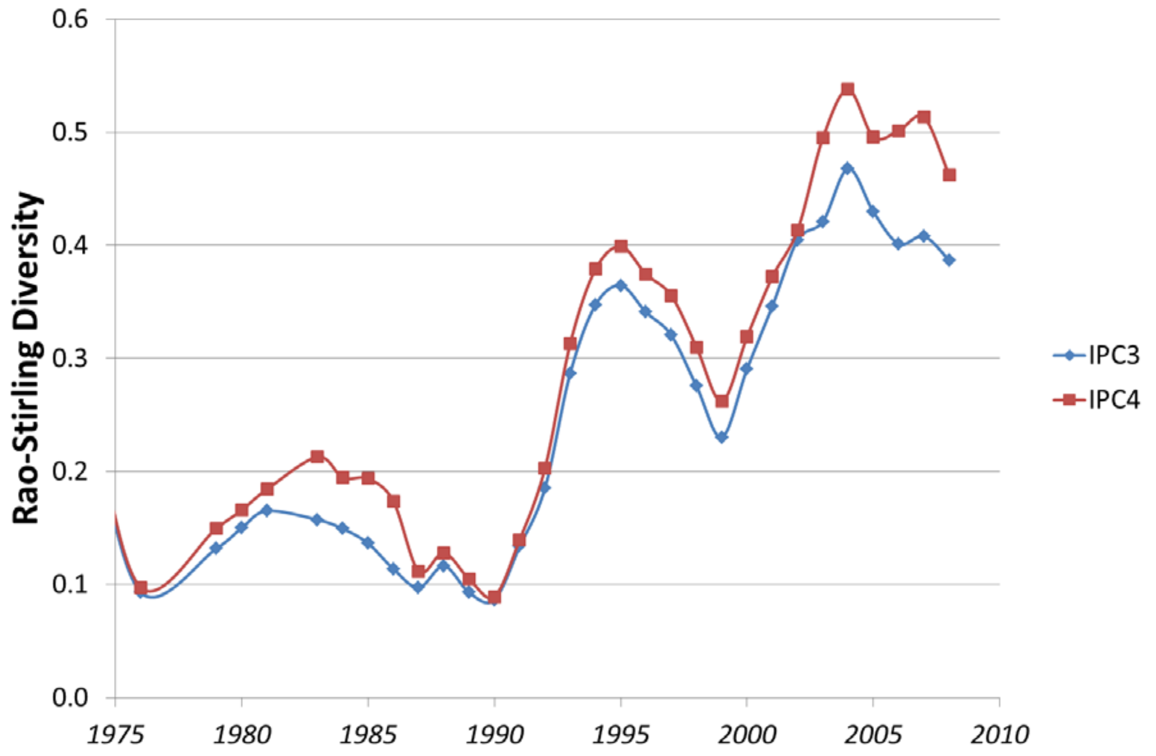


Figure 4 shows the development of Rao-Stirling diversity in the IPC-based maps during the entire period. The figure suggests that the technology was developed in three cycles. Two of the valleys, i.e., the period of convergence in the late 1980s and the latest convergent period, correspond with breakthroughs in the efficiency of thin-film solar cells (Green *et al.*, 2013). Combining the maps with split-screens (at http://www.leydesdorff.net/photovoltaic/cuinse2/dualmix.html) for each consecutive year, we suggest specifying these cycles as follows (Shafarman & Stolt, 2003):

1. an early cycle during the 1980s which is almost exclusively American; after initial development of the technology at Bell Laboratories in the ’70s, Boeing further developed the solar cells using these materials;
2. a second cycle during the 1990s that includes transatlantic collaboration and competition with Europe; and
3. a third and current cycle—the commercial phase—in which American-Japanese collaboration, on the one side, and collaboration *within* Europe, on the other, prevail.

The volume of patents continued to increase more smoothly, but with an increasing (above-exponential) rate during the most recent years (Figure 1). The pronounced articulation of these cycles in terms of Rao-Stirling diversity came as a surprise to us. As the material technology becomes mature, other technologies such as spraying the thin film on carrier materials may become crucial.

Conclusions

The maps of patents in different dimensions are instrumental to understanding the complex dynamics of innovation by providing different projections of these dynamics. We distinguished in this study between IPC-based maps that show the technological organization of the patents in a vector space, the geographic maps as overlays to Google Maps, and the social networks—not shown here—that can be overlaid to the geographic map, but can also be studied in themselves using graph-theoretical instruments such as spring-embedded layouts.

At the theoretical level, we thus aim to address what Griliches (1994) called "the computer paradox" from a methodological angle: ever more data—nowadays, one would say "big data"—are stored in ever larger repositories, but the logic of these repositories is institutional, whereas the logic of innovation is based on the transversal recombination of functions at interfaces (e.g., supply and demand). The relabeling using the Y-tag in CPC, however, provides an opportunity to follow delineated technologies within and across databases: recent agreements of EPO and USPTO with the Chinese, Korean, and Russian patent offices to use also CPC in the near future show an increased awareness to coordinate the data in a networked mode.

As innovations relate at structural interfaces—between selection environments—one can consider them from different perspectives such as market opportunities or technological novelty. Using a single (theoretical) term such as "diffusion," is then foreseeably insufficient without specification of the different systems of reference: diffusion can be defined in terms of markets/industries, geographies, or also technologies—in terms of branching and recombination (Arthur, 2009). As we argued, patents provide an (albeit imperfect) lens to this complex dynamics.

The use of Rao-Stirling diversity in this study (Figure 4) can be considered as a case in point: the literature pointed us to considering variety versus the loss of variety in shake-out phases as central to techno-economic developments (Anderson & Tushman, 1990), but the data allowed us to operationalize this in relation to the new instruments. The extension beyond two maps to be recombined follows as a progressive research agenda for quantitative innovation studies (Rotolo *et al*., in preparation).

Acknowledgements

We are grateful to comments on a previous version of the manuscript by Jan Youtie (Georgia Tech) and Wilfried van Sark (Utrecht University).

References:

Anderson, P., & Tushman, M. L. (1990). Technological discontinuities and dominant designs: a cyclical model of technological change. *Administrative Science Quarterly, 35*, 604-633.

Arthur, W. B. (2009). *The Nature of Technology*. New York, NY: Free Press.

Breschi, S., & Lissoni, F. (2009). Mobility of skilled workers and co-invention networks: an anatomy of localized knowledge flows. *Journal of Economic Geography, 9*(4), 439-468.

Green, M. A., Emery, K., Hishikawa, Y., Warta, W., & Dunlop, E. D. (2013), Solar cell efficiency tables (version 41). *Progress in Photovoltaics: Research and Applications, 21*(1), 1–11; doi: 10.1002/pip.2352

Griliches, Z. (1994). Productivity, R&D and the Data constraint. *American Economic Review, 84*(1), 1-23.

Heimeriks, G., Alkemade, F. & Leydesdorff, L. (2014). The Evolution of PV Technology, Paper to be presented at 5th International Conference on Sustainability Transitions, Utrecht, 27-29 August 2014.

Jaffe, A. B. (1986). Technological Opportunity and Spillovers of *R&D*: Evidence from Firm's Patents, Profits, and Market Value. *American Economic Review, 76*(5), 984-1001.

Jaffe, A. B., & Trajtenberg, M. (2002). *Patents, Citations, and Innovations: A Window on the Knowledge Economy*. Cambridge, MA/London: MIT Press.

Kay, L., Newman, N., Youtie, J., Porter, A. L., & Rafols, I. (in press). Patent Overlay Mapping: Visualizing Technological Distance. *Journal of the American Society for Information Science and Technology*, preprint at arXiv:1208.4380.

刘壮 (Liu, Zhuang) and 卢兰兰 (Lu, Lanlan) (2011). Development of Cu(InGa)Se2 (CIGS) Thin-Film Solar Cells, *Bulletin of Advanced Technology Research*, 5(8) 11-13.

Leydesdorff, L., & Bornmann, L. (2012). Mapping (USPTO) Patent Data using Overlays to Google Maps. *Journal of the American Society for Information Science and Technology, 63*(7), 1442-1458.

Leydesdorff, L., Kushnir, D., & Rafols, I. (2012). Interactive Overlay Maps for US Patent (USPTO) Data Based on International Patent Classifications (IPC). *Scientometrics*, doi: 10.1007/s11192-012-0923-2.

Ponds, R., Van Oort, F., & Frenken, K. (2010). Innovation, spillovers and university–industry collaboration: an extended knowledge production function approach. *Journal of Economic Geography, 10*(2), 231-255.

Porter, A. L., & Cunningham, S. W. (2005). Tech mining. *Competitive Intelligence Magazine, 8*(1), 30-36.

Rao, C. R. (1982). Diversity: Its measurement, decomposition, apportionment and analysis. *Sankhy : The Indian Journal of Statistics, Series A, 44*(1), 1-22.

Rotolo, D., Rafols, I., Hopkins, M. M., & Leydesdorff, L. (in preparation). Mapping *De-Facto* Governance in the Case of Emerging Science and Technologies. available at http://arxiv.org/abs/1303.4672.

Scheu, M., Veefkind, V., Verbandt, Y., Galan, E. M., Absalom, R., & Förster, W. (2006). Mapping nanotechnology patents: The EPO approach. *World Patent Information, 28*, 204-211.

Schoen, A., Villard, L., Laurens, P., Cointet, J.-P., Heimeriks, G., & Alkemade, F. (2012). *The Network Structure of Technological Developments; Technological Distance as a Walk on the Technology Map*. Paper presented at the Science & Technology Indicators (STI) Conference 2012 Montreal.

Shafarman, W. N., & Stolt, L. (2003). Cu (InGa) Se2 Solar Cells. In A. Luque & S. Hegedus (Eds.), *Handbook of photovoltaic science and engineering* (pp. 567-616). Hoboken, NJ: Wiley.

Stirling, A. (2007). A general framework for analysing diversity in science, technology and society. *Journal of the Royal Society Interface, 4*(15), 707-719.

Unold, T., and Kaufmans, C. A. (2012). *Chalcopyrite Thin Film Materials and Solar Cells*, Chapter 1.18 in W. G. J. H. M. van Sark (Ed.) *Photovoltaic Technology*, Volume 1 in A. Sayigh (Ed.) *Comprehensive Renewable Energy*, Elsevier, United Kingdom, pp. 5-11

Veefkind, V., Hurtado-Albir, J., Angelucci, S., Karachalios, K., & Thumm, N. (2012). A new EPO classification scheme for climate change mitigation technologies. *World Patent Information*, *34*(2), 106-111.

Scopus and Web-of-Science 2012 compared in terms of aggregated journal-journal citation relations: global maps and interactive overlays[1,2]

Loet Leydesdorff*, Félix de Moya-Anegón** and Wouter de Nooy*

*loet@leydesdorff.net; W.deNooy@uva.nl
Amsterdam School of Communication Research (ASCoR), University of Amsterdam, Kloveniersburgwal 48, 1012 CX Amsterdam (The Netherlands)

**felix.moya@scimago.es
CSIC, SCImago Research Group, Calle Albasanz 26, Madrid 28037 (Spain)

Introduction

We compare the networks of aggregated journal-journal citation relations as provided by the Journal Citation Reports (JCR) 2012 of the Science and Social Science Citation Indexes (SCI and SSCI) with similar data for 2012 based on Scopus. First, we develop basemaps and overlays for the two sets separately. Second, we match journal names across databases to assess the overlap.

Data

The data for Scopus 2012 was extracted from the Scopus database (1996-2012) in October 2013 (Leydesdorff *et al*., in press). Since single citations are aggregated in the JCR under "All others," we discarded these values and pursued the analysis with the 2,688,731 remaining links which contain 36,748,156 citation relations.

Table 1: Descriptive statistics of the data.

	Scopus 2012	JCR 2012 (SCI + SSCI)	JCR SCI	JCR SSCI
N of journals	20,172 *	10,936	8,471	3,047
Citation links	(6,672,033) 2,688,731 **	2,350,491	2,122,083	253,320
Sum of citations	(40,731,458) 36,748,156 **	37,759,948	35,721,660	2,454,015
Self-citations	2,898,006	3,248,968	3,049,332	298,637

* The N of journals is 20,554 for the period 1996-2012
** corrected for single citation links.

[1] The full paper is forthcoming as: Leydesdorff, L., de Moya Anegón, F., & de Nooy, W. (in press), Journal Maps and Interactive Overlays of Scopus and Web-of-Science 2012: The two aggregated journal-journal citation networks compared *Journal of the Association for Information Science and Technology*; a preprint is available at http://arxiv.org/abs/1404.2505 .

[2] We are grateful to Lykle Voort of the Amsterdam computer center SARA for his support. Some of this work was carried out on the Dutch national e-infrastructure with the support of the SURF Foundation.

JCR data were harvested from two JCR files for the SCI 2012 and SSCI 2012, respectively. The two files were first merged. The category "All others" is denoted as missing values.

Methods

The mapping method is analogous to the one applied previously to the aggregated set of Scopus 1996-2012 data published by Leydesdorff et al. (in press), and to the map based on JCR 2011 used by Leydesdorff et al. (2013). However, the two maps for 2012 (and the underlying matrices) can also be compared to each other.

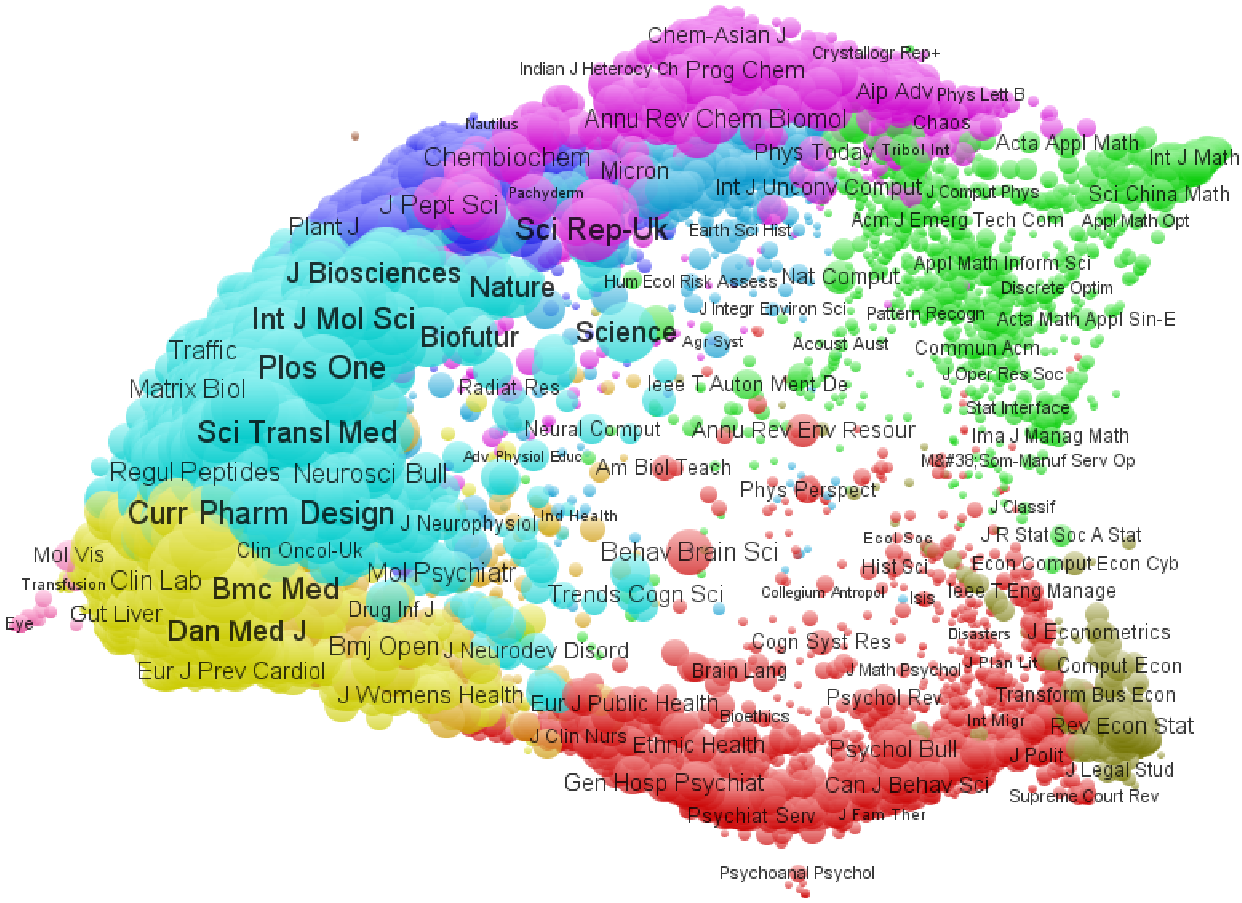
Table 2: Statistics used for the visualization in VOSviewer

	JCR-WoS 2012	Scopus 2012
Giant component	10,549	18,160
After correction for visual outliers	10,546	18,154
N of clusters (Blondel et al., 2008)	12	65
N of clusters (VOSviewer)	11	47
Modularity Q	0.557	0.694

Global maps

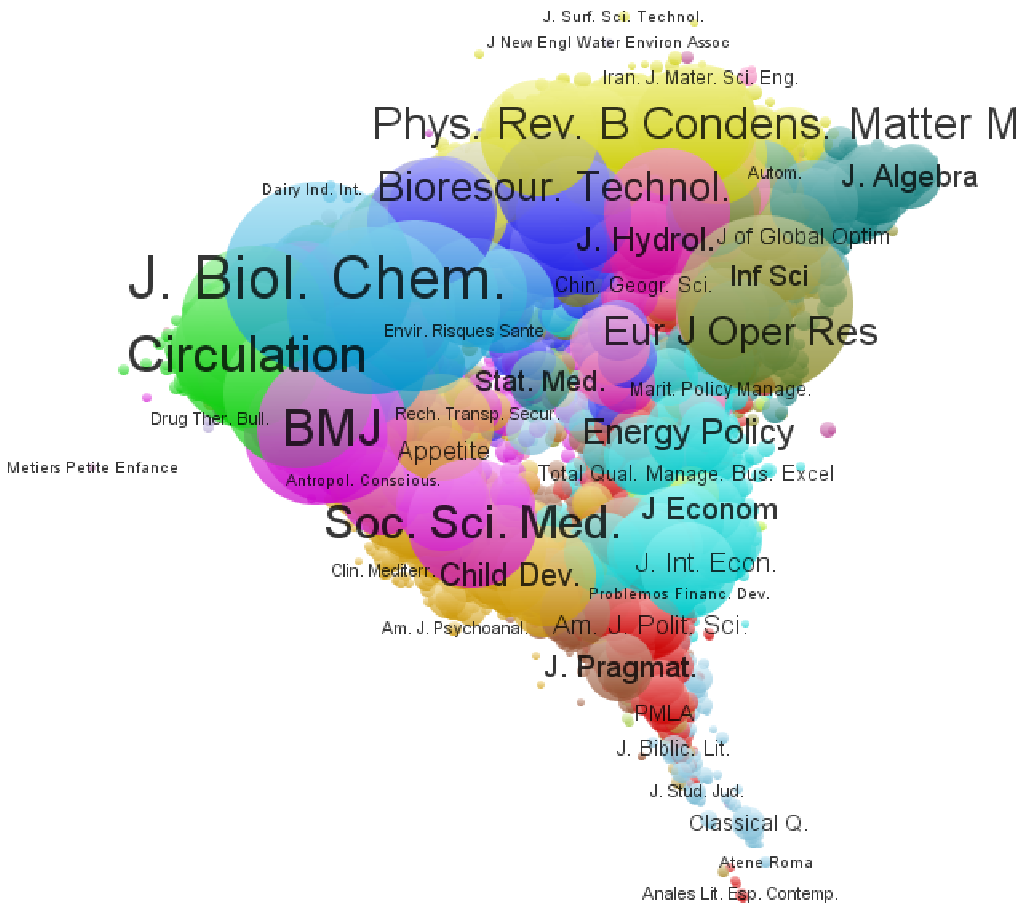
Figure 1 shows the base map for the 10,546 journals (96.4%) included in the largest component of JCR 2012. The shape and coverage is very similar to the map for 2011 (Leydesdorff et al., 2013, Fig. 1 at p. 2575). This reproduction of a base map in two different years—using the same methods—provides confidence in the validity of the technique and the reliability of the data.

Figure 1: Citing patterns of 10,546 journals in JCR 2012 visualized as a base map; cosine > .2; colors correspond to 11 communities distinguished by VOSviewer; available for webstart at http://www.vosviewer.com/vosviewer.php?map=http://www.leydesdorff.net/journals12/jcr12.txt



The map based on Scopus data 2012 (Figure 2) is also not so different from the previously published map based on aggregated Scopus data 1996-2012 (Leydesdorff *et al.*, in press: Figure 3). The tail of the humanities journals at the bottom right is lacking from the JCR-based maps, while the A&HCI is not included in JCR.

Figure 2: Citing patterns of 18,154 journals in Scopus 2012 visualized as a base map; colors correspond to 42 communities distinguished by VOSviewer; available for webstart at http://www.vosviewer.com/vosviewer.php?map=http://www.leydesdorff.net/scopus12/scopus12.txt.



Interactive overlay maps

The base maps can be used to position sets of documents (e.g., portfolios) in terms of the disciplinary composition. The routines provide Rao-Stirling diversity values for the sets under study relative to the respective maps.

In previous studies, we used datasets generated by Rafols *et al*. (2012) in which the Science and Technology Policy Research Unit (SPRU) at the University of Sussex was compared with the London Business School (LBS). These same sets of documents are used as the example in this study (e.g., Figure 3).

Figure 3: Scopus-based overlay map 2012 of journal publication portfolios from 2006 to 2010 of the Science and Technology Policy Research Unit SPRU at the University of Sussex (*N* = 268).

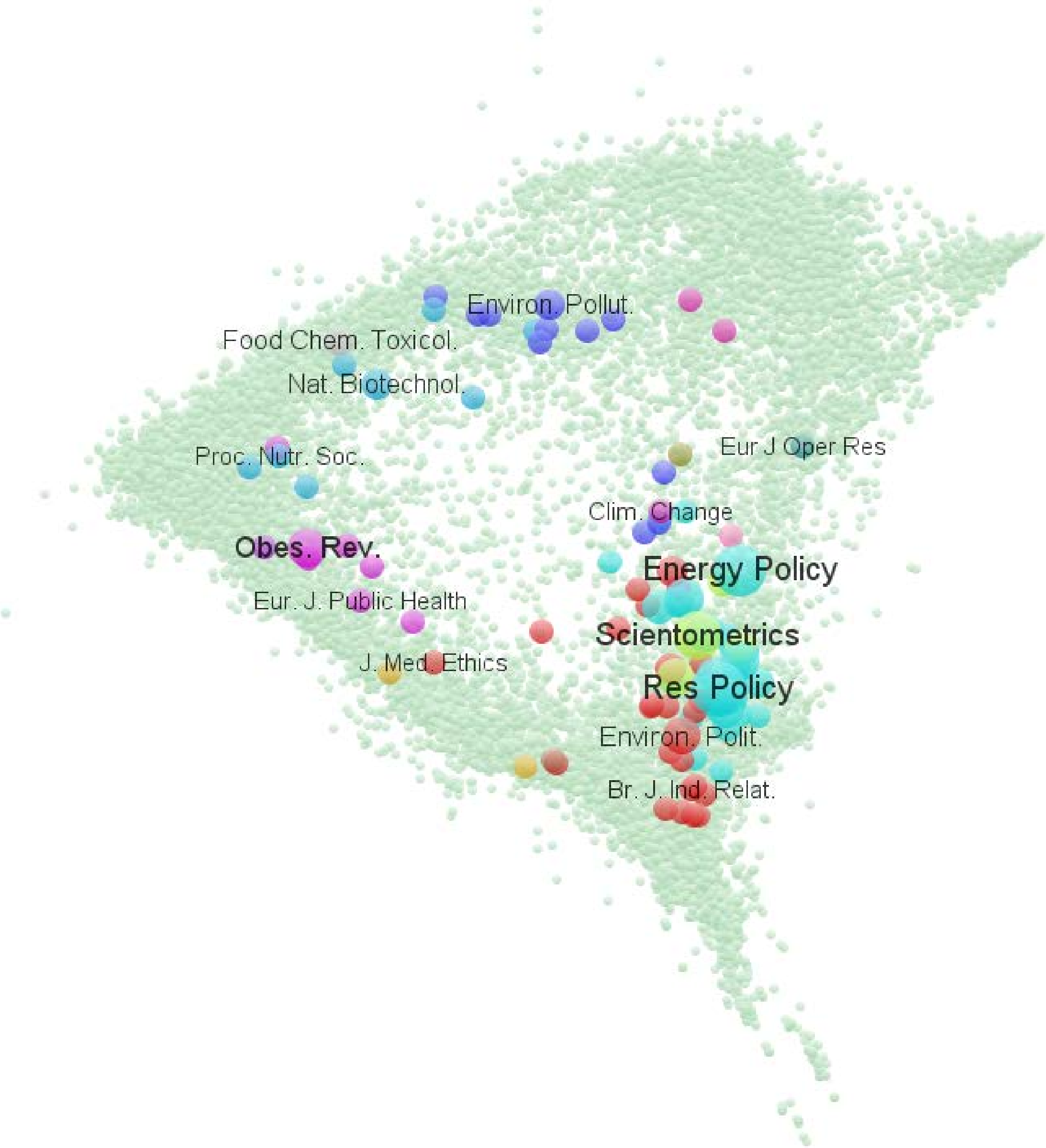


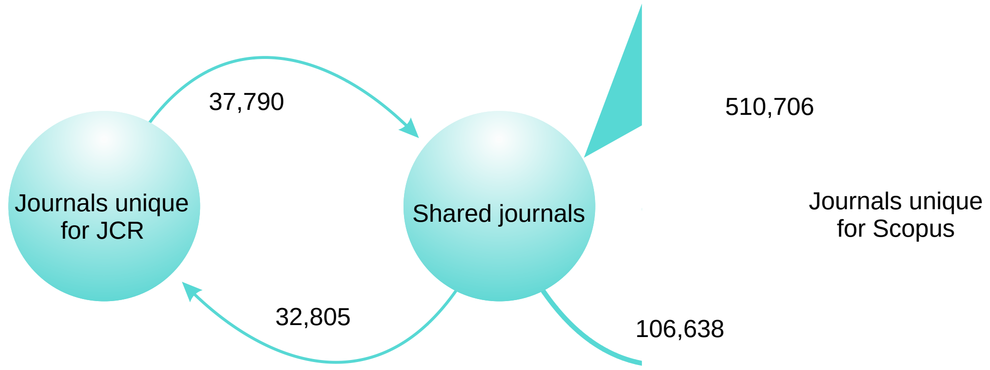
Table 3: Rao-Stirling diversity for SPRU and LBS documents (2006-2010) in both the 2011 and 2012 maps based on annual JCR data, and the two Scopus maps.

	JCR 2012 (a)	**JCR 2011 (b)**	***N WoS***	**Scopus 2012 (c)**	**Scopus 1996-2012 (d)**	***N Scopus***
SPRU	0.2170	0.2175	155	0.1219	0.1489	268
LBS	0.0918	0.0922	348	0.0863	0.0917	715

Overlap between databases

Using fuzzy-string matching and ISSN numbers, we were able to match 10,524 journal names between the two sets. An Excel file with lists of matched and unique journals in Scopus and WoS, is available online at http://www.leydesdorff.net/journals12/all_journals.xlsx.

Figure 4: Citation relations among shared and unique journals in JCR (left side) and Scopus (right side).



The 10,524 journals matched between JCR and Scopus comprise 96.3% of all JCR journals and 51.2% of all journals in Scopus. Citation flows point from journals that are unique to Scopus to journals shared by both databases (Figure 4), suggesting that the shared journals are the more important ones. Citation flows are more balanced between shared and unique journals in JCR.

Conclusion

The basemaps are available for interactive usage at http://www.leydesdorff.net/journals12 (WoS) and http://www.leydesdorff.net/scopus12 (Scopus). The user can overlay downloads from either Scopus or WoS, and generate maps in VOSviewer. In the full paper, we add a network analysis of the two citation matrices; we also compare journal ranks in these two environments.

References

Leydesdorff, L., de Moya-Anegón, F., & Guerrero-Bote, V. P. (in press). Journal Maps, Interactive Overlays, and the Measurement of Interdisciplinarity on the basis of Scopus data, *Journal of the Association for Information Science and Technology*; at http://arxiv.org/abs/1310.4966

Leydesdorff, L., Rafols, I., & Chen, C. (2013). Interactive Overlays of Journals and the Measurement of Interdisciplinarity on the basis of Aggregated Journal-Journal Citations. *Journal of the American Society for Information Science and Technology*, *64*(12), 2573-2586.

Rafols, I., Leydesdorff, L., O'Hare, A., Nightingale, P., & Stirling, A. (2012). How journal rankings can suppress interdisciplinary research: A comparison between innovation studies and business & management. *Research Policy*, *41*(7), 1262-1282.

Newspaper coverage of genetically modified foods in China[1]

Yuxian Liu* and Cong Cao**

*yxliu@tongji.edu.cn
Library of Tongji University, Tongji University, Siping Street 1239, Shanghai, 200092 (China) &
SPRU, School of Business, Management and Economics, University of Sussex, Falmer, Brighton, BN1 9SL (UK)

**Cong.Cao@nottingham.ac.uk
School of Contemporary Chinese Studies, University of Nottingham, Jubilee Campus, Nottingham, NG8 1BB (UK)

Introduction

Biotechnology, especially that used in agriculture to modify genetic compositions of crops, involves both potential benefits and risks for individuals and society as a whole. Possible benefits of genetically modified (GM) crops include input characteristics such as higher yields, wider growing conditions (better tolerance to different temperatures, soils, and so on) and increased resistance to pests and diseases as well as output characteristics such as higher nutritional content, improved food quality including taste, and added medical properties. However, concerns have also been raised that GM crops could lead to risks to human health, the environment, and public welfare. As such, portrayed as either "angel" or "evil", agricultural biotechnology has evoked heated debates in many parts of the world, including China, one of the countries that have devoted to significant resources into the technology.

As a form of formal mass media, newspapers have played an important role in the debate, not only presenting opinions of various stockholders but also helping shape the formulation and change of the policy, along with various interest groups that have a stake in the debate. Though Du & Rachul (2012) find that the most coverage of GM-based foods in two of the major Chinese newspapers – *People's Daily* and *Guangming Daily* – between January 2002 and August 2011 was positive. some other studies have different findings. Research has found that the risks associated with GMOs have been found to be amplified in *Southern Weekend*, *Chinese Youth Daily* and so on, thus giving Chinese a negative impression on GMOs (; Hou & Peng 2011; Yang 2012;). It is noticed that three studies have been conducted by scholars at the Huazhong Agricultural University, the institution to which the biosafety certificate for two strains of GM rice were granted by the Ministry of Agriculture's Biosafety Committee in November 2009.

In order to undertake a comprehensive study of how GM foods have been covered in the Chinese mass media, we have selected five newspapers which are perceived to represent the interests of different stakeholders in the controversy and debates around GM foods. They are: *People's Daily* (PD); *Science and Technology Daily* (ST); *Farmers' Daily* (FD); *Southern Weekend* (SW); and *Outlook Weekly* (OW).

[1] This work was supported by the U.S. National Science Foundation (SES–1115319) and the National Natural Science Foundation of China (71173154).

Major themes

We analyze specific topics about GM foods that have been discussed in these newspapers.

Table 1. Major themes of different newspapers

	PD	ST	Farmer	SW	OW	Total
Progress and achievement	12	74	36	1	3	126
Food Security	16	58	23	3	2	102
Debate and risk	10	55	8	13	15	101
Policy and governance	10	43	19	8	11	91
Benefit	9	32	28	0	1	70
Plant	2	28	22	2	0	54
Commercialization	0	16	25	1	1	43
Marker system	8	19	11	1	1	40
Trade war	9	6	5	0	0	20
History	2	8	4	1	0	15
Comparison	0	2	3	0	0	5
Staple food modified	0	0	3	1	1	5
Total	78	341	187	31	35	672

Coverage of GM technology in these newspapers was most around the progress and the achievement of GM, followed by food security, debate and risk, policy and governance. In particular, *People's Daily* concerned more about the food security problem, *S&T Daily* and *Farmers' Daily* paid more attention to the progress and achievements, *Southern Weekend* and *Outlook Weekly* analyzed the risk of GMOs.

Presentation of the sources

Table 2 the tones of the sources

	Objective	Neutral	Supportive	Total	%
Industries	17	23	54	94	11.3
Officials	25	48	79	152	18.3
Scientists	52	55	375	482	58.1
Humanists	17	19	15	51	6.2
Representatives of consumer organization	32	9	9	50	6.0
Total	143	154	532	829	100.0
%	17.2	18.6	64.2		

In general, the views presented in the newspapers under study were more supportive to the introduction of GM foods with over a half (64.2%) named sources having a positive attitude toward the development of GMFs in China. 18.6 percent of the persons were neutral and only 17.2 percent appeared to be negative about the development of GM technology in China. .

It is understandable that scientists involved in research and commercialization of GM crops were very supportive than other stockholders. Corporate representatives were also positive with a slightly less than half of them expressing views of support to the development of GM in China. Government officials seemed to be neutral, and scholars of humanities were less likely to be supportive with most being neutral and 30 percent absolutely objective to the development of GM in China. .

Conclusion

Through analyzing the coverage of GM foods in five leading Chinese newspapers between 2000 and 2012, this study finds that that different newspapers presented views of their own constituencies with issues of food security, debate and risk, policy and governance, in addition to reports on the progress and achievement of GM technology. The coverage tended to be supportive to the development of GMOs. As a whole, the coverage corresponded to major events related to the development of GM foods in China and may have helped shape the change of policy in research and commercialization toward GMOs in China.

References

Du, L. & Rachul, C. (2012). Chinese newspaper coverage of genetically modified organisms. BMC Public Health. 12. 326.

Hou, D. & Peng, G. (2011). Risk communications of GMO and construction of public commons: content analysis on the 176 reports from 15 newspapers. Today's Mass Media. (9). 30−31.

Yang, Y. (2012). The phenomenon of de-science formed in the process of construction of GMO: an empirical study based on the coverage of newspapers. News Hobbyists. (1), 5−6.

Zhao, J. (2012). The problems of coverage of GM crops in China and the solutions. Forum of Chinese Association of Science and Technology. (2). 191−192

Characterizing collaborations with bibliometric measures: enhancing descriptive indicators and co-authorship data

Diana Lucio-Arias* and Sandra Carolina Rivera[**]

[*] *dlucio@ocyt.org.co*
Colombian Observatory of Science and Technology, Cra 15 37-59, Bogotá,D.C., 111311 (Colombia)

[**] *crivera@ocyt.org.co*
Colombian Observatory of Science and Technology, Cra 15 37-59, Bogotá,D.C., 111311 (Colombia)

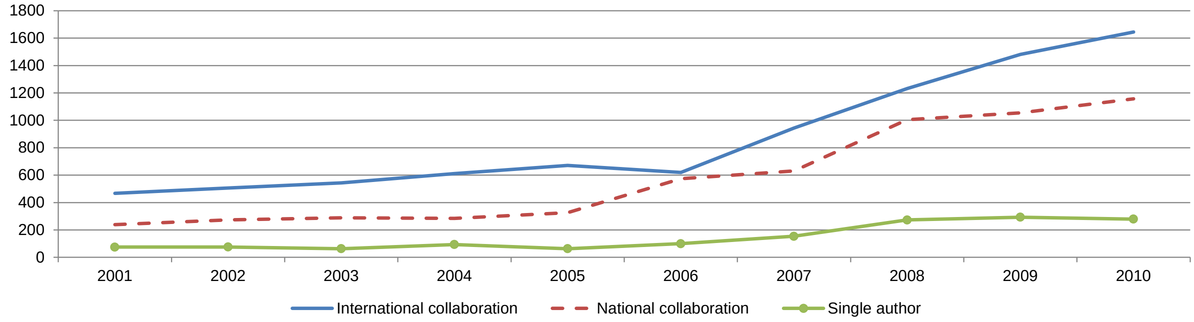
Introduction

Collaboration has become a relevant topic in the establishment of scientific policy agendas. To increase collaborations without a further reflection of the meaning of the practice, its gains and implications for each of the parties involved, simplifies the complexity of a rather heterogeneous practice where many factors take a toll. The increase interest in collaborations from the policy perspective seems to respond to efforts to increase research capacities by joining distributed efforts under shared interests. In developing countries, like Colombia, collaboration is regarded as a potential mean for knowledge and technology transfer, for "globalizing" research topics and agendas and, at the end, of improving the quality, and visibility, of research results. As a consequence many studies have emphasize on the relation between scientific collaboration and research performance (e.g. Bordons, Gomez, Fernandez, Zulueta, & Mendez, 1996; Lee & Bozeman, 2005).

The political discourse is supported by the reality, especially when it becomes simplified in terms of indicators. A transition to a highly collaborative nature of the scientific activity was already noted by Price when studying the networks of scientific papers (Price, 1963). Collaboration opens a window of opportunity for authors of many countries, and this has resulted in changes in the geographic distribution of science, particularly from the second half of the 20th century (Wagner & Leydesdorff, 2005; Leydesdorff, 2013).

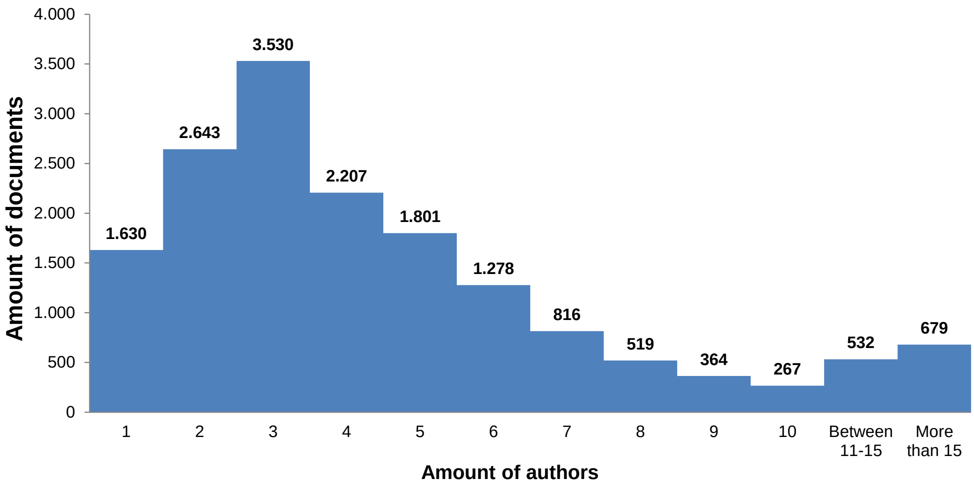
In Colombia, the increment of papers in collaboration, and un general of scientific papers, in the Web of Science and Scopus databases must be addressed under the connotations of the greater interest of the WoS to increase its coverture of regional topics, specially upon the increased competency posted by ElSevier with the Scopus database. Particulary since 2006 (see Figure 1).

Figure 1. Scientific documents from authors affiliated to Colombian institutions in the Web of Science, 2001-2010



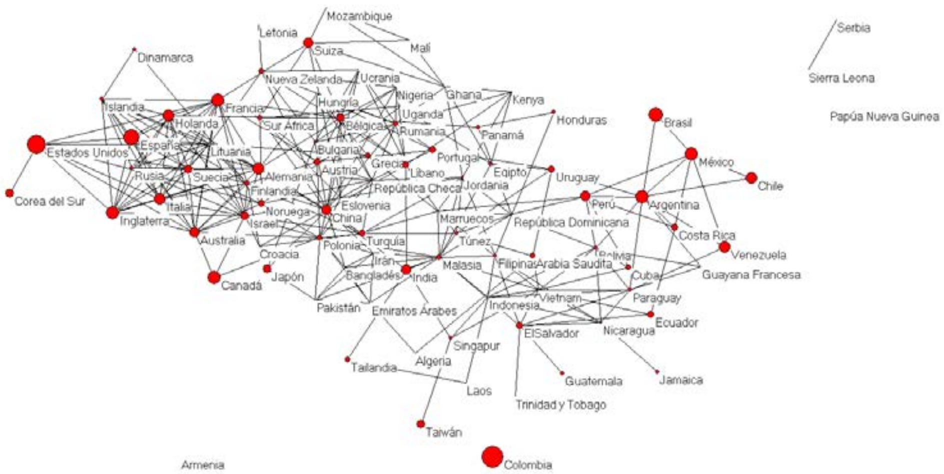
In fact, for this period, papers in international collaboration grew the most followed by national collaborations. Single authored papers presented the slowest growth rate for the aforementioned period. Most papers were written in collaboration between 2 and 4 authors, see Figure 2.

Figure 2. Number of authors per document in scientific documents from authors affiliated to Colombian institutions in the Web of Science, 2001-2010



As can be expected, in most collaborations researchers from the United States were participating followed closely by researchers from Spain (Figure 3).

Figure 3. Network representation of collaborations in Colombian Scientific Papers (2001-2010)



The flattened representations presented hereby say little about the collaborations conducted, the underlying practices, the benefits and implications for the parties involved. This is particularly worrisome since recent debates have argued that the political initiatives to increase collaboration without a further reflection on the different types of collaborations possible has meant a slow transition from research topics relevant to the context of developing countries to research in domains alien to the national realities but with better articulation to current international research topics (Arellano Hernández, Arvinitis & Vinck, 2012). These

debates have come to a point of forecasting the possible loss of national sovereignty in the definition of research agendas.
What we propose in the poster is a deeper look at research collaboration distinguishing between cognitive categories and using social network analysis to characterize collaborations, their durability in time and the impacts on research capacities in Colombia.

References

Arellano Hernández, A., Arvanitis, R. & Vinck, D. (2012). Global connexity and circulation of knowledge: Aspects of Anthropology of Knowledge in Latin America. *Revue d'anthropologie des connaissances*, 6(2).

Bordons, M. I., Gómez, I., Fernández, M. T, Zulueta, M. A., Méndez, A. (1996). Local, domestic and international scientific collaboration in biomedical research. *Scientometrics*, 37(2), 279-295.

Kreimer, P. (2006). Dependientes o integrados. La ciencia latinoamericana y la división internacional del trabajo. *Nómadas-CLACSO*, 24, 199-212.

Lee, S., & Bozeman, B. (2005). The Impact of Research Collaboration on Scientific Productivity. *History & Philosphy of Science*

Price, D. J. de Solla. (1963). *Little Science, Big Science*. New York: Columbia University Press.

Wagner, C. S., Leydesdorff, L. (2005). Network structure, self-organization and the growth of international collaboration in science. *Research Policy*, 34: 1608-1618.

Russian Universities Collaboration with Domestic and Foreign Funding Agencies[1]

Markusova V.A.*, Libkind A.N.*, Mindeli L.E.**, Noyons E.***

markusova@viniti.ru.
*All Russian Institute for Scientific and Technical Information of the RAS, Russia
Usievicha 20, Moscow 125190, Russia

** Institute for the Study of Science of the RAS. Russia
Butlerova 12, Moscow, 117485, Russia

*** Centre for Science and Technology Studies, the Netherlands
Wassenaarseweg 62A 2333AL Leiden Zuid Holland

Introduction

Russian science policy and Russian bibliometric performance were the subjects of many papers (Graham L., Wilson, 2004, Lewison & Markusova 2010). Reform of two main Russian research sectors, the Russian Academy of Sciences and the Higher Education Sector (HES) has been going on for the last ten years with the government shifting its attention and financial resources toward the HES. Bibliometric performance of the RAS and the HES played a very important role in this reform. Nowadays Russian government science policy is directed towards encouragement of competitive funding. The number of grants awarded to an organization is estimated as an indicator of economic performance.

In the last decade scholarly scientometrics journals have published significant amount of papers on acknowledgement analysis (Cronin 1993, Lewison 1995, Tiew 1999, Markusova, 2001, Wang & Shapira, 2011 and others). The goal of our empirical project was to give an overview of various funding agencies' (FA) activities supporting the HES; to identify leading universities by number of publications and level of research supported by FA; to examine universities' publications supported only by foreign FA and their subject category’s priorities.

Methods

The data for this study have been derived from Thomson Scientific resources: Science Citation Index-Expanded (SCI-E) from Web of Science (WoS) and Journal Citation Reports-Science Edition (JCR)-2010. All research documents (article, letter, note, and review) with at least one Russian address and indexed between 2009-2011 were downloaded with Thomson Scientific permission (download was performed in March 10, 2013). Publications were assigned to a country and Russian institutes based on the address which appears in a paper. Five percent of the records were excluded from analysis due to lack of data.

A more than 86,600 bibliographic records were downloaded from the SCI-E (AD=Russia and PY= 2009-2011). 18,500 records contained the information about FA support of HES. FA names were verified by special software and then checked manually. The result of verification

[1] This paper is a part of the project supported by grant 12-03-00070 and grant 14-03-00333 of the Russian Humanities Foundation

was a list that contained 1,090 FA names or organizations.

Bibliometric indicators: research output (RO) and its share supported by funding agency; RO distribution by university, subject category (SC), country; citation per paper; impact factor (IF); mean-weighted IF (MWIF); and research level (RL) of a university. Research level (RL) of a university's publications in a specific SC is a ratio between a mean-weighted IF of these publications and an aggregated IF of this SC by JCR

Results and Discussion

Total Russian RO for period 2009-2011 consisted of 86,737 records. Among 1.500 Russian universities, 467 universities contributed papers to WoS, and among them publications from 352 universities were supported by FA.

Table 1. Bibliometric statistics of Russian publications for 2009-2011

Research output of:	**2009**	**2010**	**2011**
Total Russia	29,097	27,945	29,689
Higher Education Sector (HES)	12,433	12,122	13,447
Share of HES in total Russia RO (%)	42.7	43.4	45.3
HES published in foreign journals	5,221	5,262	5,956
Share of RO published in foreign journals (%)	42.0	43.4	44.3
HES RO supported by FA	5,546	6,073	6,876
Share of RO HES supported by FA	44.6	50.1	51.1
Share of HES RO supported by Russian FA	87.9	88.3	90.2
Share of HES RO supported by Foreign FA (from total HES RO - %)	16.6	17.8	16.1
Share of HES supported by foreign FA (from HES RO supported by all FA - %)	37.3	40.2	34.0
HES RO supported by foreign FA and published in foreign journals	1,627	1,766	1851
Share of HES supported by foreign FA and published in foreign journals	78.6	82.0	85.4

A 6.5% growth of publications supported by FA was observed between 2009 and 2011. The share of HES RO supported by 119 Russian FA did not change and was very high (above 88.0%). About 35% of papers were supported by foreign FA, mainly in collaboration with Russian FA.

Our data revealed that the publications supported by the RFBR were published with the collaboration of 577 foreign FA; seventeen of them contributed no fewer than 150 papers. All these publications demonstrated a significantly higher citation scores per paper. The leading funding agency by RO collaborating with the RFBR was the German Research Foundation (DFG) (765 papers), followed by the NSF (409 papers) and U.S. Civilian Research and Development Foundation (378 papers)

Bibliometric indicators of fifteen leading universities by RO are presented in Table 2. .

Table 2. Bibliometric indicators of seventeen leading universities.

Columns: 1 – University's name; 2 – RO of university funded by FA; 3 – Share of funded RO (%); 4 – Citations share of funded RO (%); 5 - Number of citations per a paper of total university RO; 6 - Number of citations per a funded paper; 7– Mean weighted impact factor (MWIF) of total university RO; 8 - MWIF of funded RO;

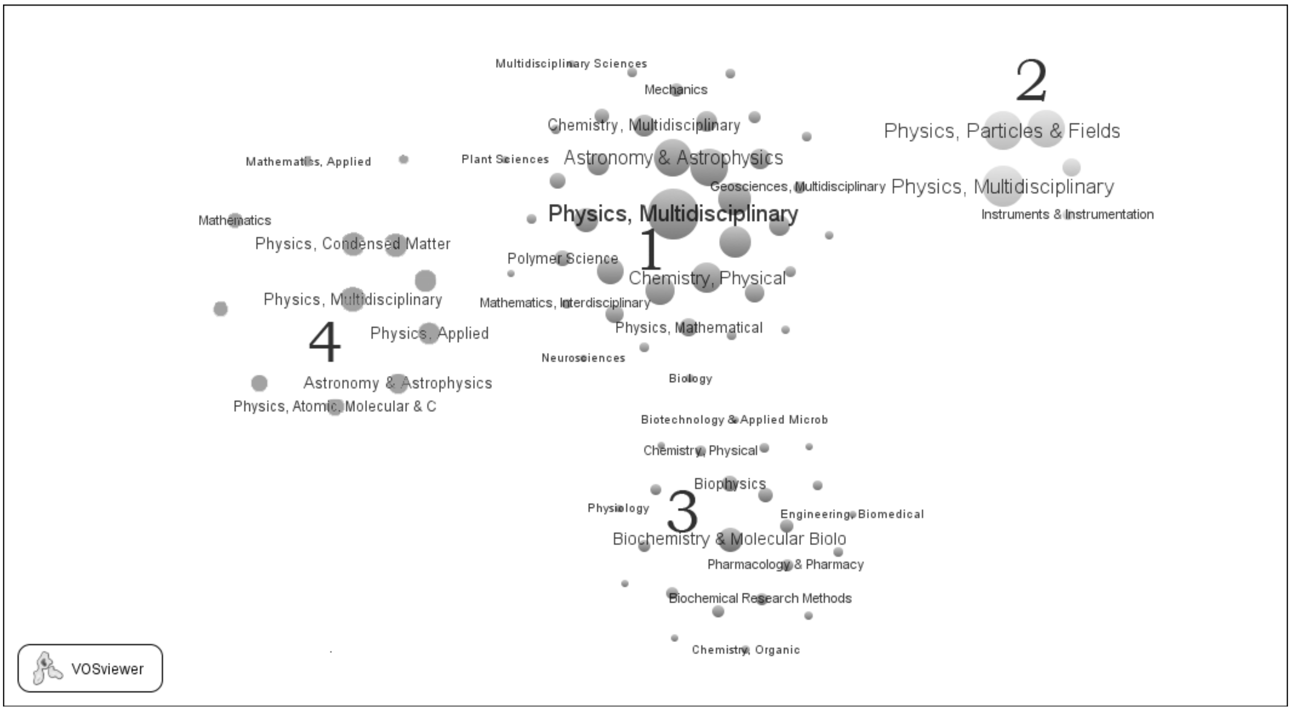
1	2	3	4	5	6	7	8
M.V. Lomonosov Moscow State University	6057	61,8	75,9	4,6	5,7	1,8	1,9
Saint Petersburg State University	1637	59,4	74,7	5,0	6,3	1,8	2,1
Novosibirsk State University	1085	69,3	73,9	3,9	4,1	1,8	1,9
Moscow Institute of Physics and Technology	582	66,6	76,5	3,6	4,2	1,7	1,9
B.N.Yeltsin Ural Federal University	477	54,5	70,5	2,8	3,7	1,2	1,4
Kazan (Volga Region) Federal University	460	59,7	75,1	3,7	4,7	1,7	2,0
N.I. Lobachevsky State University of Nizhniy Novgorod	437	67,5	75,2	2,3	2,6	1,2	1,3
Southern Federal University	435	52,3	68,5	2,7	3,5	1,3	1,5
Moscow Engineering Physics Institute (MEPhI)	421	47,7	77,4	8,9	14,5	2,0	2,8
Tomsk State University	367	59,2	68,1	2,3	2,6	1,0	1,2
Siberian Federal University	317	60,0	76,8	3,2	4,0	1,3	1,5
Saint Petersburg State Polytechnical University	309	48,7	72,3	4,3	6,4	1,6	2,2
N.G. Chernyshevsky Saratov State University	302	58,2	85,1	4,7	6,9	1,5	1,8
Voronezh State University	250	50,3	70,2	2,2	3,1	0,9	1,2
Tomsk Polytechnic University	243	51,1	72,1	2,8	3,9	1,1	1,4

Leading universities demonstrated a higher share of citations than share of funded papers. We want to emphasize that these universities citations shares for three years period (2009-2011) are significantly higher than citation shares of total Russian RO for 2008-2012 by InCites (48.04%). The Moscow State University and St. Petersburg State University occupy a special position in HES. As a consequence, there is a significant discrepancy in their total RO compared with the RO of other universities. However, by value of MWIF and Research level (RL) the first rank belongs to the Moscow Physics Engineering Institute - the National Nuclear Research University. MWIF of funded publications is slightly higher than MWIF of total university's RO.

To estimate impact of competitive funding on quality and quantity of universities, we randomly selected 85 universities located in 37 cities and 34 regions, which published at least 5 papers in WoS for the studied period. We discovered using Spearman correlation (r) that there is a significant correlation between share of papers funded by foreign FA and the MWIF of these universities' publications (r=+0.78). It was observed relatively strong influence of share of all funded papers on total university RO (r=+0.51). We found out a weak positive correlation (r=+0.006) between the share of teachers with a scientific degree and the MWIF of papers funded by all FA. The correlation between share of teachers with a scientific degree and MWIF of papers funded by foreign FA was a little bit higher (r=+0.025). Nevertheless we could assume that scientific degree does not have influence on teachers' choice to publish results in high impact journals. Our data show that paper supported by foreign FA has usually a few sponsors and a significant research team. Taking into consideration linguistic barrier it is obvious that foreign partners facilitate a Russian researcher's publication in foreign journal with high impact factor.

Analysis of 1,960 publications funded only by foreign FA allows us to identify disciplines, which attract foreign investment in Russian basic research. 606 foreign FA, located in 68 countries contributed to basic research in 183 Russian universities. Disciplinary priorities was focused on "hard sciences".The leading foreign FA was the German Research Foundation (224 papers) followed by the NSF USA (189 papers), European Commission (179), and NIH USA (115 papers).

Visualization of subject priorities by three foreign FA and one Russian was created using software VOSviewer http://www.vosviewer.com and presented at Fig.1. Cluster 1 belongs to German Research Foundation (DFG); cluster 2 to British Science and Technology Facilities Council (STFC); cluster 3 to the National Institutes of Health (NIH); and cluster 4 to private Russian foundation Zimin Dynasty Foundation. Each cluster of a SC contains no fewer than five publications. The highest number of publication in SC was 188. As we can see, three FA are heavily focused on "hard sciences" and NIH on life sciences.



Conclusions:

Short history of government science policy towards competitive funding has proved its positive impact of Russian research community. About 25% (357) of Russian universities received competitive funding from domestic and foreign funding agencies. It was observed 6.5 % growth between 2009 and 2011 in share of RO supported by FA. The study revealed an extensive collaborative network of Russians universities with foreign FA. About 10.6% of analyzed publications were supported only by foreign FA with disciplinary priorities focused on “hard sciences”.

Our data indicate that there is a good correlation by Spearman between the share of papers funded by foreign FA and mean-weighted impact factors (MWIF) of these universities' publications (r=+0.78). Despite a very substantial difference in RO of the Moscow State University and St. Petersburg State University compared with other universities, the highest value of MWIF and research level were demonstrated by the Moscow Physics Engineering Institute-the National Nuclear Research University.

Bibliometrics has become a very important tool in Russian government science policy. Our data demonstrate the impact of competitive funding on the Higher Education Sector research activity and provide a better empirical basis for science policy.

References

Graham L.R. What have we learned about science and technology from the Russian experience? Stanford University Press: Stanford, California, 1998.

Cronin, B., McKenzie, G. & Rubio, L. (1993) The Norms of Acknowledgement in four Humanities and Social Sciences Disciplines. *Journal of Documentation*, 49, 29-43.

Lewison, G., Jeshin, D. & Anderson J. (1995) A Bibliometric Database for Tracking Acknowledgements of Research Funding, *Proceedings of the Fifth Biennial Conference of the International Society for Scientometrics and Informetrics* (pp.255-264). Rosary College, River Forest, IL, USA.

Lewison, G. & Markusova, V. (2010) The Evaluation of Russian Cancer Research. *Research Evaluation*, 19, 2, 129-144.

Markusova, V.A. (2001) Communication Network of Russian Researchers: Acknowledgement Analysis of Russian Journals “Biochemistry” and “Geochemistry”. *Scientific and Technical Information*, part.1, 9, 21-28.

Tiew, W.S. & Sen B.K. (1999) Acknowledgement Patterns in Research Articles: A Bibliometric Study Based on Journal of Natural Rubber Research 1986-1997, *Proceedings of the Seventh Conference of the International Society for Scientometrics and Informetrics (*pp.488-497). Mexico: Colima University.

Wang, J. & Shapira, P. (2011) Funding Acknowledgement Analysis: an Enhanced Tool to Investigate Research Sponsorship Impacts: the Case of Nanotechnology. *Scientometrics*, 87, 563-569. (DOI 10.1007/s11192-011-0362-5).

Wilson, C. & Markusova, V.A. (2004) Changes in the Scientific Output of Russia from 1980-2000, as Reflected in the Science Citation Index in Relation to National Politico-economic Factors. *Scientometrics*, 59, 3, 345-389.

Metrics-Based Research Assessment[1]

Henk F. Moed *, Gali Halevi **

*h.moed@elsevier.com
Informetric Research Group, Elsevier, Radarweg 29, 1043 NX Amsterdam (The Netherlands)
**g.halevi@elsevier.com
Informetric Research Group, Elsevier, 360 Park Av. South, New York, NY 10011 (USA)

The multi-dimensional research assessment matrix

Introduction and base principles

When building a research assessment process, one has to decide which methodology should be used, which indicators to calculate, and which data to collect. Therefore, one should address the following key questions as their answers determine which methodology and types of indicators should be used. Each question relates to a particular aspect of the research assessment process.

- What is the unit of the assessment? A country, an institution, a research group, an individual, or a research field or an international network?
- Which dimension of the research process must be assessed? Scientific-scholarly impact? Social benefit? Multi-disciplinarity? Participation in international networks?
- What are the purpose and the objectives of the assessment? Allocate funding? Improve performance? Increase regional engagement? Which “meta assumptions” can be made on the state of the units of assessment?

The key principle is that the unit of assessment, the research dimension to be assessed, and the purposes of the assessment jointly determine the type of indicators to be used. An indicator may by highly useful within one assessment process, but less so in another. The aim of this paper is to further this principle by taking into account new bibliometric and non-bibliometric indicators, a series of aggregation levels, impact sub-dimensions, and by focusing on the objectives and policy background of the assessment.

Potential usefulness and limitations of 10 frequently used indicators

Table 1 summarizes the description of main types of indicators and gives some of the strong points and limitations of 7 publication- and citation-based indicators, a patent-based indicator and two altmetrics.

Units of assessment and the role of metrics in general

Table 2 presents the potentialities and limitations of the use of metrics for five units of assessment at different aggregation levels. Most limitations relate to the network structure among units of assessment, and underline that a particular unit must be viewed within the context of the network in which it takes a part. For instance, individual research papers are not isolated entities, but can be viewed as elements of publication oeuvres of research groups; citations to a single key paper may aim to acknowledge the total oeuvre (Moed, 2005).

[1] This article is a summary of a full research article by H.F Moed and G. Halevi, entitled “The Multidimensional Assessment of Scholarly Research Impact”, accepted for publication in the Journal of the Association for Information Science and Technology (JASIST) and available at http://arxiv.org/abs/1406.5520.

Table 1: Potentialities and Limitations of 8 Frequently Used Bibliometric and 2 Altmetrics Indicators.

Indicator	***Potentialities; strong points***	***Limitations***
Number of published articles	This is a useful tool to identify lagging research units if the metric's value is below a certain (subject field dependent) minimum	If numbers exceed a certain minimum level, differences between them cannot be interpreted in terms of performance
Number of citations	Useful for weighting individual publications. Reveals impact of the total collection of a research group's articles, disregarding how citations are distributed among cited articles	Depends upon subject field and age of (cited) publications. Depends upon the size of the group's publication volume
Citations per article	Reveals influence relative to size of publication volume	Strongly depends upon subject field and age of cited articles, and also upon type of document (e.g., normal article versus review).
Normalized citation rate	Takes into account type (e.g., review, full length article), subject field and age of cited article	Field delimitation must be sound. Should be used with special caution when comparing units with very different publication volumes or active in highly specialized subjects
Indicators based on Citation percentiles- (e.g., top 10 %)	Focuses on the most important publications; does not use the mean of (skewed) citation distributions; normalizes outliers	Maps all actual values onto a 0-100 scale; one may lose the sense of underlying absolute differences, and undervalue extraordinary cases
Journal impact factor and other journal metrics	The quality or impact of the journals in which a unit has published is a performance aspect in its own right	Journal metrics cannot be used as a surrogate of actual citation impact; impact factors are no predictors of the citation rate of individual papers
H-Index	Combines an assessment of both quantity (nr. papers) and impact (citations). Tends to be insensitive to highly cited outliers and to unimportant (uncited) articles	Its value is biased in favor of senior researchers compared to juniors; actual impact of the most cited papers hardly affects its value
Number of patents	Inventions may be disclosed in patents; patent data is available at a global level	Not all inventions are patentable or actually patented. The number of patents filed differs across countries because of legislation or culture, and also across subject fields
Full text article download counts	Are available almost immediately after publication; may reveal use or value that is not expressed in citations, impact upon scholarly audiences from other research domains or upon non-scholarly audiences	Downloaded articles may be selected according to their face value rather than their value perceived after reflection;
Mentions in social media	Are immediately available after publication; may reveal impact upon scholarly audiences from other research domains or upon non-scholarly audiences	Scientific-scholarly and societal impact are distinct concepts. One cannot measure scientific-scholarly impact with metrics based on social media mentions.

Researchers tend to operate in teams and therefore an assessment of their individual performance should take this into account. Non-bibliometric indicators may be used as a way to reflect more personal achievements, such as invitations for lectures at international conferences or at seminars in prestigious institutions. Universities in countries with a strong research infrastructure outside the university system, tend to gain less visibility in international university rankings than universities in countries in which research is mainly concentrated in the academic sector.

Table 2: Main Units of Assessment and the Role of Metrics.

Unit of Assessment	*Metrics Potentialities*	*Metrics Limitations*
Individual article	Metrics reveal differences in significance between articles and may identify key articles	Individual articles are not isolated entities but rather elements of publication oeuvres; different types of articles exist.
Individual author	Metrics reveal differences in impact between individuals	Most research articles are the result of team work and are multi-authored. How do we then assess the role of an individual in a team?
Research group	The research group is the core research entity, at least in science	Social sciences and humanities do not always show a group structure as in science
Research Institution	Metrics show status and impact of research institutions	Institutions may specialize or be more general, and have specific functions in a national research system; large differences may exist within institutions
Country	Metrics unravel the structure of national research systems	Aggregate data may conceal differences between a country's research institutions

Research dimensions and its principal indicators

The variety of impact dimensions is presented in Table 3 which distinguishes the various types of research impact, and gives typical examples of indicators that may be used to assess these. The two main categories are scientific-scholarly and societal impact. The term 'societal' embraces a wide spectrum of aspects outside the domain of science and scholarship itself, including technological, social, economic, environmental, and cultural aspects. The list of indicators includes the 10 metrics that are given special attention in this paper, and also a number of other indicators, partly derived from the AUBR Report (2010), but it does not claim to be fully comprehensive.

A distinction can be made between purpose and objective of an assessment. A purpose has a more general nature, and tends to be grounded in general notions (e.g., "increase research performance"), whereas objectives are more specific, more formulated in operational terms (e.g., "stimulate international publishing"). Objectives are grounded in assumptions on how they are relate to the general purpose (e.g., "it is assumed that by stimulating international publishing, research performance increases, at least at the longer run").

The policy relevance of both assessment purposes and objectives follows from what may be termed as a "meta assumption" on the state of the units of assessment, which in turn, is based on a Meta-analysis of these units. For instance, "stimulating international publishing" as an objective in a national research assessment exercise makes sense from a policy viewpoint only if there are good reasons to believe that the level of international publishing among a country's researchers is relatively low compared to their international counterparts. Similarly, assessing whether an academic staff member is "research active" or not, seems to make sense only of there is evidence that a non-negligible part of staff hardly carries out research.

Table 3: Types of Research Impact and Indicators

Type of impact	Short Description; Typical examples	Indicators (examples)
Scientific-scholarly or academic		
Knowledge growth	Contribution to scientific-scholarly progress: creation of new scientific knowledge	Indicators based on publications and citations in peer-reviewed journals and books
Research networks	Integration in (inter)national scientific-scholarly networks and research teams	(inter)national collaborations including co-authorships; participation in emerging topics
Publication outlets	Effectiveness of publication strategies; visibility and quality of used publication outlets	Journal impact factors and other journal metrics; diversity of used outlets;
Societal		
Social	Stimulating new approaches to social issues; informing public debate and improve policy-making; providing external users with useful knowledge; Improving people's health and quality of life; Improvements in environment and lifestyle;	▪ Citations in medical guidelines or policy documents to research articles ▪ Funding received from end-users ▪ End-user esteem (e.g., appointments in (inter)national organizations, advisory committees) ▪ Juried selection of artworks for exhibitions ▪ Mentions of research work in social media
Technological	Creation of new technologies (products and services) or enhancement of existing ones based on scientific research	Citations in patents to the scientific literature (journal articles)
Economic	Improved productivity; adding to economic growth and wealth creation; enhancing the skills base; increased innovation capability and global competitiveness; uptake of recycling techniques;	▪ Revenues created from the commercialization of research generated intellectual property (IP) ▪ Number patents, licenses, spin-offs ▪ Number of PhD and equivalent research doctorates ▪ Employability of PhD graduates
Cultural	Supporting greater understanding of where we have come from, and who and what we are; bringing new ideas and new modes of experience to the nation.	▪ Media (e.g. TV) performances ▪ Essays on scientific achievements in newspapers and weeklies ▪ Mentions of research work in social media

"International publishing" may relate to the level of the quality criteria applied by editors and referees in the review of submitted manuscripts, or to the geographical location of authors, members of the editorial or referee board, and/or readers of a journal. The following definition would include both dimensions: international publishing is publishing in outlets that have: (1) rigorous, high-standard manuscript peer review; and (2) international publishing and reading audiences.

Bibliometric studies found that the journal impact factor is a proxy of a journal's international status. For instance, Sugimuto et al. (2013) reported that acceptance rates of manuscripts submitted to scientific journals negatively correlate with the journals' impact factors, suggesting that journals with rigorous referee systems tend to generate higher impact than others. If an analysis of the state of a country's science concludes that a substantial group of researchers tends to publish predominantly in national journals that are hardly read outside the country's borders and do not have severe rigorous peer review, it is in the view of the authors of this paper, defendable to use the number of publications in the top quartile of journals according to citation impact as an indicator of research performance. In this manner one is able to discriminate between those researchers whose research quality is sufficiently high to publish in international, peer reviewed journals, and those who are less capable of doing so. This issue is further discussed in Section 2.

But if in internationally oriented, leading universities one has to assess candidates submitting their job application, it is questionable whether it makes sense comparing them according to the average citation impact of the journals in which they published their papers, using journal impact factors or other journal indicators. Due to self-selection, the applicants will probably publish at least a large part of the papers in good, international journals. Other characteristics of the published articles, especially their actual citation impact, are probably more informative as to the candidates' past research performance and future potential than indicators based on journal metrics are.

A second example relates to the use of publication counts. In order to identify academic staff that is not research active, it is reasonable to consider the publication output of the staff under assessment, and identify those whose output is below a certain – subject field dependent – minimum. But if one has to assess candidates submitting their job application to a leading research university, it hardly makes sense to compare them according to their publication counts. Due to self-selection, they will probably all meet a minimum threshold. In other words, while there are good reasons to believe that journal metrics or publication counts are appropriate indicators to identify the bottom of the quality distribution of research staff, they have a limited value if one aims to discriminate in the top of that distribution.

These examples illustrate that the choice of indicators depends not only upon the overall purpose of the assessment, but also upon the specific objectives, and on the Meta view on the state of the units of assessment. These factors are best be characterized by the term "policy context". Therefore, the conclusion is that the selection of indicators in an assessment depends upon the unit of assessment, the research aspect to be assessed, and very much on its policy context.

Discussion and conclusions

Meta-analysis

It was stated that a meta-analysis of the "state of the units of assessment" determines the methodology and indicators to be applied in an assessment process. It must be noted that bibliometric indicators and other science metrics may – and actually do - play an important role in the empirical foundation of such a Meta view. Metrics are essential tools on two levels: in the assessment process itself, and on the Meta level aimed to shape that process. Yet, their function in these two levels is different. In the first they are tools in the assessment of a particular unit, e.g., a particular individual researcher, or department, and may provide one of the foundations of evaluative statements about such a unit. At the second level they provide insight into the functionality of a research system as a whole, and help draw general conclusions about its state assisting in drafting policy conclusions regarding the overall objective and general set-up of an assessment process.

A Meta level analysis can also provide a clue as to how peer review and quantitative approaches might be combined. For instance, the complexity of finding appropriate peers to assess all research groups in a broad science discipline in a national research assessment exercise may urge the organizers of that exercise to carry out a bibliometric study first and decide on the basis of its outcomes in which specialized fields or for which groups a thorough peer assessment seems necessary. One important element of the Meta-analysis is a systematic investigation of the effects of the assessment process, both the intended and the unintended ones.

Statistical considerations
The observation that the usefulness of journal impact factors and publications counts so strongly depends upon a meta view of the units to be assessed, can also be grounded in statistical considerations. If in a particular study a positive (linear or rank) correlation is found to hold between two variables, it does not follow that it holds for all sub-ranges of values of the variables. Whether or not a sample of the two variables can be expected to correlate in a particular study, very much depends upon the range of values obtained by the units in the sample.

For instance, Sugimoto et al. (2013) examined the relationship between journal manuscript acceptance rates and 5-year journal impact factors, and found in a sample of 1,325 journals a statistically significant linear correlation coefficient between these two measures. But, most importantly, the study also found that, when dividing journals into quartiles according to their acceptance rates and analyzing correlation coefficients within quartiles, the correlation coefficients between acceptance rates and impact factors were much lower and not significant. This shows that the application of journal metrics or publication counts to assess the comparative performance of researchers who publish on a regular basis in international journals cannot be sufficiently justified by referring merely to earlier studies reporting on observed positive correlation between these measures and peer ratings of research performance. What is not defendable in the view of the authors is the use of such indicators simply because they are relatively easy to calculate and readily available.

The authors of this paper share the critique of the use of journal metrics in the assessment of individual researchers. Indeed, it does not make sense to discriminate in a group of research active researchers publishing in good journals between high and low performers on the basis of weighted impact factors of the journals in which they published their articles. On the other hand, it does not follow that the use of this type of indicator is invalid under all circumstances.

Policy considerations
Research assessments methodologies cannot be introduced in practice at any point in time, and do not have eternal lives. In the previous section it was stated that under certain conditions it is defendable to use publication counts and journal metrics as one of the sources of information in individual assessments. But one may argue that it is fair to maintain a time delay of several years between the moment it is decided to use a particular assessment method or indicator on the one hand, and the time at which it is actually used, on the other. In this

way, the researchers under assessment have the opportunity to change their publication behavior – to the extent that they are capable of doing that.

In the view of the authors of this paper it is wise to change an assessment method radically every 5 to 10 years. Two considerations may lead to such a decision. First, a meta-analysis may reveal that the overall state of the units of assessment has changed in such a manner, that the old methodology is either irrelevant or invalid. Secondly, any use of assessment methodologies and indicators must be thoroughly monitored in terms of its effects, especially the unintended ones. Severe negative effects such as manipulation of metrics may lead to the decision to abandon a method, and establish a new one, even though bibliometric can to some extent detect and correct for such behavior (Reedijk & Moed, 2008).

What is an acceptable "error rate"?

Regarding the – either negative or positive – effects of the use of metrics or any other methodology in research assessment, one may distinguish two points of view. One may focus on its consequences for an individual entity, such as an individual scholar, a research group or institution, or on the effects it has upon scholarly activity and progress in general. A methodology, even if it provides invalid outcomes in individual cases, may be beneficial to the scholarly system as a whole. Cole and Cole expressed this notion several decades ago in their study of chance and consensus in peer review of proposals submitted to the National science Foundation (Cole, Cole & Simon, 1981).

Each methodology has its strengths and limitations, and is associated with a certain risk of arriving at invalid outcomes. As Martin (1996) pointed out, this is true not only for metrics but also for peer review. It is the task of members from the scholarly community and the domain of research policy, and not of the authors to decide what are acceptable "error rates" and whether its benefits prevail, based on a notion of what is a fair assessment process. Bibliometricians and other analysts of science and technology should provide insight into the uses and limits of the various types of metrics, in order to help scholars and policy makers to carry out such a delicate task.

References

AUBR Expert Group (2010). Expert Group on the Assessment of University-Based Research. Assessing Europe's University-Based Research. European Commission – DG Research. EUR 24187 EN. Retrieved November 25, 2013.

Cole, S., Cole, J. R., & Simon, G. A. (1981). Chance and consensus in peer review. Science, 214(4523), 881-886.

Martin, B.R. (1996). The use of multiple indicators in the assessment of basic research. Scientometrics, 36 (3), 343-362.

Moed, H. F. (2005).Citation analysis in research evaluation. Dordrecht (Netherlands): Springer. ISBN 1-4020-3713-9, 346 pp.

Reedijk, J., Moed, H.F. (2008). Is the impact of journal impact factors decreasing? Journal of Documentation 64, 183-192.

Sugimoto, C. R., Larivière, V., Ni, C., & Cronin, B. (2013). Journal acceptance rates: A cross-disciplinary analysis of variability and relationships with journal measures. Journal of Informetrics, 7(4), 897-906.

Towards an alternative framework for the evaluation of translational research initiatives[1]

Jordi Molas-Gallart*, Pablo D'Este*, Óscar Llopis** and Ismael Rafols*, ***

*jormoga@ingenio.upv.es, pabdescu@ingenio.upv.es, i.rafols@ingenio.upv.es
INGENIO (CSIC-UPV), Universitat Politècnica de València, Spain

**osllocor@upvnet.upv.es
Université de Bordeaux, France

***i.rafols@sussex.ac.uk
[3]SPRU, Science and Technology Policy Research, University of Sussex, Brighton, UK

Introduction

The pathways between basic science and clinical practice and health outcomes are multifaceted and complex. The analysis of these pathways has become of interest to the biomedical research community and public health agencies. Researchers and funding agencies are concerned with the ways in which scientific breakthroughs and evidence-based clinical findings are converted into practices with beneficial health impacts, including, but not limited to, therapies and medical guidelines. This interest is largely driven by the perception that many promising results from basic science in biomedicine have not systematically contributed to medical treatments and, ultimately, health care improvements. In response, a wide range of publicly-funded initiatives have been set up with the aim to address this problem. As the main aim of these initiatives is to facilitate the "translation" of scientific discoveries into beneficial applications and practices, many of these initiatives have been branded as "Translational Research" (TR).

Translational Research has become a very popular term applied for instance, to large research programmes, research activities and, even, academic journals. Consequently, it has been the subject of fast growing interest, mainly from biomedical scholars and institutions. Often the more popular a policy concept, the more ambiguous it becomes. This has clearly been the case with Translational Research. A debate has emerged about the models of research that are to be considered "translational" and the nature and characteristics of a putative TR discipline. Consequently, the ways in which TR should be analysed, and more specifically the approaches to the evaluation of TR programmes are also the subject of debate.

Models for the assessment of translational research

In a context of ambiguity about the type of activities to be considered as TR, evaluation approaches and practices can play an important role in determining what actions and outcomes are conceived in practice to be relevant and significant, and in so doing shaping the future nature of TR initiatives. This paper discusses the dominant approaches to TR

[1] IR acknowledges support from the US National Science Foundation (Award #1064146 - "Revealing Innovation Pathways: Hybrid Science Maps for Technology Assessment and Foresight"). The findings and observations contained in this paper are those of the authors and do not necessarily reflect the views of the National Science Foundation.

evaluation and proposes an alternative evaluation framework, which would have implications both for TR evaluation processes and for the future shaping of TR programmes.

A dominant approach is to focus on outputs generated at different points of the "translational research continuum" and when they are achieved. A focus on the "what" and "when" implies a TR evaluation approach that attempts to identify results and how these differ from what would have been achieved in the absence of the initiatives under assessment. It needs to be emphasized that this focus on outputs may be derived from an explicit view of TR that sees it as addressing "translational gaps" along a "translational research continuum", or may emerge without an explicit "theory" of the processes and objectives of TR. Research is measured against success criteria revolving around the generation of outputs that are no different from those that may have been generated in a traditional research context, and this may be occurring in the absence of an explicit programme theory. Note that, in this case, the TR objectives may be defined by the evaluation strategy chosen.

Our alternative is to focus, instead, on the "how", on the processes of collaboration and exchange that can be attributed to TR initiatives. To this end we develop an alternative TR evaluation framework that focuses on understanding the processes of change and their outputs across the divides that hinder the application of the capabilities and knowledge generated by basic biomedicine to health care. The extant literature attributes the low level of practical application of biomedical research to a variety of causes, including the divide between the interests and skills of basic scientists on the one hand and clinical scientists on the other, the growing difficulties of communication among both fields as biomedical research becomes more complex and specialized, and the existence of institutional barriers.

One perspective on the assessment of TR assumes that the key indicator to assess TR initiatives is the time it takes for the different translational gaps to be bridged and, therefore, for research to be translated into treatments and other measures improving health. Such time lag is also the indicator taken by Trochim and his colleagues (2011) when developing a generic evaluation model that could provide the basis for a shared approach to TR evaluation. They propose a flexible solution focusing on what they view as the final TR objective: the reduction of the time it takes to develop new clinical practices and drugs that reach patients.

Following a generic linear TR model, they propose to identify "markers" in the translation process and assess the time that it takes for outputs to move across markers. There is flexibility in the identification of such markers, and therefore there is no need to adopt beforehand one model of translational research instead of another. There is also flexibility in the direction of the activity across markers, allowing for both "bench to bed", and "bed to bench" directions. Yet, the approach focuses on the outputs of TR and on the time it takes for the output of a specific activity to be translated into a different type of output identified in another marker. In other words, this form of evaluation is concerned by TR outputs rather than the way in which such outputs are achieved.

The alternative is to focus on how TR programmes affect the way in which research objectives are defined, research is conducted, and its results applied in practice. We can assert that TR initiatives attempt to address problems in the organization and management of biomedical research by bridging the divide between different actors involved in the development of new drugs, therapies, diagnostics or public health practices. The different groups include, for instance, doctors and patients involved in the identification and definition of therapeutic and health problems, researchers defining and addressing relevant fundamental research challenges, and clinicians and doctors developing and testing solutions. The separation among these actors takes various forms: the different groups belong to different organizations, follow different implicit and explicit rules, and respond to different sets of incentives and performance criteria. These conflicting logics can make it difficult to align the objectives among the parties, and to establish clear and fluid lines of communication. This type of separation results in a difficulty to communicate needs and results across communities separated by institutional and organisational boundaries.

Five dimensions of proximity

TR initiatives can then seek to reduce some of the divides among biomedical innovation actors. TR would then take place in networks of diverse actors, such a basic research, clinical doctors, general practitioners, regulators, etcetera. It is important to emphasise the networked nature of the social interactions: basic research, for instance, can be influenced by insights from general practitioners and from regulators, without the mediation of clinical doctors.

We propose that these interactions are less than optimal because the distances that separate these different groups make the interactions difficult. Following Boschma (2005), we can state that learning processes and knowledge exchange interactions are facilitated and strengthened by five forms of proximity: **cognitive, social, organisational, institutional and spatial**.

A degree of c*ognitive* proximity - i.e. the extent to which actors share a similar knowledge base - is a prerequisite for interactive learning, as it facilitates effective communication and a common reference space to process and transfer complex information and knowledge. However, as pointed out by Nooteboom (2000) and Boschma (2005), both too much and too little cognitive proximity can be detrimental to innovation and learning processes. A high degree of cognitive proximity between actors may lead to the exchange of irrelevant, redundant information due to a lack of variety of the knowledge sources; while too little cognitive diversity may lead to information exchange that cannot be adequately understood by the interacting actors, rendering communication ineffective.

Social proximity refers to relations between actors generally built on common experience, friendship and kinship and which can improve communication. *Organisational* proximity refers to the governance structure shaping interactions between actors. High organisational proximity is often associated with a hierarchical structure governing the interactions between actors, while low organisational proximity is generally associated with flat governance structures or arms' length interactions between actors. *Institutional* proximity refers to the norms, rules and values that influence how actors behave; a large institutional distance may impose serious impediments to fruitful learning interactions if the behaviour of interacting actors responds to different, potentially conflicting, sets of incentives or values. For example, universities and firms have considerable institutional distance because their incentives and norms differ significantly. Finally, *geographical* proximity refers to the spatial or physical distance between actors. This matters in knowledge dynamics because spatial co-location favours the exchange of knowledge that is complex or difficult to transfer (i.e. tacit knowledge).

The operationalisation of these dimensions in terms of quantitative indicators will be addressed in future studies. While not yet developed in this paper, we think that scientometric and social network approaches will be useful to track these proximities. For example, geographical and organisational proximities can be derived from the affiliations of publications, and some cognitive proximities from co-word or co-citation analyses.

All these types of proximity are inter-related. Some may complement each other, while others may act as substitutes. For instance, Howells (2002) argues that geographical proximity facilitates face-to-face interactions, favouring trust-based relationships and knowledge exchange, suggesting a reinforcing effect of spatial proximity on social proximity. In contrast, some proximity dimensions may substitute each other: barriers for knowledge exchange through large geographical distances (spatial distance) might be overcome if interacting partners share a well-defined and honed division of labour (i.e. organisational proximity).

Coming back to TR initiatives, these can explicitly or implicitly address perceived distance problems along one or more of these analytical dimensions. They can for instance establish ways to improve communication and understanding between patients, clinicians and researchers (addressing cognitive distance), they may try to establish better coordination across different organisations involved in the research and application process (addressing organisational distance), align their incentives rules and norms (addressing institutional distance), or improve trust (addressing social distance). In other words, TR initiatives can be described as aiming to bridge the gaps among the actors involved in biomedical research and the application of its results by directly reducing the distance among the actors in one or more of the five analytical dimensions.

How to think of proximities in evaluation

The focus on processes that underpins the evaluation approach we suggest here is based on the postulate that to understand the effect of TR initiatives we need to learn about how they affect the ways in which research, its objectives and the application of its results are designed and conducted. An evaluation strategy that focuses only on measuring outputs cannot offer information on how the initiative under evaluation has contributed to the observed outputs. When, as it is the case with TR, there is ambiguity about what differentiates this from of research from other forms of research, the need to understand how interventions operate in practice and what processes they trigger is particularly important. We have explored in this paper an avenue to develop a process-based approach to the evaluation of TR initiatives.

Evaluation frameworks are not neutral in relation to the objectives of an initiative. The way in which a project is evaluated will affect how it is conducted and, at least, part of the objectives that the performers will be aiming at. Focusing on specific outputs can implicitly suggest an intervention rationale that is not concerned about the organisation of research, and the way in which specific “translational gaps” are addressed. The proximities framework we are proposing can help focus attention on the way research is conducted and the specific aspects that an initiative is intended to address. These aspects may be explicit in the definition of the intervention, but they can also be implicit in the way the initiative is implemented. In the latter case, the framework can also be used to explore and develop a “programme theory” for a TR; that is to explore its rationale. The cases above show how we can use the framework to describe both the goals of TR initiatives and the way such goals are expected to be attained.

By adopting this approach we are proposing that the immediate goal of TR initiatives is to address a problem of distance separating different groups involved in the TR process. The “translational gaps” appear because of excessive distance in one or more significant dimensions. The groups involve in the translational process have cognitive differences, are institutionally separated and, therefore, follow different rules, face different types of incentives, and they are often geographically dispersed. Yet, some flexibility must be built into the definition of an initiative and its evaluation to reflect the fact that increased proximity will not always be desirable. For instance, cognitive distance can pose a problem but the same can be said of the overlaps generated by excessive cognitive overlap; cognitive proximity will be positive only up to a certain extent. A specific programme theory will need to reflect this problem and the interpretation of evaluation results will have to be sensitive to this potential problem if there is a possibility that it may become relevant.

The programme theory of a TR initiative will define the expectations about whether and how changes in proximity in one or more dimensions caused by the intervention will trigger shifts in the other dimensions, and the effects of these changes on the development and application of beneficial goods and services. These effects will be mediated by changes in the way in which research is carried out. Increased proximity can result in increased collaboration among groups involved in the different tasks that constitute the TR process (the definition of fundamental and clinical research objectives, research, and the application of its results). We can expect changes in proximity to generate new interactions across groups, like for instance, between research performers and the diverse users and beneficiaries of the research results, where knowledge is moving back and forth through various channels, not in the linear bedside to bench continuum but within networks.

We can thus define further building blocks of a TR programme theory. An intermediate outcome of increased proximities can be the generation of complex interactions among different groups that become partners in a single TR process. An analysis of intermediate outcomes in terms of interactions among the participants in the TR process needs to consider the variety of actors directly involved and affected by a TR initiative. Although this may vary across initiatives, it is important to take into account that there is a broad variety of potential stakeholders: basic researchers, clinical researchers, technologists, practitioners (doctors, nurses,…), public health and private industry managers, patients. The ways in which stakeholder groups interact can be traced and analysed using instruments developed for the evaluation of the socio-economic impact of research, like for instance those developed by the EU-funded SIAMPI project (Molas-Gallart & Tang, 2011), which focus on the processes of collaboration that can be linked to an initiative.

Our framework does not determine the research techniques to be employed; these will need to fit the specific circumstances of each initiative under assessment. The activities supported by a TR initiative will be different, implemented against different contexts and having different targets and objectives. For instance, the research techniques applied to an initiative that focuses mainly on cognitive issues, will be different from those applied to one that addresses institutional differences.

Finally, as the adequacy of a specific research technique will depend on the specific TR evaluation problem confronted and its context, it follows that the outputs of TR evaluations will not, and should not, be directly comparable in terms of either failure or success. Calls for an approach that will be based on a single set of research techniques yielding measurable and comparable indicators of TR "output" are, from the perspective we are developing, out of place. Indicators aimed at capturing each of the dimensions discussed will need to be tailored to the goals and contexts of each specific TR context. An evaluation approach that focuses on processes will aim at providing detailed information of the effects of an initiative starting at the level of those groups directly involved in it. But the way in which this information is shaped, and the indicators on which it is based will depend on the type of initiative, its objectives and the types of proximities the programme is designed to address.

The full working paper is available here:
http://www.ingenio.upv.es/sites/default/files/working-paper/2014-03.pdf

References

Boschma, R. (2005). Proximity and innovation: a critical assessment. *Regional studies*, 39(1): 61–74.

Howells, J. R. (2002). Tacit knowledge, innovation and economic geography. *Urban studies*, 39(5-6): 871–884.

Molas-Gallart, J., & Tang, P. (2011). Tracing “productive interactions” to identify social impacts: an example from the social sciences. *Research Evaluation*, 20(3): 219–226.

Nooteboom, B. (2000). *Learning and innovation in organizations and economies*. Oxford: University Press.

Trochim, W., Kane, C., Graham, M. J., & Pincus, H. A. (2011). Evaluating Translational Research: A Process Marker Model. *Clinical and Translational Science*, 4(3): 153–162.

Buckland, M. & Gey, F. (1994). The relationship between recall and precision. *Journal of the American Society for Information Science*, 45, 12-19.

The consequences of retractions for co-authors: scientific fraud and error in biomedicine.

Philippe Mongeon * and Vincent Larivière**

*philippe.mongeon@umontreal.ca; ** vincent.lariviere@umontreal.ca
École de bibliothéconomie et des sciences de l'information, Université de Montréal, C.P. 6128, Succ. Centre-Ville, Montréal, QC. H3C 3J7 (Canada)

**Observatoire des sciences et des technologies (OST), Centre interuniversitaire de recherche sur la science et la technologie (CIRST), Université du Québec à Montréal
CP 8888, Succ. Centre-Ville, Montréal, QC. H3C 3P8, (Canada)

Abstract
In the last decade, major cases of scientific fraud (e.g. Hendrik Schön, Diedrick Stapel, Eric Poehlman and Yoshitaka Fujii) have shocked the scientific community. Such frauds account for more than half of the publications retracted from the scientific literature, which have increased tremendously in the past few years. In the biomedical field, fraud can have consequences not only for the research community, but also for the public. It is a serious deviance from the norms of science, and it most likely ends the career of researchers who get caught doing it. However, researchers rarely work alone, and some of the consequences are presumably shared by their co-authors, although no empirical evidence of this has been provided so far. To evaluate the nature and extent of these shared consequences, we measured the productivity, impact and collaboration of authors who retracted papers between 1996 and 2006. We divided authors in groups according to their rank on the retracted papers' authors list and the cause of retraction (fraud or error) and compared the results for each group to those of a randomly selected control group. We found that retractions do have consequences for the career of co-authors, mostly in terms of scientific output, which are more important in cases of fraud than errors. Furthermore, first authors are generally affected more strongly by retractions than the other co-authors of the retracted publications.

Introduction
The number of retractions has skyrocketed in the last few years (Cokol, Ozbay, & Rodriguez-Esteban, 2008; Steen, 2011), mostly in the biomedical field (Grieneisen & Zhang, 2012) going from 20 retractions a year during the 90s to more than 500 in 2012 and in 2013. According to Fang, Steen and Casadevall (2012), scientific fraud (data fabrication, data falsification and plagiarism) accounts for more than half of those retractions. Previous research has mostly focused on the rise of retractions (Cokol et al., 2008; Steen, 2011), it's causes (Fang et al., 2012; Steen, Casadevall, & Fang, 2013), the ongoing citations of retracted papers (Furman, Jensen, & Murray, 2012; A. Neale, Northrup, Dailey, Marks, & Abrams, 2007; A. V. Neale, Dailey, & Abrams, 2010; Pfeifer & Snodgrass, 1990). Others have investigated and discussed the prevalence of scientific fraud (Fanelli, 2009; Sovacool, 2008; Steen, 2011), ways to prevent, detect and act upon it (Steneck, 2006), and its potential consequences for science in general and for the public (Steen, 2012). A few studies have looked at the consequences of fraud within disciplines (e.g. Azoulay, Furman, Krieger, & Murray, 2012) and within research teams (e.g. Jin, Jones, Lu, & Uzzi, 2013).

A researcher found guilty of fraud will most likely see his scientific career decline, or even come to an end. However, researchers rarely work alone, as science is becoming more and more collaborative (Wuchty, Jones, & Uzzi, 2007); a long lasting trend that is observed in almost all disciplines. Authorship confers symbolic capital as well as responsibility (Biagioli, 1999), but defining who did what and who is responsible for specific parts of the work is made more complex by this collaborative context (Biagioli, 1998; Cronin, 2001). Furthermore, the coexistence of these two trends (the increase of retractions and collaboration) may result in an exponential increase of the researchers with a retraction in their record. This brings into light the importance of investigating how the consequences of scientific fraud are shared by co-authors. Indeed, it is assumed that other authors of the fraudulent article also suffer collateral effects of the retraction (Bonetta, 2006), but no research has yet provided empirical data giving a complete account of these shared consequences.

Retractions can occur for different reasons, the most common being fraud or error. While fraud is an serious deviation from the core values and the purpose of science, there is a general agreement that honest mistakes are normal in the course of science, and that they “must be seen not as sources of embarrassment or failure, but rather as opportunities for learning and improvement” (Nath, Marcus, & Druss, 2006). Therefore, we would expect retractions for fraud to have more impact on researchers’ careers than retractions for error. Also, the specific contribution of authors to a specific paper is reflected in the order by which authors are listed. In the biomedical field, this distribution is typically U-shaped (Pontille, 2004) meaning that the first and last authors are supposedly those who contributed the most to the work, and thus receive more credit for it. Last authors are also typically senior researchers with tenure that are managing research laboratories, which puts them into a less precarious position than first authors, who are typically PhD Students, post-docs or junior researchers. This is reflected in the results of a study by Jin, Jones, Lu and Uzzi (2013), who showed that fraud has less impact on future citations of eminent co-authors. We would, thus expect the effect of a retraction to vary according to the researchers’ rank in the list of authors of the retracted paper.

In this study, we measured the pre- and post-retraction productivity, scientific impact and collaboration of all the co-authors of papers retracted in PubMed between 1996 and 2006, in order to provide answers to the following questions: Do retractions have an impact on the co-authors in terms of productivity, scientific impact, and collaboration? If so, how does this impact varies according to the retraction cause (fraud vs error), and according to the author’s rank in the retracted paper’s authors list?

Methods

Retractions sample

We used PubMed to gather all publications that were retracted between 1996 and 2006, which were then found in the Web of Science for further analysis, keeping only those published in biomedical and clinical medicine journals (n = 443). Using data from Azoulay et al. (2012) we identified the articles that were retracted for fraud (n = 179) or error (n = 114) co-authored by a total of 1,098 researchers.

We then created a control group by randomly selecting, for each of the 443 articles retracted between 1996 and 2006, a non–retracted article with the same number of authors, published in the same issue of the same journal. This provided us with a list of 1,862 distinct authors.

Co-authors sample
Using data by Azoulay et al. (2012) or looking at the retraction notices, we found 79 authors who were identified as responsible for 159 of the 179 fraud cases. The 66 distinct authors of the remaining 20 fraud cases were excluded from the sample in order to ensure that no fraudulent researchers remained. We also excluded of our sample 3 authors who were identified as responsible for 5 cases of error.

Finally, we divided the authors in three groups (first, middle, and last authors) according to their rank in the authors list of the retracted papers. Table 1 shows the distribution of authors within each group.

Table 1. Sample of authors.

	Fraud	Error	Control	Total
First authors	45	108	411	564
Middle authors	346	366	1,046	1,758
Last authors	77	102	405	584
Total	468	576	1,862	2,906

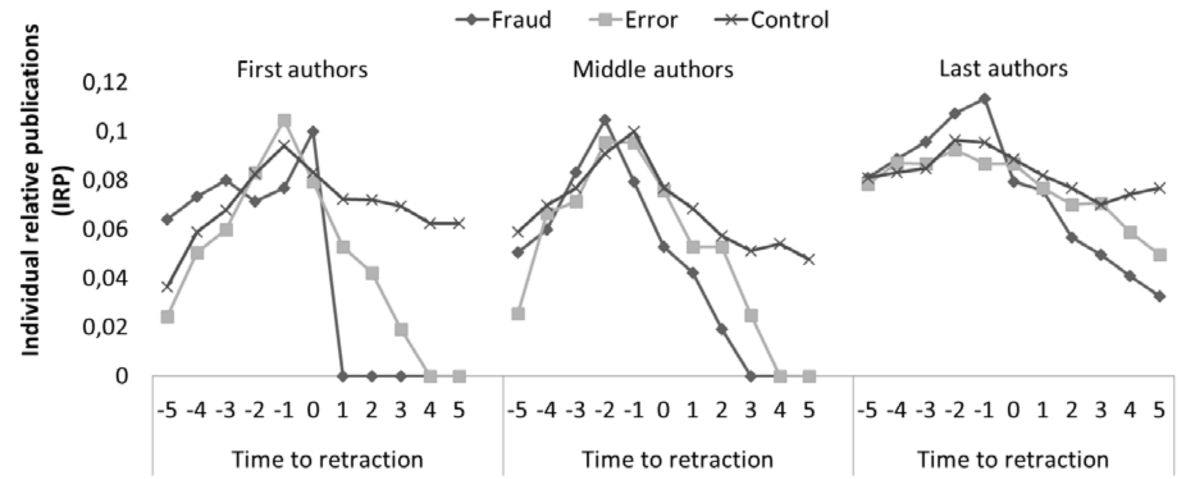
For all remaining authors, we searched the WoS for all articles, reviews and notes published in the five years preceding and following the retraction. For each paper found, the publication year was normalized by time to retraction (T). For authors with multiple retractions on different years, we gathered papers from 5 years before the first retraction to 5 years after the last one. In those cases, T = 0 for years between the first and last publication, inclusively. After author name disambiguation, we obtained a total of 15,333 distinct articles for the fraud and error groups, and 55,036 distinct articles for the control group.

Indicators
To measure the effect of retraction on the output of researchers, we used the individual relative productivity (IRP) calculated for each year by dividing the number of publications on that year by the total number of publications over the ten years period. We used the average relative citations (ARC) to measure scientific impact. Two other indicators were used to assess scientific impact: the number of highly cited papers (top 5% of the discipline), and the number of papers published in top journal (top 5% of the discipline). Thirdly, collaboration was assessed using the average number of authors, institutions and countries on the researchers' publications, all normalized by discipline.

Results

Scientific output
Figure 1 shows that retractions cause an important decrease in scientific output for all co-authors, no matter the reason for retractions. Also, for first and last authors, frauds seem to have more impact than errors, which is not the case for middle authors. First authors who retracted a paper for fraud seem to suffer a much bigger decline in scientific output than middle and last authors who retracted papers for the same reason. Furthermore, for all groups except last authors with a retraction for error, the differences in the median output between the pre- and post-retraction periods were found, using a Mann-Withney U-test, to be significantly different than the differences observed for the control groups ($P < 0.05$).

Figure 11. Median individual relative publications from five years prior to five years after the retraction.



Scientific impact

Table 2 shows the variation observed between pre- and post-retraction period for the 3 indicators of scientific impact. Since, authors must have published in both the pre- and post-retraction periods in order to compare their impact for those periods, those who had no publications in either the pre or post-retraction period were excluded for this part of analysis. The number of authors in the resulting sub-sample is indicated in table 2. Also, since many authors do not publish top papers, the 3rd quartile (and not the median) is used for that indicator.

Table 2. Difference between pre- and post-retraction average relative citations, proportion of top papers and publications in top journals.

			ARC (median)		Top papers (3rd quartile)		Top journals (median)	
		N	Var. (%)	Sig	Var. (%)	Sig	Var. (%)	Sig
First authors	Fraud	28	-8,7	.986	-13,2	.447	-33,6	.599
	Error	83	-4,0	.274	0,0	.517	-100,0	.792
	Control	354	-9,7	-	-17,6	-	-53,9	-
Middle authors	Fraud	253	-18,7	.092*	-42,1	.018**	-66,7	.001***
	Error	276	0,6	.013**	7,1	.164	-43,7	.286
	Control	860	-7,1	-	-18,0	-	-45,5	-
Last authors	Fraud	64	-14,0	.524	-17,6	.445	-39,9	.517
	Error	89	11,0	.108	10,3	.601	-25,4	.673
	Control	382	-1,0	-	0,0	-	-28,6	-
All authors	Fraud	345	-17,6	.056*	-23,5	.004***	-53,5	.003***
	Error	448	2,0	.003***	1,8	.126	-43,0	.226
	Control	1596	-7,5	-	-17,6	-	-44,8	-

Notes: P-values shown are the result from a Mann-Withney U-test, comparing the fraud and error groups with the control groups.
* $P < 0.1$; ** $P < 0.05$; *** $P < 0.01$

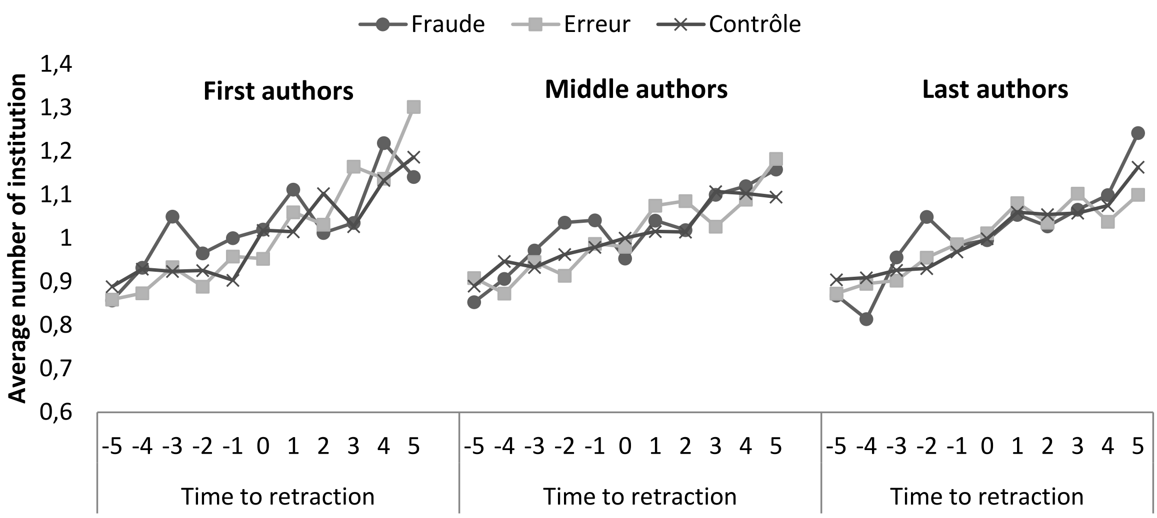
We see, in table 2, that for first and last authors, the differences observed between the fraud or error groups and the control groups are not statistically different. This may be due to the small size of this sub-sample. However, for the larger sub-sample of middle authors who retracted

for fraud, decreases observed for all three measures of impact are significantly more important than the decreases observed for the control groups.
Interestingly, for first, middle and last authors, retractions for error seem to have a positive impact on average relative citation and the proportion of top papers, in comparison with the control group. However, this is only statistically significant in the case of middle authors. This result may still be linked to a Lu, Jin, Uzzi, & Jones (2013), who showed that self-reported retractions (most likely errors) led to an increase in citations for the authors' previous work. Our results suggest that this might also be the case for the authors' ulterior work. Furthermore, the proportion of publications does not follow a similar trend. This would indicate that this increase of citations and top papers is not simply an effect of having more papers published in top journals.

Due of the small size of the first and last authors subsamples, it might be interesting to look at aggregated results for all authors. While these results are obviously influenced by the weight of the middle authors, we can say that, in general, retractions for fraud have a significant negative impact on citations, top papers and publications in top journals, and that errors have a significant positive impact on citations.

Collaboration
In the third part of our analysis, we looked at the impact of retraction on co-authors' collaboration, also using the sub-sample of authors with at least one publications in both the pre and post-retraction periods (see table 2 above). Figure 2 shows that retraction doesn't seem to have any significant impact on the inter-institutional collaboration level of co-authors. Similar results were obtained looking at the number of authors and number of countries per paper (not shown). Thus, we conclude that retractions do not appear to have any general effect on the collaboration practices of co-authors.

Figure 2. Average number of institutions per paper from five years prior to five years after the retraction.



Discussion

The results presented here show that co-authors do share the consequences of fraud. However, it is mostly the output of researchers that is affected, while the decline of the different measures of scientific impact decline appears to be less important, and the effect on collaboration, null. We expected that error would have little or no impact on co-authors' careers. However, our results show that errors do have important consequences (though not as important as cases of fraud) for collaborators in terms of publications. These results might be partly explained by the fact that retractions occur generally in cases of major errors that invalidate the findings as a whole, while minor error leads most likely to corrections. Also, our results seem to confirm that the extent of the impact of retraction is related to the position of the author in article's authors list. One unexpected finding was the positive impact that retraction for error seemed to have on the citations received by the author's subsequent work. More research will be necessary to confirm and fully understand this phenomenon.

The effect of having participated in a case of scientific fraud goes way beyond a decrease in papers or loss in scientific impact. Some consequences can be psychological (i.e. scientists losing trust in science, colleagues and institutions) or a waste of research efforts and funds. The case of Hendrik Schön, in physics, provides a good example of this waste of efforts: he forged 'ground-breaking' results that many other researchers around the globe were eager to reproduce and build upon, leading to much wasted funds and time, and the discovery of the fraud led a few discouraged scientists (mostly PhD and postdoctoral students) to abandon the idea of pursuing a career in research (Reich, 2009). Moreover, the many cases of fraud that are discovered almost every day are most likely the tip of the iceberg: in the United States, allegations of fraud received by U.S. Office of Research Integrity (ORI) have increased to a point where only a small proportion can actually be investigated (Nature News, 2013). It is, thus, likely that the number of cases will keep rising and that more and more collaborators will see their career compromised.

References

Azoulay, P., Furman, J. L., Krieger, J. L., & Murray, F. E. (2012). Retractions. *NBER Working Paper Series*, *18449*.

Biagioli, M. (1998). The instability of authorship: credit and responsibility in contemporary biomedicine. *FASEB Journal : Official Publication of* the *Federation of American Societies for Experimental Biology*, *12*(1), 3–16.

Biagioli, M. (1999). Aporias of scientific authorship: credit and responsibility in contemporary biomedicine. In *The science studies reader* (pp. 12–30).

Bonetta, L. (2006). The Aftermath of Scientific Fraud. *Cell*, *124*(5), 873–875.

Cokol, M., Ozbay, F., & Rodriguez-Esteban, R. (2008). Retraction rates are on the rise. *EMBO Reports*, *9*(1), 2.

Cronin, B. (2001). Hyperauthorship : A postmodern perversion or evidence of a structural shift in scholarly communication practices? *Journal of the American Society for Information Science and Technology*, *52*(7), 558–569.

Fanelli, D. (2009). How Many Scientists Fabricate and Falsify Research? A Systematic Review and Meta-Analysis of Survey Data (How Many Falsify Research?). *PLoS ONE*, *4*(5), e5738.

Fang, F. C., Steen, R. G., & Casadevall, A. (2012). Misconduct accounts for the majority of retracted scientific publications. *Proceedings of the National Academy of Sciences*, *109*(42), 17028–17033.

Furman, J. L., Jensen, K., & Murray, F. (2012). Governing knowledge in the scientific community: Exploring the role of retractions in biomedicine. *Research Policy*, *41*(2), 276–290.

Grieneisen, M. L., & Zhang, M. (2012). A Comprehensive Survey of Retracted Articles from the Scholarly Literature. *PLoS ONE*, *7*(10), e44118.

Jin, G. Z., Jones, B., Lu, S. F., & Uzzi, B. (2013). The Reverse Matthew Effect: Catastrophe and Consequence in Scientific Teams. *NBER Working Paper Series*. Retrieved from http://www.nber.org/papers/w19489

Nath, S. B., Marcus, S. C., & Druss, B. G. (2006). Retractions in the research literature: misconduct or mistakes? *The Medical Journal of Australia*, *185*(3), 152–154.

Nature News (2013). Seven days: 26 April–2 May 2013.http://www.nature.com/news/seven-days-26-april-2-may-2013-1.12899#/trend

Neale, A., Northrup, J., Dailey, R. K., Marks, E., & Abrams, J. (2007). Correction and use of biomedical literature affected by scientific misconduct. *Science and Engineering Ethics*, *13*(1), 5–24.

Neale, A. V., Dailey, R. K., & Abrams, J. (2010). Analysis of citations to biomedical articles affected by scientific misconduct. *Science and Engineering Ethics*, *16*(2), 251–61.

Pfeifer, M. P., & Snodgrass, G. (1990). The continued use of retracted, invalid scientific literature. *JAMA, The Journal of the American Medical Association*, *263*(10), 1420.

Pontille, D. (2004). *La signature scientifique : une sociologie pragmatique de l'attribution*. Paris: CNRS.

Reich, E. S. (2009). *Plastic fantastic : how the biggest fraud in physics shook the scientific world*. New York: Palgrave Macmillan.

Sovacool, B. K. (2008). Exploring Scientific Misconduct: Isolated Individuals, Impure Institutions, or an Inevitable Idiom of Modern Science? *Journal of Bioethical Inquiry*, *5*(4), 271–282.

Steen, R. G. (2011). Retractions in the scientific literature: is the incidence of research fraud increasing? *Journal of Medical Ethics*, *37*(4), 249–53.

Steen, R. G. (2012). Retractions in the medical literature: how can patients be protected from risk? *Journal of Medical Ethics*, *38*(4), 228–32.

Steen, R. G., Casadevall, A., & Fang, F. C. (2013). Why Has the Number of Scientific Retractions Increased? *PLoS ONE*, *8*(7), e68397.

Steneck, N. (2006). Fostering integrity in research: Definitions, current knowledge, and future directions. *Science and Engineering Ethics*, *12*(1), 53–74.

Wuchty, S., Jones, B. F., & Uzzi, B. (2007). The Increasing Dominance of Teams in Production of Knowledge. *Science* , *316* (5827), 1036–1039.

Public-private co-publications beyond university-industry links[1]

Fernanda Morillo*, Borja González-Albo* and & Luz Moreno*

* *fernanda.morillo@cchs.csic.es; borja.gonzalezalbo@cchs.csic.es; luz.moreno@cchs.csic.es*
Centre for Human and Social Sciences (CCHS), Spanish National Research Council (CSIC), Albasanz 26–28, 28037 Madrid, Spain

Introduction

Collaboration is nowadays a major concern for most scientists as it offers the possibility of facing significant challenges and of undertaking various lines of research, with a greater chance of success and higher accuracy (Breschi & Cusmano, 2004). The output of this cooperation usually surpasses the sum of its parts, creating synergy. Accordingly, several funding organisations are interested in promoting scientific partnerships and especially public-private collaboration, because it facilitates knowledge and technology transfer to industry. When thinking in this promotion, policy-makers usually consider the private sector in a similar way that FAO (2014): "enterprises, companies or businesses, regardless of size, ownership and structure". Although this definition includes a wide range of entities, it does not comprise academia, research institutions and philanthropic foundations. In line with the interest of funding bodies, there are numerous studies analysing technology transfer between university and industry (e.g., Abramo et al., 2009; D'Este et al., 2013).

However, knowledge transfer may happen beyond the limits of university-industry links. When a more inclusive definition of the private sector is considered, it is possible to observe how other organisations play an important role in this transfer. For OECD (2001), private sector comprises "private corporations, households and non-profit institutions serving households". In this sense, the objective of this work is to analyse public-private co-publications considering not only industry, but also other private organisations who also contribute to the knowledge exchange and the advancement of science. Although this is only an exploratory study based on the Spanish results, it is expected to observe similar characteristics than those found in previous works, regarding the areas and centres involved in knowledge transfer between public and private sector.

Methods and materials

We downloaded Spanish documents included in the Web of Science databases (2008-2012) and identified the public and private Spanish organisations responsible for those documents, following the OECD definition (2001). This identification was possible thanks to the information provided by the organisations' web pages and/or through email answers. On a general basis, centres mainly funded by governments or administrations were included in the public sector and the rest of the centres were considered in the private sector. With this data, we produced some bibliometric results for those identified organisations in each area, based on the Current Contents Connect Editions (number of articles, percentage of articles in the JCR first quartile of each discipline, type of collaboration, relative impact factor -RIF- and relative citations -RC- compared to the Spanish ones), studying co-authored publications between public and private institutions (network analysis with Pajek) and comparing this output with the total for Spain.

[1] This work was supported by the Spanish Ministry of Science and Innovation (Grant CSO2011-25102).

Results and discussion

Public-private co-publications represented less than 7% (14,665 articles) of the total Spanish production in the analysed period, but they were more than half of the private sector output. These documents were included in similar journals than those of the Spanish output (47% versus 49% of articles in the first quartile). As expected, when there was international collaboration (27% of the output), the results for impact were better. In general, public-private co-publications had no greater impact than the total Spanish output, with the exception of those documents with international collaboration. Although other authors usually study university-industry collaboration (e.g., Abramo et al., 2009; D'Este et al., 2013), they also find no relationship between scientific excellence and commitment to the industry.

Regarding areas (Table 1), impact factor and citations were similar or lower than those of the total of Spain (RIF and RC ≤ 1), with the only exceptions of Mathematics (higher RIF) and Social Sciences (higher RIF and RC). Private collaboration in Mathematics included mainly technological centres and high-tech firms, chiefly in statistics and applied mathematics. While in Social Sciences it involved hospitals and pharmaceutical firms, mostly in socio-medical specialities.

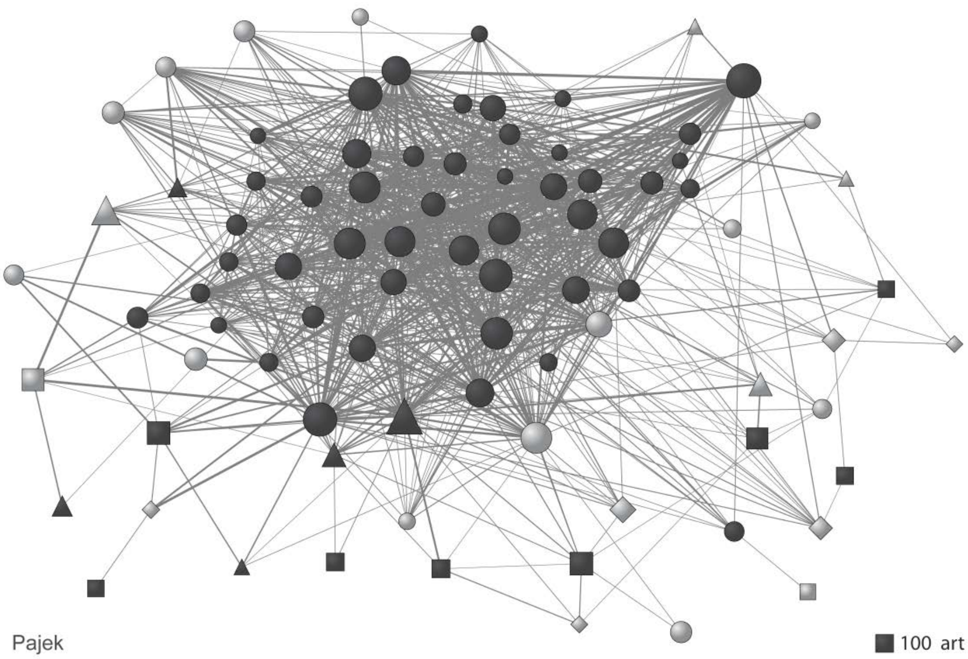
Table 1. Spanish public-private co-publications by areas (WoS 2008-2012, only articles).

Areas	**Art**	**%**	**RIF**	**RC**	**% of Spain**
Agriculture, Biology & Environmental Sciences	2300	15.68	0.92	0.83	5.31
Life Sciences	3725	**25.40**	0.95	0.87	8.54
Socials Sciences	1230	8.39	**1.14**	**1.13**	5.66
Physics	857	5.84	0.83	0.66	2.18
Arts & Humanities	80	0.55	**--**	**--**	1.29
Engineering Technology	2551	17.4	0.89	0.79	5.25
Mathematics	248	1.69	**1.20**	0.89	1.98
Clinical Medicine	6546	**44.64**	1.00	0.92	13.00
Multidisciplinary Sciences	301	2.05	0.88	0.75	6.80
Chemistry	1111	7.58	0.85	0.62	3.46

No RIF or RC are included for Arts & Humanities, due to their lesser relevance.

Even though in our results firms represented 36% of the private publications, there was a greater participation of the health sector (38%) and some presence of non-profit organisations (18%, included technology centres, very involved with small and medium enterprises). RIF and RC were lower than the Spanish ones, except for firms and the health sector. In addition, some differences between areas could be found. Whereas most of them presented low results, there were few exceptions in the Social Sciences and Clinical Medicine areas. These areas were also the most applied ones, a fact that is consistent with the conclusions of Perkmann & Walsh (2009). Focusing on collaboration within the Clinical Medicine area, the most productive one, it was observed that most of the relations were found in the public health sector, which was also connected to the private health sector, universities or firms (Figure 1).

Figure 1. Spanish public-private co-publications in Clinical Medicine (WoS 2008-2012, 70% around average, only articles).



Dark nodes: public organisations; light nodes: private organisations. Ellipses: health sector; diamonds: firms; boxes: universities; triangles: non-profit organisations.

Conclusions

When studying public-private collaboration, usually only university and industry are considered, standing out in areas closely related to manufacturing, where most public funds are allocated. If a broader analysis is applied, other areas outstand, such as Clinical Medicine, in which hospitals and pharmaceutical firms are the leaders. Indirect or alternative interactions can be seen, which may be more effective for improving the system in all aspects (Tether & Tajar, 2008). The public health sector has a key role in addressing socially significant problems and through this work, we can see important connections with private health, universities and firms.

References

Abramo, G.; D'Angelo C.A.; DiCosta, F. & Solazzia, M. (2009). University-industry collaboration in Italy: A bibliometric examination. *Technovation*, 29(6-7), 498–507.

Breschi, S. & Cusmano, L. (2004). Unveiling the texture of a European Research Area: Emergence of oligarchic networks under EU framework programmes. *International Journal of Technology Management*, 27(8), 747-772.

D'Este, P.; Tang, P.; Mahdi, S.; Neely, A. & Sánchez-Barrioluengo, M. (2013). The pursuit of academic excellence and business engagement: is it irreconcilable? *Scientometrics*, 95(2), 481–502.

FAO (2014). Food and Agriculture Organization of the United Nations: Private Sector. Retrieved January 20, 2014 from: http://www.fao.org/partnerships/private-sector/en/.

OECD (2001, November 20). OECD Glossary of Statistical Terms - Private sector Definition. Retrieved January 21, 2014 from: http://stats.oecd.org/glossary/detail.asp?ID=2130.

Perkmann, M. & Walsh, K. (2009). The two faces of collaboration: impacts of university-industry relations on public research. *Industrial and Corporate Change*, 18(6), 1033–1065.

Tether, B.S. & Tajar, A. (2008). Beyond industry–university links: Sourcing knowledge for innovation from consultants, private research organisations and the public science-base. *Research Policy*, 37(6-7), 1079–1095

Women and science in Russia: a historical bibliometric analysis

Adèle Paul-Hus*, Rébecca L. Bouvier*, Chaoqun Ni**, Cassidy R. Sugimoto**, Vladimir Pislyakov***and Vincent Larivière****

*adele.paul-hus@umontreal.ca; rebecca.l-bouvier@umontreal.ca
École de bibliothéconomie et des sciences de l'information, Université de Montréal
C.P. 6128, Succ. Centre-ville, Montréal, QC, H3C 3J7 (Canada)

**chni@indiana.edu; sugimoto@indiana.edu
School of Informatics and Computing, Indiana University, 1320 E. 10thSt, Bloomington, IN. 47405 (USA)

***pislyakov@hse.ru
Library, National Research University Higher School of Economics, Myasnitskaya, 20, Moscow, 101000 (Russia)

**** vincent.lariviere@umontreal.ca
École de bibliothéconomie et des sciences de l'information, Université de Montréal
C.P. 6128, Succ. Centre-ville, Montréal, QC, H3C 3J7 (Canada) and
Observatoire des Sciences et des Technologies (OST), Centre Interuniversitaire de Recherche sur la Science et la Technologie (CIRST), Université du Québec à Montréal, CP 8888, Succ. Centre-Ville, Montréal, QC, H3C 3P8 (Canada)

Introduction

Women in science

Gender disparities persist in several areas of society, and scientific research is no exception. Differences between men and women in science appear in terms of productivity, speciality, collaboration and scientific impact (Larivière et al., 2013). Although the position of women in Western society has improved greatly in the last century, numerous studies confirm that gender disparities in science remain, including in the United States (Xie & Shauman, 2003), Québec (Larivière et al., 2011), Russia (Lewison & Markusova, 2011), Poland (Suchanska & Czerwosz, 2013), Italy (Abramo, D'Angelo & Caprasecca, 2009) and France (De Cheveigné, 2009). This study seeks to describe the evolution of the place of female researchers in Russia, taking into account the socioeconomic, political and historic context of the country, which was marked by the fall of the USSR in 1991.

Whereas Lewison and Markusova (2011) provided evidence of a gender gap in Russia, based on bibliometric data for three non-consecutive years (1985, 1995 and 2005), the present article proposes to corroborate these results and study the situation over a larger time window, with data from 1973 to 2012. We thus seek to evaluate the place of women in the Russian scientific research system in the various disciplines and how this position has evolved during the last forty years in terms of their proportion of the published research output and scientific impact.

Science in Russia

The end of the communist regime induced deep changes to Russian science and technology. By 1992, science had entered a profound crisis. For several years, the budget allocated to scientific research decreased constantly and, thus, scientists had difficulties obtaining the equipment essential to pursue their research. Russian science survived in large part through the financial support of international funding – such as that provided by the Hungarian-American billionaire George Soros or European Union programs (e.g., INTAS). In these conditions, many male researchers left Russia or changed careers, leaving more positions for women in scientific research (Lewison & Markusova, 2011). Moreover, the demilitarization reform initiated in 1992 resulted in the layoff of a significant proportion of Russian scientists. Staff working in research halved between 1992 and 1999, leading to a decrease of scientific publications and less international visibility (Milard, 2009). The same bibliometric trend persisted later on (Kotsemir, 2012; Pislyakov & Gokhberg, 2008).

During the Cold War, the Soviet Union partly succeeded in establishing Russian as an international scientific language. Thus, its scientific production was mostly published in Russian. However, a rapid and complete shift toward Russians publishing in English occurred in 1991, resulting in a greater visibility of Russian science at the international level (Kirchik, Gingras & Larivière, 2012).

Sources and methods

Data for this study are drawn from Thomson Reuters' Web of Science database (Science Citation Index Expanded, Social Sciences Citation Index and Arts and Humanities Citation Index). All articles, notes and reviews published between 1973 and 2012 are included in the analysis. Papers taken into account contained at least one institutional address situated in Russia (or USSR before 1991) for a total of 1,059,939 papers. Given the well-known limitations of data on the Social sciences and Humanities (Archambault et al., 2006; Larivière et al., 2006) – especially for non-English speaking countries and, particularly, Russia (Savelieva & Poletayev, 2009) – these were excluded from the analysis (except Psychology which is situated halfway between the social sciences and the natural sciences). The NSF categorization (based upon the Science and Engineering Indicators (National Science Foundation, 2006)) was adopted instead of WoS categories since the former classifies each journal into only a single specialty and discipline, which avoids possible double counting of papers during analysis. Additionally, NSF categorization provides a hierarchical structure of two levels (discipline and specialty), which allows for analysis at different levels of aggregation.

Based on the characteristics of Russian surnames, which contains gender-specific suffixes[1], it was possible to determine genders for each authorship. Surnames which did not meet those criteria were excluded from the selected data. As a result, over the 1973-2012 period, 89% of papers contained at least one author to whom a gender was assigned. The analysis of male and female researchers' relative contribution to published papers is based on the proportion of papers published by authors of each gender for whom gender could be assigned. The number of papers is obtained by fractional counting where each author is given $1/x$ count of the authorship, with x representing the number of authors for which gender was identified in the given paper (Larivière et al., 2013).

We also compared the scientific impact of male and female researchers using the average of relative citations (ARC). ARC provides field-normalized citation rates, thus allowing the comparison of data between the different specialities that have otherwise different citation practices. More specifically, the number of citations received by a given paper is divided by

[1]Suffixes associated to male gender: -ov, -in, -ev, -ky, -kii, -kiy, -yi, -ny, -oy, -oi, except -tsoi and -tsoy. Suffixes associated to female gender: -ova, -ina, -eva, -aia and –aya.

the average number of citations received by articles in the same discipline published that year. An average of relative citations (ARC) greater than 1 indicates that an article is cited above the world average for the same field, and an ARC below 1 means that it is cited below the world average. Citation measures used for this analysis include all citations received by a given paper, from its publication year to the end of 2012.

Results

Research output

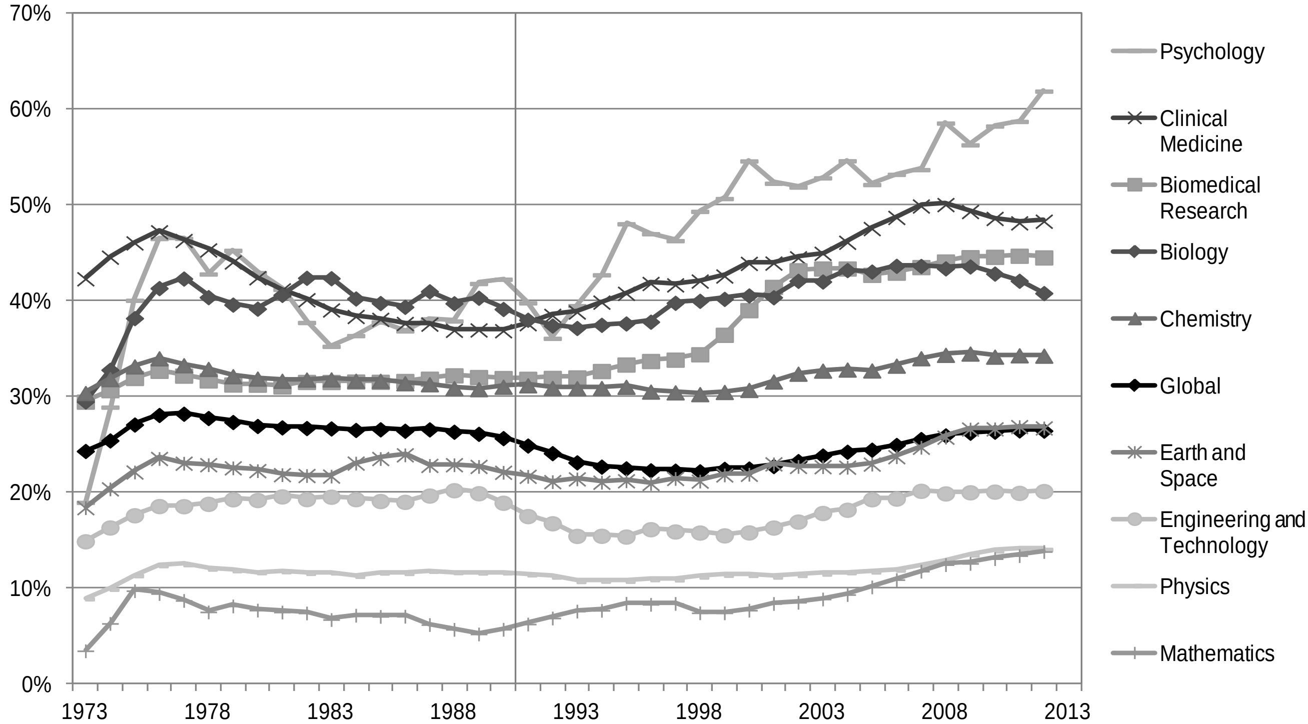
To assess the place of Russian women in science, we evaluated their relative contribution to all papers that were published in Russia in each of the selected disciplines, between 1973 and 2012. Figure 1 shows that women's proportion of fractionalized authorship is lower than that of men in all disciplines except Psychology. All disciplines taken together, women account for less than 30% of fractionalized authorship over the studied period. However, for Psychology, the contribution of women to published articles averages 45%, reaching more than 50% after 2000, making it the most gender-equal discipline of those in the analysis. One of the explanations for this result may be that a majority of Russian Psychology papers are published in two Russian journals. Indeed, these national journals account for 74% of Russian papers published in this discipline after 2000, where women account for 59% of fractionalized authorship against a proportion of 46% in the rest of foreign Psychology journals indexed in the database. Women are thus overrepresented in the Russian journals in terms of fractionalized authorship in Psychology, between 2000 and 2012. On the other hand, areas in which Russia has been historically very active – such as Mathematics, Physics and Engineering and Technology – are traditionally male dominated (Xie & Shauman, 2003, p. 33). Our results show that, in these disciplines, women represent less than 20% of fractionalized authorships.

Variations in the proportion of female authorship can be observed over time. Between 1973 and 1976, we note an increase in female relative contribution in all disciplines. The inclusion process of Soviet journals to the Science Citation Index during these years could be a contributing factor to this increase. The gender gap being less significant in the national journals than in the foreign ones, then the inclusion of national journals in the database should lead to an effect like that shown in Figure 1. However, data of the years preceding 1973 would be necessary in order to better understand the observed increase in the proportion of female scientific output between 1973 and 1976.

From 1991 onwards, we observe a rise of the women's proportion of fractionalized authorship in Psychology, Clinical Medicine, Biology and Biomedical Research. Unsurprisingly, several of the specialties of Psychology as well as of the two medical disciplines (Clinical Medicine and Biomedical Research) are related to domains historically considered "feminized" and "care" areas of research (Witz, 1992). Mathematics is the only other discipline where we can see a slight increase in female relative contribution to scientific output after 1991. In a difficult economic position, the Russian state could not support science anymore, a large number of male scientists left Russia to continue their research abroad, which might explain part of this increase (Lewison & Markusova, 2011). On the other hand, we see after 1991 a significant decline of female relative contribution in Engineering and Technology. However, one should keep in mind that after 1991, our statistics lose all papers from other USSR republics, except the Russian Federation. If the authors' gender structure in these republics differed from that in Russian Federation, their removal might also disturb at this point the curves in Figure 1.

Figure 1 also shows, from 2008 onwards, stagnation in the female proportion of authorship in all disciplines, except Psychology – the discipline in which their proportion of the output is the largest. We should nonetheless acknowledge the fact that the total number of Russian papers in Psychology is relatively small with an average of 110 published papers per year for the 1973-2012 period, compared to an average ranging between 875 papers per year in Biology and 7527 papers per year in Physics.

Figure 12.Women's fractionalized authorships, by discipline, 1973-2012. A 3-year moving average was applied on all disciplines to enhance the readability of the figure.



Scientific impact

Figure 2 shows the evolution, between 1973 and 2012, of the relative scientific impact of Russian papers, according to the gender of the first author. It shows that, despite important variations in the overall impact of Russian papers, the difference between men and women remains relatively stable throughout the period, except after the fall of the Soviet Union in 1991, where it seems to widen. This historic period is also associated to a transition of the main publication language of Russian researchers which shifted from Russian to English (Kirchik et al., 2012). Therefore, the scientific impact of articles published after 1991 in Russia increases substantially, as articles written in English have a broader readership and, thus, a larger international impact, than papers published in Russian. As proposed by Lewison and Markusova (2011), this increasing difference can be attributed to the lesser propensity of women to publish in English, as compared to their male counterparts. One can also notice the decrease in scientific impact of Russian papers between 1973 and 1990, which is likely due to the economic decline of the USSR, initiated in 1971 (Freeze, 2002), as well as the fading impact of Russian language in science accompanied by the increase of the Soviet journals in the database.

Figure 13. Average of relative citations of Russian papers, by gender of the first author, 1973-2012

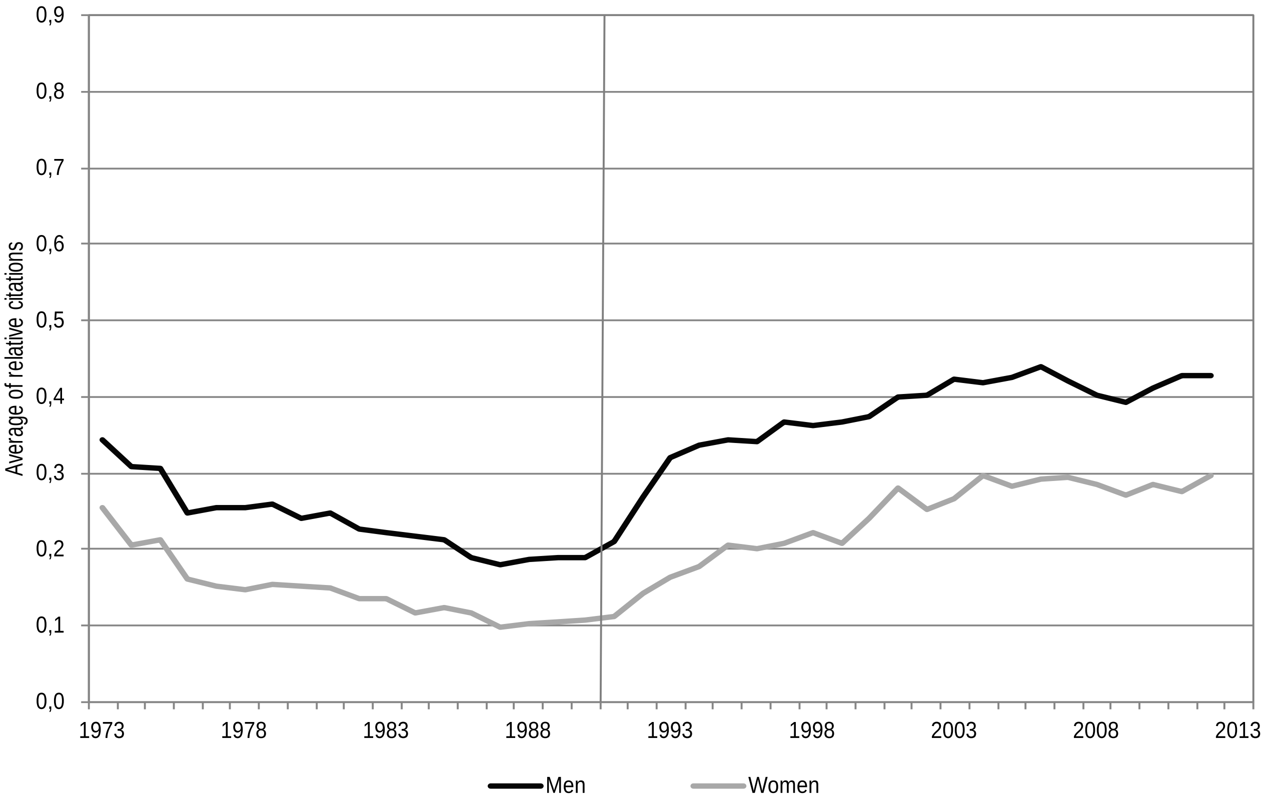
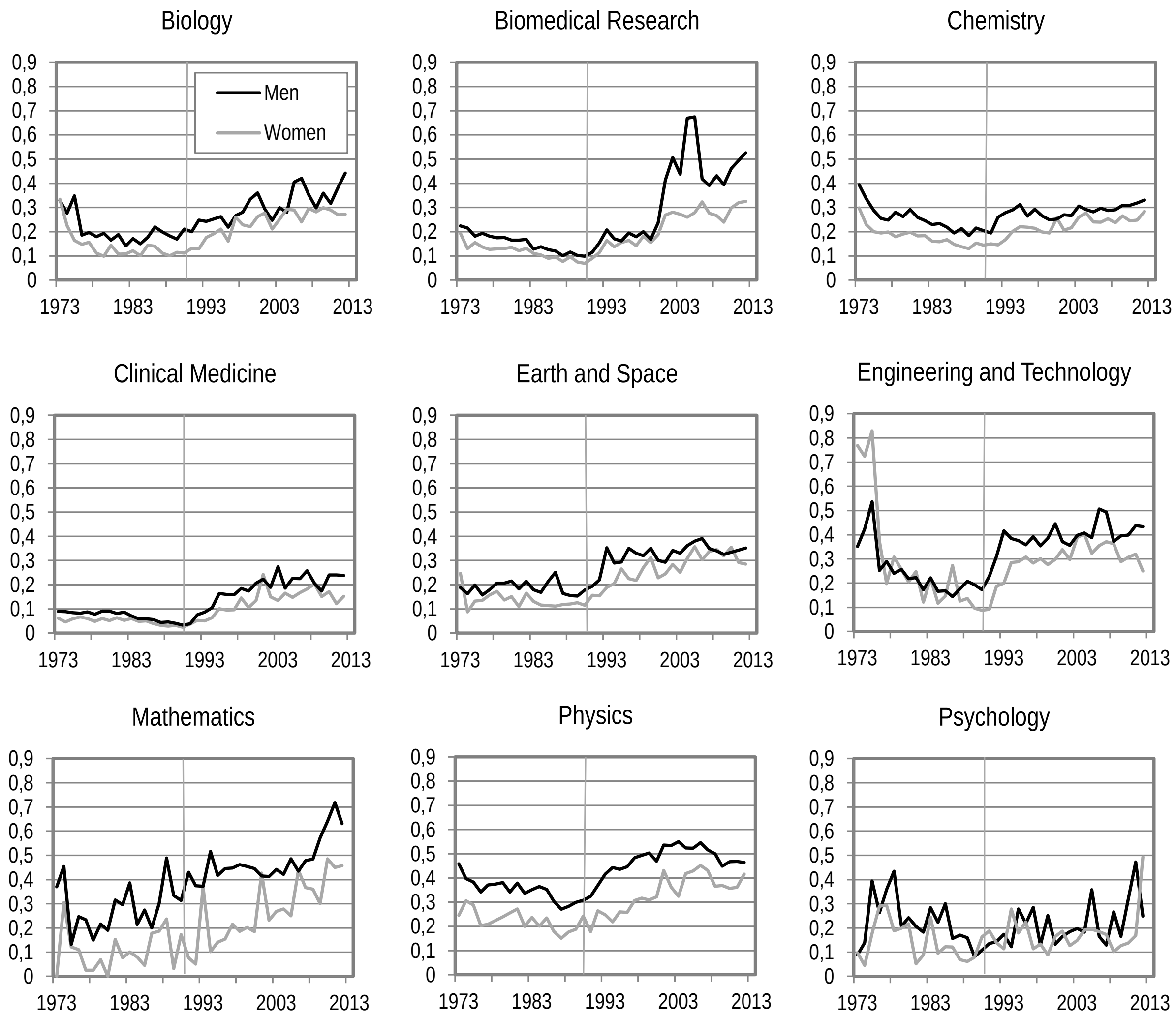


Figure 3 presents the evolution of the scientific impact of men's and women's first-authored papers by discipline. It shows, for each discipline, an increase of the scientific impact of Russian papers after the fall of the Soviet Union in 1991; a trend which is likely due to the transition of the language of scientific publications from Russian to English. The extent of the gender gap in terms of impact varies greatly by discipline. In Biology, Chemistry, Earth and Space sciences and Physics, the difference between men's and women's impact remains consistent over time, with men's impact being higher globally. In Biomedical Research and Clinical Medicine, articles published by men show a slightly higher relative impact until the 1990s for Clinical Medicine and the beginning of the 2000s for Biomedical Research. After that, men's papers' impact grows rapidly, increasing the gender gap in terms of impact.

Engineering and Technology is the only discipline where articles published by women have an impact similar to that of men, before the collapse of the USSR in 1991. This bibliometric trend could suggest that the increased need of researchers in military areas, during the arms' race period of the Cold War, was mostly filled by women. After 1991, the gap between both genders widens and male author's impact surpasses that of female authors. Mathematics and Physics are both disciplines in which Russia has specialized, and Figure 3 confirms the disparity in terms of scientific impact between men and women in these traditionally female underrepresented domains, as the lower proportion of women in these fields might have an effect on their scientific impact. The largest difference is found in Physics and remains stable over time. In Mathematics, however, ARC values show considerable annual variations and women's impact reaches men's impact a few times throughout the period. Nonetheless, the limited number of articles published in Mathematics can likely explain the significant variations seen from one year to another. Although it may appear contradictory, it is in Mathematics and Physics that women's papers have the highest impact, as a consequence of the highest overall ARC of Russian papers in these disciplines.

As observed with genders' contribution to the Russian research output (Figure 1), Psychology is also the most gender-balanced discipline in terms of scientific impact, with male researchers' impact being only slightly greater than that of female authors. However, after 1991, women's impact increases to reach that of men. We should nonetheless acknowledge the fact that the total number of Russian papers in Psychology is relatively small which explains the significant variations observed from one year to another.

Figure 14. Average of relative citations of Russian papers, by gender of the first author, by discipline, 1973-2012



Conclusion

Our analyses of Russian productivity and scientific impact over the last 40 years clearly show that gender parity is far from being achieved. Women remain underrepresented in terms of relative contribution to scientific output across disciplines, although it is in Mathematics and in Physics, both research areas in which Russia has specialized, that we observed the greatest gap (Figure 1). The Soviet Union's fall in 1991 is associated, in some disciplines, with a slight increase of the relative contribution of female authors; increase that could be explained by a "brain drain" of male researchers that followed the fall. Our results also show that, while it is in Psychology, Clinical Medicine and Biomedical Research that women's contribution to research is the most important, it is in Mathematics and Physics, the most traditionally male disciplines, that they have the highest impact (Figure 3).

After 1991, we observe an increase of both men's and women's papers' scientific impact (Figure 2). Although the impact of women's scientific output significantly increases after the fall of the USSR, the gap between both genders remains stable over time for most of the disciplines. As a result, we cannot interpret this increase as an improvement of the women's relative influence in Russian science.

The patterns presented here are not specific to Russia. As demonstrated in recent study (Larivière et al., 2013), gender disparities in science remain widespread across the world. Over the 2008-2012 period, men accounted for more than 70% of fractionalized authorship worldwide, which approximately coincides with our results for Russia (Figure 1, 'Global'). Scientific impact of women is also invariably less strong than that of their male counterparts, as articles published by female authors attract fewer citations. As the Russian government has taken a more interventionist approach since 2006 and has increased the funding for science, it seems that women's proportion of the Russian scientific community has flattened. Time will tell if their proportion will start to increase or decrease again.

References

Abramo, G., D'Angelo, C. A. & Caprasecca, A. (2009). Gender differences in research productivity: A bibliometric analysis of the Italian academic system. *Scientometrics, 79*(3), 517-539.

Archambault, É., Vignola-Gagne, É., Côté, G., Larivière, V. & Gingras, Y. (2006). Benchmarking scientific output in the social sciences and humanities: The limits of existing databases. *Scientometrics, 68*(3), 329-342.

De Cheveigné, S. (2009). The career paths of women (and men) in French research. *Social studies of science, 39*(1), 113-136.

Freeze, G. (2002). *Russia: a history*: Oxford University Press.

Kirchik, O., Gingras, Y. & Larivière, V. (2012). Changes in publication languages and citation practices and their effect on the scientific impact of Russian science (1993–2010). *Journal of the American Society for Information Science and Technology, 63*(7), 1411-1419.

Kotsemir, M. N. (2012). Publication Activity of Russian Researches in Leading International Scientific Journals. *Acta Naturae, 4*(13), 14-34.

Larivière, V., Archambault, É., Gingras, Y. & Vignola-Gagné, É. (2006). The place of serials in referencing practices: Comparing natural sciences and engineering with social sciences and humanities. *Journal of the American Society for Information Science and Technology, 57*(8), 997-1004.

Larivière, V., Ni, C. Q., Gingras, Y., Cronin, B. & Sugimoto, C. R. (2013). Global gender disparities in science. *Nature, 504*(7479), 211-213.

Larivière, V., Vignola-Gagné, É., Villeneuve, C., Gélinas, P. & Gingras, Y. (2011). Sex differences in research funding, productivity and impact: an analysis of Québec university professors. *Scientometrics, 87*(3), 483-498.

Lewison, G. & Markusova, V. (2011). Female researchers in Russia: have they become more visible? *Scientometrics*, *89*(1), 139-152.

Milard, B. (2009). L'héritage soviétique dans la nouvelle organisation de la science en Russie. *Revue d'anthropologie des connaissances, 2*(3), 391-411.

National Science Foundation. (2006). Science and Engineering Indicators. Chapter 5: Academic Research and Development. Data and Terminology.

Pislyakov, V. & Gokhberg, L. (2008). Assessing the relative standing of Russian science through a set of citation and publication indicators. *Excellence and Emergence. Book of Abstracts. 10th International Conference on Science and Technology Indicators.* (pp. 400-403). Vienna: ARC.

Savelieva, I. M. & Poletayev, A. V. (2009). Publications of Russian humanity scholars abroad: A sociometric analysis. *Journal of Educational Studies, 4*, 199–217.

Suchanska, M. & Czerwosz, E. (2013). *Women in technical universities in Poland*. Paper presented at the AIP Conference Proceedings.

Witz, A. (1992). *Professions and Patriarchy*: Routledge.

Xie, Y. & Shauman, K. A. (2003). *Women in science: Career processes and outcomes* (Vol. 26): Harvard University Press Cambridge, MA.

Within- and between-department variability in individual productivity. The case of Economics

Antonio Perianes-Rodrigueza, and Javier Ruiz-Castillo[b]

antonio.perianes@uc3m.es
[a] Departamento de Biblioteconomía y Documentación, Universidad Carlos III, SCImago Research Group, Antonio Perianes

jrc@eco.uc3m.es
[b] Departamento de Economía, Universidad Carlos III, Javier Ruiz Castillo *()*

Introduction

There are two types of research units whose performance is usually investigated in one or several scientific fields: individuals (or publications), or larger units such as universities or entire countries. In contrast, the information about the university departments (or research institutes) is not easy to come by (Van Raan, 2005). This is important because, in the social sciences, university departments are the governance units where the demand for and the supply of researchers determine an equilibrium allocation of scholars to institutions. This paper uses a unique dataset consisting of all individuals working in 2007 in the top 81 Economics departments in the world according to the Econphd university ranking (2004).

The allocation of researchers to departments takes place under different institutional scenarios in different countries of the world. Consider first countries where hiring and promotion procedures are essentially guided by meritocratic practices and competitive market forces. Let us think, for example, of the U.S. and, to a large extent, Canada or the UK. The demand side for first job contracts consists of a set of departments initially ordered in terms of a number of observable variables, such as research performance, wages, research facilities, geographical location, and prestige. Job offers are not tended at random among recent PhDs. On the other hand, self-selection from the supply side strongly affects the workings of this market. Taking into account a number of personal characteristics, such as the University where she graduates, the adviser and the other faculty members writing her recommendation letters, and the characteristics of her dissertation and job market paper, each recent PhD applies to the highest ranked sub-set of departments where she thinks she has a chance of being hired. In this way, search costs for departments are economized: each department can focus their attention on its set of self-selected candidates. Taking into account department needs, the credentials supplied by each candidate, as well as the results of interviews and seminars, each department makes a set of offers among the pool of its prospective candidates. Some offers are eventually accepted by some PhDs in all departments every year.

This process reveals a good deal of information to all parties concerned. The self-selection acting from the supply side of the market facilitates an efficient matching between applicants and departments. Nevertheless, strong doses of uncertainty still pend over the outcomes in this annual market. Not even the young participants are at all sure about their long-run "quality", and hence it is not obvious to anyone whether each recent PhD has been assigned to the "right" department. The tenure process serves to dispel some of these uncertainties. After a careful review, tenure is offered in each department to some of the individuals on tenure-track after a maximum period of, say, six years. In parallel, mobility across departments in response to meritocratic and competitive market forces provides another adjustment mechanism. Some people move towards better departments, and some others move in the opposite direction. In the absence of new elements –such as substantial variations in departments' total resources– this complex process can be conjectured to reproduce the initial department ranking.

In other non Anglo-Saxon countries, where less flexible public sector hiring and promotion practices play a dominant role, meritocratic and competitive forces may play a lesser role in determining final outcomes. Nevertheless, in a cross-section of world elite departments in a given field dominated by Anglo-Saxon countries, as we have in this paper, we can assume for the sake of the argument that the equilibrium allocation of individuals to departments captured in our sample does approximately reproduce some initial department ranking.

Be it as it may, this paper contributes to the formulation of a demand and supply equilibrium model for researchers by investigating two key stylized facts for our set of elite world Economics departments in 2007: the within- and between-department variability of several characteristics of productivity distributions or, in other words, the following two empirical questions:

1. Do we expect faculty members in a given department to have all similar productivities around the department mean?

2. If department productivity distributions are not uniform, do we expect these distributions to be similar across departments?

Naturally, in the absence of a formal model for the labor market as a whole in the entire field, it is not easy to come up with sensible conjectures to these questions. As a first move in this direction, this paper studies empirically these issues for 81 top Economics departments. We obtained information about the publications in the periodical literature for the 2,705 economists working in these departments in 2007. We could not find information about a person's education and/or publications in 50 cases, and there are 175 faculty members without any publication at all. Therefore, we focus on the remaining 2,530 faculty members with at least one publication that constitute what we call the population as whole.

Let the individuals be indexed by i, where i = 1, 2,.., 2,530. For every i, we measure individual productivity as a quality index, Q_i, that weights differently the articles published in four journal equivalent classes, where the first three classes consist of five, 34, and 47 journals, respectively, while the fourth consists of all other journals in the periodical literature. The four classes are assigned weights equal to 40, 15, 7, and 1 point, respectively (see Albarrán *et al.*, 2014, for further details concerning the construction of this index, as well as the comparison of our sample with the field of Economics as a whole). Given the way the data was selected, it is not surprising that we are working with a very productive sample.[1]

[1] We also measure individual productivity as the number of publications until 2007. The robustness of our results using both measures can be seen in the Working Paper version of this paper, Perianes-Rodriguez & Ruiz-Castillo (2014), hereafter PRRC.

Characteristics of the productivity distribution for the population as a whole

Basic characteristics

For the productivity distribution $\boldsymbol{Q} = (Q_1, \dots, Q_i, \dots, Q_{2,530})$, we are interested in two basic characteristics: the mean, and the individual variability within the distribution in question. Two aspects of the latter are investigated: the productivity inequality, measured by the coefficient of variation (*CV* hereafter), and the skewness of the distribution for which we follow the Characteristic Scores and Scale (CSS hereafter) approach (see PRRC for a second skewness measure using an index robust to extreme observations). The following two *characteristic scores* are determined at any aggregation level: μ_1 = mean productivity, and μ_2 = mean productivity for individuals with productivity greater than μ_1. Consider the partition of the distribution into three broad classes: (i) individuals with low productivity smaller than or equal to μ_1; (iii) fairly productive individuals, with productivity greater than μ_1 and smaller than or equal to μ_2, and (iii) individuals with remarkable or outstanding productivity greater than μ_2. The information about the main characteristics of distribution $\boldsymbol{Q}$ for the population as a whole is in Table 1.

Table 1. Characteristics of productivity distribution *Q*. Results of the CSS approach for the entire population

		Percentage of individuals in category:			Percentage of total articles in category:		
Mean	***CV***	**1**	**2**	**3**	**1**	**2**	**3**
307.3	1.30	69.2	20.0	10.8	24.2	32.2	43.6

Two comments are in order. Firstly, the productivity inequality according to the *CV* is 1.3, a very high figure indicating that the standard deviation is 1.3 times greater than the mean. Secondly, distribution $\boldsymbol{Q}$ is considerably skewed: the percentage of people with below average productivity is approximately 19 points to the right of the median, and 10.8% of the total population are responsible for 43.6% of all index points. These figures are comparable to what we find for the population of scholars in Economics & Business in Ruiz-Castillo & Costas (2014) for a much larger population. This parallelism reflects the fractal nature of productivity distributions in our field.

Individual variability within- and between- departments

We now turn towards the two questions raised in the Introduction for the partition of distribution $\boldsymbol{Q}$ into the 81 departments. Table A in the Appendix of PRRC presents the results for the mean productivity and the *CV* in each department, while Table B presents the results of the CSS approach for all departments. The average over all departments, and the coefficient of variation of these characteristics are in in Table 2.

Table 2. Average (coefficient of variation) over 81 Departments for different characteristics of productivity distributions. Results of the CSS approach (*Q*)

		Percentage of people in category			Percentage of total articles in category		
Mean	***CV***	**1**	**2**	**3**	**1**	**2**	**3**
294.6 (0.55)	1.04 (0.27)	62.8 (0.14)	22.6 (0.29)	14.7 (0.31)	25.3 (0.25)	32.2 (0.25)	43.3 (0.21)

The first conclusion is that productivity distributions at the department level are far from uniform: there is a high productivity inequality, and the majority of departments are clearly skewed to the right. Moreover, the high coefficients of variation in Table 2 indicate that productivity inequality and the skeweness of productivity distributions are very different across departments. Therefore, although we find large within-departmental variability, the productivity inequality and the degree of skeweness of productivity distributions is very different across departments.

Finally, in PRRC we investigate in detail the importance of differences between department productivity distributions in the measuring framework introduced in Crespo *et al.* (2013a) with the purpose of analyzing the effect on overall citation inequality of differences in production and citation practices across scientific fields. The conclusion is that the effect on overall productivity inequality that can be attributed to differences in the 81 productivity distributions in Economics (29%) is clearly greater than the corresponding effect attributable to differences in citation distributions across a large number of Web of Science subject categories (Crespo *et al.*, 2013a, b, Li *et al.*, 2013, and Waltman & Van Eck, 2013). However, the part of these differences that can be attributed to scale factors in our dataset is of a comparable order of magnitude (84%) to the same phenomenon in the context of sub-field citation distributions.

Characteristics of productivity distributions after age normalization

Since Lotka (1926), individual productivity datasets typically consist of a cross-section of researchers of different age in a given moment of time. However, human capital models suggest a humped-shaped progression of individual research productivity with academic age because the stock of human capital needs to be built up at the beginning of the career while, due to the finiteness of life, no new investment offsets depreciation and net investment declines (eventually) over time (Diamond, 1984). Consequently, the productivity of two scientists of different age in a given field is, in principle, non-comparable. Fortunately, our dataset has information on both individual researchers publications and their academic age, Age_i, i = 1,…, 2,530, where Age_i is the number of years since the completion of their Ph.D. and 2007.

Denote by $\boldsymbol{Q/Age} = (Q_1/Age_1, \dots, Q_i/Age_i, \dots, Q_{2,530}/Age_{2,530})$ the distribution of individual productivity after age normalization. We begin by asking: what types of changes in the ordering of individuals and departments are generated by age normalization? Firstly, it is observed that individuals are very much affected: more than 50% of all individuals experience re-rankings of more than 250 positions, and almost 60% of them experience changes in the relative indicators of productivity greater than 0.20. Secondly, the ranking of departments is also greatly altered (see PRRC for details). However, as can be observed in Table 3, age normalization does not change very much the characteristics of the productivity distribution for the population as a whole. Comparing with Table 1, there is simply a moderate decrease in both productivity inequality, measured by the *CV* and the skeweness of the distributions.

Table 3. Characteristics of productivity distribution *Q/age*. Results of the CSS approach for the entire population

		Percentage of individuals in category:			Percentage of total articles in category:		
Mean	***CV***	**1**	**2**	**3**	**1**	**2**	**3**
14.9	0.93	65.0	22.0	13.0	28.5	32.7	38.8

Next, we should answer the two questions raised in the Introduction. Firstly, does the variability within department productivity distributions change when productivity is normalized by academic age? Taking into account the information summarized in Table 4, the answer is: not very much. On average, both productivity inequality, and the skeweness of productivity distributions are somewhat smaller after age normalization.

Table 4. Average (coefficient of variation) over 81 Departments for different characteristics of productivity distributions. Results of the CSS approach (*Q/age*)

		Percentage of people in category			Percentage of total articles in category		
Mean	***CV***	**1**	**2**	**3**	**1**	**2**	**3**
14.2 (0.49)	0.77 (0.25)	59.0 (0.13)	24.7 (0.24)	16.3 (0.29)	30.3 (0.24)	32.1 (0.21)	37.9 (0.18)

Secondly, does between-department variability change when we consider productivity per year? Differences across departments are now considerably increased. In comparison with Table 3, the coefficients of variation in Table 4 indicate that, although mean productivity differences are somewhat reduced, the between-department variability experienced by both productivity inequality, and the skeweness of productivity distributions is clearly greater after age normalization. The large differences across department productivity distributions according to the CSS approach are illustrated in Figure 1 (see also Table D in the Appendix in PRRC).

Finally, how is the effect on overall productivity inequality attributable to productivity differences across departments affected by the normalization of individual productivity by academic age? As reported in detail in PRRC, this effect increases from 29% to 36%. However, the part of these differences that can be attributed to scale factors is of a similar order of magnitude before and after age normalization.

Conclusions

The matching of individuals and university departments in any scientific field results from a complex equilibrium between the demand for and the supply of researchers at different stages in their career. As a first step towards the development of a formal model of this process, this paper has investigated some of the characteristics of productivity distributions for a population of 2,530 individuals with at least one publication who were working in 81 top Economics departments in 2007.

For the population as a whole, the productivity inequality and the skewness of distribution ***Q*** before and after age normalization are of the same order of magnitude as the figures for the much larger population of scholars in Economics & Business in Ruiz-Castillo & Costas (2014). In relation to the partition of the population into the 81 departments, the main findings are the following two.

(i) Department productivity distributions are far from uniform. In other words, within each department, individuals have very different productivity.

(ii) There is not a single pattern of productivity inequality and skewness at the department level. On the contrary, productivity distributions are very different across departments. Consequently, the effect on overall productivity inequality of differences in productivity distributions across the 81 departments is greater than the effect attributable to differences in production and citation practices across 172 or 219 sub-field citation distributions. Interestingly enough, to a large extent these differences –however important– are accounted for by scale factors well captured by departments' mean productivities.

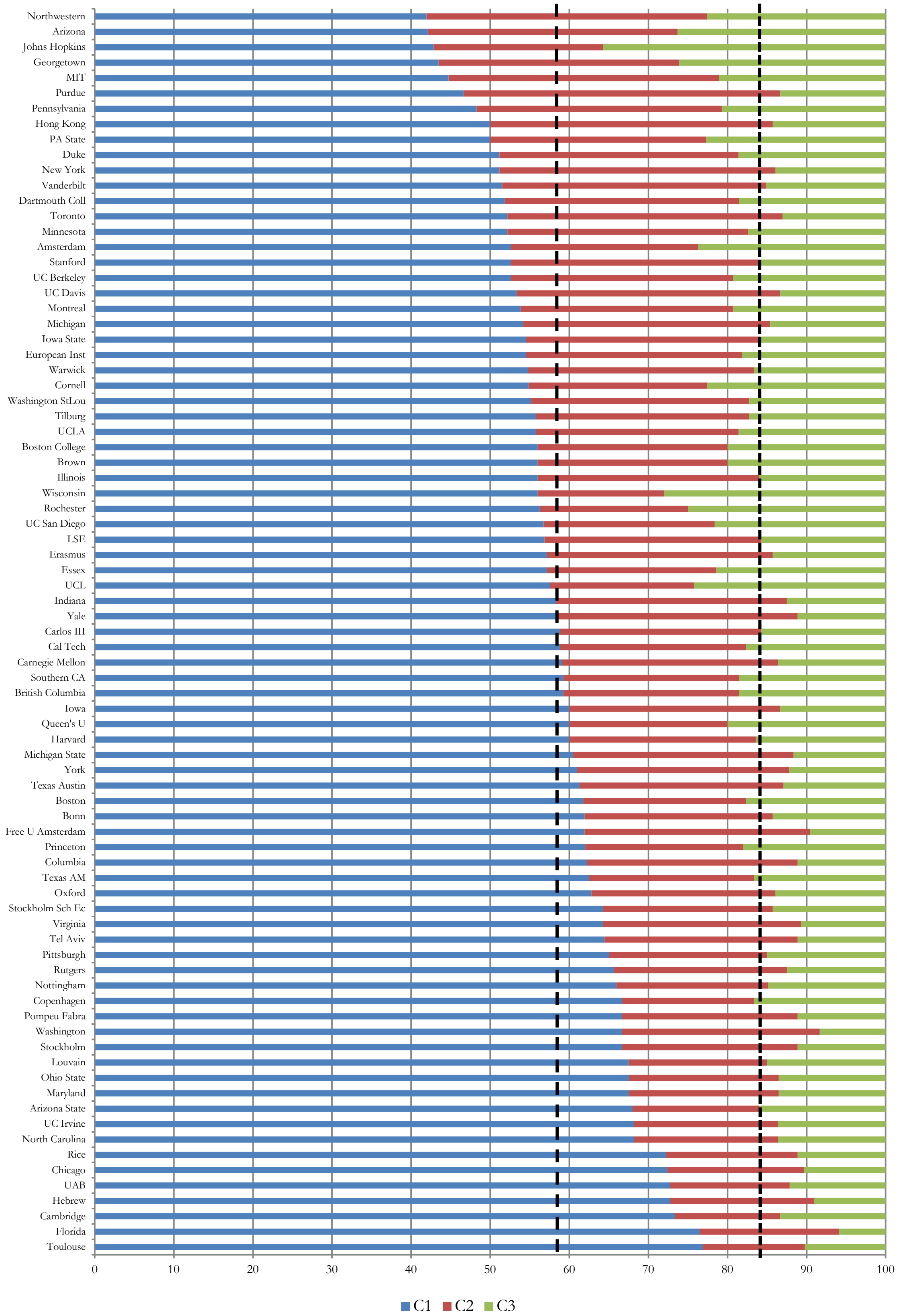


Figure 1. The partition of departments' productivity distributions into three categories according to the CSS technique. Individual productivity = quality index points per year per person (Distribution *Q/Age*)

The conclusion is that, both before and after age normalization, any theory about the interaction between demand and supply forces for researchers must cope with the following two features: large within-department individual productivity variability, and strong differences between department productivity distributions.
Between-department productivity heterogeneity goes against the considerable similarity between: (a) productivity distributions across broad scientific fields, (b) citation distributions across scientific fields at different aggregation levels, and (c) country citation distributions within certain broad scientific fields. As pointed out by a referee, the large doses of between-department heterogeneity may be due in part to statistical fluctuations combined with the relatively small number of researchers by department (see the evidence in this respect in PRRC). Therefore, a natural question to ask is whether the aggregation of departments into countries in our dataset leads us to recover some similarity. As documented in PRRC, this is essentially what we find when we partition the sample into seven countries and a residual category. The conclusion is that a high degree of departmental heterogeneity is compatible with considerably greater country homogeneity.

The above results are necessarily provisional in at least four important respects. Firstly, we conjecture that, at least part of the within- and between-department variability reported in the paper, may very well be due to the fact that the quality of the institutional and personal information provided by our Internet sources is admittedly very uneven and subject to error. Secondly, it should be recalled that the nexus between productivity and age is highly non-linear. Furthermore, Albarrán *et al.* (2014) have shown that this relationship is much weaker for remarkably productive scholars than for the rest of the elite included in our sample. Under these conditions, the simple age normalization used in this paper leaves much to be desired. The residuals of a regression of productivity on age and other control variables might provide a promising avenue for a tailor-made individual adjustment for every individual in the sample. Thirdly, given the skewness of the citation distribution of articles in any journal, including an important percentage with zero citations, Seglen's (1992, 1997) seminal contributions caution us about the wisdom of judging the quality of individual publications –as we have done in this paper– by the citation impact of the journal where they have been published. Therefore, one way to improve upon the results presented in this paper is to introduce productivity measures based on the citation impact directly achieved by each individual publication. Finally, our results only refer to the field of Economics. Before formally modeling the interplay of demand and supply of researchers at the department level, it is advisable to extend the coverage of the issues studied in this paper to other scientific fields.

References

Albarrán, P., J. Crespo, I. Ortuño, and J. Ruiz-Castillo (2011), "The Skewness of Science In 219 Sub-fields and A Number of Aggregates", *Scientometrics*, 88: 385-397.

Albarrán, P., Carrasco, R., and Ruiz-Castillo, J. (2014), "The Elite in Economics", mimeo, Universidad Carlos III.

Crespo, J. A., Li, Yunrong, and Ruiz-Castillo, J. (2013a), "The Measurement of the Effect On Citation Inequality of Differences In Citation Practices Across Scientific Fields", *PLoS ONE* 8(3): e58727.

Crespo, J. A., Herranz, N., Li, Yunrong, and Ruiz-Castillo, J. (2013b). The Effect on Citation Inequality of Differences in Citation Practices at the Web of Science Subject category Level, *Journal of the American Society for Information Science and Technology*, 65: 1244-1256

Diamond, A.M. (1984), "An Economic Model of the Life-Cycle Research Productivity of Scientists", *Scientometrics*, 6: 189-196.

Econphd.net rankings (2004), http://econphd.econwiki.com/rank/rallec.htm.

Li, Y., Castellano, C., Radicchi, F., and Ruiz-Castillo, J. (2013), "Quantitative Evaluation of Alternative Field Normalization Procedures". *Journal of Informetrics*, 7: 746– 755.

Lotka, A. (1926), "The Frequency Distribution of Scientific productivity", *Journal of the Washington Academy of Sciences*, 16: 317-323.

Perianes-Rodriguez, A. & Ruiz-Castillo, J. (2014), "Within and across department variability in individual productivity. The case of Economics", Working Paper 14-04, Departamento de Economía, Universidad Carlos III (http://hdl.handle.net/10016/18470).

Ruiz-Castillo, J., and Costas, R. (2014), "The Skewness of Scientific productivity", Working Paper 14-02, Departamento de Economía, Universidad Carlos III (http://hdl.handle.net/10016/18286).

Seglen, P. (1992), "The Skewness of Science", *Journal of the American Society for Information Science*, 43: 628-638.

Seglen, P. (1997), "Why the Impact Factor of Journals Should Not Be Used for Evaluating Research", *British Medical Journal*, 314: 498-502.

Van Raan, A.F.J. (2005), "Fatal Attraction: Ranking of Universities by Bibliometric Methods", *Scientometrics*, 62: 133-143.

Waltman, L., & Van Eck, N.J. (2013), "A systematic empirical comparison of different approaches for normalizing citation impact indicators", *Journal of Informetrics*, 7: 833–849.

Using bibliometrics in research evaluation and research support - Academic librarians as professional providers of bibliometric services[1]

Sabrina Petersohn*

*Sabrina.petersohn@gesis.org
Computational Social Science, GESIS Leibniz Institute for the Social Sciences, Unter Sachsenhausen 6-8, Cologne, 50667 (Germany)

Introduction

One of the most important research areas within the field of scientometrics is bibliometrics for science policy and management (Glänzel, 2003). The rise of evaluative bibliometrics has been accompanied by a growing number of different types of actors serving the need for quantitative data production and analysis in research evaluation (Gläser & Laudel, 2007). To date, however, there is a lack of systematic empirical evidence on who these users of bibliometrics are and how they relate to and interact with the academic field of scientometrics. Yet the need for knowledge about their professional roles, skills, and competences has been amply demonstrated during debates on the standardization of bibliometric indicators in recent ISSI and STI conferences (Sirtes & Waltman 2013, Wouters, Glänzel & Gläser et al., 2013). The present study aims to fill this empirical gap. It explores an important group of actors who use bibliometric methods in their daily work practice: academic librarians.

Objective of the study and theoretical framework

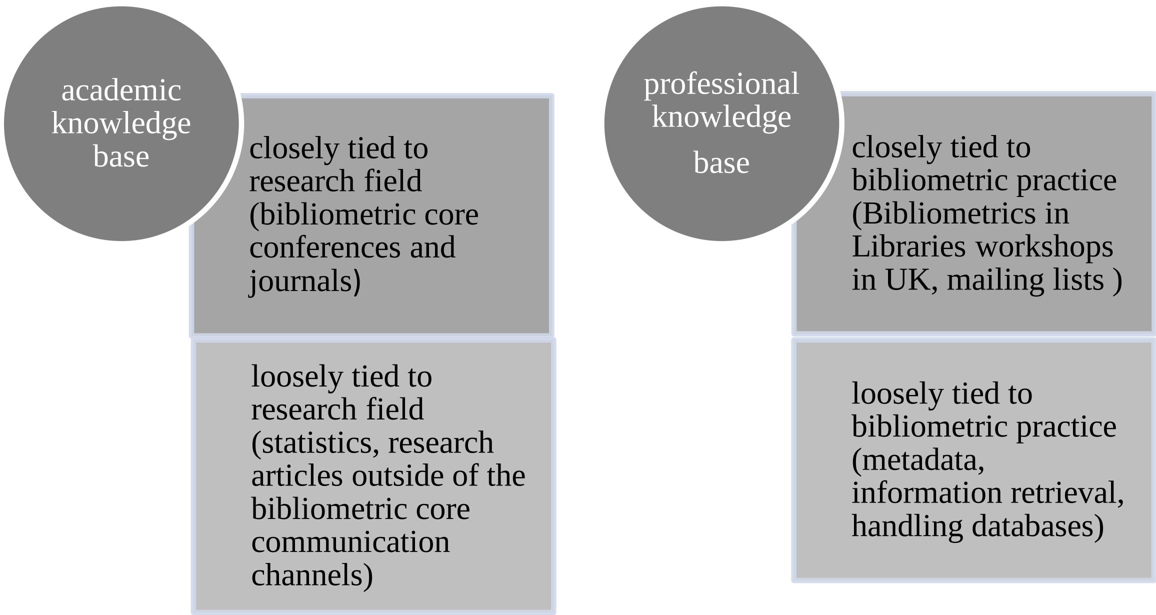
Historically, bibliometrics has been primarily used to facilitate collection development and journal evaluation in academic libraries (Gross & Gross, 1927). While traditional bibliometric applications continue to exist, a shift to using bibliometrics in supporting research assessment has been discerned (Corrall, Kennan & Afzal, 2013). Practitioners and scholars have begun to endorse evaluative bibliometrics as a promising new service area (Ball & Tunger, 2006; Gumpenberger, Wieland & Gorraiz, 2012). Yet, it remains unclear whether the emerging trend to use bibliometric methods in research libraries indicates the development of a stable expert group of scientometric practitioners outside the research field.
The present study aims to assess whether bibliometric services for research support and assessment at research libraries constitute a professional practice. It does so by applying Abbott's sociological theory of professions (Abbott, 1988) as a conceptual framework to research evaluation expertise outside the research field of bibliometrics. To account for differences in national research evaluation and library systems, Germany and Great Britain are studied in a comparative perspective.

A professional task is constituted via a specific framing of the problem by an occupational group and a set of three professional mechanisms (diagnosis, inference and treatment) applied to solve this specific problem. What differentiates occupations from professions is that the latter claim an exclusive right to solve the professional task at hand based on their abstract knowledge and their specific framing of the problem. This cognitive claim of jurisdiction has

[1] This work is supported by grant Nr. 01PY13013B from the German Federal Ministry of Education and Research (BMBF).

to be complemented by a social claim to jurisdiction in the workplace as well as the legal and the public arena. To account for necessary preconditions of the formation of scientometric expertise among academic librarians Abbott's framework has been refined in the light of empirical findings. Next to the abstract academic knowledge base, a professional knowledge base is found to exist. It consists of professional competencies and skills, which are related to varying degrees to the bibliometric core knowledge. The study particularly focuses on these knowledge bases and their interrelations, how they inform the working practices and relate to the cognitive and social claims on bibliometrics made by academic librarians.

Table 1. Knowledge bases related to bibliometrics in academic librarianship



Data and method

To examine bibliometric working practices in research libraries and to assess the type of claims made, expert interviews with 27 British and German information professionals have been conducted. In addition to these interviews, documents such as power point presentations, institutional websites describing the library service and opinion papers in scientific and practitioner journals practitioners have been collected.

The interviews and documents are subject to a qualitative content analysis (Schreier, 2012), a systematic, rule-driven and theory-guided method to unravel underlying themes in texts.

Hypotheses and expected outcomes

The working hypothesis postulates that the professional knowledge base may function as an intermediary knowledge system if the transfer from the academic knowledge base into professional practices is restricted. It is thus expected that academic librarians are not able to put a full jurisdictional claim on bibliometric expertise.

Preliminary findings reveal that the social claim put forward in the public markedly differs from the claim inside the research organization as the workplace which doesn't put as strong an emphasis on the sole capacity of academic librarians to embrace bibliometric services in their research support activities.

Only a part of the experts have received training in bibliometrics such as in the courses offered by the CWTS and regularly consult scientometric core journals. Many of the librarians see their profession in the position to claim jurisdiction on bibliometrics on the basis of metadata knowledge, information retrieval skills and proficient handling of databases. The cognitive claim thus seems to rest more on the wide professional knowledgebase than on the abstract knowledge of the research field scientometrics.

References

Abbott, A. (1988). *The system of* professions. *An essay on the division of expert labour*. Chicago, London: University of Chicago Press.

Ball, R. & Tunger, D. (2006). Bibliometric analysis – A new business area for information professionals in libraries? Support for scientific research by perception and trend analysis. *Scientometrics*, 66, 561-577.

Corrall, S., Kennan, M. A. & Afzal, W. (2013). Bibliometrics and Research Data

Management Services: Emerging Trends in Library Support for Research. *Library Trends* 61, 636-674.

Glänzel, W. (2003). Bibliometrics as a research field. Retrieved October, 17 2012 from http://www.cin.ufpe.br/~ajhol/futuro/references/01%23_Bibliometrics_Module_KUL_BIBLIOMETRICS%20AS%20A%20RESEARCH%20FIELD.pdf.

Gläser, J. & Laudel, G. (2007). The Social Construction of Bibliometric Evaluations. In R Whitley, J. Gläser, J. (eds.), *The Changing Governance of the Sciences. The Advent of Research Evaluation Systems*. *Sociology of the Sciences Yearbook Vol. XXVI.* (pp. 101-120) Dordrecht: Springer.

Gumpenberger, C., Wieland, M. & Gorraiz, J. (2012). Bibliometric practices and activities at the University of Vienna. *Library Management* 33, 174-183.

Gross, P.L. K. & Gross, E.M. (1927). College libraries and chemical education. *Science*, 66(1713), 385-389.

Schreier, M. (2012). *Qualitative content analysis in practice*. London: Sage Publications.

Sirtes, D. & Waltman, L. (2013). Bibliometric Evaluation Standards Debate. In *Hinze, S. & Lottmann, A. (Eds.),* Translational twists and turns: Science as a socio-economic endeavor. *Proceedings of STI 2013 Berlin – 18th International Conference on Science and Technology Indicators*. Berlin: iFQ.

Wouters, P., Glänzel, W. & Gläser, J. et al. (2013). The dilemmas of performance indicators of individual researchers: an urgent debate in bibliometrics, *ISSI Newsletter* 9, 48-53.

Using Google Scholar in research evaluation of social science programs, with a comparison with Web of Science data

Ad Prins*, Rodrigo Costas**, Thed van Leeuwen** and Paul Wouters**

* *info@adprins.nl*
Ad Prins Support in ResearchManagement, Lage der A 31, Groningen, 9718 BL (The Netherlands)

** *rcostas@cwts.leidenuniv.nl; leeuwen@cwts.nl; p.f.wouters@cwts.leidenuniv.nl*
CTWS Centre for Science and Technology Studies, Leiden University, Wassenaarseweg 62A, Leiden, 2333 AL (The Netherlands)

Introduction

In the Netherlands, research evaluation is organized under the combined responsibility of the Association of Universities (VSNU), the National Research Council (NOW), and the Royal Academy of Arts and Sciences (KNAW). Their combined responsibility results in a Standard Evaluation Protocol, (SEP) that describes in detail the organization of research assessment, the various aspects taken care off during research assessments, and the indicators that should be part of the reporting by the committee. In the assessment cycles, that cover a six year time span, including an international assessment as well as an internal mid-term review, peer review is the guiding principle.

In the SEP, the usage of quantitative measures, such as bibliometric indicators is not compulsory, however, in many assessment cycles mainly in the natural, life and medical sciences, bibliometric indicators are introduced to support the work of the review committee. As it is well known that in the social sciences, the humanities, and law, the application of bibliometric indicators is of relatively lesser value, due to the lower degree of coverage of the systems that form the basis for bibliometric analyses (van Leeuwen, 2013), in most of the Dutch assessments in the SSH and Law domains bibliometrics was not applied. In the past, the field of psychology applied bibliometrics, just as the fields of economics and business & management. (Nederhof, 2006) These fields stand out among the SSH and Law domains, as the communication among scholars in these domains has shifted more and more towards journals publications. However, from the SSH domains a strong concern with respect to the design and organization of research assessment has led to the report “Judging research on its’ merits” (KNAW, 2005), which initiated a further thinking among the scholars in these domains on how to further elaborate the preferred way of assessing research in the SSH and Law domains. Two advisory councils were installed, and this led to two reports, one for the humanities (KNAW, 2012), and one for the social sciences (KNAW, 2013). These two reports have strongly influenced the new SEP, that has to be applied from 2015 onwards. An important shift in this new SEP is a lesser focus on productivity, and a wider focus on the impact of scholarly activities, not only in the scientific realm, but also on societal impact.

As the publication cultures differ in the social sciences and humanities (Hicks, 2004, Nederhof et al, 2010, and van Leeuwen et al, forthcoming), impact cannot be established in the regular, journal based electronic databases normally used for bibliometrics (e.g. Web of Science or Scopus). An alternative for the traditional journal-based systems is Google Scholar (hereafter referred as GS). Although this system has been studied before (Harzing 2008, Kousha 2008, 2011), to the best of our knowledge there are no examples of its use in a real life assessment procedure.

In this paper we report on the application of GS based metric in the formal assessment of research programs in the fields of Education and Pedagogical Sciences, and in Anthropology, and on a meta-analysis on the comparison of the results based on Google Scholar and WoS.[v] Finally we discuss some issues with regard to methods in relation to the context of the assessments.

Data and methods

The assignments by deans of the participating faculties to use GS in an evaluative bibliometric context was proposed due to concerns about the representation of SSH outputs in Web of Science (hereafter referred to as WoS)[1].

Education and Pedagogical Sciences (hereafter referred as Ed/Ped) comprised 13 programs of six universities over the evaluation period 2006-2011. Anthropology comprised five programs of an equal number of universities, over the evaluation period 2004-2012. The selection of publications differs slightly for both cases. In the case of Ed/Ped, each program was asked to send in 10 publications per year (60 publications per program). Program directors were asked to send in highly valued or highly cited publications, possibly including also books. In the case of Anthropology, the selection was based on 5 – 10 most cited publications for each year to be evaluated, related to the size of the program. A reduction of numbers of selected publications was chosen for, assuming that small programs are less likely than large programs to produce equal numbers of highly cited publications. Checking for publications that were listed in more than one program as double entries, the resulting numbers of publications were 774 for Ed/Ped (with 6 double entries) and 328 for Anthropology (four double entries).

Data collection for the publications was based on keywords of title and author, allowing for various spellings. As doubts has been cast on the reliability of GS information (Jacso, 2012) information was retrieved for the full second order GS citing data (i.e. enabling a check on the citing sources). The selection criterion was that the citing source should be verifiable, meaning that the source should be traceable in terms of a proper working URL of websites of journals, publishers or other location. Other citing sources, in particular those without proper URL, might still be valid if checked individually, but were nevertheless taken out of the data set. This was also the case with sources with defective data such as improper year of reference in comparison to the publication date of the cited reference. The net certified citations were 22887 (89,8% of gross total of Ed/Ped), and 8092 (89,7% of gross total for Anthropology).

In a further analysis of data quality, performed during the later meta analysis, specific sources of the citing data have been investigated. This analysis was based on the specific URL of each of the citing publications. Information provided by GS is based on the indexes produced by crawling specific internet sources such as electronic academic journals, academic books (Google Books), websites of academic publishers, and internet repositories such as www.jstor.org, www.cairn.info, http://papers.ssrn.com or www.academia.edu. GS indexes also university libraries, as well as academic societies, governments and other sources. The

[1] It is important to mention here the involvement of the research directors of the faculties as stakeholders within their field(s) of expertise in choosing a selection base for the publications to be analyzed.

majority of these sources contain verifiable meta data, either the proper (post print) academic publication itself, or the meta data of pre-prints or the version of record available repositories or university libraries. However, as is also noted by Jacso, some sources such as university libraries may contain also other referring publications, such as theses by PhD's or master theses and repositories might also include conference papers and reports by research institutes. These citing sources might therefore be considered to have a somewhat wider range of reliability in terms of academic status. The meta data were therefore classified, based on the available URL revealing characteristics of the citing source, such as the publisher or university. Classification was possible for the majority of the URL's, as these frequently shared common characteristics such as websites identifiably owned by publishers like Sage, Elsevier, Oxford or Cambridge, or from university libraries. The second order data have been classified as coming from (a) verifiable academic sources (including academic journals, academic publishers of volumes etc. and academic books), (b) university libraries, (c) repositories not identifiable as university libraries, (d) other sources than the above, including academic societies, government sites, blogs and personal webpages of researchers. In a number of cases, the available URL did not share common characteristics, occurring only once in the database. For efficiency reasons these were classified as “unknown”.

Results

Table 1 shows that in both fields the academic citing sources account for more than half of the total volume of cites. Also, the volume of “unknown” sources, which may contain both “exceptional” and customary citing sources, is fairly low (11% and 7%).

Table 10: Sources for citations in GS for Anthropology and Ed/Ped

Source	Anthrop #	Anthrop %	Ed/Ped #	Ed/Ped %
Academic sources	4573	56,5%	14470	63,2%
University Libraries	1677	20,7%	4573	20,0%
repository other than ULs	616	7,6%	1236	5,4%
Other Sources	327	4,0%	1085	4,7%
Unknown	899	11,1%	1523	6,7%
Total	8092	100,0%	22887	100,0%

The results also show that the fields of Ed/Ped and Anthropology differ with regards to the document types of the cited output. Whereas in Ed/Ped the share of journal articles in the total set of publications is almost 90%, in the case of Anthropology the share of journal articles is 58%, with higher percentages for books, volumes and chapters. Differences in publication cultures are even more apparent in the volume of citations per publication types. Whereas in the case of Ed/Ped journal articles on average are the most cited publication type, in Anthropology books are the more cited type (table 2). These differences are not due to a single or a few outliers, as figures 1 and 2 show the median volume of cites per document type is higher for books in Anthropology and higher for journal articles in Ed/Ped. Even though books in Ed/Ped receive considerable attention, for the Anthropology programs books are both important forms of output as well as important means for receiving scientific impact.

Table 11: Citations per cited document type

	Anthropology			Ed/Ped		
	Cites #	Pubs #	Cites p publ	Cites #	Pubs #	Cites p publ
Books (Monographs)	1818	44	41,3	700	28	25,0
Chapters in volume	613	38	16,1	788	42	18,8
Journal articles	3885	187	20,8	21256	695	30,6
Other	357	19	18,8	109	7	15,6
Edited Volumes	1419	39	36,4	34	2	17,0
Total	8092	327	24,7	22887	774	29,6

Figure 1: Distribution of cites per document type in Anthropology

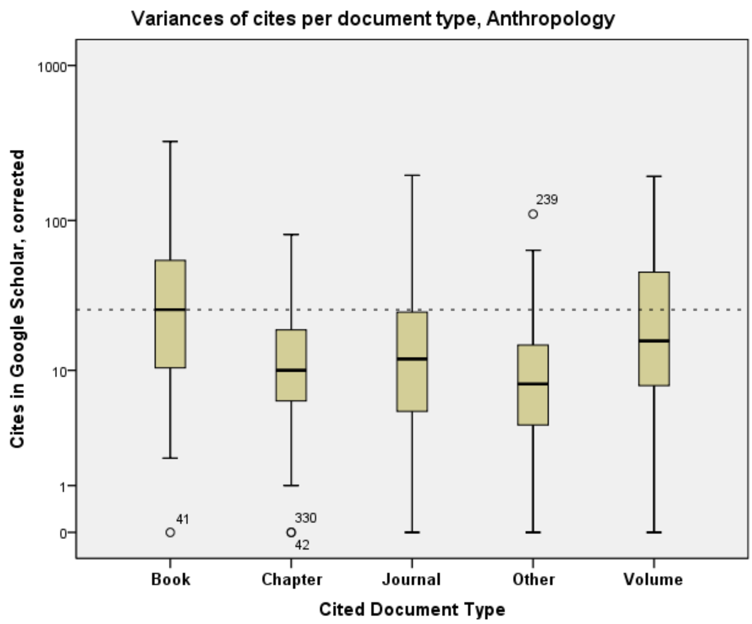
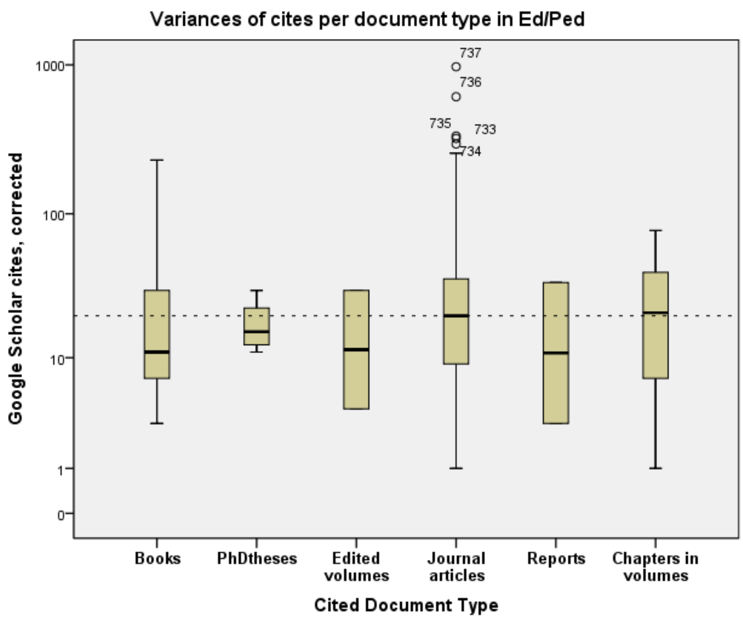
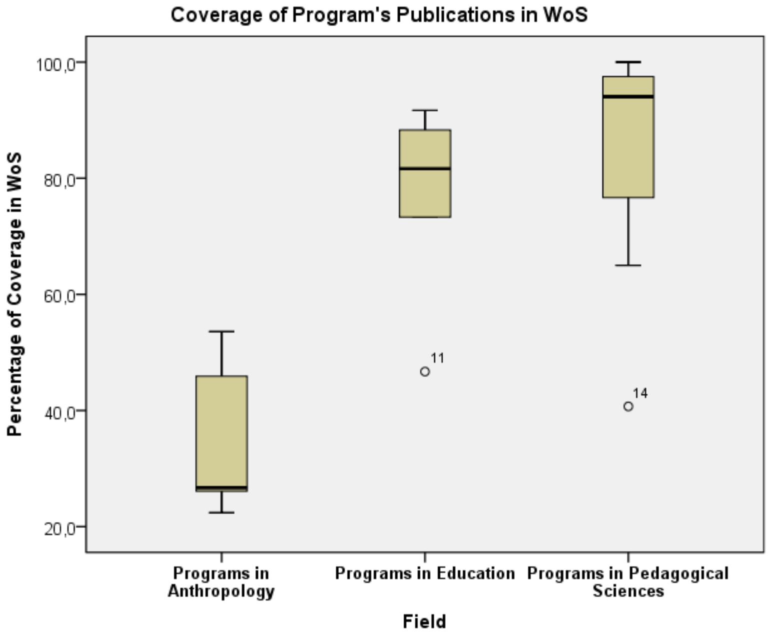


Figure 2: Distribution of cites per document type in Education and Pedagogical Sciences



A comparison of the results for GS with data retrieved from WoS shows large differences for the studied publications in their coverage. Whereas over 80% of the publications in most programs in Ed/Ped were published in journals covered by WoS, this percentage fell to an average of 37.5% for the programs in Anthropology. Relatively lower coverage were also noted for the two programs in Ed/Ped on theory, history and philosophy in Ed/Ped, represented as extreme values with O in figure 3. (figure 3)

Figure 3: WoS Coverage of publications per program in three fields



A further comparison of citations from GS with those from WoS indicate higher levels of citation information provided by GS for both fields. This is also true if only GS citations from identifiably academic sources, such as academic journals, publishers and books are considered (Table 3).

Table 12: Total cites in Google Scholar and Web of Science for two fields

Field	Cites in Google Scholar	Cites in Google Scholar, Academic sources	Cites in WoS
Anthropology	8092	4573	1097
Education & Pedagogical Sciences	22887	14470	8870

The numbers of citations per publications show a fair correlation for GS and WoS both in Ed/Ped and Anthropology. However, the correlation for Anthropology is based on a strongly reduced set, as only 37.5% of the publications were covered in WoS. In figures 4 and 5, the individual programs are correlated and the differences observed among them lead to the conclusion that programs vary with regard to how their citations are calculated based on GS or WoS.

Figure 4 Scatterplot WoS and GS citations for programs in Anthropology

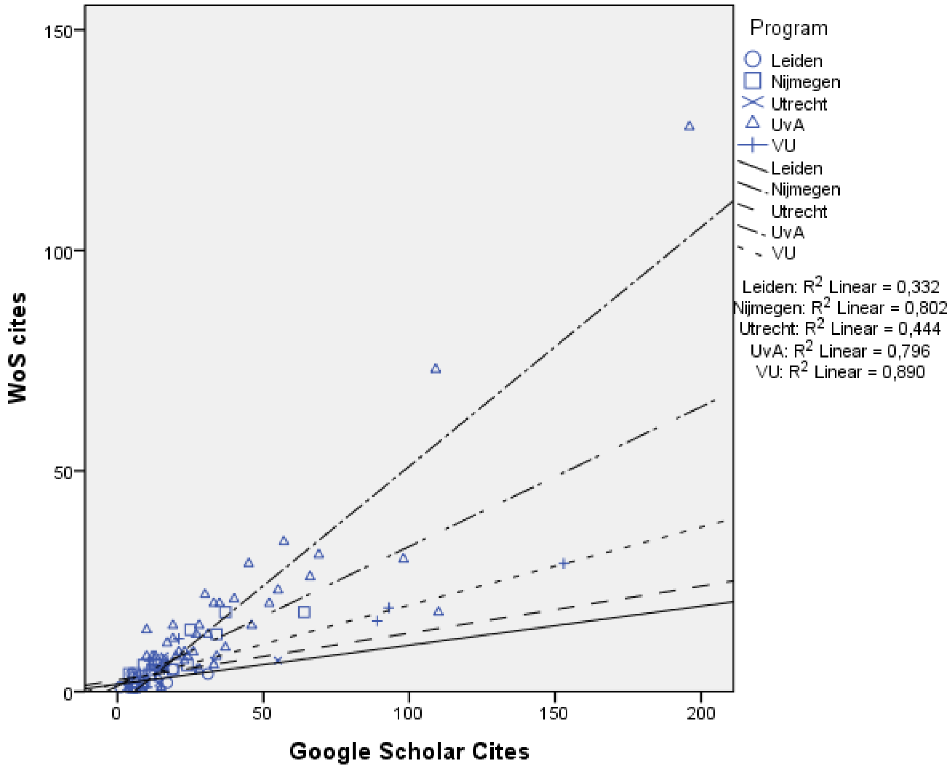
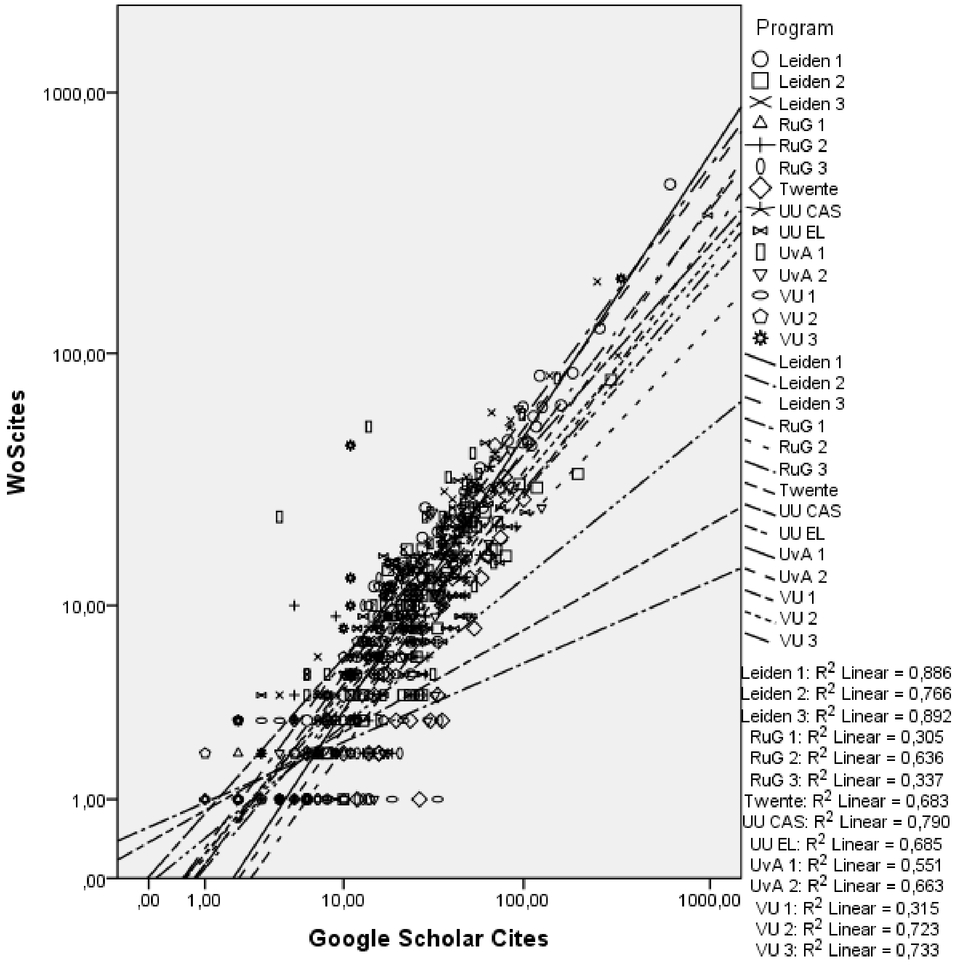


Figure 5: Scatterplot WoS and GS citations for programs in Ed/Ped



Discussion

Our results indicate that indeed it is possible to perform bibliometric studies for evaluation purposes using GS, both with regards to data collection and data reliability, once data are based on selected publications and cleaned for erroneous data.

The comparison of GS results with WoS results indicates that it is fruitful to use GS for fields with lower degrees of coverage in WoS (Van Leeuwen, 2013), in particular fields that produce more diverse types of output than articles in journals included in WoS. As we show, in Ed/Ped and even more so in Anthropology other types of publications are important means of communication, receiving considerable impact according to GS which is missing in WoS.

In contrast to claims by critics of GS that the results are very unreliable (Jacso, 2012), the information in GS, once retrieved on the basis of existing publication data and cleaning of citation sources, indicate acceptable levels of reliability in terms of source. Also, the volume of information that can be retrieved for Anthropology increases considerably to levels comparable with the results for programs in Education and Pedagogical Sciences. There are however several issues regarding how GS results are to be used in the context of assessments.

1. Workload and data limitations

In contrast to WoS, GS data are to be retrieved with quite considerable effort, in particular if the analysis is to be based on second order data. These data are essential in establishing the traceability of citations and the source of the citation. Recent limitations of search results to 20 per query, imposed by the GS engine contributes to this situation. The workload thus imposes limitations to how many publications can be investigated, and influences the design of robust bibliometric analysis, since labor intensive studies are costly. Also, although GS indexes are based on the available meta data of publishers and repositories, including page numbers, issues, author lists and journal title, few of this information is provided to the end user of GS, thus making the correct identification of the publications harder. WoS is in this respect a more precise source, be it that its precision is not very relevant for fields such as Anthropology.

2. Possibilities for field normalization

One important limitation is that GS provides as yet very limited opportunities for field-normalized indicators (Wouters & Costas, 2012). For this study, attempts have been made for GS based field normalization in part based on PoP (Publish or Perish) data for journals (Harzing, 2008) (data not shown). Although technically feasible, these attempts are as yet rather unsatisfactory since the data were based on averages of citations per paper per year, whereas the selection base comprised highly cited papers. Also, using tools for journal data such as PoP does not allow for a traceability check as performed in this study. Even though the percentage of non-traceable citations was small, the comparison might still be biased. In the case for Ed/Ped Sciences attempts have been made to include information of other sources such as Scimago Journal Rank (SJR) (SCImago, 2007). This is possible, but it leads to complicated procedures and methodological issues.

3. The definition of citations

Using GS implies a shift in the definition of what may count as a citation. Whereas the citations in WoS are based on references in academic journals (and increasingly also in other academic sources), the criterion is the academic nature of these forms of publications as established in this database. In GS however, citations also may include references from scientific reports, PhD theses and also student theses. Once using GS, the results inevitably include these citations too, of which one may argue that these suffice or not as tokens of academic recognition. Whether this shift is accepted in view of changing views about assessment standards such as in the new SEP remains to be seen.

4. The selection base

The selection base for publications to be analyzed is obviously relevant to the results and to the methods to be used. In the Ed/Ped case, the selection was performed by program leaders. This led to a selection of highly cited papers but included also publications that were most likely deemed very relevant to the program, but possibly not highly cited. As an indication, 4.5% of the selected publications were not found to be cited at all in GS. This situation of inefficiencies in selecting research outputs has been also observed elsewhere (Abramo et al, 2014). More importantly, although technical issues - such as workload - impose limitations in selecting higher volumes of publications, the selection base is crucially related to the questions to be addressed in the assessment.

5. Questions to address in the assessment
The shifting ideas about assessment goals for research programs may lead to more variegated bibliometric questions, which in turn may require different research designs such as a focus on specific publications typical for the mission of institutes, or contextual bibliometrics. As the precision and transparency of data shows limitations and the workload is also high, the use of GS will impose restrictions to the possibilities to answer the desired assessment questions.

Conclusions

One of the crucial factors for not applying bibliometry in the social sciences is the coverage of output in Web of Science, which is low for fields such as Anthropology and mediate for Education or Pedagogical Sciences. There is reasonable evidence to consider GS as a valuable source for the analysis of certain fields of science, particularly in the Social Sciences and perhaps also in the Humanities, in providing more information based on a broader set of publication types. However, attention should be given to data reliability. To use GS in the context of evaluation, various ways for benchmarking or field normalization have to be worked out, for instance on the basis of available journal data, to address the issues of research assessments. These are not only technical problems, but they are also issues dependent on the questions raised in assessments. Moreover, the application of GS may find important limitations in fields that rely on even higher volumes of non-journals sources than in the case of the current programs in Anthropology.

References

Abramo, G., D’Angelo, C.A. & Di Costa, F. (2014). Inefficiency in selecting products for submission to national research exercises. *Scientometrics*. 98, 2069–2086

Harzing, A. & Van der Wal, R. (2008) Google Scholar as a new source for citation analysis, *Ethics in Science and Environmental Politics*, 8, 61-73.

Hicks, D. (2005) The Four Literatures of Social Science. In Moed, H. et al, *Handbook of Quantitative Science and Technology Research*. Springer Netherlands, 473-496.

Jacsó, P. (2005). Google Scholar: the pros and the cons. *Online information review*, *29* (2), 208-214.

KNAW (2005) Judging research on its’ merits, Amsterdam

KNAW (2013) Towards a framework for the quality assessment of social science research, Amsterdam.

KNAW (2012) Quality indicators in the humanities, Amsterdam

Kousha, K. & Thelwall, M. (2008) Sources of Google Scholar citations outside the Citation Index: A comparison between four science disciplines, *Scientometrics*, 74 (2), 273-294.

Kousha, K., Thelwall, M., & Rezaie, S. (2011). Assessing the citation impact of books: The role of Google Books, Google Scholar, and Scopus. *Journal of the American Society for Information Science and Technology*, *62* (11), 2147-2164

Nederhof, A.J. (2006) Bibliometric monitoring of research performance in the social sciences and the humanities: A review. *Scientometrics* 66 (1), 81-100.

Nederhof, A. J., van Leeuwen, T. N., & van Raan, A. F. (2010). Highly cited non-journal publications in political science, economics and psychology: a first exploration. *Scientometrics*, *83*(2), 363-374.

SCImago. (2007). SJR — SCImago Journal & Country Rank. Retrieved between September 10 and October 10, 2012, from *http://www.scimagojr.com*

Van Leeuwen, T.N. (2013) Bibliometric research evaluations, Web of Science and the Social Sciences and Humanities: a problematic relationship?, Bibliometrie – Praxis & Forschung, 2013, http://www.bibliometriepf.de/article/view/173

Van Leeuwen, T.N., Van Wijk, E. & Wouters, P.F. (Forthcoming) Bibliometric analysis of output and impact based on CRIS data: A case study on the registered output of a Dutch university.

Wouters, P., & Costas, R. (2012) Users , narcissism and control – tracking the impact of scholarly publications in the 21 st century. In M. Van Berchum, & K. Russell (Eds.) Retrieved from http://sticonference.org/Proceedings/vol2/Wouters_Users_847.pdf

[v] The first part of the actual bibliometry has been performed by Ad Prins. The second part has been performed by all present authors.

Methodological Proposal for Evaluating Performance of Social Science Researchers in Mexico[1]

Cristina Restrepo-Arango* and Jane M. Russell**

*crestrepoarango@gmail.com
Programa de Posgrado en Bibliotecología y Estudios de la Información. Universidad Nacional Autónoma de México, Unidad de Posgrado, Edificio H, 1er. Nivel, H-104. Circuito de Posgrado, Ciudad Universitaria. Coyoacán, México D.F. C.P. 04510 (Mexico)

**jrussell@unam.mx
Instituto de Investigaciones Bibliotecológicas y de la Información, Universidad Nacional Autónoma de México, Unidad de Posgrado, Edificio H, 1er. Nivel, H-104. Circuito de Posgrado, Ciudad Universitaria. Coyoacán, México D.F. C.P. 04510 (Mexico)

Introduction

Many questions have been raised in the specialist literature as to the use of bibliometric techniques to assess the research outputs of social scientists especially those from the developing world. Although bibliometric indicators have the advantage of objectivity the fact that the Web of Science is the primary source for data collection implies a bias towards journals published in English over other languages, such as Spanish (Etxebarria & Gómez-Urunga, 2010; Mali, 2010). Also, this form of metric evaluation does not take into account that the academic career of a researcher is defined by many variables related to his/her "cultural capital" linked to professional prestige and peer recognition (Luz, 2005). Therefore, for the social sciences the use of bibliometrics is recommended in combination with qualitative techniques that take into account "the research object, the applied methodologies and the academic communication structure" (Moed, Luwel, & Nederhof, 2002). These aspects make the social sciences clearly different from the natural and applied sciences (Luwel et al., 1999), so much so that in the social sciences the book is preferred means of communication (Etxebarria & Gómez-Urunga, 2010) as well as conference proceedings. Therefore, it is necessary to reconsider evaluation methods, as well as the variables that are regularly taken into account in the social science fields.

For a country such as Mexico where social science research is essential for economic and social development, there is a pressing need for an alternative method of evaluation; a model where scientific production and citations in mainstream journals are just two elements in assessing scientific performance and not the central measures as in the natural sciences. Our proposal integrates variables of scientific production with three other groups of variables: teaching activities; professional involvement; and scholarly recognition to build a holistic view of scientific achievement in the social sciences taking into consideration that research is a social process inserted in a specific cultural context and influenced by the behavior of the individual researchers. Our main research question is what are the relative impacts of the four main groups of variables on the scientific performance of a researcher in the Mexican social sciences? The overall aim is to propose a model for assessing scientific performance and apply it to a selected sample of social science researchers in Mexico. As far as we know, this

[1] This work was supported by a grant for the first author from the National Council of Science and Technology (CONACyT), Mexico.

is first time such a model has been proposed to measure the performance of this group of scientists.

Methodology and Data

The target population consists of 2,038 researchers in Sociology, Economics, Political Science, Law and Demography who are members of the Social Sciences Area V of the Mexican National Researchers System (SNI)[2]. Membership is a way of rewarding academic performance and is given at different hierarchical levels: Levels I, II, III and Emeritus. A stratified random sample of 227 (10%) researchers at the different levels of membership and fields consisted of: Sociology (81), Economics (62), Political Science (39), Law (38), and Demography (7).

The data were collected from the CVs requested via email. Each set of criteria was assessed according to the group of variables in Table 1. Criteria were constructed on the basis of previous studies as well as preliminary analysis of a representative sample of CVs.

The occurrence of each criterion included in the sets was calculated. Then, based on the frequencies and ranges of these scores, points were assigned and a data matrix created for each criterion extracted from the CVs. SPSS software version 20 for Windows was used to process the data. This was done in order to test the modeling of the data which was associated with the following problems: high values in the two highest SNI categories (Level III and Emeritus) and low values in Level I, in other words researchers who had published books and those who had not; null values that need to be statistically transformed to run properly. Analysis of results of multivariate techniques by multiple correspondence analysis and principal components analysis is under evaluation.

Our research in progress is in the data collection phase and is 90% complete. This phase has been slow for several reasons. Firstly, not all researchers responded to the request to provide their CVs making it necessary to constantly update the sample. Lack of uniformity in the presentation of CVs was another drawback, not all of them included comprehensive information on all four groups of variables and not all were entirely up to date. Each CV was carefully examined for any mention of activities relating to the individual criteria from the start of their scientific careers. Public access to unified formats of researcher's CVs as occurs in other Latin American countries such as Brazil, would have reduced certain of the limitations associated with the use of CVs as the data source.

Table 1. Proposed Criteria and Variables

Groups of Variables	**Sets of Criteria***
Scientific production	Author or co-author scholarly book.
	Author or co-author scholarly book chapter.
	Compiler, editor, scholarly publication.
	Article in foreign or national refereed scholarly journal.
	Article in foreign or national non- scholarly journal.
	Article in foreign or national newspaper.
	Other publications (brochures, reports, etc.)
	Book reviews.
	Preface, foreword, introduction of scholarly book.
	Translations of scholarly books or articles.

[2] http://www.conacyt.gob.mx/sni/paginas/default.aspx

	Paper presented at foreign or national event.
Teaching activities	Teaching at secondary, preparatory, specialization, undergraduate, master´s and Ph. D. levels.
	Thesis director at specialization, undergraduate, master's and Ph. D levels.
	Thesis committee at undergaduate, master's and Ph. D levels.
	Reader of thesis at undergraduate, master's and Ph. D levels.
	Coordinator and organizer of events, congress, seminars, etc.
	Visiting professor or researcher.
Professional involvement	Consultant in others institutions.
	Administrative post or member of institutional committees.
	Membership of professional bodies (scientific societies, professional associations, etc.)
	Reviewer of scholarly journal article, book chapter, book and research project o research grant.
	Member of editorial board.
	Member of research group or research project.
	Commentator, chair, book presentation, etc. in event.
	Interviewed in radio or television program.
	Invited speaker in foreign or national event.
Scholarly recognition	Awards presented for academic papers.
	Awards for distinguished service (emeritus, distinctions, etc.)
	Tributes for academic career.
	Research grants obtained by public competition.
	Scholarships (graduate, postdoctoral and research).
	Honorable mentions and academic incentives.
	Doctorate Honoris Causa.

* The occurrence of each criterion was obtained from the researchers CVs. For example, author scholarly book was counted separately from co-author scholarly book; these were grouped together in the table to economize space.

Preliminary Findings

Our preliminary data reveal important differences between researchers in the various category levels of the SNI. For example, the emeritus members surpass the other categories in the number of scholarly books, on average 14 at this level; also in articles published in national newspapers, on average 206 and in scholarly recognition, average of 4. Researchers in Level II stand out in the number of articles published in peer-reviewed foreign scholarly journals with an average of 12; number of undergraduate courses taught, average of 14; number of courses at doctoral level, average of 8 and the supervision of doctoral theses, average of 8. The averages of these indicators for Level II researchers are slightly lower than those of Level III except for those criteria where the emeritus members excel. Researchers in Level I show lower percentages in all criteria.

We also found differences with respect to the fields of the researchers. For example, those in Law stood out in the average number of scholarly books published, 6 in all; in Economics, the most salient finding was an average of 6 articles in foreign peer-reviewed scholarly journals while those in Political Science published on average 40 newspaper articles. Most undergraduate courses were given in the three areas of Political Science, Demography, and Economics, with an average of 12 courses each. At postgraduate level Demography came out top with 11 doctoral courses and 23 at master's level. Political Science, and Sociology supervised the greatest number of undergraduate theses, both with 8. Scholarly recognition was most prevalent in Law, average of 4 per researcher.

In general, the examination of the data collected up until now shows differences with respect to the levels and areas of the SNI, in particular those in Level I who are starting out on a research career show lower average levels of all activities than their peers in the three other categories. Level III scientists are mature researchers with high productivity in all the variables studied and aspiring to reach Emeritus level. Those already with Emeritus status are more concerned with publishing scholarly books, probably as a means of establishing their authority and consolidating their careers. Not surprisingly this group has the highest average in the scholarly recognition category.

References

Etxebarria, G. y Gómez-Urunga, M. (2010). Use of Scopus and Google Scholar to measure social sciences production in four major Spanish universities. *Scientometrics,* 82 (2), 333-349.

Luwel, M., Moed, H. F., Nederhof, A.J., De Samblanx V., Verbrugghen, K., & van der Wurff, L.J. (1999). Towards indicators of research performance in the social sciences and humanities: An exploratory study in the fields of law and linguistics at Flemish Universities. Brussels: Waamse Universitaire Raad.

Luz, M. T. (2005). Prometeu Acorrentado: análise sociológica da categoria produtividade e as condições atuais da vida acadêmica. PHYSIS: Revista Saúde Coletiva, 15 (1), 39- 57.

Mali, F. (2010). Policy issues of the international productivity and visibility of the social sciences in central and eastern European countries. Sociologija I proctor, 48 (3), 415-435.

Moed, H.F., Luwel, M. & Nederhof, A. J. (2002). Towards research performance in the humanities, Library Trends, 50(3), 498-520.

Analysing human resources and knowledge production in Colombian universities[1]

Sandra Rivera [*], Marcela Galvis [**]

[*] *crivera@ocyt.org.co*
Human Resources in Science and Technology, Colombian Observatory of Science and Tecnology, Carrera 15 #37-59 B, Bogotá, (Colombia), Facultad de Ciencias Económicas, Universidad Nacional de Colombia

[**] *mgalvis@ocyt.org.co*
Human Resources in Science and Technology, Colombian Observatory of Science and Tecnology, Carrera 15 #37-59 B, Bogotá, (Colombia), Friedrich Schiller University Jena, Germany

Introduction

Researchers' scientific profile differs widely among different research fields, scientific communities and the institutions they are part of. These differences determine the scientific production and the character of the scientific communities built by these researchers (Dietz & Bozeman; 2005). The aim of this paper is to characterize the human resources working in research groups in 81 universities in Colombia[2], in terms of the variables that define their scientific profile (socio-demographic variables, skills, education level and institutional links) compared to their scientific production.

To achieve this goal, we propose to define and characterize a dataset of Colombian researchers, based on the Curriculum Vitae (CV) registered in the ScienTI platform. After, we build a profile using demographic variables, scientific field and scientific production for each researcher[3]. Then, using a sequence diagram, we identify a model of structural equations which allows to include multiple cross-dependency relationships in one equations (Bollen, 1998).

In addition, we intend to analyse the network of relationships of the researchers in order to represent the scientific community working in universities and their evolution in the last decade, which is determined by the profile of the human resources working in research groups defined by the structure of scientific work in the S&T system recognized in Colombian institutions. Preliminary results show that the patterns of production differ widely among subjects and profiles related to the scientific trajectory of the researchers.

Conceptual framework

Cañibano and Bozeman (2009) review different approaches to analyse human capital and suggest to overcome the "product paradigm" and to focus on the "capacity paradigm". In this paper, we subscribe to the capacity view of human resources in order to understand the

[1] This work was supported by the Colombian Observatory of Science and Technology and the doctoral thesis "Strategies of knowledge generation in Colombian universities", financed by Colciencias and the National University of Colombia.
[2] A research group is defined as "a group of people who do research and jointly generate knowledge in one or several topics, according to an agenda in the medium to long terms" (Colciencias, 2013)
[3] We only take into account researchers who have had scintific production in the last two years, in this case they are called "active researchers"

characteristics of the Colombian scientific communities inside universities using indicators of the scientific and academic trajectories, understood as the set of processes and relationships between the individuals and the institutions, which are mediated by the research groups as a specific form for scientific work, adopted by Colombian universities.

In this sense, this research in progress paper integrates elements from the capacities approach and the study of the researcher's academic careers, using the concept of the life cycle in order to evaluate the determinants of the scientific production and the communities that are established to facilitate this process. The capacities approach proposed by Cañibano & Bozeman (2005) emphasises the interaction between agents, processes, organizational forms and contexts, complemented with the analysis of the curriculum vitae (CV).

Another aspect that must be taken into account when evaluating scientific capabilities is the cumulative nature of knowledge, since, as pointed out by Furman & Stern (2006, page 1), "(…) the cumulative nature of knowledge production has been recognized as central to the process of economic growth"

The analysis of scientific careers is a multivariate process determined by the human capital capacities and the scientific knowledge generation (Cañibano & Bozeman, 2009). The use of the CV allows to identify the researcher´s ways of working, and methodologically it provides information about different dimensions of the researchers' activities; it allows to analyse individual capacities through its scientific background, especially through the ties and the interactions established with peers, creating collective capacity (Cañibano, et al., 2008; Lepori & Probst, 2009).

In Colombia previous studies have found that the development of scientific and academic communities, has to take into account the existence of interactions among researchers, institutions, knowledge and partners (Jaramillo & Forero, 2001; Villaveces & Jaramillo, 2004; Villaveces & Forero, 2007; Bucheli et al. 2012).

Structural equations models
Structural equation models (SEMs), are multivariate regression models. They differ from other multivariate models because in a SEM the response variable in one regression may be a predictor in another equation; these structural equations are meant to represent causal relationships among the variables in the model (Fox, 2002). SEM is used to test relationships between observed (measured) and unobserved (latent) variables and also relationships between two or more latent variables.

Boomsma (2000) recommends a way to structure papers that use the tools of SEMs in the analysis. We intend to implement this model in the analysis; the process starts with a substantial problem that should have a theoretical framework and a population under study. In our case, the theoretical framework is given by the theories of human capital, capacity building and life cycles, explained before, and the population under study corresponds to researchers working in one of the 81 universities and whose CV is registered in ScienTI platform.

After that, a set of models has to be proposed and a raw sample data is determined out of the population under analysis. Then the model's characteristics and measurements and a data characteristics sample matrix should be defined, later, the model estimation procedure is structured. Finally there should be a model selection and evaluation followed by conclusions and discussion. This research in progress paper only incorporates these first two stages of the process, e.g. until the definition of the model.

Methodology

This work follows four phases:

1. Definition and characterization of a dataset of Colombian researchers based on the Curriculum Vitae (CV) registered in the ScienTI platform. 20.740 researchers (Rivera et al, 2013).
2. Identify institutional links and knowledge products derived from their activities.
3. Identify collaboration using co-authorship analyst registered in database ScienTI during 2002-2012;
4. Identify the model of structural equations: identify dependent and independent variables and build a sequence diagram to show the relationships among them.

The variables extracted from the CV are shown in Table 1.

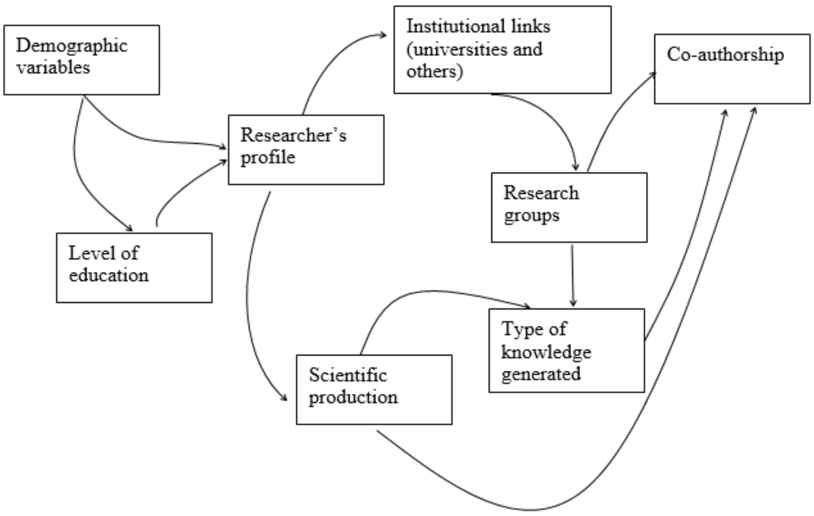
Table 1. Variables extracted from the CV.

Information	**Contents**
Education level	Trajectory of education
	Classification of each level by research fields
	Type of institution
Activities	History of activities: including research, teaching, consulting, administration, technical support, etc.
Scientific production	Type of production: New scientific knowledge, New technological knowledge, Thesis advisory, Social appropriation of knowledge
	Classification by language and country of publication
Institutional links	Links of each researcher to universities
	Links of each researcher to groups

Source: author's proposal based on Colciencias (2013)

Some observed variables include the level of scientific production, the maximum level of education achieved by the researchers, demographic variables such as age and sex, the institutional links (groups and universities) and co-authorship. We use a sequence diagram to show the theorized relationships in the model.

Diagram 1: sequence diagram of the relationships in the model



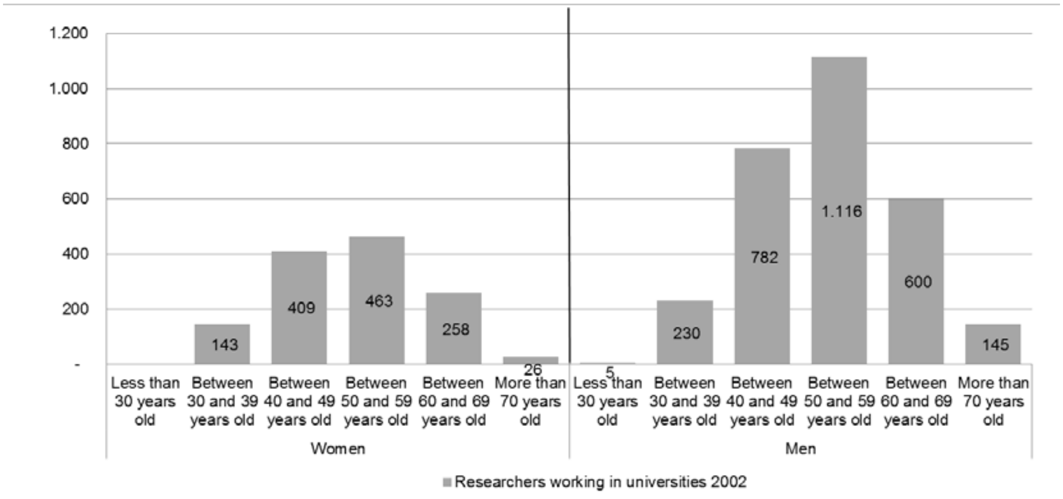
Source: author's proposal

Data
As stated before we use a dataset of researchers working in Colombian universities based on the Curriculum Vitae (CV) registered in the ScienTI platform, along 2000-2011. We found 20.740 active researchers in this period. The population is characterized in Figure 1, results are depicted for 2002-2006-2010 to show how the population evolves as well as the age and sex composition. The size of the population differs each year, depending on the activity of the researchers. It includes 6.743 researchers in 2002, 12.909 in 2006 and 15.539 in 2010.

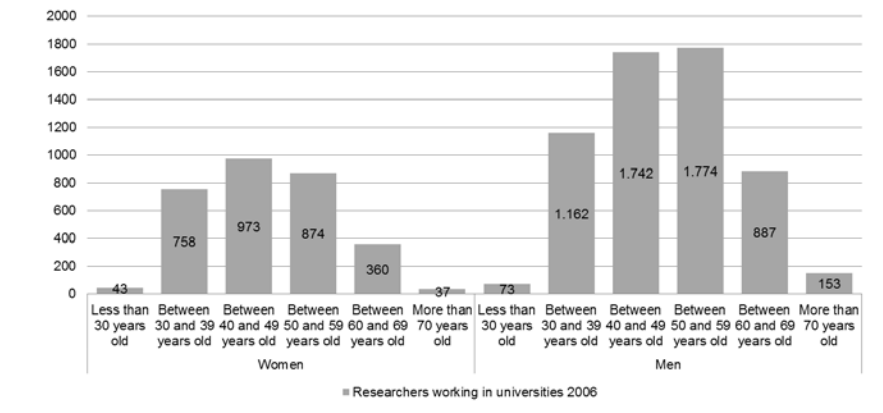
The age composition of the population changes along the period of analysis, Figure 1 shows that the population is becoming younger. Also, the sex composition is changing over the period. As a result, even though knowledge is cumulative, the researchers that support this knowledge change over time, in this sense the new patterns of education, skills and institutional links are a characteristics of universities that determine the scientific production as will be seen in the next graphs.

Figure 1: Evolution of the Population by age range and sex

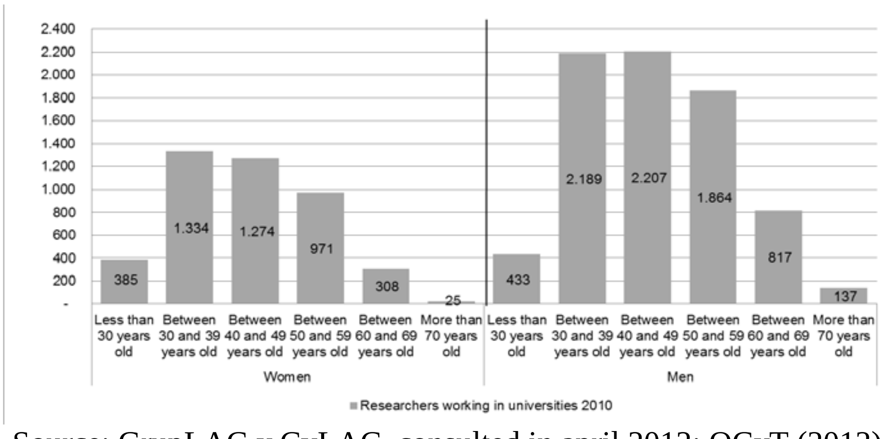
a. Year 2002.



b. Year 2006.



c. Year 2010.



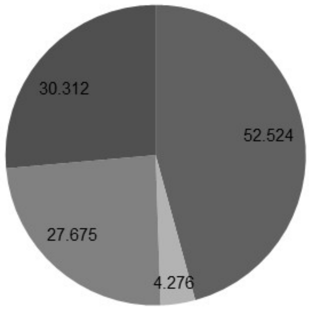
Source: GrupLAC y CvLAC, consulted in april 2012; OCyT (2012). Calculations: OCyT.

When considering the percentage distribution of the researchers by research fields; we find that most of the researchers are working in Social sciences (26,51%), compared to 20,64% in Natural sciences, 18,62% in Humanities, 12.63% in Medical and health sciences, 11,44% in Engineering and only 5,89% in Agricultural sciences.

Preliminary Results

The data show that the researchers differ in their age composition, sex and research field. The notion of trajectory helps explain these differences in terms of the life cycle theory. The total number of scientific results produced by the researchers is 114.787, they can be classified in four categories, defined by the measurement model applied by Colciencias (Colciencias, 2008)[4].

Figure 2: Scientific production of the researchers, by type 2000-2011*

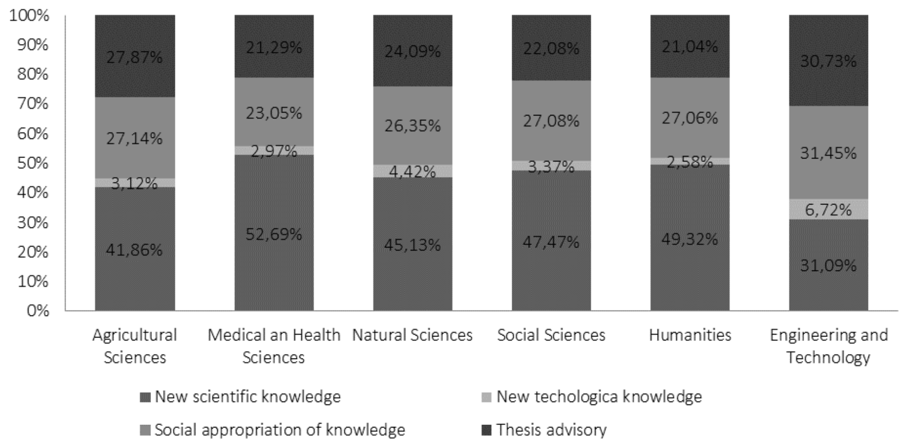


Source: GrupLAC y CvLAC, consulted in april 2012; OCyT (2012).
Calculations: OCyT.
* The number of researchers reporting each type of product is distributed like this: thesis advisory, 4.107 researchers; social appropriation of knowledge, 5.750 researchers; new scientific knowledge, 9.230 researchers and new technological knowledge, 1.387 researchers.

Most of the production is concentrated in the first category. However, the patterns differ when we look deeper into the data. The scientific field (in this papers we selected OECD fields) is one of the variables that helps to build the profile of the researchers, as shown in Diagram 1, we propose that the scientific production is explained by the researcher's profile which at the same time is influenced by demographic variables and the level of education. This hypothesis is founded in the data shown in the next figures. Regarding the research fields, engineering has the smallest proportion of new scientific knowledge among its results, favouring technological knowledge and thesis advisory. On the contrary, Medical and health sciences show the highest proportion of new scientific knowledge in its results (Figure 3).

[4] New scientific knowledge includes Articles, Research books, Research book chapters, Working papers and other related products. New technological knowledge includes products such as Industrial designs, Patents, Pilot plants, Industrial processes, Prototypes, etc. Thesis advisory includes the advisement of undergraduate, master's or PhD thesis. Finally, Social appropriation of knowledge includes the participation (as speaker) in events such as Conferences, and Seminars, and the writing of articles to communicate science in newspapers or other means. For a detailed description see Colciencias "Modelo de Medición de Grupos de Investigación Científica, Desarrollo Tecnolóigco e Innovación Año 2008" (Measurement model of research groups, year 2008) http://web.www3.unicordoba.edu.co/sites/default/files/anexo_1_modelo_medicion_de_grupos_2008.pdf. Currently the model is changing, and since 2013 the individual researher is given more weight in the measurment, before, the model privileged colective results produced in research groups.

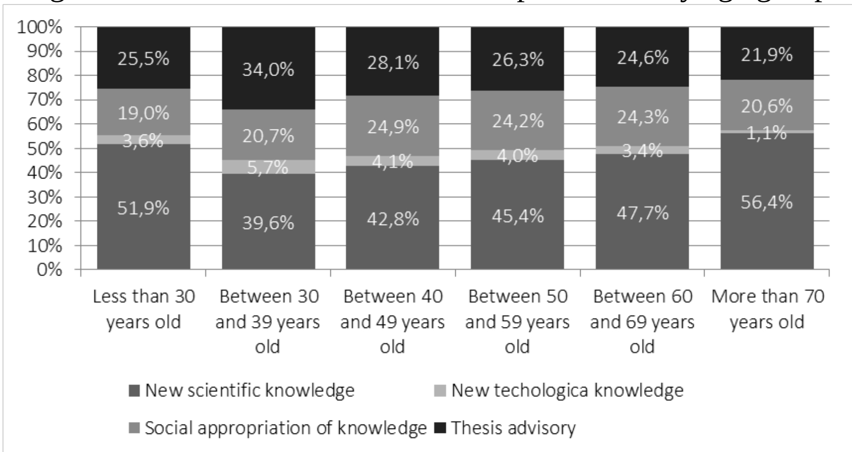
Figure 3: Distribution of the scientific production by research fields



Source: GrupLAC y CvLAC, consulted in april 2012; OCyT (2012).
Calculations: OCyT.

In addition, younger researchers are more concentrated on new scientific knowledge production, and researchers with 30-39 years old produce more thesis advisory compared to the former, in contrast, researchers from 40 to 69 produce more appropriation results compared to the other age groups. These movements could indicate the different moments in the life cycle in which each generation of researchers is. The analyses presented here are descriptive, in order to have a deeper understanding of the causal relationships we will require to estimate a model correlating the institutional links, research field and in general the profile of the researchers with the scientific production.

Figure 4: Distribution of the scientific production by age group



Source: GrupLAC y CvLAC, consulted in april 2012; OCyT (2012).
Calculations: OCyT.

Conclusions

Preliminary results show that the patterns of production differ widely among subjects with different demographic characteristics related to the scientific trajectory of the researchers like their age, sex and research field; further studies will have to determine the correlations among these variables and also with the level of education achieved and the rest of the variables proposed in this paper.

Even though we do not test the model, the data show that this could be plausible way to explain the interactions among the variables. Given the complexity and interdependency, SEM seems to be a good options when analysing this kind of datasets. More exploration and the testing of each relationship will be required in order to validate the hypotheses.

In Addition, further work will require to understand how the institutional links (working in a group inside a university) favours or limits the levels of scientific production and collaboration with other institutions. Traditionally, in the context of Colombia's STI system, the measurement of the scientific results has privileged research groups, since the policy was been oriented toward strengthening their role as organizational forms for the scientific work. In the last year, the measurement of scientific production has changed towards the role of individual researchers. In this sense, the proposal to measure scientific capacities using correlations between the researcher's profiles and their scientific production, will be useful to understand the configuration of scientific communities inside Colombian universities (which concentrate 90.1% of the capacity in the country) (Ocyt, 2013).

Another topic to be explored is the heterogeneity, not only in research fields, but also in the type of products, which will allow to establish groups of universities according to the type of knowledge they produce.

References

Bollen, K. A. (1998). Structural equation models. John Wiley & Sons, Ltd.

Boomsma, A. (2000). Reporting analyses of covariance structures. Structural equation modeling, 7(3), 461-483.

Bucheli, V., Díaz, A., Calderón, J. P., Lemoine, P., Valdivia, J. A., Villaveces, J. L., & Zarama, R. (2012). Growth of scientific production in Colombian universities: an intellectual capital-based approach. *Scientometrics, 91*(2), 369-382.

Cañibano, C., & Bozeman, B. (2009). Curriculum vitae method in science policy and research evaluation: the state-of-the-art. Research Evaluation, 18(2), 86-94.

Cañibano, C., Otamendi, J., & Andújar, I. (2008). Measuring and assessing researcher mobility from CV analysis: the case of the Ramón y Cajal programme in Spain. Research Evaluation, 17(1), 17-31.

Dietz, J. S., & Bozeman, B. (2005). Academic careers, patents, and productivity: industry experience as scientific and technical human capital. Research Policy, 34(3), 349-367.

Fox, J. (2002). Structural equation models. Appendix to an R and S-PLUS Companion to Applied Regression.

Jaramillo, H., & Forero, C. (2001). Interacción entre Capital Humano, Capital Social y Capital Intelectual. El Contexto del Nuevo contrato Social. V taller Iberoamericano e Interamericano de indicadores de Ciencia y Tecnología, Montevideo.

Lepori, B., & Probst, C. (2009). Using curricula vitae for mapping scientific fields: a small-scale experience for Swiss communication sciences. Research Evaluation, 18(2), 125-134.

Rivera-Torres, S., Pallares, C., Cárdenas-Osorio, J., & Galvis-Restrepo, M. (2013). Análisis del recurso humano vinculado a grupos de investigación en universidades colombianas: una aproximación a los indicadores de trayectorias científicas y tecnológicas. En Observatorio Colombiano de Ciencia y Tecnología, *Observando el Sistema Colombiano de Ciencia Tecnología e Innovación: sus actores y sus* productos (págs. 105-146). Bogotá: Observatorio Colombiano de Ciencia y Tecnología.

Villaveces, J. L., & Forero, C. (2007). Cincuenta años de ciencia en Colombia (1955-2005). Ciencia y tecnología en Colombia, una política coherente desde los sesenta. Bogotá, D.C.: Arfo Editores e Impresos Ltda.

Villaveces, José Luís & Hernán Jaramillo (2004): La Acumulación de Capital de Conocimiento; Paris, Seminario Internacional sobre Indicadores de C&T en el Contexto de la Sociedad de la Información; INIST-CNRS, RICyT, marzo del 2004.

The impact of a few: The effect of alternative formulas for recruiting talent in a non-competitive system[1]

Nicolas Robinson-Garcia*, Evaristo Jiménez-Contreras* and Clara Calero Medina**

*elrobin@ugr.es; evaristo@ugr.es
EC3 Evaluación de la Ciencia y de la Documentación Científica, Universidad de Granada, Colegio Máximo de Cartuja, s/n, Granada, 18071 (Spain)

** clara@cwts.leidenuniv.nl
Centre for Science and Technology Studies, Leiden University (The Netherlands)

Introduction

National university systems such as the Spanish or the italian have been critized for having low levels of governance and being non-competivie (Abramo, Cicero & D'Angelo, 2011). Here, differences between universities are not as significant as differences between individuals within universities. This is the case of Spain, where mobility within universities is seriously constrained, showing high rates of inbreeding due to their restrictive employment conditions (Navarro & Rivero, 2001). These conditions discourage foreign researchers (Pickin, 2001) and prevent Spanish researchers from returning to their home country. In Spain, the Catalan regional government created in 2001 the *Institució Catalana de Recerça i Estudis Avançats* (ICREA)for attracting senior researchers with an international research background. In this paper we intend to explore the influence exerted by researchers hired through this program.

ICREA establishes collaboration agreements with any research institution in Catalonia. Every year ICREA offers a closed number of positions for which candidates must have previously come in terms with their potential host institution. These are evaluated through a peer review process undertaken by five panels each of them corresponding to a scientific field, and the ones with the highest ratings, despite the selection panel or the host institution, are then selected. Candidates are expected to create new research lines and develop and lead their own research groups. Figure 1 briefly resumes the selection process.

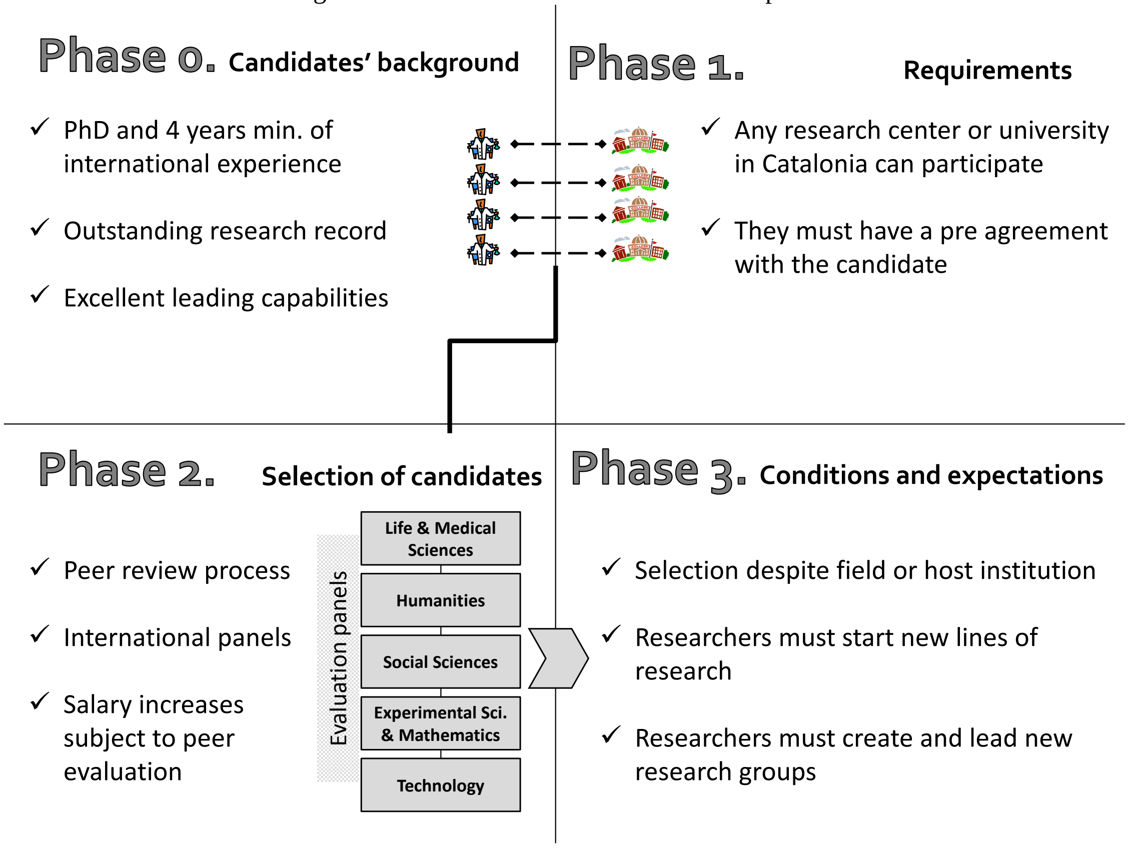
In this study we focus on the strategies followed by universities to allocate these researchers and their impact in specific disciplines. Our research questions are:

- Can we perceive through the number of publications strategic policies when allocating ICREA researchers by universities? Are they reinforcing specific fields?

- Are these researchers making a difference in terms of highly cited papers?

[1] Nicolás Robinson-García is currently supported by a FPU (Formación de Profesorado Universitario) grant of the Spanish Ministerio de Economía y Competitividad.

Figure 1. Flowchart of the ICREA selection process



Material and methods

We use the dataset of the Leiden Ranking (Waltman et al., 2012). We identify ICREA publications through the address field and identified a subset of publications linked to ICREA and Catalan universities. We analyzed five research areas. We focused on all subject categories for each university with at least 50 publications and selected those where ICREA publications represented at least 20% of the share. We identified highly cited papers in each field. Highly cited papers are publications cited equal or more than the 90th percentile limit of the citation distribution.

Results

Catalan universities produced 77,547 publications in 2002-2012. They represent 4,6% of the total share. Table 1 shows the total output by university and area, along with the distribution by areas of ICREA publications. Figure 2, shows that ICREA publications are distributed similarly in most universities. In figure 3 we show subject categories by university. In figure 4 we look at the share of highly cited papers for ICREA publications, not ICREA publications and the overall for the universities and subject categories shown in figure 3.

Table 1. Output of Catalan universities and distribution of ICREA publications for the 2002-2012 time period

University	Biomedical & Health Sci		Life & Earth Sci		Maths & Computer Sci		Natural Sci & Engineering		Social Sci & Hum	
	Pub	%ICREA	Pub	%ICREA	Pub	%ICREA	Pub	%ICREA	Pub	%ICREA
Barcelona	16350	1.72	5840	2.35	1332	4.05	9513	6.83	2538	3.23
Autónoma Barcelona	12385	1.89	5368	4.79	2114	5.44	5588	**15.44**	2283	5.43
Politecnica Catalunya	792	2.53	1719	0.93	3851	0.91	4912	2.75	298	1.68
Rovira i Virgili	1706	1.88	1265	7.75	833	0.36	2419	6.12	539	2.97
Pompeu Fabra	2720	**13.42**	630	**19.52**	707	**13.15**	291	14.09	1228	**9.20**
Girona	968	1.14	1220	2.05	536	1.12	1470	7.28	389	0.00
Lleida	850	2.00	1356	3.76	262	0.00	477	0.00	209	0.48

Note: Higher shares of ICREA output are highlighted in bold.

Figure 2. Distribution of ICREA publications by broad fields

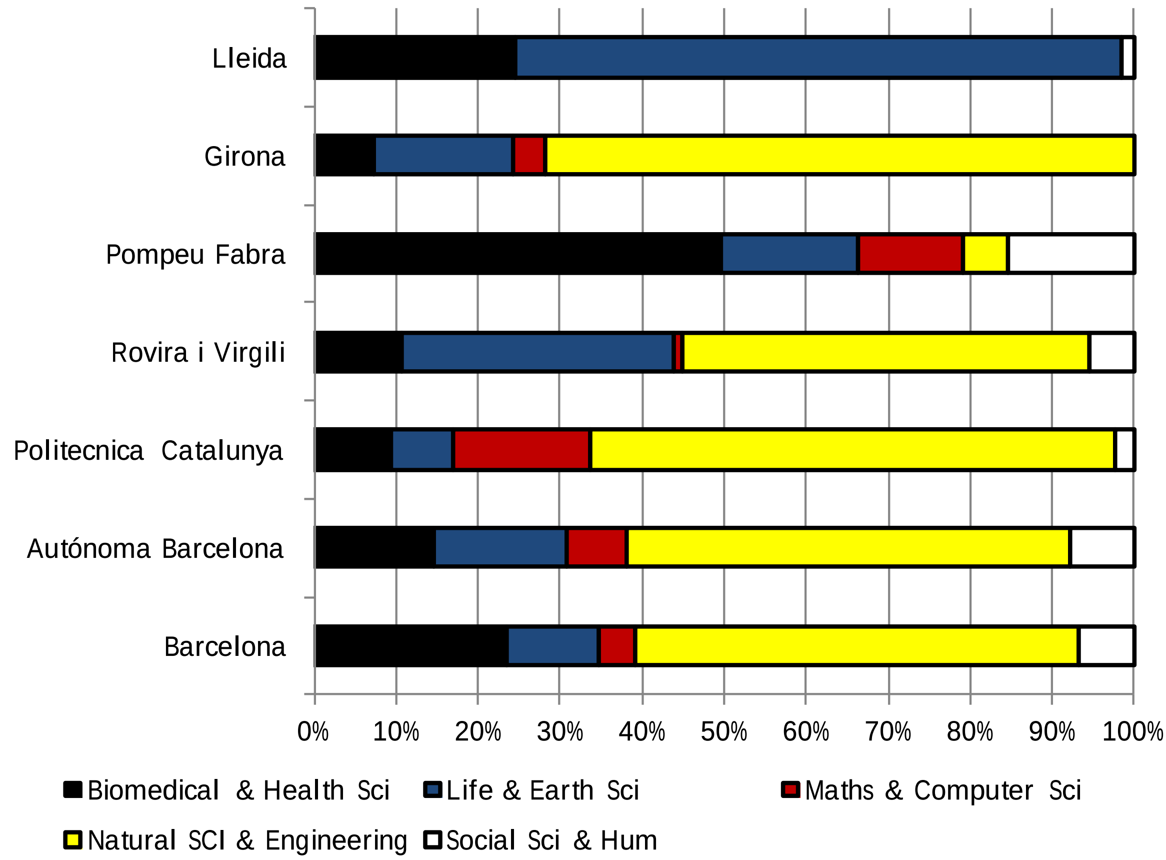


Figure 3. Share of ICREA publications by university and by subject categories for the 2002-2012 period. Only subject categories with ≥ 50 publications and ≥ 20% of ICREA output are shown

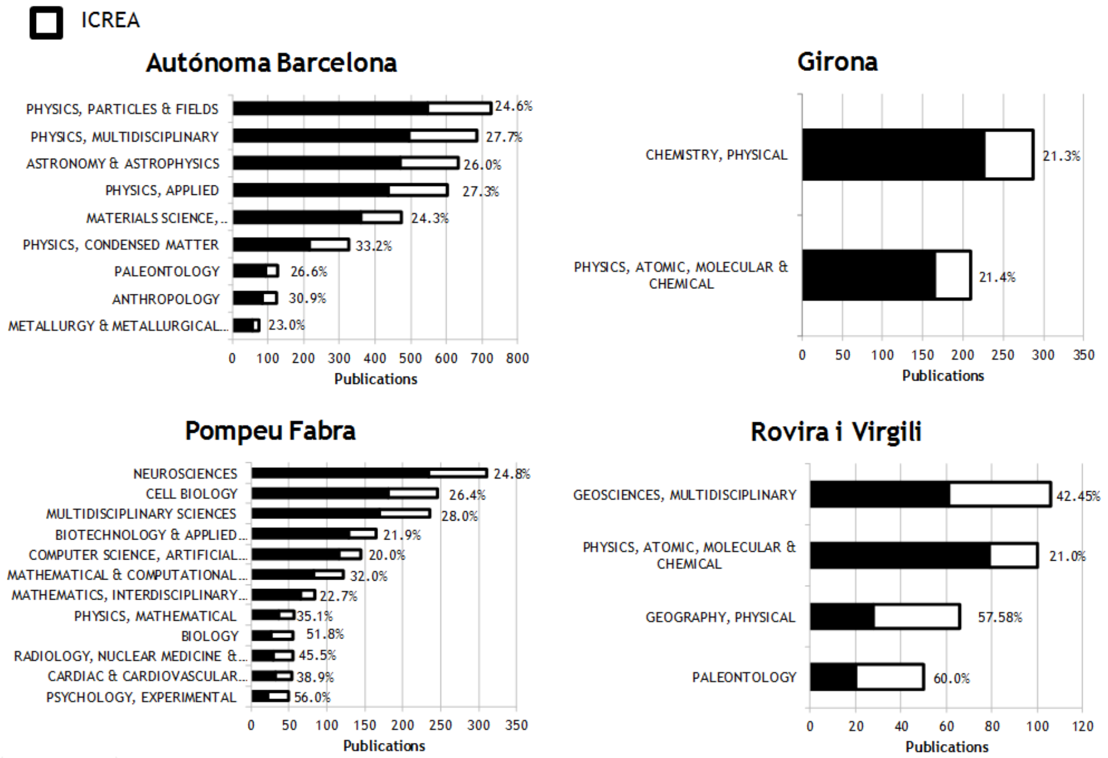
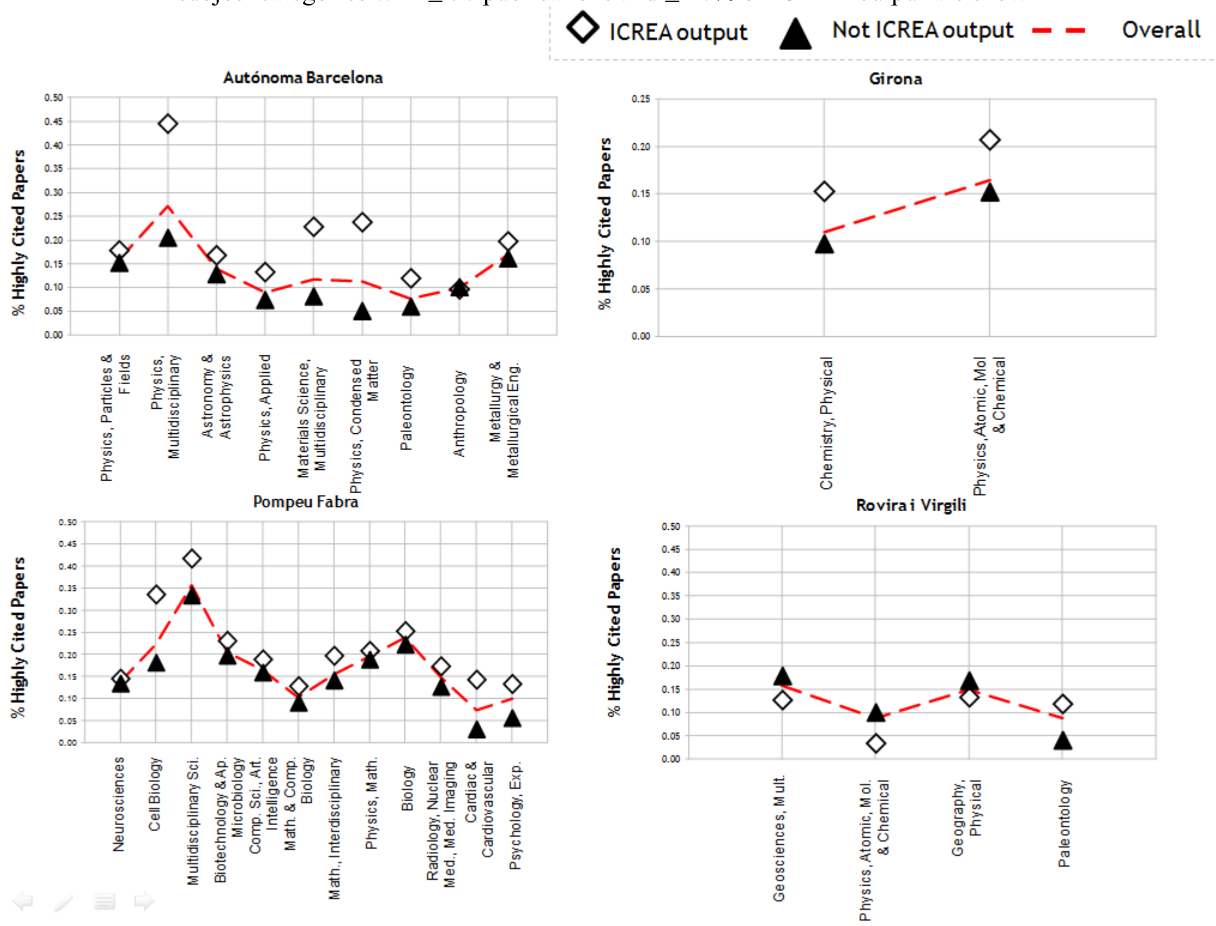


Figure 4. Percentage of highly cited papers by university and subject categories for the 2002-2012 period. Only subject categories with ≥ 50 publications and ≥ 20% of ICREA output are shown



Conclusions

The main conclusions of this poster are:

- Formulas such as ICREA allow Spanish universities, which present low levels of governance (Aghion, et al., 2010), a chance to develop research policy strategies to reinforce certain departments and scientific fields.

- ICREA output shows higher values when focusing on the share of highly cited papers.

- Further research should include bio data for ICREA researchers, in order to further understand the implications of such policy within the regional university system, and compare ICREA with other similar recruiting formulas implemented in other countries.

References

Abramo, G., Cicero, T., & D'Angelo, C.A. (2011). The dangers of performance-based research funding in non-competitive higher education systems. *Scientometrics*, 87, 641-654.

Aghion, P., Dewatripont, M., Hoxby, C., Mas-Colell, A., & Sapir, A. (2010). The governance and performance of universities: Evidence from Europe and the US. *Economic Policy*, 25, 7-59.

Jiménez-Contreras, E., Moya-Anegón, F., & Delgado López-Cózar, E. (2003). The evolution of research activity in Spain: The impact of the National Commission for the Evaluation of Research Activity (CNEAI). *Research Policy*, 123-142.

Navarro, A., & Rivero, A. (2001). High rate of inbreeding in Spanish universities. *Nature, 410*, 14.

Pickin, S. (2001). Spain is a closed culture to foreign researchers. *Nature*, *410*, 627.

Technopolis Group (2011). *Evaluation of ICREA 2001-2011. Final Report*.

Waltman, L., Calero-Medina, C., Koosten, J., Noyons, E., Tijssen, R.J.W., van Eck, N.J., van Leeuwen, T.N., van Raan, A.F.J., Visser, M.S., & Wouters, P. (2012). The Leiden Ranking 2011/2012: Data collection, indicators and interpretation. *Journal of the American Society for Information Science and Technology*, *63*(12), 2419-2432.

Investigation of Partition Cells as a Structural Basis Suitable for Assessments of Individual Scientists[1]

Nadine Rons*

*Nadine.Rons@vub.ac.be
Research Coordination Unit, Vrije Universiteit Brussel (VUB), Pleinlaan 2, B-1050 Brussels (Belgium)

Abstract
Individual, excellent scientists have become increasingly important in the research funding landscape. Accurate bibliometric measures of an individual's performance could help identify excellent scientists, but still present a challenge. One crucial aspect in this respect is an adequate delineation of the sets of publications that determine the reference values to which a scientist's publication record and its citation impact should be compared. The structure of partition cells formed by intersecting fixed subject categories in a database has been proposed to approximate a scientist's specialty more closely than can be done with the broader subject categories. This paper investigates this cell structure's suitability as an underlying basis for methodologies to assess individual scientists, from two perspectives:
(1) Proximity to the actual structure of publication records of individual scientists: The distribution and concentration of publications over the highly fragmented structure of partition cells are examined for a sample of ERC grantees;
(2) Proximity to customary levels of accuracy: Differences in commonly used reference values (mean expected number of citations per publication, and threshold number of citations for highly cited publications) between adjacent partition cells are compared to differences in two other dimensions: successive publication years and successive citation window lengths.
Findings from both perspectives are in support of partition cells rather than the larger subject categories as a journal based structure on which to construct and apply methodologies for the assessment of highly specialized publication records such as those of individual scientists.

Introduction

Leading excellent researchers have become the focus of an increasing number of funding programs worldwide, in a context of growing global competition to recruit the best. Grantees are selected by peer review, generally accepted as the primary methodology to evaluate research quality, but nevertheless under pressure from criticisms. These include issues related to the methodology itself as well as to resources (over-sollicitation of experts) and to workload (high numbers of applications to be evaluated in a same round). Quantitative indicators that are a good proxy for quality as perceived by peers can be valuable complements to peer reviews in evaluation procedures. Advanced bibliometric indicators in particular have been applied at the levels of research groups, university departments and institutes (van Raan, 2005). Still, several decades after the introduction of the Web of Science in its earliest form, it remains a challenge to develop bibliometric indicators that can be adequately applied to publication records as 'small' and specialized as those of individual scientists. Recently multiple conference sessions have been explicitely dedicated to

[1] This paper is related to work carried out in the framework of the research theme 'Long term investments in top research' (2007-2013) of the Vrije Universiteit Brussel in the Flemish interuniversity Centre for Research & Development Monitoring ECOOM.

methodological and ethical aspects of individual-level evaluative bibliometrics (14th Conference of the International Society for Scientometrics and Informetrics, Vienna; 18th International Conference on Science and Technology Indicators, Berlin; 5th Biennial Atlanta Conference on Science and Innovation Policy, Atlanta). Among efforts to develop novel and better adapted methodologies for the assessment of individual scientists was the partition based field normalization method, using cells smaller than subject categories as reference sets to more closely fit highly specialized publication records (Rons, 2012). These partition cells and the larger subject categories both are journal-based structures. Among methods to distinguish between fields (Schubert & Braun, 1996), also paper-based methodologies can produce reference sets representing scientific specialties. Depending on the methodology these may however be available for certain disciplines only, or may require intensive calculation or manual efforts. One way to proceed is the arrangement of individual papers in fine-grained classification schemes maintained by certain research domains, e.g. in the Chemical Abstracts database used in a bibliometric analysis by Neuhaus and Daniel (2009). Another is to use advanced algorithms to generate approximations of specialties, involving e.g. bibliographic coupling (Kessler, 1963), co-citation (Small, 1973), direct citation, co-words (Callon et al., 1983), or a combination. The journal-based and paper-based methodologies offer different advantages. In particular, paper-based structures can offer high precision, and journal-based structures high stability. Partition cells, as journal-based structures more subdivided than subject categories, combine higher precision with stability. Directly determined by the fixed subject categories, the partition cells are available as a more subdivided basis for all disciplines and for various methodologies. Besides the normalization context in which they were originally proposed, they can for instance also be applied in the identification of highly cited publications (Rons, 2013). Progress in domain delimitation however responds to just one of the challenges faced in comparative bibliometric assessments of individual scientists. Other issues include data accuracy, approaches to interdisciplinary research and to multidisciplinary and general journals in journal based reference standards, and the variation in bibliometric characteristics among equally distinguished scientists. A methodology needs to cope with these various issues to be able to produce results that are strongly correlated with peer judgments on individual scientists. Such results could mean a significant support in evaluation procedures, as they would help evaluators concentrate on candidates in a crucial range or dedicate more of their time to the most complex dossiers, and thus perform their tasks more efficiently. Gaging the actual improvements in accuracy contributed by using partition cells rather than subject categories in various methodologies requires further investigation. This paper focuses on the ability to more closely fit reference sets to publication records of individual scientists, and on the level of accuracy attained as compared to customary levels for other variables.

Data and Methodology

The investigations were conducted using publication and citation data from Thomson Reuters' online Web of Science (WoS), and for article type documents only (i.e. no multiple types) given their prominent role and status in the process of knowledge creation and dissemination. The investigated partition cells are formed by the fixed structure of overlapping WoS subject categories, such that each cell contains all publications associated to exactly the same combination of subject categories (Rons, 2012). These cells form reference sets of an intermediate size between journals (relatively narrow) and entire subject categories (relatively wide). Reference sets at different levels of aggregation generate different reference values to which a performance can be compared. The level of aggregation for instance strongly influences which publications belong to the top-cited groups (Zitt, Ramanana-Rahary &

Bassecoulard, 2005). Broad discipline-based systems (such as the 60 subfields developed by the groups in Leuven and Budapest; Glänzel & Schubert, 2003) were found to be a sufficiently accurate basis for assessments at country and institution level (Glänzel et al., 2009). Compared to these levels, individual scientists have much more specialized research profiles. The structure of partition cells, introduced in the specific context of highly specialized publication records, offers a more subdivided basis of reference sets that are neither so wide as to include strongly diverging publication and citation practices from unrelated areas, nor so narrowly fit as to reflect a researcher's potential bias towards certain publication media.

ERC Grantee Publication Records

From the WoS, all articles were collected that were authored by a sample of grantees from the first European Research Council's ERC Advanced Grants call (2008). The ERC Advanced Grants are aimed at exceptional, established, scientifically independent research leaders with a track-record of significant research achievements in the last 10 years (http://erc.europa.eu/advanced-grants). The selection is based on international peer review with excellence as the sole criterion, using a system of 25 discipline-based panels of high-level scientists. For this bibliometric investigation, grantees were observed from two ERC panels representing domains with strongly different publication productivity and citation characteristics:

- 21 grantees in the panel 'Mathematical foundations' (slower characteristics);
- 14 grantees in the panel 'Fundamental constituents of matter' (faster characteristics).

For the identification of the grantees' articles among those of homonyms, cv-information was used next to the data available in the WoS. For each grantee, the distribution of articles over partition cells was calculated in the 8-year publication period preceding the call (2000-2007). The mean number of citations per article, in a 5-year citation window for each publication year, was calculated per grantee as a rough indication of citation levels (a more thorough investigation of citations in view of indicator design or verification being beyond the scope of this paper).

Levels of Accuracy

It is customary good practice to compare observed citation-based values for a studied entity to reference values for the same publication year and citation window. Differences between reference values in subsequent publication years or citation window lengths, indicate the influence on results of using a reference time frame just one year off the appropriate one. Similarly, differences can be observed between reference values in adjacent partition cells associated to a same subject category, indicating the influence on results of using a reference research area just next to the appropriate one. These differences and their relative magnitudes were studied for two commonly used reference values:

- The mean expected number of citations per article (*e*), used in the general standard field normalized citation rates (Braun & Glänzel, 1990; Moed et al., 1995);
- The threshold number of citations for outstandingly cited articles (*T*) as determined by the methodology of Characteristic Scores and Scales (Glänzel & Schubert, 1988), used for the identification of highly cited publications, which have since long been regarded as bibliometric emanations of research excellence at the level of individual scientists (Garfield, 1986).

Differences between pairs of reference values were calculated in three dimensions:

(1) In subsequent publication years (pairs 2005-2006 and 2006-2007);

(2) In subsequent citation window lengths (pairs 3-4 and 4-5 years, including the publication year);
(3) In adjacent partition cells C (pairs C_i-$C_{i;j}$ and $C_{i;j}$-C_j, where C_i, $C_{i;j}$ and C_j contain the articles in journals assigned respectively to subject categories i only, i and j combined, and j only.
The investigations were conducted in two domains with strongly different citation characteristics related to the two ERC panels in the 'ERC Grantee Publication Records' section above, for the following cells:
- C_{M}, $C_{\mathrm{M;MA}}$ and C_{MA} in the domain of Mathematics, where M and MA respectively stand for the subject categories 'Mathematics' and 'Mathematics Applied'.
- C_{AA}, $C_{\mathrm{AA;PPF}}$ and C_{PPF} in a sub-domain of physics, where AA and PPF respectively stand for the subject categories 'Astronomy & Astrophysics' and 'Physics, Particles & Fields'.
For each of the two reference values e and T, 36 pair-wise comparisons of positive reference values x and y were made in each of the three dimensions, calculating the absolute relative difference r_e and r_T as r=2*|x-y|/(x+y).

Results and Discussion

ERC Grantee Publication Records

Figures 1 and 2 show the distribution and concentration of articles in the observed 8-year publication period for each grantee in the two panels, indicating the shares of articles per cell. For each grantee the top share is highlighted. The cells listed are all those that contain articles by at least one grantee (excluding for instance 2% of the subject category 'Mathematics' in Figure 1, and 13% of the subject category 'Physics, Particles & Fields' in Figure 2). Within the highly fragmented structure of cells, the grantees' articles are strongly concentrated in one or a few cells, and totally absent or very limitedly present in the other cells that are associated to the same subject categories. The partition cells therefore offer a more accurate structural basis than the larger subject categories to delimit the publication environment in which to position these scientists’ performances. For instance, some grantees publish the top share of their articles in a cell associated to exactly one subject category, and no articles in any of the other cells associated to that subject category (M_1, M_2, M_3, M_{13}, M_{15}). One grantee publishes the top share of his articles in a cell associated to a combination of two subject categories, and only 3% in all other cells associated to either of these two subject categories (F_3).
Figures 1 and 2 further demonstrate a strong diversity in publication profiles among grantees who were evaluated in a same disciplinary panel. Even grantees who have been co-authors, presumably at least partly working on closely related topics, may publish their top shares of articles in different cells (F_1 and F_9). For some grantees who do have highly similar distributions of articles over cells, productivities and citation levels differ by a factor 2 (M_1 and M_2; M_6 and M_7). Similar observations of publication records with very different characteristics for equally distinguished scientists in a same discipline were made by Sugimoto and Cronin (2012) for six information scientists. Among the factors that play a role in such variation may be contextual issues regarding the individual scientist (local environment, personal choices), and issues regarding the organization of scientific literature in the scientist's domain (e.g. the presence of interdisciplinary or multispecialty journals or database categories).

Figure 1: Distribution of articles over partition cells for grantees from the ERC Advanced Grant call 2008, panel 'Mathematical foundations'.

Grantee:	M_1	M_2	M_3	M_4	M_5	M_6	M_7	M_8	M_9	M_{10}	M_{11}	M_{12}	M_{13}	M_{14}	M_{15}	M_{16}	M_{17}	M_{18}	M_{19}	M_{20}	M_{21}
Number of articles 2000-2007:	11	6	6	22	25	20	11	14	20	10	27	19	12	98	7	19	62	20	15	22	31
Cell C_X with X:	Share of articles per cell																				
ACS;CSAI;MA																5%					
ACS;MA												11%							20%		3%
BRM;BAM;CSIA;MCB;SP														1%							
B																	2%				
B;MCB																	24%				
Bp;EB											4%										
BF;E;MIA;SSMM																			7%		6%
CCS;RS;S											4%										
CB																	2%				
CSAI																				23%	
CSAI;CSC;EEE																				5%	
CSAI;CSIA;EB;RNMMI																				9%	
CSAI;EEE																				9%	
CSHA;CSIS;CSSE;CSTM									5%												
CSIS														2%							
CSIS;EEE														3%							
CSIA;CSSE;MA																	11%				
CSIA;EB;EEE;ISPT;RNMMI																				5%	
CSIA;MA																	2%				
CSIA;MIA;Mc;PFP																5%					
CSIA;Mc																11%					
CSIA;PM																37%		10%		5%	
CSSE;MA														1%							
CSSE;MA;M									5%				8%	6%							
CSSE;ORMS;MA									5%												
CSTM														2%							
CSTM;M														1%							
CSTM;MA														6%							
CSTM;MA;L									5%					1%							
CSTM;MA;M																		5%		5%	
CSTM;M;SP									15%					13%							
E;MIA;SSMM																					3%
EB;RNMMI																				9%	
EEE;O;ISPT																				5%	
EMd;MA											4%										
EMd;MIA											4%										
EMd;MIA;Mc											19%										
HM											4%										
MCB																	2%				
M	100%	100%	100%	95%	88%	85%	82%	21%	60%		4%	53%		48%	29%		6%		20%		16%
MA											56%	11%		10%		26%	32%	30%	13%	14%	6%
MA;M				5%	4%	5%	9%		5%	20%		16%		3%	29%		8%	5%	27%		6%
MA;Mc																	2%				
MA;PM																		5%			3%
MA;PM;SP																					13%
MA;PMd;PM																		10%			
MA;SP																					16%
MIA																	2%				
MIA;Mc												11%									
MIA;PM											4%					5%		10%			
MIA;PMd;PM																	2%				
Mc																5%					
Mc;PFP																		5%			
MS					8%												2%				6%
N																				5%	
N;Ni;RNMMI																				9%	
Oc;CN																	2%				
ORMS;MA																					3%
PFP;PM																	2%	5%			
PM						10%	9%	71%		60%			33%		43%		2%	5%			
PMd										10%											
PMd;PM													8%								
PMd;PPF;PM								7%		10%											
RNMMI														1%							
SP													50%	1%				10%	13%		16%
T;EM;Mc																5%					
	Mean number of citations per article in a 5-year citation window (including the publication year):																				
	6	11	7	6	4	4	11	9	8	9	8	6	6	5	12	16	8	9	8	19	7

Subject categories:

ACS: Automation & Control Systems
B: Biology
BAM: Biotechnology & Applied Microbiology
BF: Business, Finance
Bp: Biophysics
BRM: Biochemical Research Methods
CB: Cell Biology
CCS: Cardiac & Cardiovascular Systems
CN: Clinical Neurology
CSAI: Computer Science, Artificial Intelligence
CSC: Computer Science, Cybernetics
CSHA: Computer Science, Hardware & Architecture
CSIA: Computer Science, Interdisciplinary Applications
CSIS: Computer Science, Information Systems
CSSE: Computer Science, Software Engineering
CSTM: Computer Science, Theory & Methods
E: Economics
EB: Engineering, Biomedical
EEE: Engineering, Electrical & Electronic
EM: Engineering, Mechanical
EMd: Engineering, Multidisciplinary
HM: Humanities, Multidisciplinary
ISPT: Imaging Science & Photographic Technology
L: Logic
M: Mathematics
MA: Mathematics, Applied
Mc: Mechanics
MCB: Mathematical & Computational Biology
MIA: Mathematics, Interdisciplinary Applications
MS: Multidisciplinary Sciences
N: Neurosciences
Ni: Neuroimaging
O: Optics
Oc: Oncology
ORMS: Operations Research & Management Science
PFP: Physics, Fluids & Plasmas
PM: Physics, Mathematical
PMd: Physics, Multidisciplinary
PPF: Physics, Particles & Fields
RNMMI: Radiology, Nuclear Medicine & Medical Imaging
RS: Respiratory System
S: Surgery
SP: Statistics & Probability
SSMM: Social Sciences, Mathematical Methods
T: Thermodynamics

Data sourced from Thomson Reuters Web of Knowledge (formerly referred to as ISI Web of Science). Web of Science (WoS) accessed online 04& 09.10.2013.

Figure 2: Distribution of articles over partition cells for grantees from the ERC Advanced Grant call 2008, panel 'Fundamental constituents of matter'.

Grantee:	F_1	F_2	F_3	F_4	F_5	F_6	F_7	F_8	F_9	F_{10}	F_{11}	F_{12}	F_{13}	F_{14}
Number of articles 2000-2007:	34	19	117	64	85	98	38	29	23	33	149	48	69	50
Cell C_X with X:	**Share of articles per cell**													
AA								7%						
AA;PMd		5%						7%	4%					
AA;PMd;PPF	3%			3%										
AA;PPF		63%		8%					26%					
Bp												2%		
CA											1%			
CMd													1%	
CMd;CP;NN;MSM;PA;PCM												6%		
CP;MSM;PA;PCM						1%								
CP;NN;MSM												2%		
CP;NST;PAMC													1%	
CP;PAMC										6%				4%
CSHA;CSIS;O;Tc					1%									
CSSE;MA											1%			
CSTM;EEE;MA											1%			
CSTM;PPF;PM			2%								5%			
ESD;PMd					1%						1%			
EEE											1%			
EEE;II								3%						
EEE;O;PA					1%	1%					5%			2%
EEE;O;Tc											3%			
EMd;II			1%											2%
II;PA					1%						1%		4%	
MSM;PA												4%	1%	
MA						1%								
MA;PM				2%										
MS	3%		1%		18%	6%	21%	3%	4%		3%	10%	6%	4%
O			3%		12%	46%	11%			18%	13%	31%	4%	24%
O;PA					2%		8%	3%		6%	1%	2%	1%	12%
O;PAMC			50%		14%	2%	16%	31%		36%	26%			18%
PA					2%	2%					2%	17%	12%	10%
PAMC										9%				
PCM			1%									8%		6%
PFP													29%	
PFP;PM						13%							7%	
PFP;PN													3%	
PM				8%							1%			
PMd	21%	21%	42%	30%	47%	28%	42%	45%	39%	24%	32%	17%	25%	18%
PMd;PM			1%				3%				4%			
PN;PPF				2%					4%				1%	
PPF	74%	11%		48%					22%					
RNMMI													3%	
T;CP											1%			
	Mean number of citations per article in a 5-year citation window (including the publication year):													
	23	47	25	22	48	34	52	89	61	22	25	71	32	31

Subject categories:

AA: Astronomy & Astrophysics
Bp: Biophysics
CA: Chemistry, Analytical
CMd: Chemistry, Multidisciplinary
CP: Chemistry, Physical
CSHA: Computer Science, Hardware & Architecture
CSIS: Computer Science, Information Systems
CSSE: Computer Science, Software Engineering
CSTM: Computer Science, Theory & Methods
EEE: Engineering, Electrical & Electronic
EMd: Engineering, Multidisciplinary
ESD: Education, Scientific Disciplines
II: Instruments & Instrumentation
MA: Mathematics, Applied
MS: Multidisciplinary Sciences
MSM: Materials Science, Multidisciplinary
NN: Nanoscience & Nanotechnology
NST: Nuclear Science & Technology
O: Optics
PA: Physics, Applied
PAMC: Physics, Atomic, Molecular & Chemical
PCM: Physics, Condensed Matter
PFP: Physics, Fluids & Plasmas
PM: Physics, Mathematical
PMd: Physics, Multidisciplinary
PN: Physics, Nuclear
PPF: Physics, Particles & Fields
RNMMI: Radiology, Nuclear Medicine & Medical Imaging
T: Thermodynamics
Tc: Telecommunications

Data sourced from Thomson Reuters Web of Knowledge (formerly referred to as ISI Web of Science).
Web of Science (WoS) accessed online 21.10.2013.

Levels of Accuracy
Table 1 shows the reference values e and T in the observed range in each dimension (publication years, citation window lengths and partition cells), and Figure 3 the absolute relative differences r_e and r_T between pairs of reference values in each dimension (successive publication years, successive citation window lengths and adjacent partition cells). The observed variation of reference values presents an example of the magnitudes that can be attained, and in particular of the relative magnitudes in the three dimensions. For the investigated suite of cells and time frame, the pair-wise differences between reference values are of a same order of magnitude in the three dimensions, with the differences for adjacent partition cells (r_e: 0-42%; r_T: 0-36%) lying higher than the differences for successive publication years (r_e: 0-18%; r_T: 0-31%) and lower than the differences for successive citation window lengths (r_e: 18-51%; r_T: 8-59%). A comparison to reference values per partition cell is therefore of comparable importance in terms of potential error generated, as a comparison to reference values per publication year and for the exact citation window observed, which is customary.
For publication records of large entities such as institutes or countries, calculating reference values per larger subject category has been experienced to generate sufficiently accurate global results, and is common practice. Such publication records are widely spread out over numerous partition cells. In such conditions, compensating effects from publications in cells 'advantaged' and 'disadvantaged' by reference values calculated at a more global level are likely to occur, and limit the overall influence on results of less accurate reference values for certain sub-domains. Figures 1 and 2 illustrate that publication records of individual scientists are concentrated in only a limited number of partition cells. In such conditions the same compensating effects are unlikely to occur, and it is equally recommendable to use reference values calculated per partition cell, as it is to use reference values calculated per publication year and for the exact citation window observed.

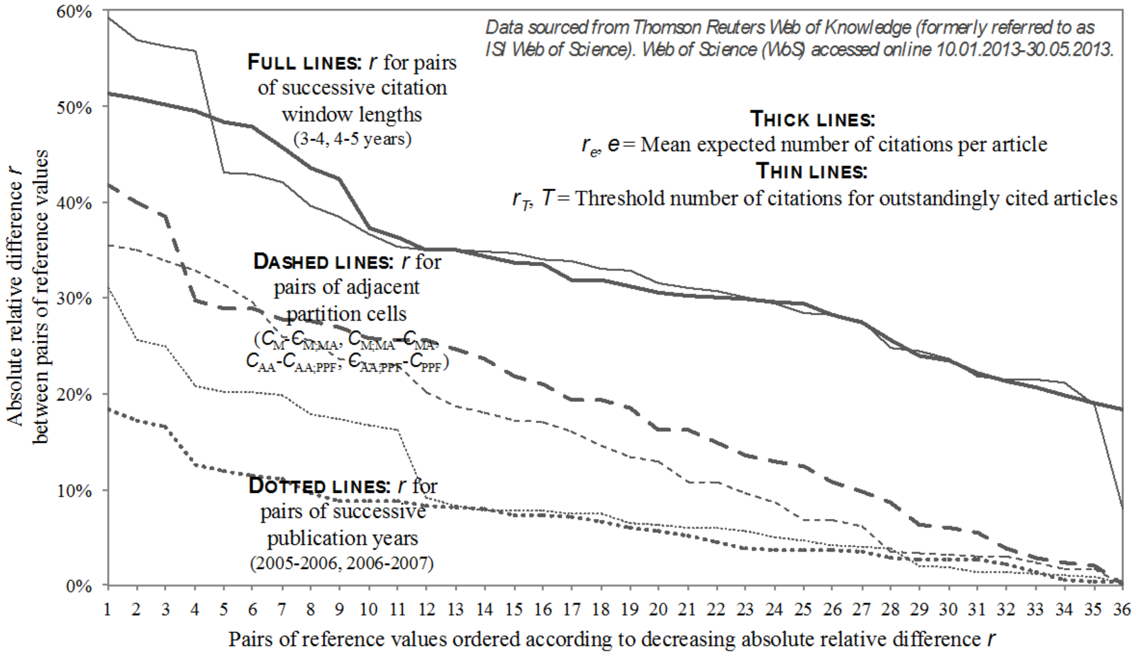
Table 1. Reference values in the observed range of publication years, citation window lengths and partition cells.

			Reference values					
Cell C_X with X:	**Publication year**	**Number of articles**	**Mean expected number of citations per article (*e*)**			**Threshold number of citations for outstandingly cited articles (*T*)**		
			Citation window length (years, incl. publication year)					
			3	**4**	**5**	**3**	**4**	**5**
M	2005	10055	1.2	2.0	2.8	6.1	8.3	11.8
	2006	10400	1.4	2.2	3.0	6.4	11.4	15.3
	2007	11410	1.4	2.2	2.9	6.5	11.6	12.5
M;MA	2005	4322	1.3	2.2	3.2	6.3	11.6	15.9
	2006	4531	1.6	2.6	3.6	8.1	12.5	16.7
	2007	5717	1.8	2.9	3.9	8.4	12.4	18.0
MA	2005	4589	2.0	3.3	4.7	8.2	14.7	20.1
	2006	4871	2.1	3.5	4.8	9.6	14.9	20.1
	2007	5288	2.3	3.7	5.0	11.8	17.7	24.6
AA	2005	8203	9.4	13.4	17.3	37.0	52.9	69.8
	2006	8798	9.8	14.2	18.0	37.8	58.0	74.0
	2007	8920	9.8	13.9	17.6	38.3	54.6	70.0
AA;PPF	2005	2398	10.3	14.2	17.7	39.7	56.3	69.8

	2006	2574	10.1	14.2	17.6	36.5	52.0	64.7
	2007	2545	10.9	14.8	18.1	39.5	55.5	68.9
PPF	2005	1780	9.0	12.1	14.6	38.7	54.6	67.5
	2006	1728	8.4	11.1	13.4	32.8	45.7	55.1
	2007	1815	9.1	12.0	14.6	34.1	45.3	57.5

Data sourced from Thomson Reuters Web of Knowledge (formerly referred to as ISI Web of Science). Web of Science (WoS) accessed online 10.01.2013-30.05.2013.

Figure 3: Absolute relative differences r_e and r_T between pairs of reference values e and T in three dimensions.



Conclusion

The results of the investigation show that (at least for the examined domains, sample of scientists and range of bibliometric variables):

- Reference sets of publications can be more closely fit to a scientist's specialty when using partition cells than when using the larger subject categories;
- The level of accuracy attained when using partition cells is comparable to customary levels of accuracy applied when calculating reference values per publication year and per citation window length.

These findings are in support of partition cells rather than the larger subject categories as a basis for methodologies for the assessment of highly specialized publication records. As more subdivided structures directly determined by the fixed subject categories, partition cells combine higher precision with stability, and are available for all disciplines. The actual effect in particular cases of applying partition cells rather than subject categories can be expected to vary with specialty and indicator, such study however being beyond the scope of the present paper. Reported examples of important differences between expected citation rates calculated for WoS subject categories versus smaller paper-by-paper classification structures indicate that the effect can be decisive (differing by a factor 2.2 for a sub-domain of neurology reported on by Bornmann et al., 2008, and by a factor 1.7 for a sub-domain of biochemistry reported on by Neuhaus & Daniel, 2009). Similar comparisons between results of different indicators based on subject categories versus partition cells thus are an important element for further studies investigating actual improvements associated to the latter. The outcomes are not only of interest in a context of assessments of individual scientists, but may also be relevant for other applications involving domain delineation and reference sets for publication records in scientific specialties, regardless of the volume of publications and the number of authors concerned.

References

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

Braun, T. & Glänzel, W. (1990). United Germany: The new scientific superpower? *Scientometrics*, 19(5-6), 513-521.

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.

Garfield, E. (1986). Do Nobel prize winners write citation classics? *Essays of an Information Scientist*, 9, 182-187.

Glänzel, W. & Schubert, A. (1988). Characteristic scores and scales in assessing citation impact. *Journal of Information Science*, 14(2), 123-127.

Glänzel, W. & Schubert, A. (2003). A new classification scheme of science fields and subfields designed for scientometric evaluation purposes. *Scientometrics*, 56(3), 357-367.

Glänzel, W. Thijs, B., Schubert, A. & Debackere, K. (2009). Subfield-specific normalized relative indicators and a new generation of relational charts: Methodological foundations illustrated on the assessment of institutional research performance. *Scientometrics*, 78(1), 165–188.

Kessler, M.M. (1963). Bibliographic coupling between scientific papers. *American Documentation*, 14(1), 10-25.

Moed, H. F., De Bruin, R. E. & Van Leeuwen, T. N. (1995). New bibliometric tools for the assessment of national research performance: Database description, overview of indicators and first applications. *Scientometrics*, 33(3), 381-422.

Neuhaus, C. & Daniel, H.-D. (2009). A new reference standard for citation analysis in chemistry and related fields based on sections of Chemical Abstracts. *Scientometrics*, 78(2), 219-229.

Rons, N. (2012). Partition-based field normalization: an approach to highly specialized publication records. *Journal of Informetrics*, 6(1), 1-10.

Rons, N. (2013). Groups of highly cited publications: stability in content with citation window length. In: *Proceedings of ISSI 2013, 14th International Society of Scientometrics and Informetrics Conference*, Vienna, Austria, 15-19 July 2013, Juan Gorraiz, Edgar Schiebel, Christian Gumpenberger, Marianne Hörlesberger, Henk Moed (Eds.), Vol. 2, 1998-2000.

Schubert, A. & Braun, T. (1996). Cross-field normalization of scientometric indicators. *Scientometrics*, 36(3), 311-324.

Small, H. (1973). Co-citation in the scientific literature: A new measure of the relationship between two documents. *Journal of the American Society for Information Science*, 24(4), 265-269.

Sugimoto, C.R. & Cronin, B. (2012). Biobibliometric profiling: an examination of multifaceted approaches to scholarship. *Journal of the American Society for Information Science and Technology*, 63(3), 450-468.

van Raan, A.F.J. (2005). Measurement of central aspects of scientific research: Performance, interdisciplinarity, structure. *Measurement*, 3(1), 1–19.

Zitt, M., Ramanana-Rahary, S. & Bassecoulard, E. (2005). Relativity of citation performance and excellence measures: from cross-field to cross-scale effects of field-normalization. *Scientometrics*, 63(2), 373-401.

Do Funding Sources Complement or Substitute? The Case of the UK Cancer Research[1]

Daniele Rotolo*, Michael Hopkins**, Nicola Grassano***

[*] *d.rotolo@sussex.ac.uk*
SPRU (Science and Technology Policy Research), University of Sussex, Brighton, BN1 9SL, (UK)

[**] *m.m.hopkins@sussex.ac.uk*
SPRU (Science and Technology Policy Research), University of Sussex, Brighton, BN1 9SL, (UK)

[***] *n.grassano@sussex.ac.uk*
SPRU (Science and Technology Policy Research), University of Sussex, Brighton, BN1 9SL, (UK)

Abstract

This paper examines complementarity of research funding sources. Previous research has extensively investigated this phenomenon at the level of the scientist and research host organisation neglecting that multiple individuals and research host organisations with their own associated funding sources are often involved in the research process to produce a single scientific output. Funding sources may complement or substitute at the level of the publication. We attempt to expand our understanding of the complementarity and substitution among funding sources by focusing on the publication as unit of analysis. We distinguish funding sources between (i) national, (ii) international, and (iii) industry. National funding sources are further classified as major and minor sources according to proportion of the overall research output they supported. The empirical analysis is performed on a sample of publications (N= 7,510) related to the cancer domain and involving UK research host organisations. Findings reveal complementarity among national and international funding sources and between the latter and industrial support. The empirical analysis does not provide evidence of complementarity between national and industrial funding sources while a strong complementary among (major and minor) national funding sources is found.

Keywords: funding systems, complementary, funding data, cancer research.

Introduction

Large science budgets are under pressure. The financial crisis has led major developed countries to revise public spending on scientific research. The Research Councils UK (RCUK), for example, have had a flat budget over the fiscal year 2012–2013 while the US National Institutes of Health (NIH) have faced a budget cut of 10% in real terms since 2010. In a time of austerity, policy makers are likely to make further cuts in many national funding systems. Yet, our understandings of the interdependencies that characterise national funding

[1] We acknowledge the support of Cancer Research UK ("Exploring the Interdependencies of Research Funders in the UK" research project) and the UK Economic and Social Research Council (award RES-360-25-0076 - "Mapping the Dynamics of Emergent Technologies"). We would also like to thank Emma Greenwood, Aoife Regan, Jon Sussex, Koonal Shah, Mark Samuels, Virginia Acha, Mike ONeill, Ismael Rafols, Ben Martin, Paul Nightingale, Kevin Dolby and Shemila Nahrajani for helpful comments in the early stages of this research, as well as the interviewee scientists who shared their experiences of the funding system with us. We also thank Josh Hutton, Frederique Lang, Philippa Crane, Christopher Farrell, Abigail Mawer, Chelsea Pateman, and Tammy-Ann Sharp for research assistance.

environments and, thus the potentially amplified effects of budget cuts, are limited. National funding systems are indeed populated by a large number of organizations which often play different roles, fund distinct types of research, but also may be strongly interdependent on other funding efforts to maximise the benefit of the research they support.

Previous studies have examined the role of funding on research productivity (e.g. Arora, David, & Gambardella, 1998; Lee & Bozeman, 2005) and collaboration (e.g. Bozeman & Gaughan, 2007; Defazio, Lockett, & Wright, 2009), and have investigated the interdependencies among different funding sources mostly in terms of the extent to which receiving a certain type of funding may attract a different funding (complementarity) or failing in receiving a certain funding may increase the likelihood of applying for another type of funding (substitution). Public and private funding sources were specifically found to complement each other (e.g. Blume-Kohout, Kumar, & Sood, 2009; Payne, 2001) whereas different types of government funding were found to be substitutive (e.g. Grimpe, 2012; Jacob & Lefgren, 2007).

Those studies however, by focusing at individual (scientist) and research host organisation levels, have overlooked the intense collective efforts science research requires to be carried out. Research usually involves more than one individual and multiple research host organisations, which often have their own associated funding sources (Adams, Black, Clemmons, & Stephan, 2005). Those sources may complement or substitute each other at the single research output level, i.e. the publication. Thus, a focus on the funding of the research host organisation or the scientist misses significant streams of funding that contribute to the advancement of particular research objectives, as pursued in scientific publications.

The present paper aims to fill this gap by examining complementarity among funding sources at scientific publication level. We consider multiple funders on a single publication to be complementary to each other, and where these co-funded outputs occur more often than expected, we ascribe this to authors' strategies to access funding directly or indirectly (e.g. via recruitment of funded co-authors) in order to achieve their scientific objectives.

The empirical analysis is conducted on 7,510 UK publications related to cancer and published in the year 2011. Preliminary findings provide evidence of both the large variety of funding sources scientists rely on to produce new scientific knowledge and the strong complementarity among those sources at the single research output level.

Theoretical Background

The role of research funding in the production and diffusion of scientific knowledge has been a central theme in policy research. A number of studies have examined the impact of various funding sources on scientists and research host organisations' productivity suggesting the presence of a positive relationship between the two (e.g. Arora et al., 1998; Defazio et al., 2009; Jacob & Lefgren, 2007; Lee & Bozeman, 2005). Jacob and Lefgren (2007), for example, estimated that recipients of NIH grants increase their productivity between 7% and 20%. Similar results were also found by Defazio et al. (2009) on a sample of researchers working on EU research networks.

The impact of funding on research productivity appears to be independent from the type of source. Industrial financial support has been also found to stimulate research productivity (e.g. Bonaccorsi, Daraio, & Simar, 2006) – though an exception is represented by Beaudry and

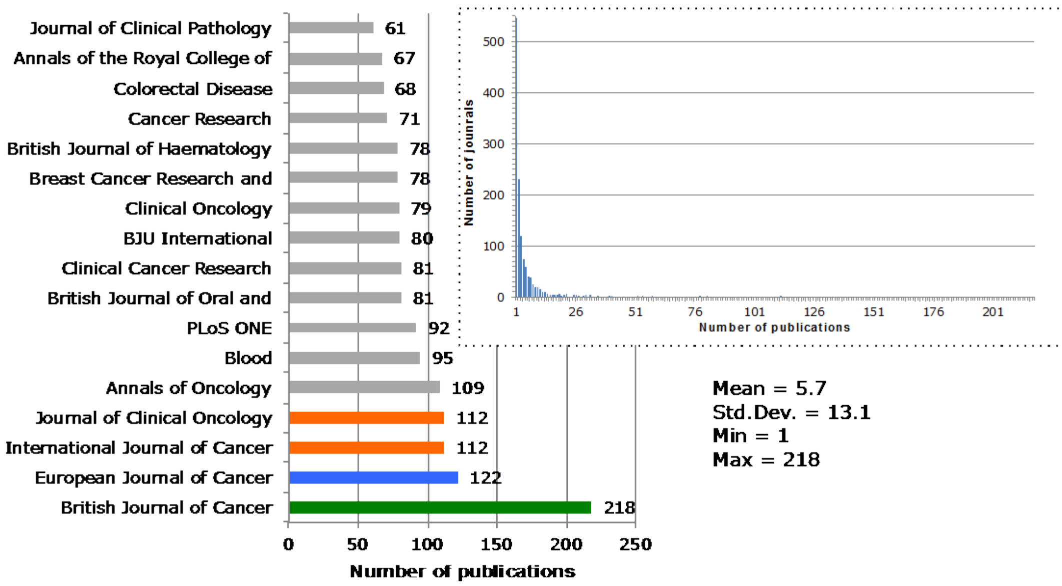
Allaoui's (2012) study where no significant relationship was found. Yet, the effect of funding on research productivity does not seem to extend to research quality (Godin, 1996; Payne & Siow, 2003).

Building on the studies pointing out to the importance of collaboration in research (Katz & Martin, 1997; Lee & Bozeman, 2005), scholars have also examined the impact of funding on collaboration patterns. For example, Adams et al. (2005), by analysing a sample of more than two million scientific papers, show that top US research universities whose scientists were awarded prestigious grants or federal funding tend to participate in teams of a larger size. Grants positively impact on both a scientist's total number of different co-authors and embeddedness into the scientific community (Ubfal & Maffioli, 2011).

Complementarity and substitution of funding sources has also been extensively investigated (David, Hall, & Toole, 2000). Scholars have studied the extent to which funding may attract additional funding as well as to what extent funding sources can be substituted. For example, Blume-Kohout et al. (2009) provide evidence that federal funding to US universities attract non-federal R&D funding. External funding may indeed attract industry support by acting as signal of high quality research (Connolly, 1997). Payne (2001) found complementary also between federal funding and private donations while Grimpe (2012) provided evidence of EU Framework Program 6 grants to be substitutive to grants from foundations and industry whereas government grants to be substitutive to foundations grants and complementary to industry funding.

Examining complementarity and substitution effects at individual (scientist) and research host organisation levels provides inevitably a limited perspective on the interaction among different funding sources. Such an approach neglects that research usually involves more than one individual and multiple research host organisations with associated funding that may complement or substitute at the single research output level, as mentioned above.

Figure 1: Journals-publication distribution (N= 7,510).



Empirical approach

Data

Our analysis focuses on cancer research in the UK, i.e. scientific publications related to cancer domains and involving at least one UK research host organisation. We focus on publications produced in the year 2011.[2] Research on the cancer domain represents a relevant setting for this research given the intense funding activity that has featured with it in most of the developed countries. Furthermore, the UK represents a exceptionally suitable context due the variety of organizations populating the funding system (Eckhouse, Lewison, & Sullivan, 2008).

Delineating a broad topic such as cancer is a complex task for scientometric analyses. Approaches using ad hoc searches for keywords (e.g. in titles and abstracts of publications) may yield a relatively high number of false positive and false negative cases (Leydesdorff, 1989). Our empirical approach instead builds the Medical Subject Heading (MeSH) classification of the MEDLINE/PubMed database (Leydesdorff, Rotolo, & Rafols, 2012). This classification provides a relatively comprehensive index with 'descriptors' that best represent each document's content. Descriptors are organised in tree like branching structure that comprehensively covers different medical domains and allow searches at various levels of specificity. Unlike keywords, which are added by authors and/or searched for by the analyst, the MeSH indexing process is a standardised process performed by examiners and offers analysts the choice of appropriate labels for their search domains.

We captured cancer related publications as those assigned to the "Neoplasms" descriptor and all related sub-levels in the MeSH tree.[3] This returned an initial sample of 115,101 documents published globally in 2011. We matched the MEDLINE/PubMed dataset with SCOPUS data for comprehensive data on authors' affiliations.[4] The match covered 98.1% of the initial sample of publications. We then identified UK publications as those involving at least one author affiliated to a UK research host organisation. The remaining unmatched records (1.9%) were manually screened and added to the dataset when a UK research host organisation was found. The resulting dataset included 7,922 publications (~6.9% of the global production). Those publications were distributed across 1,480 journals (see Figure 1). While we were not able to electronically access 130 journals due to publisher restrictions, access to full text of 7,510 publications was obtained (94.8% coverage).

For each publication we collected and coded acknowledgements to funders that authors made. Information was extracted from the following publications' headings: (i) "Acknowledgements", (ii) "Funding", (iii) "Conflicts of Interest", (iv) "Financial Disclosure", (v) "Role of Funding Sources", (vi) "Financial Supports", (vii) "Competing Financial Interests", and (ix) "Statement of interests". Acknowledgements were then read to establish

[2] To map the UK cancer funding landscape it is necessary to take a snapshot of the research and funding organisational 'eco-system' as a whole. To ensure comprehensive coverage a window of one full calendar year was selected based on the assumption that research active scientists would author at least one published paper per year and that even relatively small funding organisations would have been likely to have publications stemming from their work published during a given year. The intense data collection associated with this approach poses limits in examining cause-effect relationships. The data therefore provide only a cross-section of the funding 'eco-system'.

[3] The "neoplasms" descriptor is defined as: "New abnormal growth of tissue. Malignant neoplasms show a greater degree of anaplasia and have the properties of invasion and metastasis, compared to benign neoplasms".

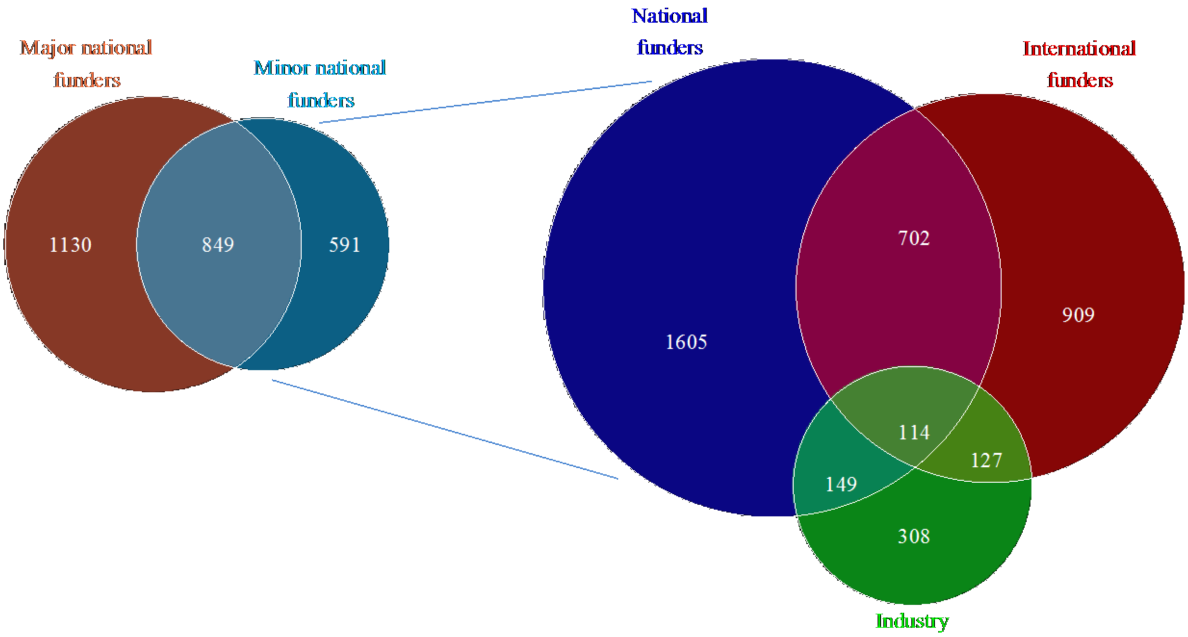
[4] The match was performed by using publications' MEDLINE/PubMed PMID and DOI.

whether they mentioned the support of funders. All sources of funding were recorded where authors indicated financial support for work leading to a publication. Other forms of support (e.g. colleagues reading drafts, helpful comments) were not coded. Declarations of conflicts of interest stemming from historic support, unrelated to the current publication, were excluded. The data collection and coding process is described elsewhere (Grassano et al., 2014; Shah et al., 2014).

Funding data: Descriptive statistics

The data on external funding sources revealed that 3,914 publications (52.1%) disclosed at least one funder in acknowledgement sections while 2,286 (30.4%) did not acknowledge external funding but made other forms of acknowledgement and 1,310 (17.5%) publications had no acknowledgement sections. Where funders were acknowledged, this may often be directly by name of organisation (in full or abbreviated form), section, funding scheme, or even by grant number alone. This required harmonising funders' names. The initial list included 4,086 names variations, which the harmonisation process reduced to 2,549 distinct funders' names. 663 UK funders, 1,579 non-UK funders, and 307 industrial funders compose the harmonised list of funders.[5] Furthermore, the average number of funding sources acknowledged in each publication is 3.3 when we focus on the sample of publications reporting at least on funding source (N=3,914).

Figure 2: Type of funding source and co-occurrence at the publication level.



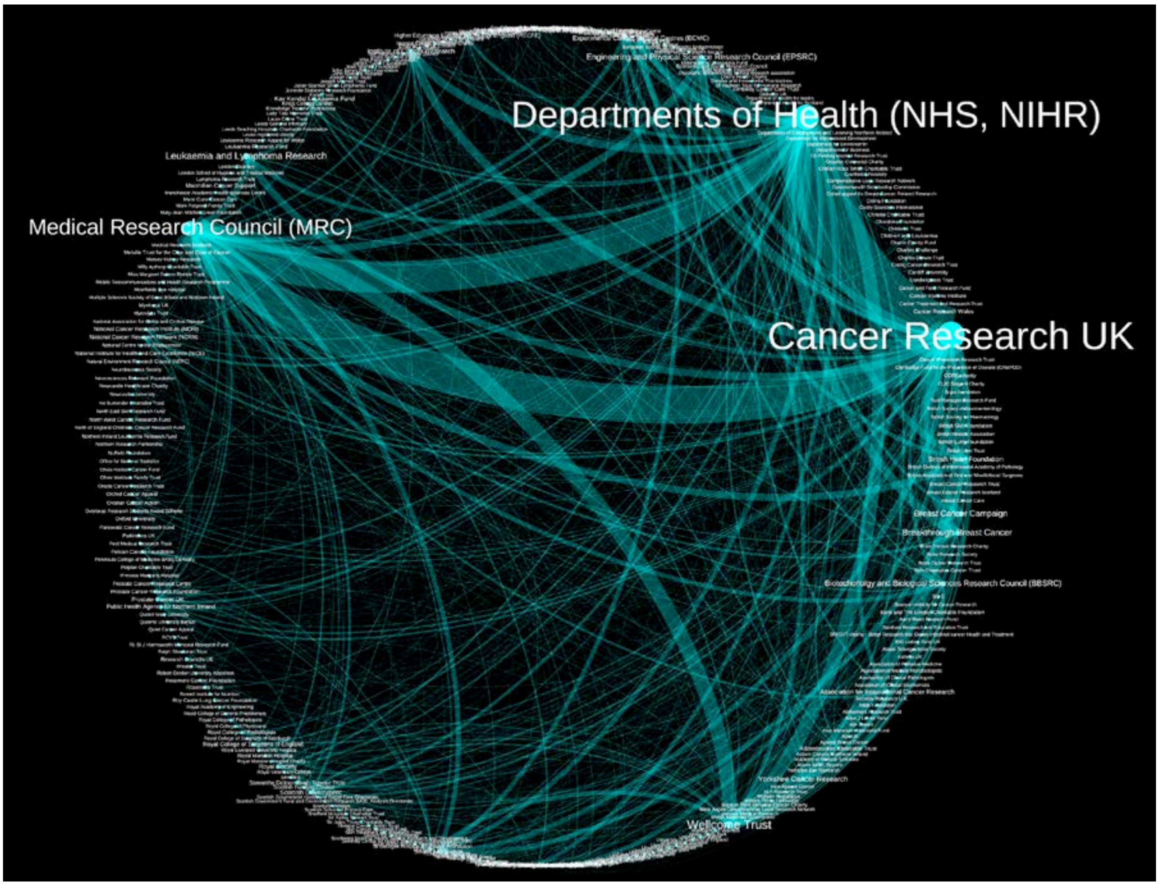
We distinguish between three types of sources: (i) national, (ii) international, and (iii) industry (see Figure 2).[6] National funding sources are also further distinguished in major and minor

[5] We performed the name harmonisation process also for authors' affiliations by considering the highest organisational level for consistency (e.g. not departments but universities). The initial list of research host organisations included 28,580 names variations. Those variations were reduced to 6,265 distinct research host organisations' names of which 1,155 and 5,110 were UK and non-UK, respectively

[6] Industrial support was kept as a separate category since most of the industrial actors supporting publications are located in multiple geographical areas and the information contained in the acknowledged sections does not allow to consistently geo-localise industrial funders.

when the national funder is acknowledged in at least or less than 2% of the publications in our sample, respectively.[7] 2,570 publications acknowledged at least one national external funding source. 591, 1,130, and 849 of these acknowledged support from minor, major, and jointly minor-major national funders, respectively. Figure 3 depicts the complex interaction among funding sources at publication level as measured by the co-occurrence of national funders in acknowledgment sections – industry is exclude here (but not elsewhere in our analysis) as are international funders (discussed elsewhere). The size of the nodes and labels is proportional to the number of publications in which a given funder was acknowledged. International funding sources were acknowledged in 1,852 publications while support from industry was reported in 698 publications.

Figure 3: Co-funding (co-occurrence) of minor and major national funders and co-funding matrix of major national funders.[8]



Empirical strategy: Variables and model specification

We analysed a number of variables that may affect the probability that certain types of external funding sources are acknowledged in a publication. Those variables are described in Table 1. Firstly, we tested for the extent to which our control variables are more likely to be associated with publications that acknowledged any source of funding. We specifically used the probit and logit estimations. We also estimated the relationships those variables have with

[7] It is worth noting that minor national funders are not necessarily of a small size since some of the large national funders may be less focused on the cancer domain, which is the focus of our analysis.

[8] Only national funders acknowledged in two or more publications are depicted. The nodes' size is proportional to the number of publications that acknowledged the given funder while the thickness of the line between two funders is proportional to the number of publications in which the two funders were jointly acknowledged.

the number of distinct acknowledged funding sources – we used both the Poisson and negative binomial estimations. Secondly, we tested for complementarity among funding sources as the extent to which certain sources significantly co-occur more than others in publications. We used a multivariate probit model (Cappellari & Jenkins, 2003) on the different categories of funding sources: national (major and minor), international, industry. The multivariate probit estimation returns correlations among the errors terms and these correlations can be used as an indication of complementarity or substitution between funding sources. For example, if a certain type of funding source is acknowledged in a publication to what extent another type of funding sources is also acknowledged or not in the same publications. Correlations among the errors terms and their statistical significance can provide evidence of these effects.

Table 1. Variables' description.

Variable	Description
National external funding source (d)	Dummy variable set to one if a national (UK) funding sources is acknowledged in the publication
Minor national external funding source (d)	Dummy variable set to one if a minor national (UK) funding sources is acknowledged in the publication — minor national funders are those acknowledged in less than 2% of the publication sample
Major national external funding source (d)	Dummy variable set to one if a major national (UK) funding sources is acknowledged in the publication — major national funders are those acknowledged in at least 2% of the publication sample
International external funding source (d)	Dummy variable set to one if an international (non-UK) funding sources is acknowledged in the publication
Industry external funding source (d)	Dummy variable set to one if an international (non-UK) funding sources is acknowledged in the publication
Highly-cited publications	Dummy variable set to one if the publication is within the top 5% of the most cited publication — the citation count is normalized by field as defined by the ISI SCs assigned to the journal were the research output is published
Research variety	Number of distinct MeSH qualifiers assigned to the publication
Research focus	Lowest level of the C04 (Neoplasms) descriptors assigned to the publication
External funding sources	Number of external funding sources
External funding source (d)	Dummy variable set to one if an external funding sources is acknowledged in the publication
Article (d)	Dummy variable set to one if the publication is an article
Review (d)	Dummy variable set to one if the publication is a review
Number of references	Number of references cited in the publication
Number of authors	Number of authors listed in the publication
Internationality (ratio)	Number of authors listed in the publication out the number of different countries involved as based on authors affiliation addresses
Affiliations variety (ratio)	Number of authors listed in the publication out the number of different affiliations
Research domains (dummy variables)	Dummy variables on research domains covered by the publication as defined by the qualifiers assigned at least 5% of the publication sample: pathology, genetics, surgery, diagnosis, drug therapy, metabolism, therapy, mortality, epidemiology, complications, etiology, secondary, radiotherapy, immunology, radiography, physiopathology, blood, and prevention and control
Productive research host (dummy variables)	Dummy variables for the most productive national (UK) research host organizations as those producing at least 1% of the publication sample: Cardiff University, Imperial College London, Institute of Cancer Research, King's College London, London Research Institute, Mount Vernon Hospital, Ninewells Hospital and Medical School, Oxford University, Queen Mary University of London, Queen's University, St George's Hospital NHS Trust, University College London, University of Aberdeen, University of Birmingham, University of Bristol, University of Cambridge, University of Edinburgh, University of Glasgow, University of Leeds, University of Leicester, University of Liverpool, University of Manchester, University of Newcastle, University of Nottingham, University of Sheffield, and University of Southampton

'(d)' refers to dummy variables.

Results

Results of the probit and logit estimations are reported in the Models 1 and 2 (Table 2). Certain research domains, such genetics, metabolism, epidemiology, and immunology, are more likely to acknowledge external funding sources. On the contrary, publications in surgery, diagnosis, therapy, complications, secondary, and radiography domains are less likely to acknowledge support from external funders. Articles and reviews are more likely to cite external funding sources and this probability also increases as the number of references increases. Finally, while publications involving a high number of authors and distinct research host organisations are more likely to have been supported by external funding sources, the number of distinct countries involved does not affect this likelihood. Most of these results are confirmed when the number of distinct funding sources acknowledged in publications is considered (Models 3 and 4).

Table 2. Regression results on the likelihood of acknowledging funding sources.

	(1) External funding source (d) (Probit)	(2) External funding source (d) (Logit)	(3) External funding sources (Poisson)	(4) External funding sources (Negative binomial)
Cardiff University (d)	0.233^{+} (0.120)	0.487* (0.199)	-0.124 (0.096)	-0.092 (0.092)
Imperial College London (d)	0.119 (0.079)	0.292* (0.137)	0.517*** (0.096)	0.275** (0.089)
Institute of Cancer Research (d)	0.412*** (0.083)	0.823*** (0.150)	0.428*** (0.064)	0.463*** (0.074)
King's College London (d)	0.069 (0.089)	0.166 (0.155)	-0.124 (0.085)	-0.029 (0.085)
London Research Institute (d)	0.598** (0.206)	1.065** (0.365)	0.171^{+} (0.091)	0.210* (0.101)
Mount Vernon Hospital (d)	-0.168 (0.228)	-0.173 (0.344)	-0.291 (0.215)	-0.124 (0.256)
Ninewells Hospital and Medical School (d)	0.112 (0.171)	0.266 (0.279)	0.430** (0.160)	0.374* (0.164)
Oxford University (d)	0.427*** (0.080)	0.897*** (0.137)	0.604*** (0.064)	0.343*** (0.069)
Queen Mary University of London (d)	0.116 (0.123)	0.238 (0.208)	-0.154 (0.120)	0.050 (0.097)
Queen's University (d)	0.494*** (0.145)	0.872*** (0.247)	0.100 (0.177)	0.241* (0.107)
St George's Hospital NHS Trust (d)	0.316* (0.143)	0.582* (0.249)	-0.459* (0.188)	-0.012 (0.134)
University College London (d)	0.372*** (0.068)	0.638*** (0.119)	0.316*** (0.065)	0.308*** (0.073)
University of Aberdeen (d)	0.553** (0.192)	1.051** (0.322)	-0.200 (0.196)	0.152 (0.134)
University of Birmingham (d)	0.443*** (0.119)	0.826*** (0.206)	0.076 (0.097)	0.188* (0.089)
University of Bristol (d)	0.796*** (0.233)	1.808*** (0.396)	0.278 (0.248)	0.767*** (0.158)
University of Cambridge (d)	0.308*** (0.083)	0.645*** (0.141)	0.699*** (0.063)	0.314*** (0.074)
University of Edinburgh (d)	0.296* (0.123)	0.563** (0.208)	0.303*** (0.079)	0.154^{+} (0.090)
University of Glasgow (d)	0.217* (0.086)	0.413** (0.146)	0.064 (0.102)	0.021 (0.083)
University of Leeds (d)	0.273** (0.083)	0.496*** (0.143)	0.080 (0.077)	0.096 (0.072)
University of Leicester (d)	0.550** (0.184)	0.951** (0.325)	0.319* (0.154)	0.510** (0.164)
University of Liverpool (d)	0.300* (0.126)	0.514* (0.219)	-0.163 (0.191)	0.009 (0.120)
University of Manchester (d)	0.577*** (0.075)	1.020*** (0.129)	0.392*** (0.116)	0.455*** (0.098)
University of Newcastle (d)	-0.016 (0.107)	0.016 (0.184)	0.020 (0.104)	-0.046 (0.105)
University of Nottingham (d)	0.557*** (0.138)	0.980*** (0.238)	0.058 (0.102)	0.103 (0.101)
University of Sheffield (d)	0.319* (0.130)	0.677** (0.207)	-0.174 (0.193)	-0.103 (0.130)
University of Southampton (d)	0.409*** (0.117)	0.733*** (0.206)	0.244* (0.109)	0.184^{+} (0.105)
Pathology (d)	-0.030 (0.072)	-0.073 (0.119)	-0.001 (0.076)	0.031 (0.069)
Genetics (d)	0.505*** (0.086)	0.837*** (0.146)	0.369*** (0.071)	0.380*** (0.067)
Surgery (d)	-0.730*** (0.086)	-1.278*** (0.151)	-0.794*** (0.218)	-0.726*** (0.169)
Diagnosis (d)	-0.258*** (0.076)	-0.460*** (0.128)	-0.333*** (0.088)	-0.285*** (0.086)
Drug therapy (d)	0.133^{+} (0.080)	0.191 (0.132)	-0.031 (0.090)	0.056 (0.090)
Metabolism (d)	0.539*** (0.093)	0.919*** (0.155)	0.289** (0.089)	0.400*** (0.075)
Therapy (d)	-0.201* (0.089)	-0.389** (0.150)	-0.198 (0.134)	-0.111 (0.130)
Mortality (d)	0.124 (0.125)	0.188 (0.217)	0.154 (0.171)	0.130 (0.146)
Epidemiology (d)	0.195^{+} (0.112)	0.336^{+} (0.183)	0.467*** (0.105)	0.154 (0.117)
Complications (d)	-0.549*** (0.103)	-0.932*** (0.179)	-0.822*** (0.144)	-0.709*** (0.139)
Etiology (d)	0.150 (0.127)	0.276 (0.208)	0.235 (0.155)	0.141 (0.117)
Secondary (d)	-0.412** (0.126)	-0.695** (0.212)	-0.636*** (0.171)	-0.603*** (0.168)
Radiotherapy (d)	-0.198 (0.155)	-0.353 (0.252)	-0.261 (0.178)	-0.173 (0.162)
Immunology (d)	0.446** (0.137)	0.688** (0.233)	0.590** (0.220)	0.499*** (0.124)
Radiography (d)	-0.395** (0.152)	-0.639* (0.250)	-0.543* (0.223)	-0.432* (0.211)
Physiopathology (d)	0.214 (0.161)	0.329 (0.274)	0.043 (0.173)	0.216 (0.165)
Blood (d)	-0.191 (0.172)	-0.321 (0.284)	0.246 (0.190)	-0.086 (0.149)
Prevention and control (d)	0.150 (0.135)	0.241 (0.225)	-0.073 (0.118)	-0.074 (0.115)
Article (d)	0.827*** (0.067)	1.368*** (0.118)	0.999*** (0.148)	0.829*** (0.131)
Review (d)	0.368*** (0.090)	0.579*** (0.159)	0.410* (0.162)	0.327* (0.146)
Number of references	0.005*** (0.001)	0.010*** (0.002)	0.005*** (0.001)	0.007*** (0.001)
Number of authors	0.055*** (0.010)	0.126*** (0.016)	0.011*** (0.002)	0.049*** (0.009)
Internationality (ratio)	0.000 (0.014)	-0.024 (0.024)	-0.023* (0.009)	-0.027* (0.011)
Affiliations variety (ratio)	0.053*** (0.015)	0.096*** (0.025)	0.019 (0.016)	0.038* (0.015)
Highly-cited publications	0.087 (0.089)	0.127 (0.154)	-0.083 (0.103)	-0.018 (0.083)
Research variety	-0.032 (0.056)	-0.057 (0.091)	0.004 (0.067)	0.015 (0.055)
Research focus	-0.048*** (0.010)	-0.088*** (0.016)	-0.003 (0.013)	-0.041*** (0.012)
Constant	-1.316*** (0.084)	-2.344*** (0.143)	-0.918*** (0.162)	-1.090*** (0.142)
Degree of freedom	53	53	53	53
Log likelihood	-3903.3	-3858.2	-14645.3	-11538.1
(Wald)χ^2	1174.1***	998.1***	2819.6***	1238.0***
Observations	7510	7510	7510	7510

Robust standard errors in parentheses clustered at journal level.
'(d)' refers to dummy variables.
$^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

The results of the multivariate probit model are reported in Table 3. The benchmark case is represented by publications not reporting any source of external funding. The multivariate probit model tests for correlation of the errors terms between the four equations. This allows testing for complementarity among the dependent variables. A positive correlation between major and minor national funding sources (*rho = 0.587, p<0.001*) is found. Major and minor national funding sources are also positively correlated with support from international funders (*rho = 0.057, p<0.05*; *rho = 0.135, p<0.001*, respectively) while they are not significantly correlated with funding from industry. Results also show a positive correlation between international and industrial funding sources (*rho = 0.094, p<0.01*).

Table 3. Multivariate probit model: Major/minor national, international, and industry funders.

	Minor national external funding source (d)	Major national external funding source (d)	International external funding source (d)	Industry external funding source (d)
Cardiff University (d)	-0.087 (0.105)	-0.045 (0.113)	0.137 (0.103)	0.263* (0.121)
Imperial College London (d)	0.093 (0.067)	0.043 (0.065)	0.089 (0.059)	0.132^{+} (0.079)
Institute of Cancer Research (d)	0.183** (0.062)	0.231*** (0.060)	0.213*** (0.059)	0.184* (0.082)
King's College London (d)	0.031 (0.078)	-0.030 (0.074)	0.014 (0.068)	-0.140 (0.093)
London Research Institute (d)	0.273* (0.129)	0.077 (0.190)	0.111 (0.163)	-0.033 (0.234)
Mount Vernon Hospital (d)	0.111 (0.163)	-0.321 (0.199)	-0.122 (0.187)	0.158 (0.196)
Ninewells Hospital and Medical School (d)	0.461** (0.168)	0.143 (0.155)	-0.050 (0.169)	0.096 (0.193)
Oxford University (d)	0.236*** (0.059)	0.320*** (0.060)	0.032 (0.069)	-0.016 (0.084)
Queen Mary University of London (d)	0.191* (0.094)	0.157^{+} (0.089)	0.009 (0.093)	0.115 (0.106)
Queen's University (d)	0.141 (0.124)	0.026 (0.125)	0.107 (0.131)	0.045 (0.151)
St George's Hospital NHS Trust (d)	-0.023 (0.136)	0.257^{+} (0.134)	-0.000 (0.136)	-0.013 (0.184)
University College London (d)	0.088 (0.059)	0.242*** (0.056)	0.167** (0.060)	0.157^{+} (0.083)
University of Aberdeen (d)	0.140 (0.148)	0.301* (0.135)	0.247^{+} (0.141)	0.036 (0.214)
University of Birmingham (d)	0.038 (0.104)	0.239* (0.094)	0.200* (0.093)	0.108 (0.119)
University of Bristol (d)	0.422** (0.133)	0.467** (0.156)	0.448*** (0.136)	0.042 (0.169)
University of Cambridge (d)	0.126^{+} (0.068)	0.142* (0.065)	0.086 (0.066)	0.187* (0.076)
University of Edinburgh (d)	0.060 (0.099)	0.100 (0.083)	0.059 (0.099)	0.272* (0.107)
University of Glasgow (d)	0.068 (0.087)	-0.030 (0.079)	0.135^{+} (0.078)	0.286** (0.098)
University of Leeds (d)	0.095 (0.083)	0.142* (0.071)	0.173** (0.063)	0.106 (0.092)
University of Leicester (d)	0.215 (0.166)	0.105 (0.144)	0.123 (0.155)	0.168 (0.176)
University of Liverpool (d)	0.166 (0.113)	0.203^{+} (0.110)	0.161 (0.125)	0.133 (0.136)
University of Manchester (d)	0.204** (0.078)	0.236*** (0.066)	0.231*** (0.069)	0.100 (0.087)
University of Newcastle (d)	0.040 (0.104)	0.019 (0.111)	-0.027 (0.103)	0.110 (0.125)
University of Nottingham (d)	0.207 (0.128)	0.192 (0.123)	0.343* (0.141)	0.458** (0.153)
University of Sheffield (d)	0.138 (0.111)	0.050 (0.105)	-0.047 (0.095)	0.074 (0.111)
University of Southampton (d)	-0.075 (0.116)	0.161 (0.106)	0.074 (0.111)	0.196 (0.128)
Pathology (d)	-0.041 (0.068)	-0.046 (0.074)	0.008 (0.070)	0.181^{+} (0.093)
Genetics (d)	0.149^{+} (0.077)	0.372*** (0.081)	0.242*** (0.072)	0.245* (0.100)
Surgery (d)	-0.470*** (0.095)	-0.606*** (0.100)	-0.404*** (0.085)	-0.272* (0.116)
Diagnosis (d)	-0.213** (0.079)	-0.128^{+} (0.077)	-0.129^{+} (0.075)	-0.179 (0.111)
Drug therapy (d)	0.036 (0.071)	0.087 (0.068)	0.175** (0.068)	0.122 (0.091)
Metabolism (d)	0.192* (0.084)	0.408*** (0.078)	0.382*** (0.078)	0.113 (0.108)
Therapy (d)	-0.138^{+} (0.080)	-0.127 (0.087)	-0.054 (0.089)	-0.050 (0.124)
Mortality (d)	0.107 (0.111)	0.023 (0.114)	0.086 (0.108)	0.148 (0.131)
Epidemiology (d)	-0.021 (0.091)	0.111 (0.103)	0.287** (0.091)	0.237* (0.117)
Complications (d)	-0.442*** (0.122)	-0.317** (0.112)	-0.362** (0.118)	-0.328* (0.162)
Etiology (d)	0.107 (0.121)	0.085 (0.102)	0.181 (0.142)	0.186 (0.141)
Secondary (d)	-0.477** (0.160)	-0.252 (0.156)	-0.270^{+} (0.162)	-0.222 (0.183)
Radiotherapy (d)	-0.031 (0.103)	-0.178 (0.135)	-0.118 (0.123)	0.123 (0.200)
Immunology (d)	0.031 (0.145)	0.116 (0.143)	0.137 (0.117)	0.479** (0.152)
Radiography (d)	-0.494*** (0.147)	-0.172 (0.156)	-0.258 (0.162)	-0.338 (0.266)
Physiopathology (d)	0.026 (0.160)	0.008 (0.165)	0.112 (0.159)	0.089 (0.201)
Blood (d)	-0.023 (0.144)	0.040 (0.144)	0.076 (0.166)	-0.474^{+} (0.269)
Prevention and control (d)	-0.021 (0.125)	0.196^{+} (0.115)	0.116 (0.111)	0.054 (0.175)
Article (d)	0.483*** (0.087)	0.519*** (0.069)	0.619*** (0.066)	0.473*** (0.089)
Review (d)	0.159 (0.101)	0.226** (0.085)	0.289*** (0.083)	0.192^{+} (0.113)
Number of references	0.003*** (0.001)	0.003*** (0.001)	0.004*** (0.001)	0.004*** (0.001)
Number of authors	0.003^{+} (0.002)	0.004 (0.003)	0.003 (0.002)	0.000 (0.002)
Internationality (ratio)	0.024** (0.008)	0.012 (0.008)	0.006 (0.008)	-0.005 (0.012)
Affiliations variety (ratio)	-0.010 (0.011)	0.022^{+} (0.012)	0.011 (0.012)	0.020 (0.013)
Highly-cited publications	0.097 (0.079)	0.019 (0.065)	0.160* (0.080)	-0.002 (0.090)
Research variety	-0.012 (0.056)	-0.024 (0.058)	-0.026 (0.056)	-0.146^{+} (0.075)
Research focus	-0.038*** (0.010)	-0.026** (0.010)	-0.032*** (0.009)	-0.018 (0.011)
Constant	-1.389*** (0.096)	-1.332*** (0.087)	-1.414*** (0.085)	-1.867*** (0.104)
$\rho_{21}, \rho_{31}, \rho_{41}$		0.587*** (0.028)	0.057* (0.027)	-0.023 (0.032)
ρ_{32}, ρ_{42}			0.135*** (0.029)	0.029 (0.030)
ρ_{43}				0.094** (0.030)
Degree of freedom	212			
Log likelihood	-13078.0			
χ^2	2301.9***			
Observations	7510			

Robust standard errors in parentheses clustered at journal level.
(d) refers to dummy variables.
$^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

As a robustness check, we also explored complementarity among funding sources by considering each source of funding both as treatment and effect. We specifically tested for the extent to which if a given source of funding is acknowledged in a publication (treatment) also another category of funding is acknowledged and vice versa. To do so, we used the nearest-neighbour matching estimators for average treatment effects (Abadie, Drukker, Herr, & Imbens, 2004). The nearest-neighbour matching estimators for average treatment effects matched each observation in the treatment group with other four observations in the control group (Abadie et al., 2004). Those observations are similar across a number of covariates, i.e. all our control variables and the remaining category of funding source.

Table 4. Nearest-neighbour matching estimators for average treatment effects.

Treatment (funding source)	Effect (funding source)	SATE
International	National	0.081*** (0.018)
National	International	0.028* (0.014)
Industry	National	0.021 (0.026)
National	Industry	-0.025* (0.009)
Industry	International	0.125*** (0.027)
International	Industry	0.040** (0.013)
Observations	7510	

SATE indicates the sample average treatment effect.
Bias-corrected matching estimators are reported.
$^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Results, reported in Table 4, support previous findings. Non-industry national and international funding sources are complementary. This is confirmed by the average treatment effects of international funding sources on national funding sources (*SATE = 0.081, p<0.001*) and vice versa (*SATE = 0.028, p<0.05*). Similarly, international funding sources are found to be complementary to industrial support when both sources are alternatively considered as treatment and effect (*SATE = 0.125, p<0.001; SATE = 0.040, p<0.01*). The results also provide evidence of a substitution effect when the treatment effect of national funding on the likelihood that a publication acknowledges industrial support is considered (*SATE = −0.025, p<0.05*). However, this effect is only partial since it is not confirmed when the support from industrial funders is considered as the treatment and support from national funders as the effect (*SATE = 0.021, p>0.1*). We therefore conclude that there is not enough empirical evidence to suggest the presence of a substitution effect between national and industry funding sources.

Conclusion

Scientists rely on a large variety of funding sources to produce new scientific knowledge and those sources show strong complementarity when their interaction is analyzed at single research output level. Complementarity is found among national funders, with combinations of national and international funders as well as where international and industry sources are accessed. However, in contrast to prior research that has focused on the author or research organisation levels, our analysis reveals no substitution effect for funding sources at the publication level. These findings have important policy implications. They specifically suggest that, given the complementarity among founding sources, cuts both at national and international level of the budgets for science may have a disproportional negative impact on national scientific research output.

References

Abadie, A., Drukker, D., Herr, J. L., & Imbens, G. (2004). Implementing matching estimators for average treatment effects in Stata. *Stata Journal*, *4*(3), 290–311.

Adams, J. D., Black, G. C., Clemmons, J. R., & Stephan, P. E. (2005). Scientific teams and institutional collaborations: Evidence from U.S. universities, 1981–1999. *Research Policy*, *34*(3), 259–285.

Arora, A., David, P. A., & Gambardella, A. (1998). Reputation and Competence in Publicly Funded Science: Estimating the Effects on Research Group Productivity. *Annals of Economics and Statistics / Annales d'Économie et de Statistique*, (49/50), 163–198

Beaudry, C., & Allaoui, S. (2012). Impact of public and private research funding on scientific production: The case of nanotechnology. *Research Policy*, *41*(9), 1589–1606.

Blume-Kohout, M. E., Kumar, K. B., & Sood, N. (2009). *Federal Life Sciences Funding and University R&D* (Working Paper No. 15146). National Bureau of Economic Research.

Bonaccorsi, A., Daraio, C., & Simar, L. (2006). Advanced indicators of productivity of universitiesAn application of robust nonparametric methods to Italian data. *Scientometrics*, *66*(2), 389–410.

Bozeman, B., & Gaughan, M. (2007). Impacts of grants and contracts on academic researchers' interactions with industry. *Research Policy*, *36*(5), 694–707.

Cappellari, L., & Jenkins, S. P. (2003). Multivariate probit regression using simulated maximum likelihood. *Stata Journal*, *3*(3), 278–294.

Connolly, L. S. (1997). Does external funding of academic research crowd out institutional support? *Journal of Public Economics*, *64*(3), 389–406.

David, P. A., Hall, B. H., & Toole, A. A. (2000). Is public R&D a complement or substitute for private R&D? A review of the econometric evidence. *Research Policy*, *29*(4–5), 497–529.

Defazio, D., Lockett, A., & Wright, M. (2009). Funding incentives, collaborative dynamics and scientific productivity: Evidence from the EU framework program. *Research Policy*, *38*(2), 293–305.

Eckhouse, S., Lewison, G., & Sullivan, R. (2008). Trends in the global funding and activity of cancer research. *Molecular Oncology*, *2*(1), 20–32.

Godin, B. (1996). The Impact of Research Grants on the Productivity and Quality of Scientific Research. *Working Paper. INRS. (Institut National de La Recherche Scientifique), Montreal, Quebec.*

Grassano, N., Hopkins, M., Rotolo, D., Hutton, J., Lang, F., Crane, P., … Sharp, T.-A. (2014). Funding data: Collection, coding, and caveats. Cancer research funding in the UK. *In Preparation*.

Grimpe, C. (2012). Extramural research grants and scientists' funding strategies: Beggars cannot be choosers? *Research Policy*, *41*(8), 1448–1460. doi:10.1016/j.respol.2012.03.004

Jacob, B., & Lefgren, L. (2007). *The Impact of Research Grant Funding on Scientific Productivity* (Working Paper No. 13519). National Bureau of Economic Research.

Katz, J. S., & Martin, B. R. (1997). What is research collaboration? *Research Policy*, *26*(1), 1–18.

Lee, S., & Bozeman, B. (2005). The Impact of Research Collaboration on Scientific Productivity. *Social Studies of Science*, *35*(5), 673–702. doi:10.1177/0306312705052359

Leydesdorff, L. (1989). Words and co-words as indicators of intellectual organization. *Research Policy*, *18*(4), 209–223.

Leydesdorff, L., Rotolo, D., & Rafols, I. (2012). Bibliometric perspectives on medical innovation using the medical subject Headings of PubMed. *Journal of the American Society for Information Science and Technology*, *63*(11), 2239–2253.

Payne, A. A. (2001). Measuring the Effect of Federal Research Funding on Private Donations at Research Universities: Is Federal Research Funding More than a Substitute for Private Donations? *International Tax and Public Finance*, *8*(5-6), 731–751.

Payne, A. A., & Siow, A. (2003). Does Federal Research Funding Increase University Research Output? *Advances in Economic Analysis & Policy*, *3*(1).

Shah, K., Sussex, J., Hernandez-Villafuerte, K., Garau, M., Rotolo, D., Hopkins, M. M., … Sharp, T.-A. (2014). *Exploring Interdependencies of research funders in the UK*. Office of Health Economics (London, United Kingdom) and SPRU-University of Sussex (Brighton, United Kingdom) for Cancer Research UK. Retrieved from www.cancerresearchuk.org/sites/default/files/interdependencies_of_funders_in_the_uk_march2014.pdf

Ubfal, D., & Maffioli, A. (2011). The impact of funding on research collaboration: Evidence from a developing country. Research *Policy*, *40*(9), 1269–1279.

The Emergence of Molecular Biology in the Diagnosis of Cervical Cancer: A Network Perspective[1]

Daniele Rotolo*, Michael Hopkins**, Ismael Rafols*** and Stuart Hogarth****

* *d.rotolo@sussex.ac.uk*
SPRU (Science and Technology Policy Research), University of Sussex, Brighton, BN1 9SL, (UK)

** *m.m.hopkins@sussex.ac.uk*
SPRU (Science and Technology Policy Research), University of Sussex, Brighton, BN1 9SL, (UK)

*** *i.rafols@ingenio.upv.es*
Ingenio (CSIC-UPV), Universitat Politècnica de València, Camí de Vera, s/n, València, 46022, (Spain)
&
SPRU (Science and Technology Policy Research), University of Sussex, Brighton, BN1 9SL, (UK)

**** *stuart.hogarth@kcl.ac.uk*
Department of Social Science, Health and Medicine, King's College London, Strand, London, WC2R 2LS, (UK)

Abstract
Cytology-based technologies have been extensively used for decades to screen for cervical cancer in women despite the large number of false negative cases these technologies may report. The rise of molecular biology, since mid-1980s, has spurred the emergence of novel screening technologies, which have significantly changed both the research landscape and clinical practices around cervical cancer. Within this context, the present paper examines how different institutional groups of actors have contributed to the emergence of molecular biology from an inter-organisational network lens. To do so, we analyse the patterns of network interactions among different groups involved in the emerge process. We specifically examine the formation of ties (dyads) as well as the extent to which organisational actors operate in different brokerage positions (triads). The analysis is based on a sample of scientific articles published over more than 30 years in the diagnosis domain of cervical cancer research. Findings provide evidence that the process of tie formation as well as the brokerage activity follow different patterns according to the considered institutional group. The process of tie formation and brokerage activity also evolve over emergence.

Keywords: cervical cancer; diagnosis; molecular biology; tie formation; network brokerage; inter-organisational networks; emerging technologies.

Introduction
Cervical cancer is one of the most common cancers among women. About 530,000 new cervical cancers occur and cause about 275,000 deaths each year. Large screening programs are generally credited with decreasing its impact, though cervical cancer still represents an important issue in less developed countries where screening is not available, which account

[1] Daniele Rotolo, Michael Hopkins, and Ismael Rafols acknowledge the support of UK Economic and Social Research Council (award RES-360-25-0076 - "Mapping the Dynamics of Emergent Technologies") in the development of the case studies. The findings and observations contained in this paper are those of the authors and do not necessarily reflect the funder's view.

for about 85% and 88% of new cases and deaths, respectively.[2] Cytology-based diagnostic technologies, such as the Pap test, have dominated screening programs for decades despite their well documented low sensitivity – the Pap test may report from 15% to 50% false negative cases. This 'mono' diagnostic approach started to be challenged only in the late 1990s due to major advancements in molecular biology juxtaposed with earlier key discoveries in pathology. Molecular biology offered the opportunity to address the limited sensitivity of cytology-based testing technologies, thus spurring the emergence of a novel stream of diagnostics.

Different groups of actors may have significantly shaped the development and adoption of molecular diagnostic technologies in cervical cancer screening domain. Research on technological change has indeed provided evidence of the strong influence actors with associated interests and visions may exert on the directionality of emerging technologies (e.g. Geels, 2002; Stirling & Scoones, 2009). Within this process, the complex networks of inter-organisational relationships play a key role for the 'problematisation' of an emerging technology (Blume, 1992). Networks provide access to knowledge and resources as well as allocate power, control, and influence (Brass, 1992). Certain actors can strengthen the system of innovation by increasing the cohesion of the network. By mediating between actors otherwise unconnected or weakly connected, these actors can also reduce the resilience of the network and allocate power and control disproportionally.

Despite the importance of inter-organisational networks in the shaping of emerging technologies, our understanding of the emergence process from a network perspective is limited. We aim to fill this gap by examining inter-organisational networks over the emergence of the molecular biology in the diagnosis domain of cervical cancer research. To do so, we distinguish between different phases of emergence (see Blume, 1992) – exploration, development, adoption, and growth – and identify different groups of actors, namely 'institutional groups'. We then examine the inter-organisational network at the level of dyads (tie formation) and triads (brokerage activity). Our empirical analysis is based on the scientific articles related to the domain of diagnosis for cervical cancer and published from 1980 to 2011. We used bibliometric data to construct the inter-organisational networks over the emergence period.

Theoretical background

Emerging technologies are important drivers of technological change as documented across a number of research streams. Emergence can be either 'constructive' or 'destructive' (Goldstein, 1999). New industries, for example, may be created, whereas existing ones may be significantly changed (Day & Schoemaker, 2000). The creative and destructive power of emerging technologies was noted for the first time by Schumpeter (1934) and subsequently further investigated by a number of scholars (for a review see Martin, 2012).

Innovation studies and evolutionary economics have framed the dynamics of emerging technologies in terms of trajectories that develop and are selected within paradigms (Dosi, 1982), while scholars in science and technology studies have emphasized the role of socio-technical regimes (Geels, 2002) including the visions, perceptions, strategies, and expectations of the involved groups of actors in shaping the directionality of technological developments (van Lente & Rip, 1998). Research has also highlighted how the direction of

[2] GLOBOCAN 2008 available at http://globocan.iarc.fr.

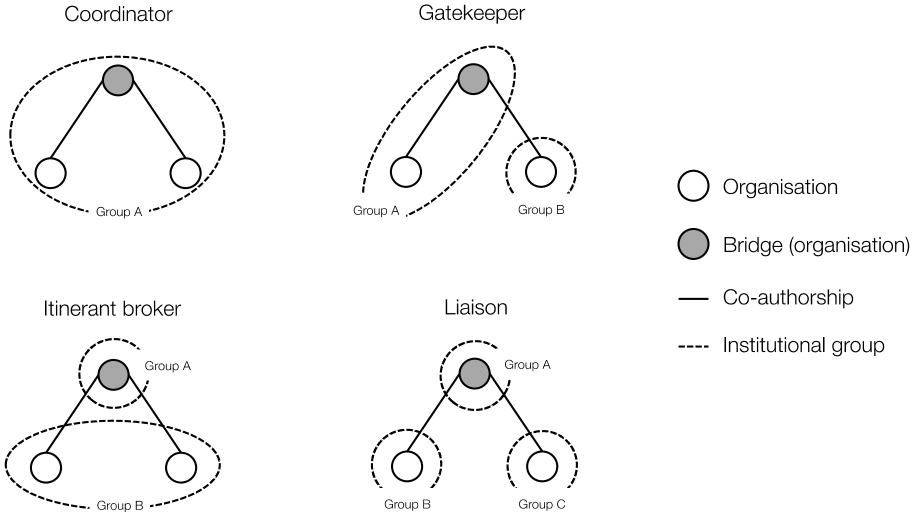
developments can be steered more easily in the early period of emergence, whereas in a later stage network-incumbents may influence the selection environment in maintaining the momentum of established technological trajectories (Collingridge, 1980). This may in turn create technological path dependence and lock-in (Arthur, 1989; David, 1985).

Sources of innovation are not embedded in one single organisational actor, but “they are commonly found in the interstices between firms, universities, research laboratories, suppliers, and customer” (Powell, 1990, p. 118). Innovation is a distributed process that relies on the coordination of a multiplicity of actors contributing to the process with different levels of involvement (Garud & Karnoe, 2003). This is also largely emphasised by the literature examining ‘systems of innovations’ (e.g. Freeman, 1995; Malerba, 2002). In the context of emerging technologies the set of relationships interconnecting the large variety of actors involved in the process of emergence assumes therefore particular importance.

Inter-organisational networks play a key role in the ‘problematisation’ of the technology (Blume, 1992). They provide both access to knowledge and resources as well as distribute power and control. Certain network configurations favour interactions among actors, others concentrate control and power on few – e.g. on those actors linking otherwise weakly connected or disconnected actors (Burt, 1992). This in turn can increase the involvement of some actors in the emergence process at the same time excluding others (Willer & Willer, 2000). Positions in inter-organisational networks have also signalling properties, which enable actors to build a reputation that extend beyond their direct ties (Gulati & Gargiulo, 1999).

To examine this facet of the emergence process, we conceptualise different organisational groups on the basis of the main institutional role of the actors comprising them. We specifically classify actors into five ‘institutional groups’: Research and Higher Education (RHE), Governmental (GOV), Hospital and Care (HC), Industrial (IND), and Non-Governmental (NGO) organisations (see Table 1).

Figure 1: Institutional groups and bridging roles.



Our main theoretical focus is at the micro level since a perspective at the level of nodes and ties can inform on the evolutionary process with more granularity than a perspective at the macro level. To this end, we first explore the formation of dyads (ties) both within and between institutional groups and then extend the dyadic perspective to network triads. We build on Fernandez and Gould's (1994) taxonomy of brokerage roles (see Figure 1). An organisation specifically acts as coordinator when it mediates between actors in the same institutional group, while it acts as gatekeeper when it screens and gathers knowledge and resources from another group and distributes them in its own group.[3] Itinerant broker (or cosmopolitan) and liaison organisations act instead as intermediaries between two organisations that belong to the same or different institutional groups, respectively.

Research setting: The case of cervical cancer

Human Papilloma Virus (HPV) infections have been identified as a necessary, but not sufficient, cause of cervical cancer (zur Hausen, 1987).[4] The first clue on the viral origin of cervical cancer can be found in a study conducted by an Italian physician, Domenico Rigoni-Stern, in 1842 (Rigoni-Stern, 1842). Rigoni-Stern produced mortality statistics of women dying of cancer in the city of Verona. This analysis pointed out that the cancer of uterus was much more common in married women and widows than in virgins and nuns. More than a century later cytologists started to recognize the presence of 'koilocytes' – cells characterized by large nuclei and large perinuclear spaces – as a manifestation of a viral infection of the genital 'condylomas'. The link between koilocytes and condylomas, hence the viral origin of cervical cancer, attracted the interest of virologists (Reynolds & Tansey, 2009). Great efforts were made in investigating a number of viruses potentially associated with cervical cancer (e.g. the herpes simplex virus) until early 1980s when a German research team led by Harald zur Hausen, at the German Cancer Research Centre, provided evidence of the strong association between cervical cancer and HPV (Clayton, 2012).[5] This discovery supported the idea that certain viruses by infecting the cell are able to change the cell's properties turning it into a cancerous cell.

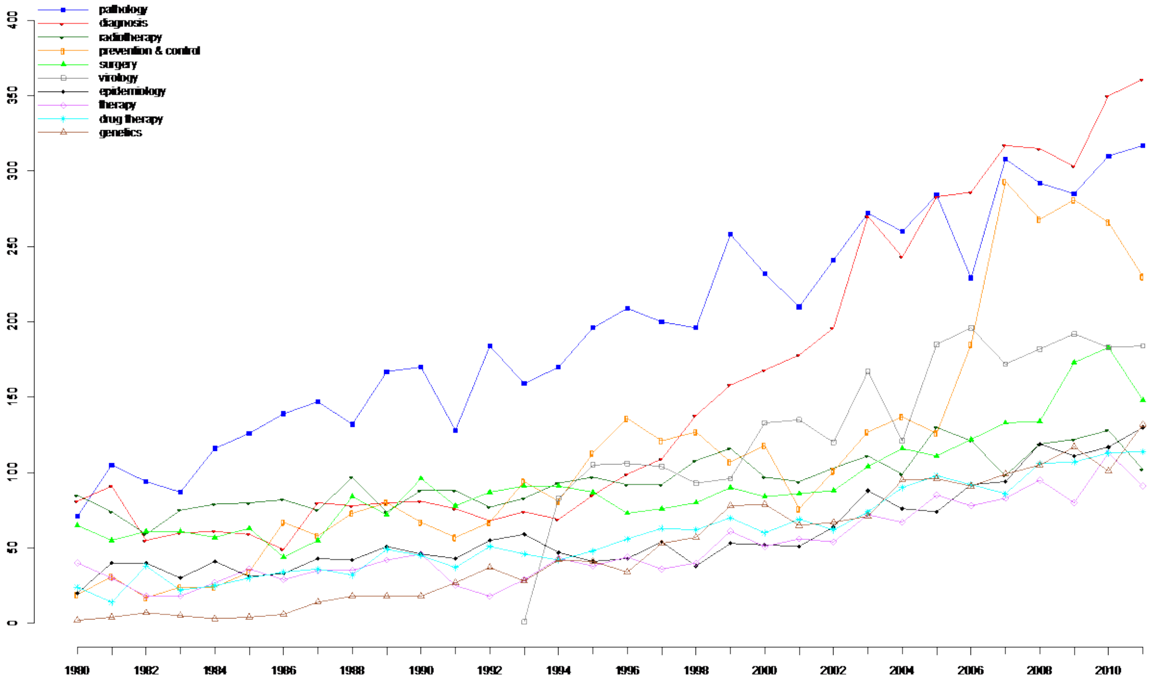
Table 1. Institutional groups and organisational actors.

Institutional group	Organisations
Research and Higher Education (RHE)	Universities, University hospitals, Medical centres (involved in teaching activities), Colleges and schools, Research centres/institutes
Governmental (GOV)	National/public agencies/bureaus, Health boards, National institutes, Registries, National laboratories, Public departments, Councils, National programmes/initiatives, Ministries, Regional centres, International organisations supported by governments
Hospitals and Care (HC)	Medical centres, Clinics, Infirmaries, Cancer centres, Medical Groups/Associates, Healthcare providers
Industry (IND)	Pharmaceutical and biotechnology firms, Consultants, Laboratories (when independent from other organisations)
Non-governmental (NGO)	Societies, Associations, Other non-profit organisations, Charities, Foundations
Other (OT)	Organisations not in the previous categories

[3] Gatekeeping and representative roles overlap when, as in our case, undirected networks are considered.
[4] HPVs are small DNA tumour viruses that belong to the family of *Papovaviridae*.
[5] zur Hausen won the Nobel Prize in "Physiology or Medicine" in 2008.

Subsequent studies deepened the understanding of the natural history of the HPV infection and cervical cancer. A persistent infection, may eventually cause cervical cancer, evolving through four stages (Schiffman, Castle, Jeronimo, Rodriguez, & Wacholder, 2007). First, the virus is sexually transmitted. Its infection manifests in inconspicuous squamous intraepithelial lesions (SIL) of the cervix. This initial stage is classified as 'Cervical Intraepithelial Neoplasia 1' (CIN1). Second, CIN1 lesions are, in most cases, cleared within 12 months. Yet, a small percentage (less than 10%) of women do not clear them – while genetic predisposition, smoking, high number of children, and viral load are the main factors in affecting the capability to clear the infection the associated mechanisms are less understood. If not cleared the infection then persists and progresses to squamous-cell carcinoma (or adenocarcinoma) of the cervix. This stage is classified as CIN2. Third, CIN3 stage occurs where abnormal cells duplicates replacing the full thickness of the cervical epithelium. Finally, the infection transforms in 'invasion' where the genome of HPV is integrated into the host's genome.

Figure 2: Publications by cervical cancer domains (top-10 assigned qualifiers).



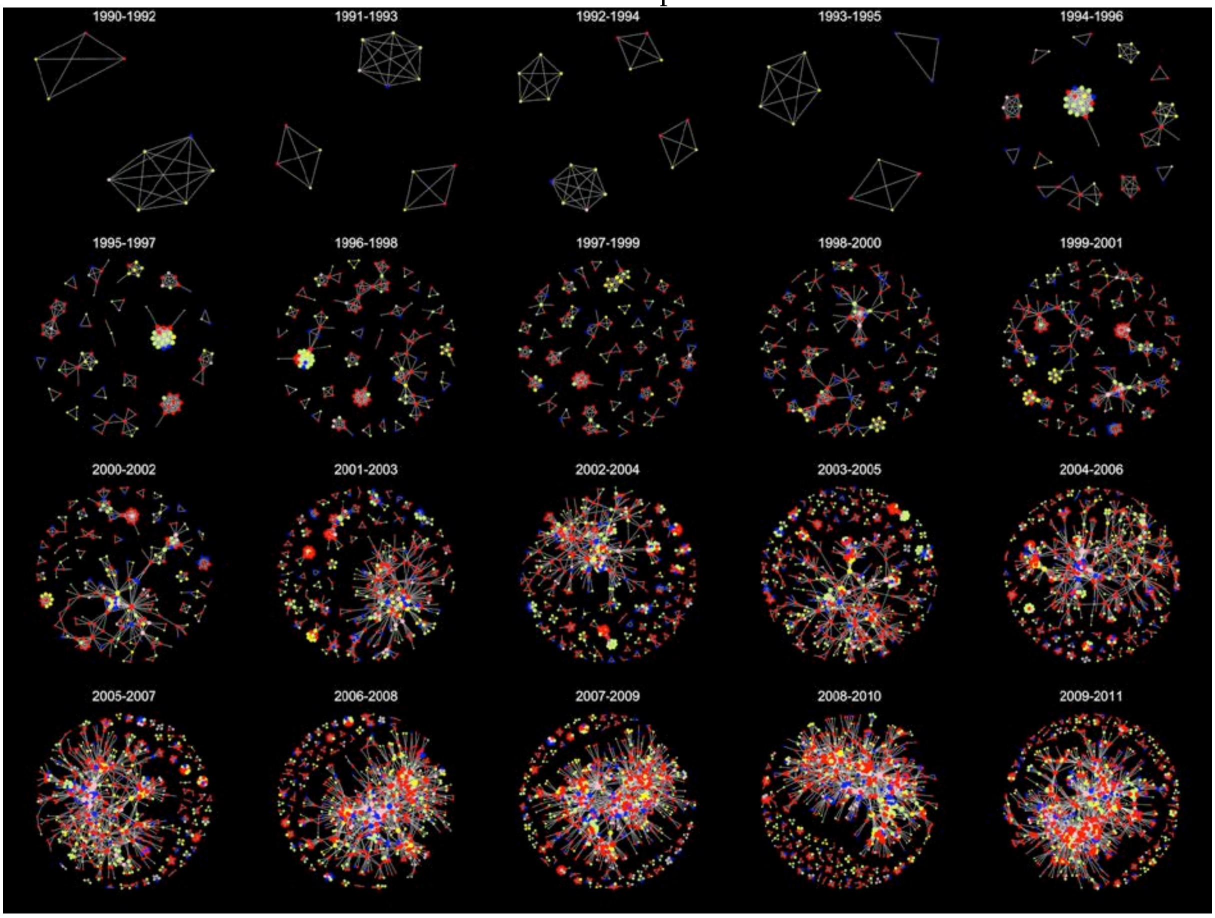
Data and methods

We collected scientific articles related to cervical cancer by querying the MEDLINE/PubMed database and using the Medical Subject Heading (MeSH) classification (Leydesdorff, Rotolo, & Rafols, 2012). We included in our sample all records to which the "Uterine Cervical Neoplasms", "Uterine Cervical Dysplasia", and "Cervical Intraepithelial Neoplasia" MeSH descriptors were assigned as major topics. Furthermore, we considered records published as scientific articles from 1980 to 2011. This returned a sample of 34,285 records.

Figure 2 depicts the number of publications over the observation period divided into major topics, as indexed for MEDLINE/PubMed. The rapid emergence of the diagnostic area provides evidence of the intense activity characterizing this area in comparison to the others. The 'diagnosis' domain has started to rapidly emerge since the early 1990s together with 'virology' and 'pathology' – the discovery of viral origin of the cervical cancer in 1980s boosted the subsequent understanding of the disease and created novel technological opportunities (Hogarth, Hopkins, & Rodriguez, 2012). For the purpose of the paper, we focused on the "diagnosis" domain obtaining a sample of 4,921 publications. The MEDLINE/PubMed dataset was matched with SCOPUS data for more comprehensive bibliographic information. This process matched 95.94% records, i.e. 4,722 scientific articles, representing our final sample.

Conceptually building on Blume's (1992) work and empirically drawing from previous studies on the case (Casper & Clarke, 1998; Hogarth et al., 2012), we identified four phases of the emergence of molecular diagnostic technologies in cervical cancer screening domain: (i) exploration (1980-1989), (ii) development (1990-1999), (iii) adoption (2000-2005), and (iv) growth (2006-2011).[6] We analysed each phase in terms of the bridging roles different institutional groups played. To do so, we first harmonised affiliations' names – 9,806 name variations were harmonised to 3,072 distinct names – and built inter-organisational collaborative networks (co-authorship) by using three-year time window (see Figure 3). We then classified the actors in the above-described five different institutional groups.

Figure 3: Co-authorship networks and institutional groups (components involving at least three nodes are depicted). Colours are assigned as in the followings: RHE=red, GOV=blue, HC=green, IND=yellow, NGO=pink, OT=grey. The size of nodes is proportional to the log2 of actors' number of publications.



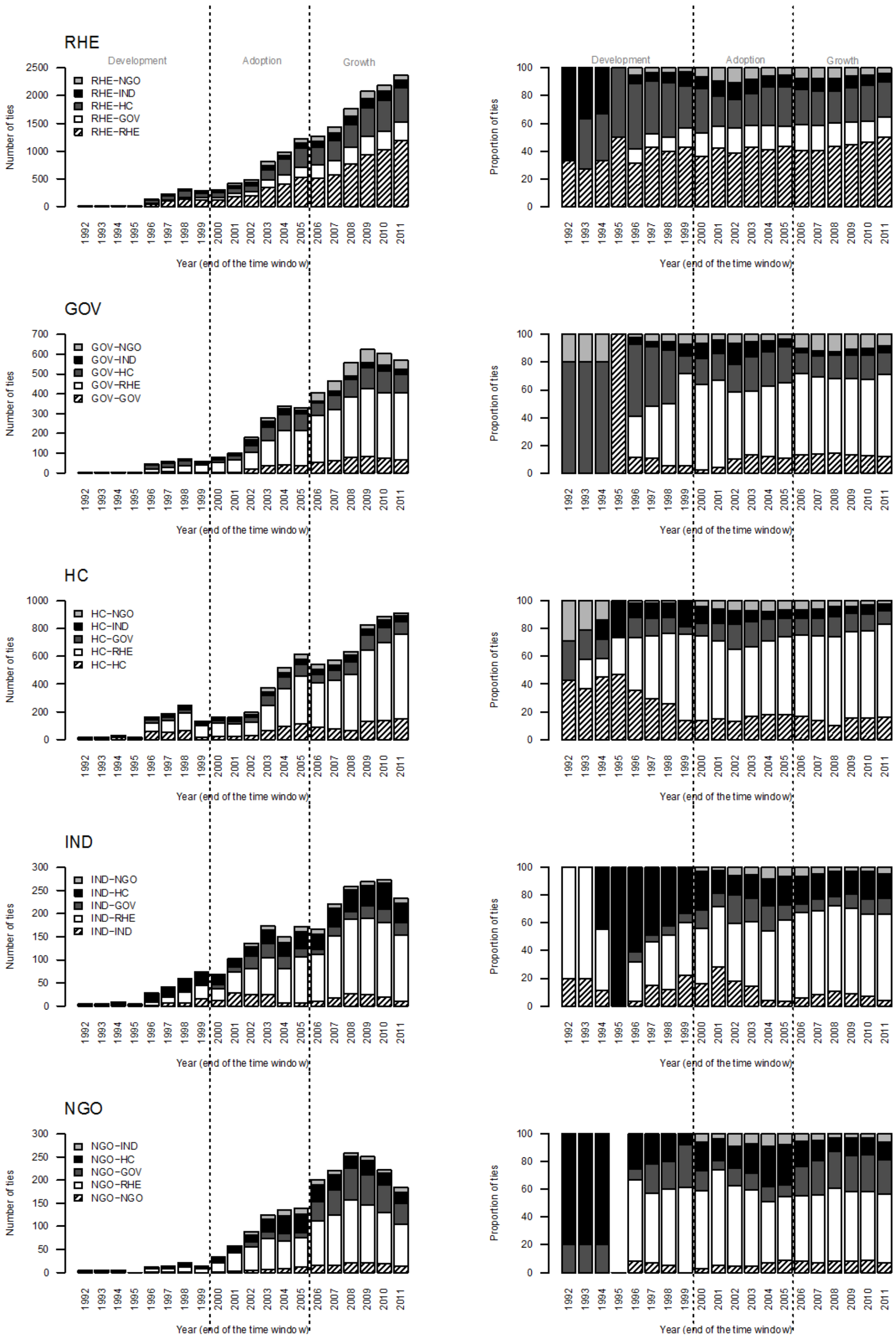
[6] It is worth noting, that the timing of the phases of the emergence we identified specifically fits for the US context. This implies that the timing of these phases may differ when other contexts, such as the European countries, are considered.

Results

Figure 4 examines the presence of patterns in the formation of intra- and inter-group ties – the absolute number of ties is reported on the left while their proportion is depicted on the right. The overall number of ties increases for each group. Yet, the GOV, IND, and NGO organisations' number of ties decreases in the last few years of the growth phase.

RHE is the group that established the highest proportion of intra-group ties (~40%) – the proportion of those ties also remains relative stable in the three phases of emergence. A large proportion of RHE organisations' ties are instead established, especially in the development phase, with the HC group. In this phase, HC organisations may have represented a critical source of samples for the development and testing of novel methodologies for the diagnosis of cervical cancer. The number of RHE-HC ties however reduced over the adoption and growth phases when a more frequent interaction with GOV, IND, and NGO groups is observed.

Figure 4: Intra- and inter-group ties.



GOV organisations' proportion of intra-group ties is relatively low (less than 20%) and stable over the observation period. This group's collaborative activity in the development phase was mainly with HC organisations. While GOV-HC reduced in the subsequent phases, the proportion of ties with the RHE rapidly increased. GOV organisations' collaborative activity with IND actors was more frequent in the adoption phase when the technology may have been in the process of being regulated and it reduced in the growth phase. The GOV-NGO proportion of ties instead increased in the growth phase.

The proportion of intra-group ties for HC actors is higher in the development phase (~35%) than in the adoption and growth phases. A large proportion of HC organisations' inter-group ties are with RHE organisations and this proportion increases over the emergence process. While the proportion of HC-GOV ties is relatively stable over the three phases, HC-IND collaborations are more frequent in the development and adoption phases than in the growth phase. HC-NGO ties are instead more frequently established in the adoption phase.

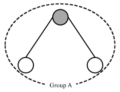
IND actors' intra-groups collaborative activity is relatively low. Yet, the analysis shows that IND-IND collaborations are more frequent in the transition between the development and adoption phases. IND-HC collaborations are dominant in the development phase. This interaction however reduces in the subsequent phases in favour of IND-RHE, IND-GOV, and IND-NGO collaborations.

A great proportion of NGO organisations' ties are established between groups. NGO-RHE ties are the most frequent type of tie, though the proportion of NGO-HC ties is relative high in the development phase. NGO organisations' collaborations with GOV reduce from the development phase to the adoption phase to increase again in the growth phase. NGO-HC ties are frequent especially in the development phase and in the transition between the adoption and growth phases, while NGO-IND collaborations occur more frequently in the adoption phase.

Brokerage roles can be characterised only for those actors that contributed to the diagnosis domain of cervical cancer research with two or more publications in a given time window – an actor that published only once, by definition, cannot act as broker since all the organisations co-authoring the publication (if any) will be directly connected in the co-authorship network. Table 2 reports the descriptive statistics on the evolution of brokerage roles for each institutional group. The inter-organisational network is too fragmented before the 1994-1996 time window, thus no brokerage activity is observed before this period.

The proportion of organisations for which brokerage roles can be evaluated increases over time until the last time three years of observation when actors' publication activity seems to reduce. This suggests that while the number of organisations in each group increases or remains relatively stable the research output seems to be less concentrated, i.e. the higher proportion of organisations contributing to the diagnosis domain with just one publication is higher. The number of organisations occupying at least one brokerage role also increases with peaks in the transition between the adoption and growth phases, when more than 70% of the organisations in each institutional group with the possibility to broker do broker. Yet, brokerage activity reduces for all groups (except for RHE organisations) in the last three years of observation.

Table 2. Evolution of brokerage roles by institutional group.

	Coordinator	Gatekeeper	Itinerant broker	Liaison
				

f organisations occupy brokerage positions.

The decrease in brokerage activity is also observed across institutional groups with most of the types of brokerage roles. However, the probability of an actor to occupy a given brokerage position differs from one institutional group to the other. The comparative chart shows RHE organisations emerging in the network mainly as gatekeepers while their coordination role starts to emerge after 2002, i.e. in the adoption phase. GOV actors are more likely to occupy itinerant broker and liaison roles in the adoption phase while they also emerge as gatekeepers while approaching to the growth phase. HC organisations' probability to occupy a brokerage role is relative similar for the positions of gatekeeper, itinerant broker, and liaison positions.

HC organisations are instead less likely to coordinate within the their group. IND actors' probability to occupy roles of itinerant broker and liaison increases over time with a peak in the transition between the adoption and growth phases and then reduces in the last few years of observation. A similar trend can be observed for the NGO group with the exception that this organisational group is also more likely to occupy gatekeeper positions.

Conclusion

The present paper delved into the key role of inter-organisational networks in shaping the emergence of novel technologies. The analysis revealed that the process of tie formation differs from one group to the other and it evolves across the different phases of emergence. For example, certain actors are more active in establishing intra-group ties (RHE organisations) while others collaborate with organisational actors belonging in other groups more frequently (GOV, IND, and NGOs). Groups also profile according to the different brokerage roles. For example, RHE organisational actors are more likely to coordinate within the group and to act as gatekeeper between the RHE group and another. GOV organisations are more likely to act as itinerant broker and liaison over the emergence, whereas their gatekeeper role increases in the transition between the adoption and growth phase. The HC, IND, and NGO groups are more active in gatekeeper, itinerant broker, and liaison roles than in coordinating within their groups. These findings may have important implications on the designing of policies capable to stimulate and support innovation by leveraging the complex inter-organisational networks, which sustain established technologies and which shape the emergence of new technologies.

References

Arthur, W. B. (1989). Competing technologies, increasing returns, and lock-in by historical events. *The Economic Journal*, *99*(394), 116–131.

Blume, S. (1992). *Insight and Industry: on the Dynamics of Technological Change in Medicine*. Cambridge, MA: MIT Press.

Brass, D. J. (1992). Power in organizations: A social network perspective. *Research in Politics and Society*, *4*(1), 295–323.

Burt, R. S. (1992). *Structural* Holes*: Social Structure of Competition*. London: Harvard Business Press.

Casper, M. J., & Clarke, A. E. (1998). Making the Pap smear into the "right tool" for the job: Cervical cancer screening in the USA, circa 1940-95. *Social Studies of Science*, *28*(2), 255–290.

Clayton, J. (2012). Clinical approval: Trials of an anticancer jab. *Nature*, *488*(7413), S4–S6.

Collingridge, D. (1980). *The social* control *of technology*. Milton Keynes: Open University Press.

David, P. A. (1985). Clio and the economics of QWERTY. *The American Economic Review, 75*(2), 332–337.

Day, G. S., & Schoemaker, P. J. H. (2000). Avoiding the pitfalls of emerging technologies. *California Management Review*, *42*(2), 8–33.

Dosi, G. (1982). Technological paradigms and technological trajectories : A suggested interpretation of the determinants and directions of technical change. *Research Policy*, *11*(3), 147–162.

Fernandez, R. M., & Gould, R. V. (1994). A dilemma of state power: Brokerage and influence in the national health policy domain. *American Journal of Sociology*, *99*(6), 1455–1491.

Freeman, C. (1995). The "National System of Innovation" in historical perspective. *Cambridge Journal of Economics*, *19*(1), 5–24.

Garud, R., & Karnoe, P. (2003). Bricolage versus breakthrough: distributed and embedded agency in technology entrepreneurship. *Research Policy*, *32*(2), 277–300.

Geels, F. (2002). Technological transitions as evolutionary reconfiguration processes: a multi-level perspective and a case-study. *Research Policy*, *31*(8–9), 1257–1274.

Goldstein, J. (1999). Emergence as a construct: History and issues. Emergence, 1(1), 49–72.

Gulati, R., & Gargiulo, M. (1999). Where Do Interorganizational Networks Come From? *American Journal of Sociology*, *104*(5), 1439–1438.

Hogarth, S., Hopkins, M. M., & Rodriguez, V. (2012). A molecular monopoly? HPV testing, the Pap smear and the molecularisation of cervical cancer screening in the USA. *Sociology of Health & Illness*, *34*(2), 234–250.

Leydesdorff, L., Rotolo, D., & Rafols, I. (2012). Bibliometric perspectives on medical innovation using the medical subject Headings of PubMed. *Journal of the American Society for Information Science and Technology*, *63*(11), 2239–2253.

Malerba, F. (2002). Sectoral systems of innovation and production. *Research Policy*, *31*(2), 247–264.

Martin, B. R. (2012). The evolution of science policy and innovation studies. *Research Policy*, *41*(7), 1219–1239.

Powell, W. W. (1990). The Transformation of Organizational Forms: How Useful is Organisation Thoery in Accounting for Social Change. In R. O. Friedland & A. F. Robertson (Eds.), *Beyond the Marketplace: Rethinking Economy and Society*. New York, NY, USA: Walter de Gruyter.

Reynolds, L. A., & Tansey, E. M. (2009). *History of cervical cancer and the role of human papillomavirus, 1960–2000* (Vol. 38). London: The Wellcome Trust Centre for the History of Medicine at UCL.

Rigoni-Stern, D. (1842). Fatti statistici relativi alle malattie cancerose che servirono di base alle poche cose dette dal dott. Giornale per servire ai progressi della patologia e della terapeutica, *2*, 507–517.

Schiffman, M., Castle, P. E., Jeronimo, J., Rodriguez, A. C., & Wacholder, S. (2007). Human papillomavirus and cervical cancer. *The Lancet*, *370*(9590), 890–907.

Schumpeter, J. A. (1934). *The Theory of Economic Development. An Inquiry into Profits, Capital, Credit, Interest and the Business Cycle*. Cambridge, MA: Harvard Univ. Press.

Stirling, A., & Scoones, I. (2009). From risk assessment to knowledge mapping: Science, precaution, and participation in disease ecology. *Ecology and Society*, *14*(2). Retrieved from http://www.ecologyandsociety.org/vol14/iss2/art14/

Van Lente, H., & Rip, A. (1998). Expectations *in Technological Developments: an Example of Prospective Structures to be Filled in by Agency*. Walter de Gruyter. Retrieved from http://doc.utwente.nl/34732/

Willer, R., & Willer, D. (2000). Exploring dynamic networks: hypotheses and conjectures. *Social Networks*, *22*(3), 251–272.

Zur Hausen, H. (1987). Papillomaviruses in human cancer. *Applied Pathology, 5*, 19–24.

Skewness for journal citation curves[1]

Ronald Rousseau*

*ronald.rousseau@uantwerpen.be
Institute for Education and Information Sciences, IBW, University of Antwerp (UA), Antwerp, B-2000, Belgium
& KU Leuven, Dept. Mathematics, Leuven B-3000, Belgium

Introduction

It is well-known that informetric data are highly skewed (Seglen, 1992). Most of the time one assumes that citation data follow (more or less) a power law (Egghe, 2005; Seglen, 1992) and for this reason exhibit a high degree of skewness. Yet, less is known about the degree of skewness (Seglen did not provide skewness values) and its possible relation with impact and other indicators. Concretely, do journals with a high impact, usually having many highly-cited publications, have a higher skewness than other ones? In this poster we try to find a (preliminary) answer to this question.

Methods

In order to obtain a reasonable spread of data we applied the following procedure. We consider all SCI journals in the JCR 2012, ranked decreasingly according to their 5-year impact factor. Now we take the natural logarithm of these 5-year data. An IF of 1 corresponds to the number zero and an IF of 54 corresponds to the number 4 (approximately). Now we consider the values 4, 3.8, 3.6 and so on ending with 0.2 and 0: 21 numbers in total. The journals included in the investigation are those for which the natural logarithm of their 5-year impact factor in 2012 is closest to these 21 numbers and which, moreover, have published at least 100 articles (publications of *article* type) in the year 2009. We collected citation data over the period 2009 – 2014 (on 27 February 2014) from the Web of Science (WoS). For each journal we determined the 5-year JIF (JCR, 2012), the diachronous impact factor, DIF5, (= mean number of citations) for the year 2009 (see (Ingwersen et al., 2001) for more information about the synchronous and the diachronous impact factor), the median number of citations (again for publications in the year 2009), the total number of *articles* in 2009 and the skewness of the citation curve of *articles* published in the year 2009. Skewness, denoted as Sk, was calculated using the following formula for skewness (m_2 and m_3 denote the second and third moment about the mean; n is the number of data):

$$Sk = \frac{\sqrt{n(n-1)}}{n-2} * \frac{m_3}{(m_2)^{\frac{3}{2}}} = \frac{\sqrt{n(n-1)}}{n-2} * \frac{\frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^3}{\left(\frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2\right)^{3/2}}$$

The factor $\frac{\sqrt{n(n-1)}}{n-2}$ is used to reduce bias when skewness is calculated from a sample. Then we calculated the Pearson correlation between these data, i.e. between JIF5, DIF5, median, Sk and number of published articles.

[1] This work was partly supported by the National Natural Science Foundation of China (NSFC Grant No 71173185).

Results

Table 1. Data: JIF5 from the JCR, other values based on WoS data.

	JIF5 2012	DIF5 2009	median	Sk	# articles
NEW ENGL J MED	50.807	207	107	2.69	319
NATURE	38.159	150.21	96	5.38	800
NAT GENET	34.520	157.01	122	2.69	209
JAMA-J AM MED ASSOC	29.273	103.22	71	2.45	211
NAT METHODS	23.231	76.479	51	2.16	117
J CLIN ONCOL	17.255	62.726	41	4.04	773
NAT NEUROSCI	16.412	68.202	52	5.79	203
ANGEW CHEM INT EDIT	13.560	38.526	28	4.43	1542
AM J RESP CRIT CARE	10.919	42.655	30.5	1.93	264
CLIN INFECT DIS	8.980	32.462	23	9.91	457
PHILOS T R SOC B	7.298	27.01	21	1.84	175
APPL CATAL B-ENVIRON	6.031	23.05	18	3.73	443
ELECTROCHEM COMMUN	4.950	21.531	15	5.88	588
CELL SIGNAL	4.060	12.59	11	1.26	180
J MOL CATAL A-CHEM	3.319	13.005	10	2.96	368
J ASIAN EARTH SCI	2.714	11.6	7	3.37	137
AM MINERAL	2.230	9.093	7	1.84	193
BIOMED CHROMATOGR	1.815	6.816	5	1.90	163
EUR ARCH OTO-RHINO-L	1.489	4.89	3	1.34	245
ARCH GYNECOL OBSTET	1.216	4.056	3	4.48	358
BUNDESGESUNDHEITSBLA	1.005	2.237	1	9.63	152

Answering one of the research questions we see that skewness takes values between 1 and 10. Recall that the skewness of the normal distribution is zero and is two for an exponential distribution. Hence all these citation curves are positively skewed, and even often highly skewed. In order to provide some intuitive idea about the observed skewness values we compare them with those of the familiar Poisson distribution. As the skewness of a Poisson distribution with parameter μ (= mean μ) is $1/\sqrt{\mu}$, a skewness of 1 corresponds to a Poisson distribution with parameter 1, while a skewness of 10 would correspond to a Poisson distribution with parameter 1/100.

Pearson correlation coefficients between the data shown in Table 1 are shown in Table 2.

Table 2. Pearson correlation coefficients.

Pearson	JIF5	DIF5	median	Sk	# articles
JIF5	1.00	0.99	0.96	-0.08	0.14
DIF5		1.00	0.98	-0.08	0.09
median			1.00	-0.08	0.09
skew				1.00	0.25
# articles					1.00

The main observations derived from Table 2 are:
1) Skewness is totally uncorrelated to the impact factors and to the median number of citations. This is illustrated in Fig.1. This figure moreover illustrates the fact that data points (JIF5 values) were evenly chosen on a logarithmic scale.
2) Skewness is weakly correlated to the number of articles.
3) The median is always smaller than the mean (=DIF5) as expected for positively skewed data.
4) The 5-year diachronous impact factor for the year 2009 (using only publications of article type) is almost perfectly, positively, correlated with the 5-year synchronous impact factor (using ‘citable’ items) for the year 2012 (see Figure 2).

Figure 1: Skewness (vertical axis) as a function of the 5-year diachronous impact factor (horizontal axis, log scale).

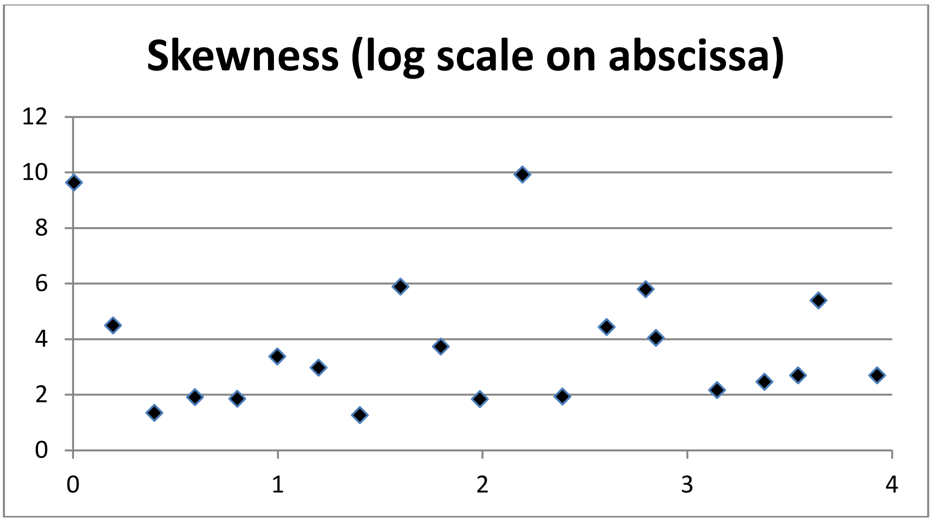
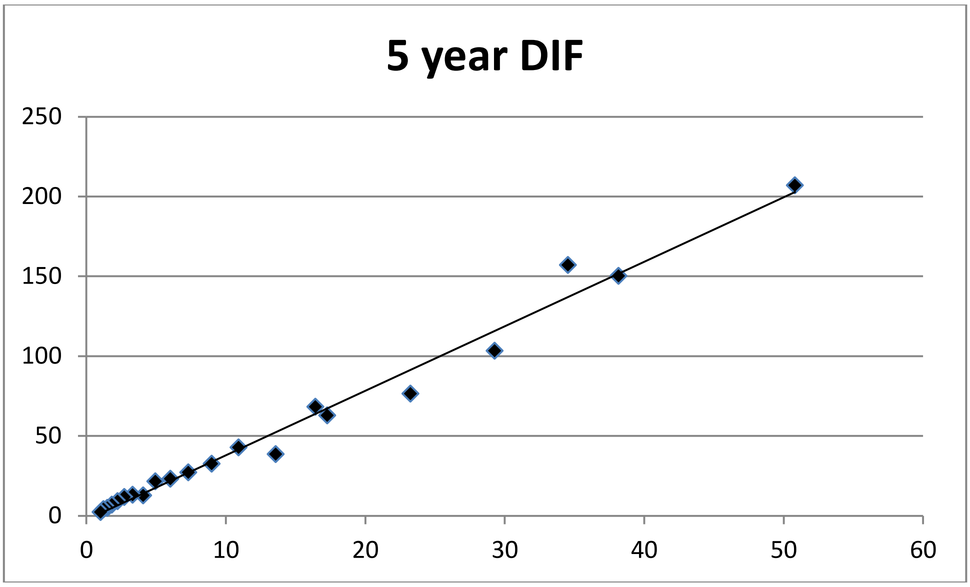


Figure 2: 5 year diachronous impact factor (vertical axis) for the year 2009 as a function of the 5-year synchronous impact factor (horizontal axis) for the year 2012.



Conclusion

We confirmed that journal citation curves are indeed highly, positively skewed and obtained values between 1 and 10. Skewness is totally uncorrelated to the 5-year impact. Finally we illustrated the fact that a synchronous impact factor (JIF5) based on one citation year (2012) and five publication years (2007-2011) is highly correlated (> + 0.99) with the diachronous impact factor (DIF5) based on one publication year (2009) and 5 citation years (2009-2013+, where the + indicates that we included the months January and February of the year 2014), where the publication year is taken in the middle of the five publications year used for the calculation of the JIF5.

References

Egghe, L. (2005). Power laws in the information production process: Lotkaian informetrics. Amsterdam: Elsevier.

Ingwersen, P., Larsen, B., Rousseau, R., & Russell, J. (2001). The publication-citation matrix and its derived quantities. *Chinese Science Bulletin*, 46(6), 524-528.

Seglen, P. (1992). The skewness of science. *Journal of the American Society for Information Science*, 43(9), 628–638.

The skewness of scientific productivity

Javier Ruiz-Castillo* and Rodrigo Costas**

* *jrc@eco.uc3m.es*
Departamento de Economía, Universidad Carlos III, Madrid 126, Getafe 28903 (Spain)

** *rcostas@cwts.leidenuniv.nl*
Center for Science and Technology Studies (CWTS), Leiden University, P.O. Box 905, Leiden 2300 AX (The Netherlands)

Introduction

The productivity of individual scientists has been studied extensively since Lotka's (1926) pioneer contribution (Alvarado, 2012, counts 651 publications from that date until 2010). However, most of these contributions analyze a relatively small number of scientists and, to the best of our knowledge, do not systematically study this phenomenon using comparable data for several scientific disciplines. As in any other scientific discipline, in Scientometrics we should clearly establish the stylized facts that characterize basic concepts such as productivity distributions in all fields. Consequently, this paper studies the productivity of individual scientists in 30 broad fields using a large dataset, indexed by Thomson Reuters, consisting of 7.7 million distinct articles published in the period 2003-2011 in academic journals. Regardless of how we measure individual productivity, a study of this type must solve the following three methodological problems.

Firstly, given the well-known differences in publication and citation practices across scientific disciplines, the performance of any pair of authors can only be compared if they belong to the same field. The problem, of course, is that Thomson Reuters often assigns publications in the periodical literature to several Web of Science subject categories. To tackle this problem, in this paper we follow a multiplicative strategy where each article is counted as many times as necessary in the several categories to which it is assigned. The number of articles in the corresponding extended count is 10.4 million, or 35% larger than the number of distinct articles. On the other hand, among the many alternatives, we classify all articles into 30 of the 35 broad fields distinguished in Tijssen *et al.* (2010).

Secondly, the assignment of articles to individual authors is known to be plagued with formidable difficulties. In this paper, we solve this problem using the algorithm recently developed by Caron & van Eck (2014), according to which there are approximately 9.6 million distinct researchers associated to the 7.7 million distinct articles of the dataset. In the extended count, this number is raised to 17.2 million individuals, a 79% increase.

Thirdly, a fundamental difficulty is how to confront the phenomenon of multiple authorship present in different degrees in all fields. In this paper, we follow Nicholls's (1989) recommendation of using what is known as the *complete* or *whole count*, namely, a multiplicative strategy in which any article co-authored by two or more persons is wholly assigned as many times as necessary to each of them. As a consequence, the total number of articles becomes 48.2 million, or 2.8 times larger than the 17.2 million authors in the extended count.

Having information on articles, authors, and citations, we measure individual productivity in two ways: number of publications in the period 2003-2011 per person, and mean citations per article per person in that period. Our main concern in this paper is to investigate whether the characteristic skewness of both types of productivity distributions is similar across scientific fields with vastly different size, average productivity, productivity inequality, and average number of authors per publication.[1] We analyze the skewness in each field for two samples: the entire population, and what we call *productive authors*, namely, the subset of scientists whose productivity is above their field average.

Productivity as the number of articles per person

In this Section, we define individual productivity as the number of distinct articles written by each individual independently of the number of authors involved.

The Skewness of Productivity Distributions

We assess the skewness of productivity distributions by means of the Characteristic Scales and Scores (CSS hereafter), a size- and scale-independent technique, first used in Scientometrics in Schubert *et al.* (1987).[2] In our case, the following two *characteristic scores* are determined for every field: μ_1 = mean number of publications for the entire productivity distribution, and μ_2 = mean number of publications for authors with a number of articles above μ_1. Consider the partition of the distribution into three broad classes: (i) authors with low productivity that publish a number of articles smaller than or equal to μ_1; (ii) fairly productive authors, with productivity greater than μ_1 and smaller than or equal to μ_2; (iii) authors with remarkable or outstanding productivity greater than μ_2. The average (the standard deviation), and the coefficient of variation over the 30 fields of the percentage of authors in the three classes, as well as the corresponding percentages of the total number of citations accounted for by each class appear in row I in Table 1.

[1] To save space, the distribution of fields by size, average productivity, productivity inequality measured by the coefficient of variation, and average number of authors per publication can be found in the Working Paper version of this paper, Ruiz-Castillo & Costas (2014), denoted in the sequel by RCC.

[2] In RCC, we also summarize the skewness of productivity distributions by means of skewness indices that are robust to extreme observations. For reasons of space, such indices are excluded from this paper.

Table 1. The skewness of two types of productivity distributions according to the CSS approach. Average, standard deviation, and coefficient of variation over 30 fields of the percentages of individuals, and the percentages of articles (or citations) by category

	Individual productivity = number of articles per person					
	Percentage of people in category			**Percentage of articles in category**		
I. Total population	**1**	**2**	**3**	**1**	**2**	**3**
Average (Std. dev.) (6.3)	79.3 (3.4)	14.8 (2.4)	5.9 (1.2)	40.4 (7.0)	24.5 (1.8)	35.1
Coeff. of variation	0.04	0.17	0.19	0.17	0.07	0.18
II. Successful authors with above average productivity						
Average (Std. dev.) (3.5)	71.4 (2.4)	19.8 (1.7)	8.8 (1.1)	41.4 (7.0)	27.4 (1.5)	31.1
Coeff. of variation	0.03	0.09	0.12	0.10	0.06	0.11

	Individual productivity = mean citation per article per person					
	Percentage of people in category			**Percentage of total mean citations in category**		
I. Total population	**1**	**2**	**3**	**1**	**2**	**3**
Average (Std. dev.) (4.6)	71.0 (2.1)	20.7 (1.2)	8.3 (1.1)	22.6 (3.1)	40.2 (3.7)	37.2 (4.6)
Coeff. of variation	0.03	0.06	0.13	0.14	0.09	0.12
II. Successful authors with above average productivity						
Average (Std. dev.) (3.7)	71.0 (2.2)	20.3 (1.0)	8.3 (1.2)	52.0 (5.0)	27.7 (1.8)	20.3 (3.7)
Coeff. of variation	0.03	0.06	0.13	0.10	0.06	0.18

The key result is that the relatively small standard deviations and coefficient of variations in row I indicate that field productivity distributions tend to share some fundamental characteristics. Specifically, we find that, on average, 79.3% of all individuals have productivity below μ_1 and account for approximately 40% of all publications, while individuals with a remarkable or outstanding productivity represent only 5.9% of the total and account for 35% of all publications.

Productive Authors

As examined in detail in RCC, taking into account that we study the publication performance over a period of nine years, field mean productivity values are generally low. The reason is that, on average, 68% of authors in all fields have only contributed a single article during this period. Consequently, it seems relevant to study the behavior of what we call *productive authors* with above average productivity. The results of the CSS approach are in row II in Table 1. Firstly, on average, the percentage of people in category 1 is eight points smaller than before, while the percentage of successful people in categories 2 and 3 is five and three points greater than for the population as a whole. Secondly, the percentage of publications accounted for by all categories remains essentially constant. Thirdly, all standard deviations and coefficients of variation are smaller in row II than in row I, indicating that productivity distributions are now even more similar than before. Figures 1.A and 1.B illustrate the similarity of the partition of authors into the three classes for the population as a whole and the subset of productive authors.

Productivity as the mean citation per article per person

Characteristics of Productivity Distributions

Measuring productivity as the number of publications per author in a certain period has a long history in Scientometrics. However, in the dataset used in this paper it is possible to take into account each author's citation impact. Therefore, in this Section we define individual productivity as the mean citation per article per person during 2003-2011.[3] The correlation coefficient between the two measures in the entire sample is 0.02, indicating that, as we know from previous research (Costas *et al.*, 2010), the most prolific authors need not be those with the highest impact. Thus, the two concepts, although related, are best treated separately.

For the application of the CSS approach, let m_1 be mean productivity for the population as a whole, let m_2 be the mean productivity of authors with productivity above m_1, and consider again the partition of the distribution into three broad classes. The results are in row III in Table 1. Judging from the partition of authors in the three classes, the type of skewness in rows II and III in Table 1 is essentially the same. The main difference is in the way publications and citations are accounted for by the three categories. The explanation lies in the fact that category 2 includes authors with a relatively large number of publications *per capita*. Given the high correlation between number of publications and citations received per publication, which is 0.67 on average for all fields, we find that category 2 accounts for a large percentage of the sum of the values of the variable mean citation per article over all authors in the field (abbreviated as total mean citations in Table 1).

Productive Authors

Just as before, it is interesting to study productive authors with above average productivity. The results of the CSS approach are in row IV in Table 1. The comparison with the population as a whole yields a first fundamental result: on average, the partition of both populations over the three CSS categories is exactly the same. Furthermore, judging from the coefficients of variation, the similarity across fields is again the same as before. This illustrates the fractal-type nature of individual productivity distributions when productivity is measured as the mean citation per article per person. Figures 2.A and 2.B provides a graphical illustration of the situation.

[3] Admittedly, although this productivity measure is a standard summary measure of individual citation distributions, it is not an ideal indicator of individual citation impact.

Figure 1A: Partition of productivity distributions into three categories according to the CSS technique. Productivity = number of articles per person. Population as a whole

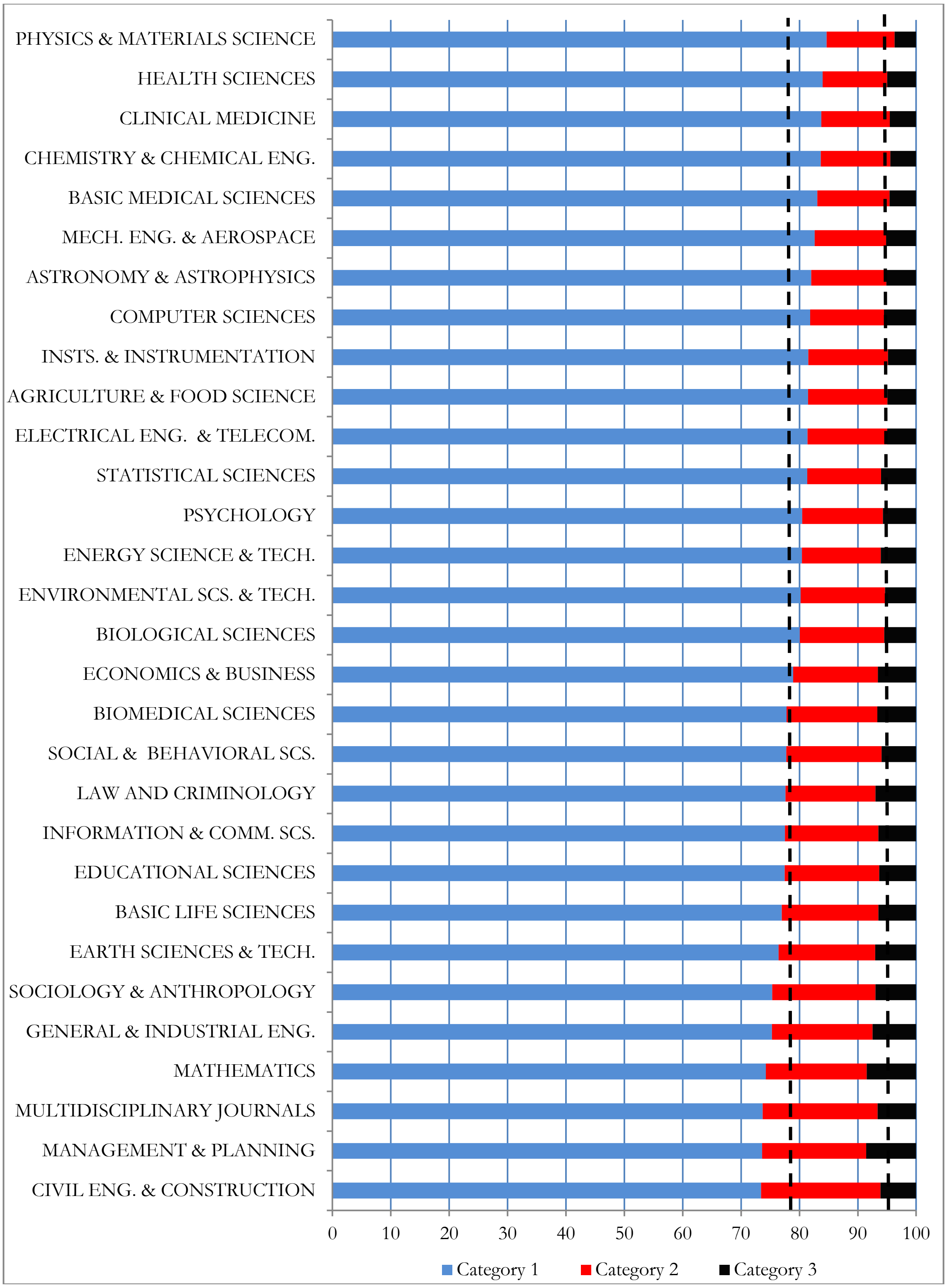


Figure 1B: Partition of productivity distributions into three categories according to the CSS technique. Productivity = number of articles per person. Successful authors with above average productivity

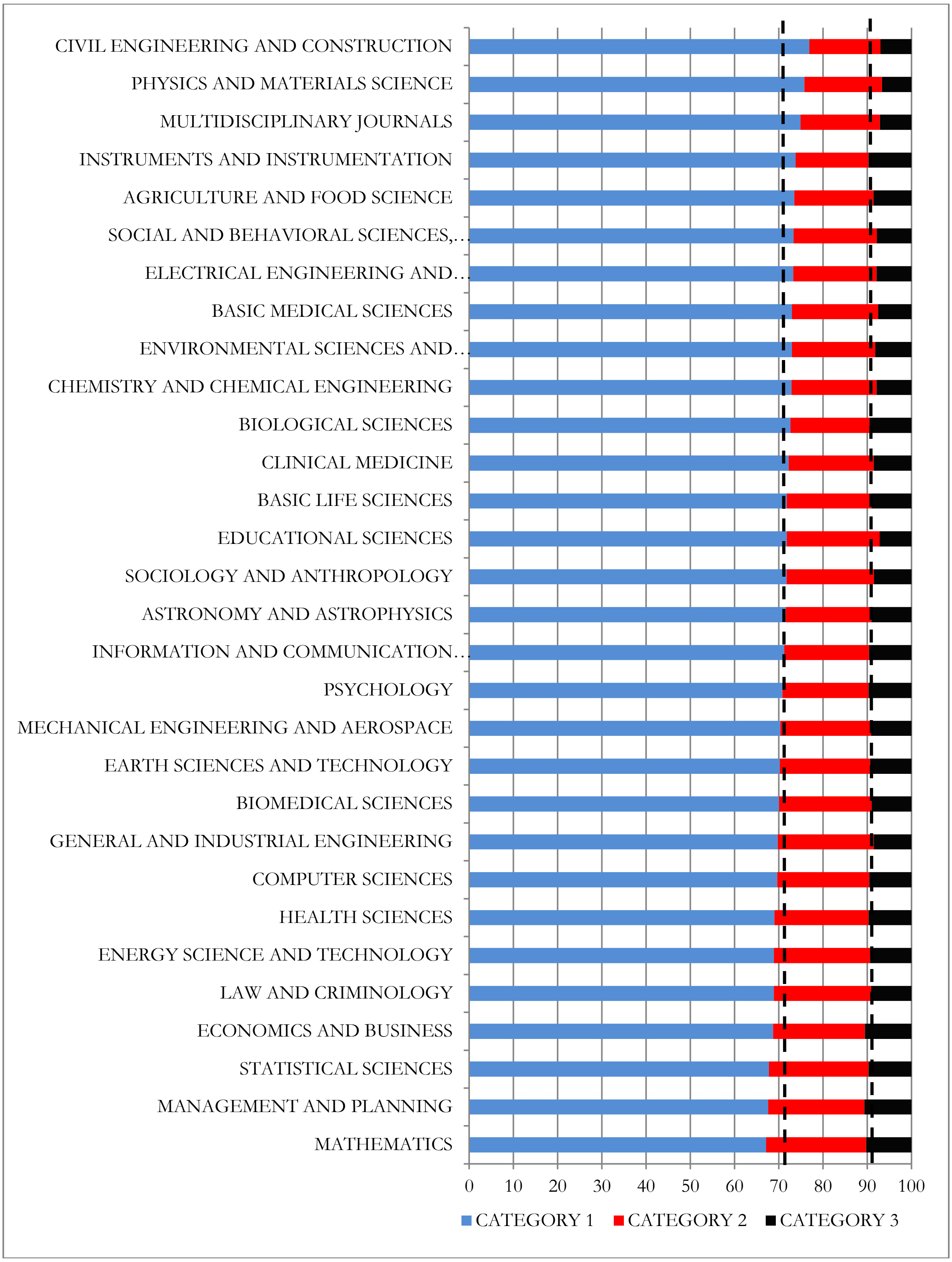


Figure 2A: Partition of productivity distributions into three categories according to the CSS technique. Productivity = mean citation per article per person. Population as a whole

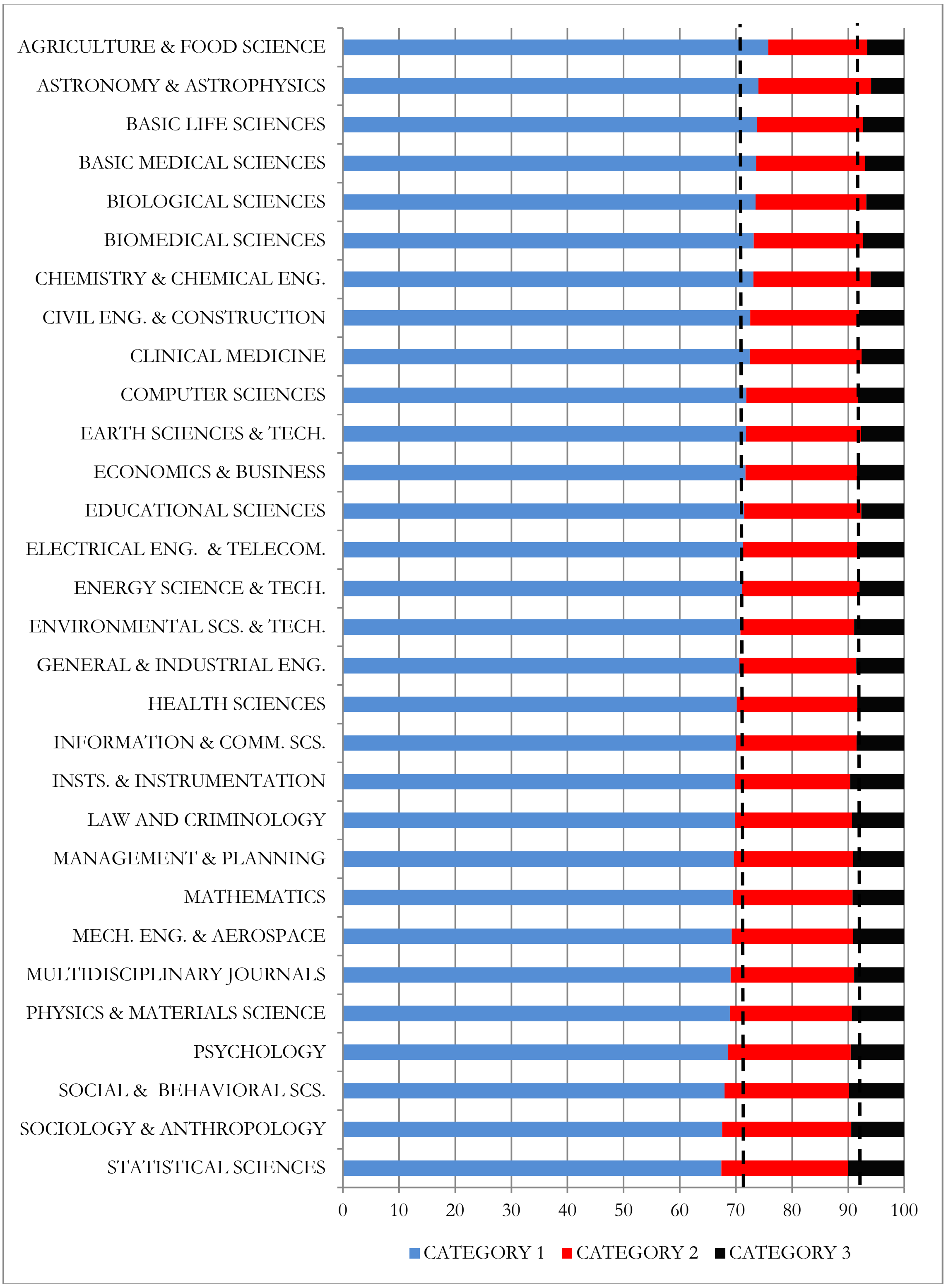
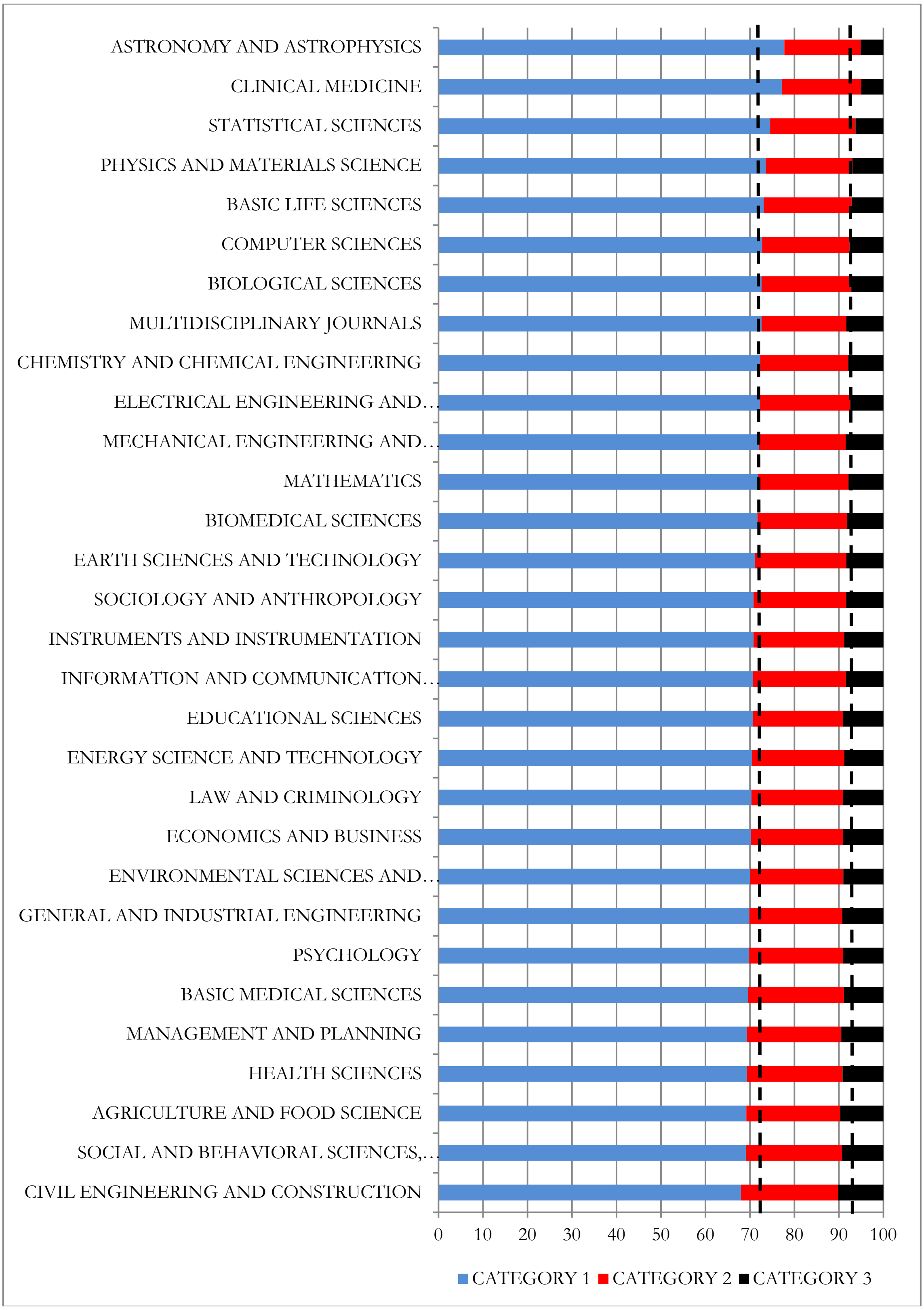


Figure 2B: Partition of productivity distributions into three categories according to the CSS technique. Productivity = mean citation per article per person. Successful authors with above average productivity



Conclusions

The main results of the paper can be summarized as follows. We have confirmed that the well-known, wide differences in production and citation practices across fields generate large differences in mean productivity across fields. However, size- and scale-independent techniques have allowed us to establish that productivity distributions are highly unequal and very similar across fields. Citation distributions exhibit a comparable skewness and similarity across scientific disciplines at different aggregation levels (Albarrán & Ruiz-Castillo, 2011, Albarrán *et al.*, 2011, and Li *et al.*, 2013). Consequently, rather than a set of models for different fields of science, we need a single explanation of within-field variation of scientists' productivity as manifested in the productivity inequality and skewness dimensions documented in this paper.

On the other hand, previous results have established that the similarity of citation distributions across scientific fields opens the way to the comparability of the citations received by articles belonging to heterogeneous disciplines (Crespo *et al.*, 2013a, b, and Li *et al.*, 2013). To explore the comparability of levels of productivity across fields in our case, we have normalized field productivity distributions by computing the ratio between mean productivities in every field and mean productivity in Chemistry & Chemical Engineering, taken as the reference field in Table 2. When productivity is measured as the number of articles per person, the similarity between columns 2 and 3 in Table 2 indicates that this normalization strategy is very promising. When productivity is measured as the mean citation per article per person, the similarity between columns 4, 5, and 6 verifies the above intuition. However, rigorously studying this normalization strategy must be left for further research.

Table 2. Comparison of individual productivities across fields taking Chemistry & Chemical Engineering as the reference field for the two productivity measures

	Number of articles per person			Mean citation per article per person		
	μ_1 (1)	μ_2 (2)	μ_3 (3)	m_1 (4)	m_2 (5)	m_3 (6)
ASTRONOMY & ASTROPHYSICS	27.2	30.6	30.6	11.1	11.7	14.3
BASIC LIFE SCIENCES	8.8	6.4	5.9	15.6	15.6	16.2
BIOLOGICAL SCIENCES	7.8	6.1	5.5	9.6	9.8	9.9
BIOMEDICAL SCIENCES	8.8	6.7	6.0	12.3	11.6	11.3
CHEMISTRY & CHEMICAL ENG.	10.0	10.0	10.0	10.0	10.0	10.0
CIVIL ENG. & CONSTRUCTION	6.1	3.5	3.5	5.8	5.8	5.1
CLINICAL MEDICINE	10.3	10.7	10.6	13.9	15.2	18.4
COMPUTER SCIENCES	7.3	5.7	4.9	3.9	4.5	4.8
EARTH SCIENCES & TECH.	9.7	7.3	6.7	7.7	7.3	6.8
ECONOMICS & BUSINESS	7.5	5.4	4.3	6.6	6.9	6.6
ELECTRICAL ENG. & TELECOM.	7.7	6.3	6.0	4.6	5.0	5.1
ENVIRONMENTAL SCS. & TECH.	7.8	6.1	5.6	8.6	8.1	7.4
GENERAL & INDUSTRIAL ENG.	5.6	3.2	2.7	4.7	4.6	4.2
INFORMATION & COMM. SCS.	5.3	3.1	2.6	6.0	6.2	6.1
MATHEMATICS	9.8	6.8	5.7	3.4	3.9	4.0
MECHANICAL ENG. & AEROSPACE	7.1	5.7	4.9	4.8	4.8	4.7
MULTIDISCIPLINARY JOURNALS	5.3	2.8	2.3	51.4	64.9	71.2
PHYSICS & MATERIALS SCIENCE	14.3	17.3	20.4	7.6	8.6	9.3
SOCIAL & BEHAVIORAL SCS.	5.1	2.8	2.3	7.7	7.5	6.8
SOCIOLOGY & ANTHROPOLOGY	5.3	2.9	2.4	6.1	5.6	5.2

Finally, we have investigated in RCC the robustness of our results to an adjusted or fractional approach to the treatment of articles co-authored by two or more persons. The conclusion is that the skewness of productivity distributions in each field, and the similarity of productivity distributions across fields when using the complete or the adjusted approach are essentially indistinguishable. Moreover, when productivity is measured as the mean citation per article per person, the comparability of mean productivities across fields in the multiplicative approach remains essentially unchanged. To save space, the discussion of possible extensions of this paper can be found in RCC.

References

Albarrán, P., & Ruiz-Castillo, J. (2011), "References Made and Citations Received By Scientific Articles", *Journal of the American Society for Information Science and Technology*, 62: 40-49.

Albarrán, P., Crespo, J., Ortuño, I., & Ruiz-Castillo, J. (2011), "The skewness of science in 219 sub-fields and a number of aggregates", *Scientometrics, 88*, 385–397.

Alvarado, R. (2012), "La colaboración de los autores en la literature producida sobre la Ley de Lotka", *Ciência da Informação*, 40: 266-279.

Caron & van Eck (2014), "Large scale author name disambiguation using rule-based scoring and clustering", CWTS Working Papers.

Costas, R., Leeuwen, T. N. Van, & Bordons, M. (2010), "A Bibliometric Classificatory Approach for the Study and Assessment of Research Performance at the Individual Level : The Effects of Age on Productivity and Impact", *Journal of the American Society for Information Science and Technology*, 61: 1564–1581.

Crespo, J. A., Li, Yunrong, and Ruiz-Castillo, J. (2013a), "The Measurement of the Effect On Citation Inequality of Differences In Citation Practices Across Scientific Fields", *PLoS ONE* 8(3): e58727.

Crespo, J. A., Herranz, N., Li, Yunrong, and Ruiz-Castillo, J. (2013b), "The Effect on Citation Inequality of Differences in Citation Practices at the Web of Science Subject category Level", *Journal of the American Society for Information Science and Technology*, 65: 1244-1256.

Li, Y., Radicchi, F., Castellano, C., and Ruiz-Castillo, J. (2013), "Quantitative Evaluation of Alternative Field Normalization Procedures", *Journal of Informetrics*, 7: 746–755.

Lotka, A. J. (1926), "The Frequency Distribution of Scientific Productivity", *Journal of the Washington Academy of Science*, 16: 317-323.

Nichols, P. T. (1989), "Bibliometric Modeling Processes and the Empirical validity of Lotka'a Law: The Case of Adjusted Counts for Multiauthorship Attribution", *Journal of the American Society for Information Science*, 43: 645-647.

Ruiz-Castillo & Costas (2014), "The skewness of scientific productivity", Working Paper 14-02, Universidad Carlos III (http://hdl.handle.net/10016/18286).

Schubert, A., W. Glänzel and T. Braun (1987), "Subject Field Characteristic Citation Scores and Scales for Assessing Research Performance", *Scientometrics*, 12: 267-292.

Tijssen, R.; Hollanders, H.; van Steen, J. (2010). *Wetenschaps en Technologie Indicatoren 2010*. Nederlands Observatorium Wetenschap en Technologie (NOWT).

Field-normalized citation impact indicators using algorithmically constructed classification systems of science

Javier Ruiz-Castillo* and Ludo Waltman**

* *jrc@eco.uc3m.es*
Departamento de Economía, Universidad Carlos III of Madrid, Madrid 126, 28903 Getafe (Spain)

** *waltmanlr@cwts.leidenuniv.nl*
Centre for Science and Technology Studies (CWTS), Leiden University, P.O. Box 905, 2300 AX Leiden (The Netherlands)

Introduction

We study the problem of normalizing citation impact indicators based on a classification system of science. In practice, the choice of the Thomson Reuters' Web of Science (WoS hereafter) classification system is often made because this system is readily available in the WoS database. However, it is unclear to what extent the WoS system provides a good basis for normalizing citation impact indicators. In fact, Van Eck, Waltman, Van Raan, Klautz, and Peul (2013) recently established the existence of heterogeneous sub-groups (corresponding to clinical and basic medical research) with different citation practices within WoS fields.

In this paper, we search for alternatives to the WoS classification system. We focus on the methodology introduced by Waltman and Van Eck (2012) for the algorithmic construction of publication-level classification systems of science (the WVE methodology hereafter). This methodology is able to handle very large datasets, and it uses a transparent clustering technique that classifies publications into clusters (i.e., fields) solely based on their direct citation relations. We consider twelve *granularity levels*, or twelve classification systems at different aggregation levels. By fixing the *resolution parameter* –that essentially determines the number of clusters at each granularity level–, we build a sequence of twelve classification systems in which the same set of publications is assigned to an increasing number of clusters.

We apply the WVE methodology to a WoS dataset consisting of 3.6 million articles published in the period 2005–2008 in academic journals, excluding trade journals, national journals, etc. For each article, we count the number of citations received within a five-year citation window. The number of clusters in the WVE sequence ranges from 390 to 73,205 in granularity levels 1 to 12. This contrasts with the 236 clusters in the WoS classification system. For further details, we refer to the working paper version of this contribution (Ruiz-Castillo & Waltman, 2014; RCW hereafter).

Which granularity level is used in practice in the calculation of normalized citation impact indicators is an important issue. As clearly argued by Zitt, Ramana-Rahari, and Bassecoulard (2005, p. 391), "*An article may exhibit very different citation scores or rankings when compared within a narrow specialty or a large academic discipline*" (see also Adams, Gurney, & Jackson, 2008). If we choose a granularity level dominated by a relatively small number of broad clusters, the danger is that the clusters are too heterogeneous in terms of citation practices. However, if we go in the opposite direction and choose a granularity level including too many clusters, some clusters (i) may mostly include the publications of a small set of closely connected authors citing each other, (ii) may be so small as to jeopardize their statistical properties, and (iii) may have artificially small mean numbers of citations, so that standard normalization procedures that use the mean number of citations as a normalization factor tend to overestimate the impact of the publications in these clusters.

As a consequence of the above issues, the evaluation of research units based on field-normalized citation impact indicators is likely to be dependent on the granularity level at which the evaluation takes place. As Zitt et al. (2005, p. 392) conclude, "*The fact that citation indicators are not stable from a cross-scale perspective is a serious worry for bibliometric benchmarking. What can appear technically as a 'lack of robustness' raises deeper questions about the legitimacy of particular scales of observation*".

In this paper, we investigate two questions. Firstly, what are the main characteristics of the twelve WVE classification systems, and how do these systems compare with the WoS alternative? Secondly, what are the consequences of using either the WoS classification system or an appropriately selected WVE classification system in the evaluation of research units' citation impact?

Characteristics of the Different Classification Systems

We study the characteristic of the WVE classification systems from three perspectives: (1) Cluster size distribution, (2) skewness of science, and (3) cluster homogeneity. Because of space limitations, only some selected results will be reported below. We refer to RCW for the full results, as well as for more details on the way in which the classification systems have been constructed.

Cluster Size Distribution

For each classification system, we sort clusters in decreasing order by size, where size is measured by the number of publications. For each decile of the cluster size distribution, we calculate the average number of publications per cluster. The results are reported in Table 1.

Table 1. Average number of publications per cluster in the partition into deciles of the cluster size distribution. Results are reported for six of the twelve WVE classification systems.

Decile	WoS	Level 2	Level 4	Level 6	Level 8	Level 10	Level 12
1	58,892	73,731	34,807	10,432	2,473	580	169
2	31,494	12	15,901	5,977	1,435	317	89
3	20,298	6	6,569	4,089	1,016	223	62
4	13,840	4	1,008	2,796	737	163	46
5	10,100	3	13	1,810	542	120	36
6	6,916	3	6	976	377	89	28
7	4,455	2	4	322	251	66	23
8	2,849	2	3	19	151	48	18
9	1,663	1	2	3	71	34	14
10	488	1	1	1	6	14	8
Clusters	236	489	613	1,363	5,119	21,849	73,205
Significant clusters	231	39	228	952	4,161	11,172	8,830
Small clusters	5	450	385	411	958	10,677	64,375

The main difference between the WoS system and the WVE granularity levels is the presence in the latter of a large number of small clusters, which we define as clusters with fewer than 100 publications. The number of small clusters in the WoS system is five, while in the WVE granularity levels it ranges from a few hundred in levels 1 to 7 to 64,375 in level 12. However, the share of publications included in small clusters varies dramatically across granularity levels. These publications represent less than 1% of the total in granularity levels 1 to 8, and range from 3.2% to 61.3% of the total in granularity levels 9 to 12. Going from level 1 to level 8, the number of *significant* clusters with at least 100 publications increases monotonically from 17 to 4,161. In the WoS system, we have 231 significant clusters.

The Skewness of Science

We study the skewness of cluster citation distributions by applying the size- and scale-independent technique known as Characteristic Scales and Scores (CSS hereafter). Consider the partition of a cluster citation distribution into three broad classes: (i) articles with a number of citations less than or equal to the mean number of citations, m_1; (ii) articles with a number of citations greater than m_1 and less than or equal to m_2, the mean number of citations of articles with a number of citations above m_1; (iii) articles with a number of citations above m_2. For each significant cluster (including at least 100 publications), we calculate the percentage of articles in each of the three CSS classes. For the WoS system and six of the twelve WVE granularity levels, the average percentages (and the corresponding standard deviations) over all significant clusters are reported in Table 2.

Table 2. Average (standard deviation) over all significant clusters of the percentage of articles in each of the three CSS classes. Results are reported for six of the twelve WVE classification systems.

	WoS	**Level 2**	**Level 4**	**Level 6**	**Level 8**	**Level 10**	**Level 12**
CSS class 1	69.0 (3.3)	70.3 (3.6)	70.3 (3.3)	69.4 (3.7)	68.3 (4.2)	67.3 (4.6)	67.2 (5.2)
CSS class 2	21.5 (2.0)	21.2 (2.6)	21.0 (2.0)	21.2 (2.2)	21.7 (2.7)	22.1 (3.0)	22.1 (3.4)
CSS class 3	9.5 (1.7)	8.5 (1.3)	8.7 (1.7)	9.3 (1.9)	10.0 (2.3)	10.6 (2.7)	10.7 (3.1)
Significant clusters	231	39	228	952	4,161	11,172	8,830

The average percentages of articles in each class –approximately equal to 69-70/21/9-10– illustrate the high skewness of cluster citation distributions, while the relatively low standard deviations show the strong similarity across clusters. These two features –high skewness and strong similarity of cluster citation distributions– are typically found in the literature on citation distributions using large WoS datasets (e.g., Albarrán, Crespo, Ortuño, & Ruiz-Castillo, 2011). However, it should be noted that, on average, cluster citation distributions at granularity levels 7 to 12 exhibit a slightly lower skewness, as well as a smaller degree of similarity across clusters than at levels 1 to 6.

Cluster Homogeneity

In Van Eck et al. (2013), the authors had *a priori* information about the possible lack of homogeneity, in terms of citation practices, of a number of fields in the WoS classification system. We do not have any information about clusters that may be insufficiently homogeneous in the WVE classification systems. Nevertheless, as explained in detail in the

Appendix in RCW, under the reasonable assumption that the degree of homogeneity of clusters increases as the granularity level increases, we can use an additively decomposable citation inequality index to approximate the degree of homogeneity at every granularity level by the ratio of between-group citation inequality and overall citation inequality. As shown in Table 3, the value of this ratio increases in approximately constant steps as we move from level 1 to level 12. It seems sensible to focus on granularity levels with at least the same degree of homogeneity as the WoS systems. This means we should focus on level 6 or higher.

Table 3. Between-group citation inequality as a percentage of overall citation inequality.

Classification system	Between-group citation inequality (as % of overall citation inequality)
WoS	15.9
1	6.8
2	8.8
3	9.7
4	11.3
5	12.8
6	15.1
7	18.8
8	20.9
9	23.8
10	27.8
11	31.1
12	34.7

The Citation Impact of Universities Under Different Classification Systems

We analyze the more than 1.8 million articles authored by the 500 universities included in the 2013 edition of the CWTS Leiden Ranking (Waltman et al., 2012; see www.leidenranking.com) in the period 2005–2008. Our evaluation criterion is the Mean Normalized Citation Score indicator (MNCS hereafter; Waltman, Van Eck, Van Leeuwen, Visser, & Van Raan, 2011). Field normalization is performed based on either the fields in the WoS classification system or the clusters in the WVE classification systems. We emphasize that due to small differences in data and methodology the results of our analysis cannot be compared directly with the official Leiden Ranking results.

We recommend using granularity level 7 or 8 within the WVE sequence because these granularity levels have less than 1% of all articles in small clusters and because they show a greater homogeneity than the WoS system while still capturing in an acceptable way the skewness of science. Since granularity levels 7 and 8 lead to very similar results, in the sequel we will focus on level 8. We note that for all classification systems universities' MNCS values and their ranks can be found in Tables A and B in the Appendix in RCW. We also refer to RCW for a comparison between granularity levels 7 and 8.

There is a strong correlation between the MNCS values obtained based on the WoS system and based on granularity level 8. The Pearson correlation equals 0.94. The Spearman correlation, which takes into account not the actual MNCS values but the ranking implied by these values, equals 0.97. However, these high correlations do not preclude the existence of substantial differences for individual universities. In particular, as can be seen in Tables 4 and 5, we find that approximately one third of the universities change ranks by more than 25 positions, while almost one third experience a difference in MNCS value greater than 0.05. A scatter plot of the MNCS values of the 500 universities is presented in Figure 1.

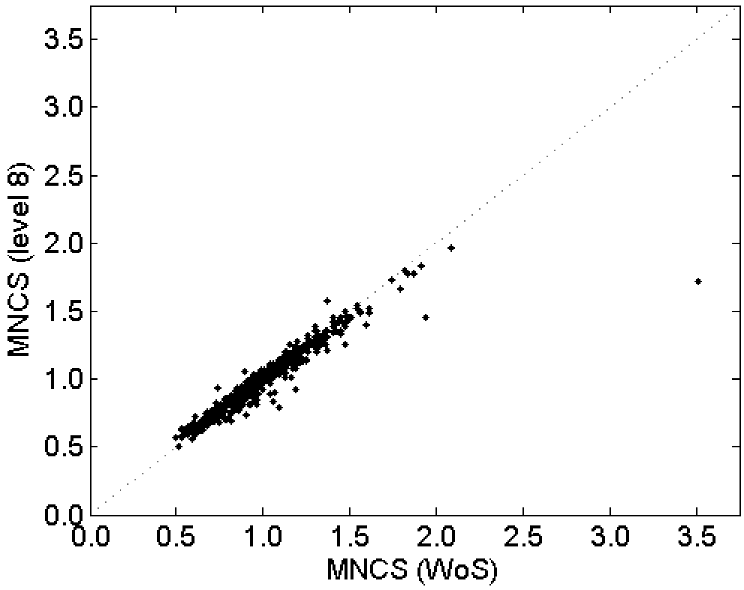
Table 4. Differences in MNCS-based university rankings between the WoS system and granularity level 8.

	First 100 universities	**Remaining 400 universities**	**All universities**
> 50 positions	2	53	55
26–50	12	101	113
16–25	13	88	101
6–15	23	97	120
≤ 5 positions	50	61	111
Total	100	400	500

Table 5. Differences in university MNCS values between the WoS system and granularity level 8.

	First 100 universities	**Remaining 400 universities**	**All universities**
> 0.20	4	3	7
> 0.10 and ≤ 0.20	8	17	25
> 0.05 and ≤ 0.10	32	97	129
≤ 0.05	56	283	339
Total	100	400	500

Figure 1. Scatter plot of the MNCS values of 500 universities obtained based on the WoS classification system and based on granularity level 8.



Among the 100 universities ranked highest based on granularity level 8, there are 11 and 5 universities that end up as, respectively, large gainers and large losers when moving from the

WoS system to granularity level 8. These universities, listed in Table 6, experience a re-ranking of more than 25 positions or a change in MNCS value of at least 0.25.

Table 6. Large gainers and losers when moving from the WoS system to granularity level 8, taking into account only the 100 universities ranked highest based on granularity level 8.

	Level 8 ranking	**Ranking difference**	**MNCS difference**
Gainers			
London School of Hygiene and Tropical Medicine	9	35	0.21
University of Saint Andrews	35	27	0.09
University College London	39	27	0.06
University of Bristol	49	26	0.06
Delft University	62	36	0.08
Queen Mary University London	65	62	0.11
Paris Tech École Polytechnic	70	32	0.06
Tech. University München	87	27	0.04
University of Stuttgart	92	54	0.08
Paris Diderot University	98	35	0.06
McMaster University	100	28	0.04
Losers			
University of Göttingen	7	6	1.78
Rice University	21	18	0.49
University Dublin Trinity College	69	46	0.21
University of Notre Dame	90	48	0.16
Lancaster University	93	36	0.11

Because of space limitations, we have reported only some selected results of our analysis. We refer to RCW for more extensive results.

Conclusions

The basic idea of citation analysis is that the number of citations of a publication reflects, in an approximate sense, the scientific impact of the publication. However, it is generally recognized that the number of citations of a publication depends not only on the impact of the publication but also on many other factors. The field in which a publication has appeared is typically seen as one of the most important factors influencing the number of citations of a publication. This is not surprising, given the fact that publications in some fields (e.g., biochemistry) on average receive about an order of magnitude more citations than publications in certain other fields (e.g., mathematics). Correcting for field-specific factors that influence the number of citations of a publication therefore is a key issue in citation analysis.

Performing an accurate correction for field-specific factors is far from trivial. In general, it requires determining for each publication in a bibliographic database the field (or the fields) to which the publication belongs. This is a problem for which there is no perfect solution. In practice, fields do not have clear-cut boundaries. Fields tend to overlap, and their boundaries tend to be fuzzy. Moreover, fields can be defined at many different levels of aggregation, and it is unclear which level is most appropriate for the purpose of normalizing citation impact indicators.

The conclusion is inescapable: any field normalization of citation impact indicators involves a certain degree of arbitrariness caused by the methodology used to define fields. In this scenario, we have developed a proposal for a normalization approach that is likely to be more accurate than the approach based on the well-known WoS classification system. In so doing, we have also provided some insight into the sensitivity of citation impact indicators to the choice of a normalization approach. Essentially, we have analyzed the uncertainty in citation impact indicators when we use classification systems at different granularity levels.

Our findings can be summarized as follows. Firstly, for the purpose of field normalization, we believe that our algorithmically constructed classification systems offer an attractive alternative to the WoS classification system. Unlike the WoS system, our algorithmically constructed systems are defined at the level of individual publications rather than at the level of entire journals. Based on the criteria we have developed, having between 2,000 and 4,000 significant clusters in an algorithmically constructed classification system seems to be a good choice. Secondly, in the case of the MNCS indicator applied at the level of universities, the sensitivity to the choice of a normalization approach turns out to be relatively small for most universities. In practice, however, there often is a tendency to pay serious attention even to rather small differences in the values of a citation impact indicator. Our results show that this introduces a serious risk of overinterpretation. For instance, in the case of the MNCS indicator applied at the university level, differences of 0.05 may well relate to the choice of a certain classification system and may therefore have little meaning in terms of actual differences in the impact of the publications of universities.

Nevertheless, before serving as definite guides in practice, our findings should be critized and validated by the wider research community. Furthermore, it is worthwhile investigating the consequences of using alternative classification systems at different granularity levels when the research units' citation impact is measured with other indicators different from the MNCS.

References

Adams, J., Gurney, K., & Jackson, L. (2008). Calibrating the Zoom – a Test of Zitt's Hypothesis. Scientometrics, *75*, 81-95.

Albarrán, P., Crespo, J., Ortuño, I., & Ruiz-Castillo, J. (2011). The skewness of science in 219 sub-fields and a number of aggregates. *Scientometrics*, *88*, 385–397.

Ruiz-Castillo, J., & Waltman, L. (2014). Field-normalized citation impact indicators using algorithmically constructed classification systems of science. Working Paper 14-3, Universidad Carlos III (http://hdl.handle.net/10016/18385).

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*, e62395.

Waltman, L., Calero-Medina, C., Kosten, J., Noyons, E.C.M., Tijssen, R.J.W., Van Eck, N.J., Van Leeuwen, T.N., Van Raan, A.F.J., Visser, M.S., & Wouters, P. (2012). The Leiden Ranking 2011/2012: Data collection, indicators, and interpretation. *Journal of the American Society for Information Science and Technology*, *63*, 2419–2432.

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*, 2378–2392.

Waltman, L., Van Eck, N.J., Van Leeuwen, T.N., Visser, M.S., & Van Raan, A.F.J. (2011). Towards a new crown indicator: Some theoretical considerations. *Journal of Informetrics*, *5*, 37–47.

Zitt M., Ramana-Rahari, S., & Bassecoulard, E. (2005). Relativity of citation performance and excellence measures: From cross-field to cross-scale effects of field-normalisation. *Scientometrics*, *63*, 373–401.

Comparison of the number of unique corresponding authors estimated using their e-mail addresses

Ayaka SAKA

saka@nistep.go.jp
Research Unit for Science and Technology Analysis and Indicators, National Institute of Science and Technology Policy (NISTEP), Kasumigaseki, Tokyo, 100-0013 (JAPAN)

Background and Purpose

In Japan in recent years, demand for accountability of investments in R&D has grown strong, and understanding the output of R&D has become an important issue. Focusing attention on scientific articles as a measurable output of R&D activities, over the past decade, Japan has faced not only a decline in share of global scientific article output but also a contraction (-1.7% annual average) in scientific article production (NISTEP, 2013, NSF, 2014).

In this decade, collaboration on S&E research papers has been increasing, with higher shares of scientific articles with institutional and international co-authorship worldwide (NISTEP, 2013, NSF, 2014). The number of authors of scientific articles has also been increasing. With the allocation of credit among authors growing more complex, it is difficult to determine authors' individual contributions.

The corresponding author (CA) is generally understood to be the person who holds the prepublication and post-publication responsibilities for an article (*Nature* journals' Authorship policies). Based on this definition, estimation of the number of unique CAs is an index for capturing the situation of actual leading researchers and for benchmarking real research activity in each country.

In this poster-presentation, the result of comparison of research activities focused on corresponding authors will be presented. Furthermore, the relation between the decline of Japanese S&E article production and CAs as actual leading researchers will be discussed.

Methods

The bibliographic information of publications was retrieved from Web of Science (SCIE, CPCI-S) at the end of 2012. On Web of Science, corresponding authors (CAs) are noted as "reprint authors." The corresponding authors' affiliation institutions and their e-mail addresses on articles were extracted and analyzed.

The numbers of articles of selected countries were measured based on all authors' affiliated institutions, using the whole counting method. The number of CA articles was counted based on the countries of the CAs' affiliated institutions. For example, in the case of an article co-authored by researchers from Japanese and U.S .institutions, if the CA's affiliated institution was located in Japan, the article was counted as a CA article for Japan.

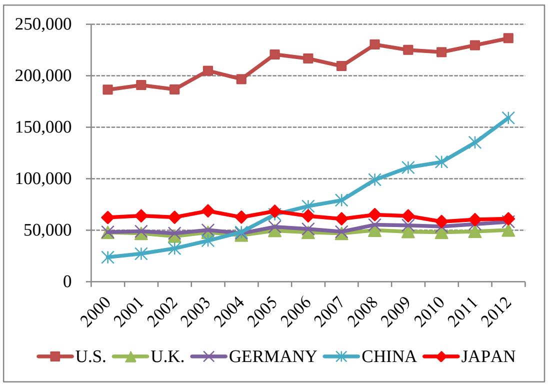
The name and e-mail address of a corresponding author is published in each paper. The unique corresponding authors (unique CAs) were identified using these e-mail addresses, based on the singularity of e-mail addresses. The number of unique CAs was counted based on the country of the CA's affiliated institution.

Results

Capturing research activity in selected countries

First, the numbers of articles and the numbers of CA articles were counted for selected countries. As shown in Table 1, only Japanese CA articles have been falling in the 2000s. Other countries show rising CA articles.

Table 1. The number of CA articles in selected countries



Note 1: This figure shows results calculated using the CA's affiliated institution.

Comparison of the percentage of CA articles within the articles of each country was also made. Over time, the trend was downwards for each country. It was low for Germany (67%) and the U.K. (63%), similar for Japan (83%) and the U.S. (79%), and somewhat higher for China (90%) in 2012.

Estimation of Unique CAs in selected countries

To identify real research activity in each country, unique CAs as actual leading researchers were isolated using their e-mail addresses.

E-mail addresses of CAs have only been captured to a certain extent in this database since 2007. E-mail address information of CAs was found for about 90% of the CA articles of each country. The time range of analysis was two periods, 2007–2009 and 2010–2012.

Table 2 shows the number of unique CAs in each country, estimated using e-mail addresses. Japan, the U.K., and Germany show similar numbers of unique CAs. The number of unique CAs for Japan was 78,076 in 2007–2009 and 73,415 in 2010–2012. Japan decreased by 4,661. Comparing the two periods, with the exception of Japan, unique CAs in each country increased.

Table 2. Comparison of the number of unique CAs estimated using e-mail addresses

	DY2007-2009	DY2010-2012
JAPAN	78,076	73,415
U.S.	286,321	292,698
U.K.	70,405	71,886
GERMANY	72,372	75,405
CHINA	100,497	140,554

Looking at the distribution of the number of papers written by unique CAs over three years, about 60% of unique CAs published only one article during the three years. For Japan, it was found that the number of unique CAs writing one article during three years fell by 4,210, accounting for 90% of the Japanese decline in unique CAs between the two periods.

Characteristics of Japanese unique CAs

Analysis poses the following two possible explanations for the drop in unique CAs. First, it is possible that as internationally co-authored papers increase, Japanese researchers are not captaining the research teams. Second, it is possible that the number of researchers performing independent research is falling.

To identify the characteristics of the decrease in unique CAs in Japan, all unique CAs were classified by the collaboration pattern of articles, such as international co-authored articles and domestic articles, and by citations in each type (Table 3).

Comparing the two periods, the number of unique CAs who were solely responsible for international co-authored papers was almost unchanged, but the number of unique CAs with domestic papers fell sharply. Additionally, the number of unique CAs with Top 1% or Top 10% highly cited papers stayed about the same. The large drop in the number of unique CAs was in CAs with "normal" papers.

Discussion

In this study, the number of unique CAs was estimated based on e-mail addresses, and research activities in selected countries were compared. My findings show that one cause of the decrease in Japan's number of papers was a decline in the number of unique CAs who were responsible for domestic or normal articles. The factors behind this decline are thought to be a change in the balance of the number of retired CAs, newcomer CAs, and CAs who could not publish one article during 3 years.

References

NISTEP. (2013) "Japanese Science and Technology Indicators 2013," Research Material-214

NSF. (2014) Science and Engineering indicators 2014

Eades JJ, Goodman R, Hada Y, editors. (2005). *The 'Big Bang' in Japanese Higher Education: The 2004 Reforms and the Dynamics of Change*. Victoria, Australia: Trans Pacific Press.

Nature journals' Authorship policies, http://www.nature.com/authors/policies/authorship.html

Table 3. Change in the number of CAs by co-authorship form and citation class

	ALL		Only domestic papers		Only international co-authored papers		Domestic papers and international co-authored papers	
	DY2007-2009	DY2010-2012	DY2007-2009	DY2010-2012	DY2007-2009	DY2010-2012	DY2007-2009	DY2010-2012
All of unique CA	78,076	73,415	63,638	58,994	7,632	7,492	6,806	6,929
CA with at least one Top1% highly cited paper	612	611	335	289	80	75	197	247
CA with at least one Top10% highly cited paper	7,270	6,678	4,695	4,114	752	676	1,823	1,888
CA with only normal paper(s)	70,806	66,737	58,943	54,880	6,880	6,816	4,983	5,041

Note 1: "Normal" papers are papers other than Top 10% highly cited papers.
Note 2: Papers are classified by affiliations of all authors.

The Complex Relationship between Competitive Funding and Performance[1]

Ulf Sandström*, Ulf Heyman** Peter van den Besselaar***

*ulfsandstrom@oru.se; ulf.sandstrom@indek.kth.se
Business Studies, Örebro University, SE-70130 Örebro (SWEDEN)

**ulf.heyman@uuadm.uu.se
Planning division, Uppsala University, SE-751 05 Uppsala (SWEDEN)

***p.a.a.vanden.besselaar@vu.nl
Department of organization sciences & Network Institute, VU University Amsterdam (NETHERLANDS)

Abstract

A growing interest for the use of international funding data in relation to scientific output highlights that efficiency at the research system level is a complex research question. As pointed out by many scholars already the OECD expenditure indicators are problematic. Not to mention the problem of how to account for research output. In this paper we suggest a method for treating both of these problems. In the present study we compare the change in scientific output with the change of funding, which to a large extent eliminates the problem of differences between countries but still requires that changes within each country is limited or possible to correct. Based on this contribution we critically discuss a new approach on the role of competitive funding developed by Abramo et al. (2012) in response to a contribution by Auranen & Nieminen (2010). Our results indicate that the level of competitive funding in a research system not at all is correlated to increases in citation performance. Additionally, we find that our data to some extent contradict the systemic relations proposed by Abramo et al.

Introduction of the problem

What is an efficient research system, how to measure efficiency and what characteristics are most important? The debate about efficiency has a long tradition in the political economics of science (for an overview, see Stephan 2012). We would argue that there are actually two problems involved, one conceptual and one empirical.

With respect to the conceptual problem, efficiency of research systems has traditionally been discussed in terms of the level of competitiveness. Competiveness is often defined as the share of basic university funds in total research funding (Abramo et al., 2012): The more institutional funding and the less project funding, the less competition. Also other systems pressures, such as new public management (NPM) and national research assessments, are associated with the level of competition (Auranen and Nieminen (2010). In this work, the authors derive characteristics that would characterize competitive (and therefore better performing) research systems, such as large variety in the quality of higher education institutions (Abramo et al., 2012). Responding to these contributions, we address first the question of the relation between the level of competiveness and performance at the national

[1] This work was supported by the Swedish "Riksbankens Jubileumsfond" (Tercentenary Foundation) project number P12-1302:1. The authors give their thanks to Staffan Karlsson at KTH Library.

research systems level: are more competitive systems performing better or not? We then discuss what that would imply for the concept of competition in research systems.

To answer the question, we are faced with the second issue distinguished above: In order to account for efficiency, one needs reliable input and output data of the science system. This is a longstanding problem. Before addressing our main research question, we will first propose solution to the data problem. After having done so in section 2, we address in section 3 the discussion about competitiveness and performance. This section ends with the specification of the questions to be answered in the paper. In section 4, we discuss in some detail the data used in this study. In section 5, we analyse the relation between funding and output. In the next three sections, we test some presuppositions found in the literature: does the share of direct university funding correlate negative with performance (section 6), do highly competitive countries have a few top universities (section 7), and is performance in Swedish universities uniformly distributed (section 8)? We end with conclusions about the meaning of the concept of competition and concerning the relation between competition and performance.

2. Measuring input s in the research system

Careful accounting for real R&D expenditures is needed when questions of input-output is put into focus. This very problem has followed, we should say haunted, the political economics of research ever since the area started during the 1980s and 1990s (Stephan, 2012; Cole and Phelan, 1999).

Robert May, UK Chief Scientist, disclosed in *Science* (1997) that UK was the most efficient country based on citations per £million government money spent on higher education research (HERD). After a critical debate (Grant & Lewison, 1997), pointing at some data problems, May gave a response based on a new indicator called the “Science Base” Expenditures on R&D (SBRD) which covered expenditures at universities and non-profit making institutions, irrespective of funding sources, and including research establishments (research institutes). UK was still the most efficient country, but Canada, Sweden, the Netherlands and Switzerland had higher ratios between output (papers) and input.

A follow-up study by David King, UK Chief Scientific Advisor, published in *Nature (2004)*, took the discussion to higher levels. King used rebased impact analysis (field normalization) and presented it to a wider audience. King noted that the OECD data for R&D expenditures gave spurious results, but sorted out per cent GDP spent on publicly funded R&D plus per cent GDP spent on higher education R&D (HERD) for his comparison. Again, UK was low on input indicators but first in all “normalized” output indicators (publications/researcher, citation/researcher, citations/unit HERD).

Two British high profile scientists gave their view but, still, it was very unclear what type of research money that should be counted. Return on investment is a serious problem and measurement issues that arise have to be discussed. The thread was taken up by Swedish researchers Jacobsson and Rickne (2004), who questioned the idea that the Swedish academic sector was bigger than in other countries. Conventional measures were considered inappropriate and the results were skewed to such an extent that figures probably had to be interpreted with care as organizational boundaries were drawn differently in different countries. Basically, they argued that the actual distribution over financial categories were the result of “different organizational choices”.

In a subsequent article Granberg and Jacobsson (2006) gave more details attacking the myth of Sweden as a well-funded research system. Monetary values were affected by structural differences, e.g. how PhD-student salaries were accounted, how their PhD-education was financed, and how buildings and office space was taken into account. They found that the HERD indicator was seriously flawed.

The discussion on these issues have continued, e.g. by Leydesdorff and Wagner (2009). They presented another analysis using macro-level indicators for funding and output. Without addressing the problematic nature of these data discussed above, they used both HERD and GOVERD (Government Expenditure on R&D) to account for differences between countries. These figures indicated that most of the countries in the northwest of Europe had about the same costs (<200,000 $PPP/publication), but there were outliers and strange results which made the analysis hard to interpret. Considerable ranges in terms of cost per publication were found also for countries that in the authors view were very much the same system.

How to solve this problem? So far, we note that on both sides of the input-output relation there are a lot of difficulties. Is it possible to measure countries research efficiency if the problem on how to proceed with databases and measurements is still unsolved? The imperative for use of statistical data is often hard to avoid (c.f. Allik, 2013).

3. Competitiveness in the research system

In this context we point at an interesting and constructive attempt to build an interesting data set for seven European countries plus Australia (selection of countries were not justified) by Auranen and Nieminen (2010). Also this time UK come out on top, and consequently, Sweden and the Netherlands were considered as poor performers with low efficiency, i.e. high cost per paper. Finland, Australia and Denmark were a group in between. Germany and Norway were close to Sweden and Netherlands.

The methodological innovation consisted in setting the issue of efficiency in relation to the ongoing changes in the research system due to pressures for better performance under the new regime defined by excellence initiatives, research assessments, and New Public Management. In their analysis Auranen and Nieminen proposed a typology of input- and output oriented core funding on one axis and share of external funding on the other axis.

Obviously, UK had been involved in RAE since many years so their system would be described as an output oriented core funding system, and, paradoxically, on top of that relatively more research money was distributed over the research funding agencies in the UK. High levels of external funding were combined with output oriented core funding. On the other hand countries like Sweden were considered as “a quite non-competitive environment” based on the fact that core funding was input oriented (student numbers, history and politics).

There are two opposed positions in a quadrant: one the one hand the diagonal positions output oriented-small share external *versus* input oriented large share external. Country-wise that would be Australia versus Finland and Sweden. The other opposed diagonal were on the one hand UK (output – large external) *versus* Norway, Netherlands, Germany and Denmark (input – small external).

However, although the analytical scheme seems interesting it does not produce interpretable results. Anomalies are commonplace; Denmark and Australia are in the wrong quadrant, Finland as well. When the authors discuss their results they ignore these anomalies and at the end they consider Sweden to be a (typical) example of a non-competitive research system, a statement we find highly improbable. Therefore, our paper aims to test some of the propositions that follow from the work of Auranen and Nieminen (2010) and especially how these have been developed by Abramo et al., (2012).

Abramo et al. (2012) formulated theoretical propositions, intrigued by the results in Auranen et al (2010), concerning the expected effects of a ‘really’ competitive academic system. They argued that over time competitive arrangements should lead to the concentration of high performing scholars between universities, i.e. the competitive process should lead to a selection of competencies and concentration in a few top universities. This, in their view, leads to a higher performance variety between universities and at the same time to a lower performance variety within each university. Low competition would lead to the opposite pattern: performance differences between universities will be small (as there is a lack of concentration of top talent) but the performance differences within universities will be large. With an empirical test they showed that Italian data confirmed the hypothesis that Italy was a non-competitive system, and they challenged other researchers to do follow-up on other countries. They saw a new competition indicator at the front line.

In the following we will test the hypothesized relation between performance and level of competition within research systems in three different ways – as presented above:

- Firstly, does the level of indirect (project) funding correlate with performance? The assumption is that systems with large General University Funds are less competitive. To test this, we will compare changes in output with changes in funding.
- Secondly, test the Abramo et al. hypothesis that highly competitive systems have a high concentration of performance in a few universities. We will use the Leiden ranking for this.
- Thirdly, we will tentatively test the Abramo et al. hypothesis that highly competitive systems may show larger performance differences between HE institutions, but low performance differences within each university. We will use Swedish university data to do this.

4. Data

Publication and citation data was collected from Web of Science and basic calculations kindly performed by the library at KTH and/or the authors. Basically, we have used the field normalized citation score (MNCS) multiplied by the fractionalized number of papers (Frac P) as a measure of scientific output.

Total funding for R&D in the higher education sector (HERD) in local currency and constant prices was chosen as measure of resource input. In most countries the vast majority of scientific articles originate from the HE-sector, the exception being countries with a large institutional sector that is not included in HERD in the OECD statistics (e.g. Italy). HERD is however by far the best measure since it includes all funding and excludes most of the R&D that results in very few papers.

Since the study concerns the rate of change in input and output it has been important to use a fairly long time period. Older data is however often of lesser quality and also longer series increases the probability of structural changes that may affect the results. The final dataset spans the period 1997-2011 and consist of 32 countries, for which economic data where present and publication data of reasonable magnitude.

There are several reasons for not using the direct indicators of bibliometric index divided by funding in PPPs. The cost of graduate students vary depending on if they receive a salary or not, renting or owning the premises results in large differences in cost, funds are to different extent recycled back to the government etc. This results in a lack of coherence between the economic data and the personnel data (which suffers from other comparability problems) in the OECD statistics. A correlation between indices of bibliometric index divided by HERD in PPP respectively full time equivalents of R&D personnel yields an R-square of 0.06 for the whole dataset, which hardly is good enough for an analysis.

If the rate of change of the variables is used instead of the direct quotients, much of the effects of structural differences between countries will be eliminated, and only changes within each country during the time period studied will affect the comparison. Indeed the R-square between change of scientific output divided by change of respectively personnel and funding rises to 0.48, which seems more reasonable in view of methodological difficulties.[2]

For analysing the performance differences between universities, we use the CWTS ranking data. For a set of countries we calculated the Coefficient of Variance (Cv) of the PP(10%) scores and the (Cv) of the MNCS scores. Consequently, the higher the differences between the universities (independent of the average level[3]), the higher the two measures for Cv will be. For the analysis at university level analysing the variety in performance levels within universities, we use disambiguated Swedish data for individual researchers at universities.

5. How research funding is related to scientific output

The relation between change of funding and change of citations for the total dataset gives an R-square of 0.42, which must be considered fairly strong in view of the large differences between countries. From their economic state and history three quite distinct groups of countries can be discerned: The fast growing emerging countries, the old OECD-countries and the countries from former Eastern Europe.

The former Eastern European countries show a lot of variation both in funding and resources and this should come as no surprise in view of the great political changes in especially the nineties. The large variability and possibly lower quality statistics explains why there is no relationship between output and input for these countries (Fig 1) and they are excluded from further analysis.

In the emerging countries it seems as if a monetary input is much more effective than in the rest of the countries and the relation between funding and citations is fairly strong (Fig 1). The fast development and high efficiency does however seem to be more of a transition state than a structural difference, since regressions made for the period 2004-2011 show much less difference between this group and the established countries.

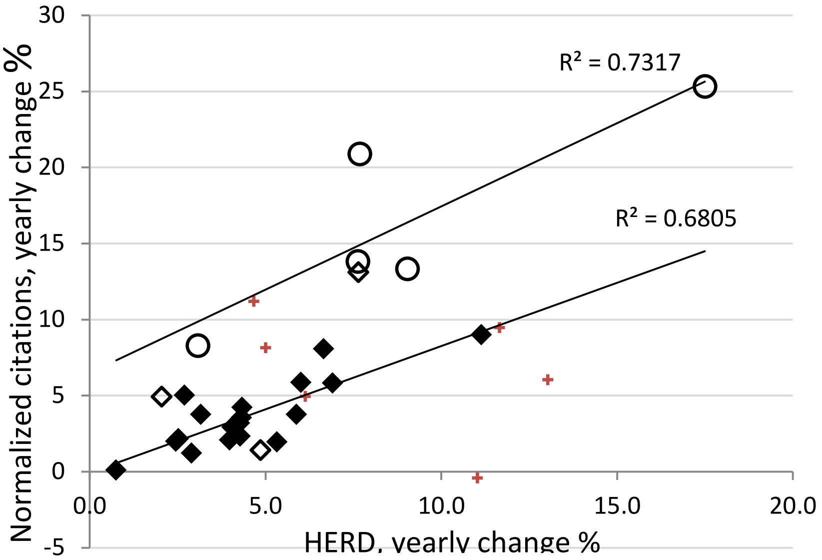


Figure 1. Relation between change in funding and change in citations.

[2] Coherence of OECD data and how to treat breaks in the time series and other methodological issues will be considered in the full paper, c.f. Maass (2003).

[3] (Cv) = standard deviation divided by the mean.

Note: Open circles denote emerging countries; small crosses former Eastern European and diamonds other OECD-countries. The open diamonds denotes countries that are excluded from the regression.

On the basis of known database problem we suggest the exclusion of some countries, US, Italy, Portugal. It is reasonable to exclude these countries when calculating a regression line, which gives a strong correlation between input and output (R-square 0.68).

Using the rate of change as measures of funding and citations thus gives the quite unsurprising result that increased funding is the main factor for increasing research output (confirming results presented by Bornmann et al., 2014). It also shows that the fast growing, especially Asian, countries tends to blend with the more established countries so that change of funding between 2004 and 2011 results in almost the same change of citations for the two groups. For the established countries the slope of the regression line increases with the shorter time span and the fit is very good (r^2=0.74) for the last available eight-year period (2002-2009) for funding data.

6. The effect of direct funding

The level of funding is a dominant factor for the development of a country's publication record. To evaluate the influence of other factors, we must eliminate the effect of funding. This can be done by comparing the rate of change in funding with the rate of change in productivity. Countries below the regression line can be deemed less efficient than countries above (Figure 1). The deviations from the estimated value (the residual) lead to an efficiency ranking of the various countries' research systems (Figure 2).

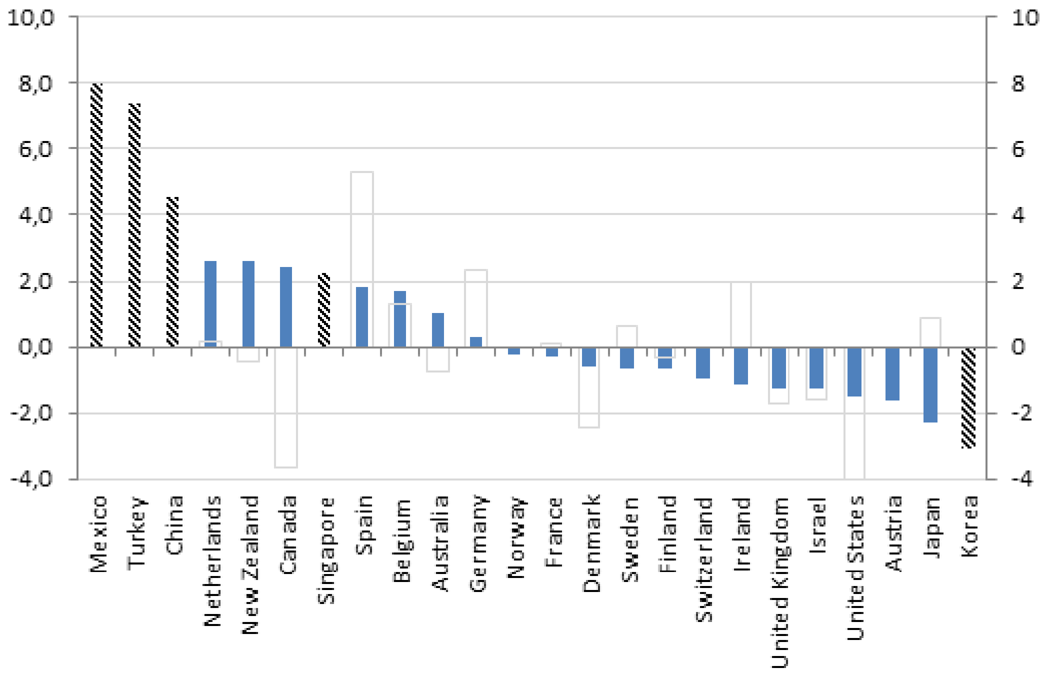


Figure 2. Efficiency of research systems, period 2002-2009

Note: Based on the residuals. Open bars represent residuals for the period 1991-1998.

The OECD data includes general university funds (GUF) which is the sum of direct government funding and the universities own funds. Here we have calculated university GUF using the figure for Civil GBAORD for General University Funds divided with the figure for HERD. The statistics is a contested area and there are probably differences in how concepts behind the statistics are interpreted in each country. Therefore, we include in Table 1 figures from a recent OECD-project (van Steen, 2012) showing the institutional funding (and level of project funding) to the HE sector. Project funding is to a large extent competitive, but since also the direct funding may have competitive components (due to PBRF) it is not a measure of competitiveness but may be used as an indicator.

Table 1. Parameters and Coefficients of variation per country

COUNTRY	INST%*	GUF%**	Residual	PBRF	Cv_MNCS	Cv_top10%
New Zealand	90	22	2,6	2001	0.457	0.501
Netherlands	80	**66**	2,6	1983	0.072	0.077
Canada	55	25,4	2,5		0.133	0.185
Spain		48,8	1,8	1989	0.097	0.172
Belgium	35	26,1	1,7	1991	0.073	0.112
Australia	47	56,8	1	1993	0.109	0.161
Germany	90	60,8	0,3		0.089	0.128
Norway	60	64,4	-0,2		0.020	0.030
France		52	-0,3		0.126	0.166
Denmark	95	59,7	-0,6		0.107	0.119
Sweden	**45**	45,9	-0,6		0.085	0.119
Finland	**45**	44,4	-0,7		0.083	0.099
Switzerland	80	65,7	-1		0.107	0.129
Ireland	50	35	-1,2		0.091	0.136
Israel	95	47,2	-1,3		0.260	0.420
United Kingdom	**35**	34,3	-1,3	1986	0.140	0.196
United States		66,4	-1,5		0.239	0.322
Austria	90	67,5	-1,6		0.079	0.120
Japan		39	-2,3		0.163	0.298

Notes:* based on van Steen (2012), p 19, **based on MSTI (2014) period 2008-2010, PBRF=year of intro. Figures in bold are added by the authors as approximations. Cv columns are based on own calculations.

There's no need for a multiple regression in order to establish that the group of "old" countries (the emerging group is too small and lacks to some extent information on GUF), when they are ordered according residuals, there's no correlation to neither GUF nor institutional funding. In the first group of countries with a positive residual there are both countries with a high GUF or a high level of institutional funding in the HE sector (based on figures presented in van Steen, 2012) and countries with low figures on these parameters.

Obviously, there seems to be more of an explanation if we look at the column for introduction of PBRF (Performance Based Research Funding), all countries, with the exception for Canada, have introduced or started to discuss PBRF during the 1990s. One interpretation would be that this has changed the publication culture in the system towards more WoS-publications and in turn this has geared the system towards higher impact (citations). There is one exception to the rule, the UK variant of RAE does not seem to imply higher efficiency, which might be due to 1) the construction based on peer review (only) and 2) the relative unimportance of direct funding in a system based in project funding.

7. Does highly competitive countries have a few top universities?

We calculated the coefficient of variation, $Cv_{,}$ for several countries, in order to test whether larger performance differences between national universities correlates to higher (positive) residuals. If higher efficiency in the research system is related to the level of competition as predicted by Abramo et al. (2012) there would be a strong correlation. Top here is actually measured in terms of relative top. In the full version we will include ‘absolute’ (international) top too.

Using the CWTS ranking, we have information per country about the share in the top 10% most cited papers (PP10%) and about the mean normalized citation score (MNCS) of each individual university. We calculate per country the Cv for PP(10%) and for MNCS – which correlate high (r=.96). Columns to the right in Table 1 shows the results, and data for Cv_MNCS in relation to institutional funding is plotted in Figure 3.

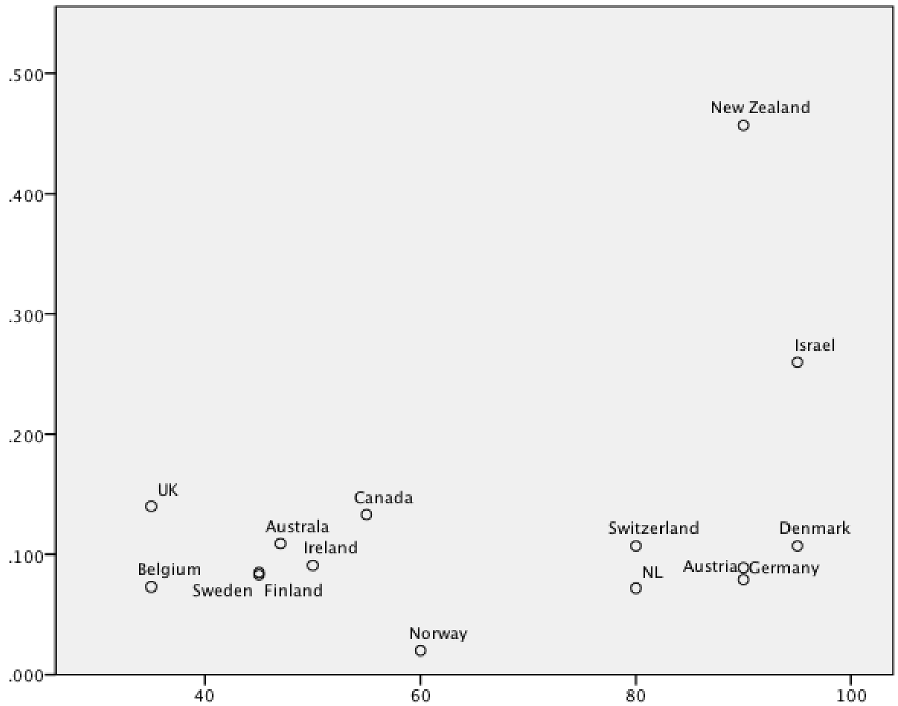


Figure 3. Performance differences (Cv MNCS) by % institutional funding

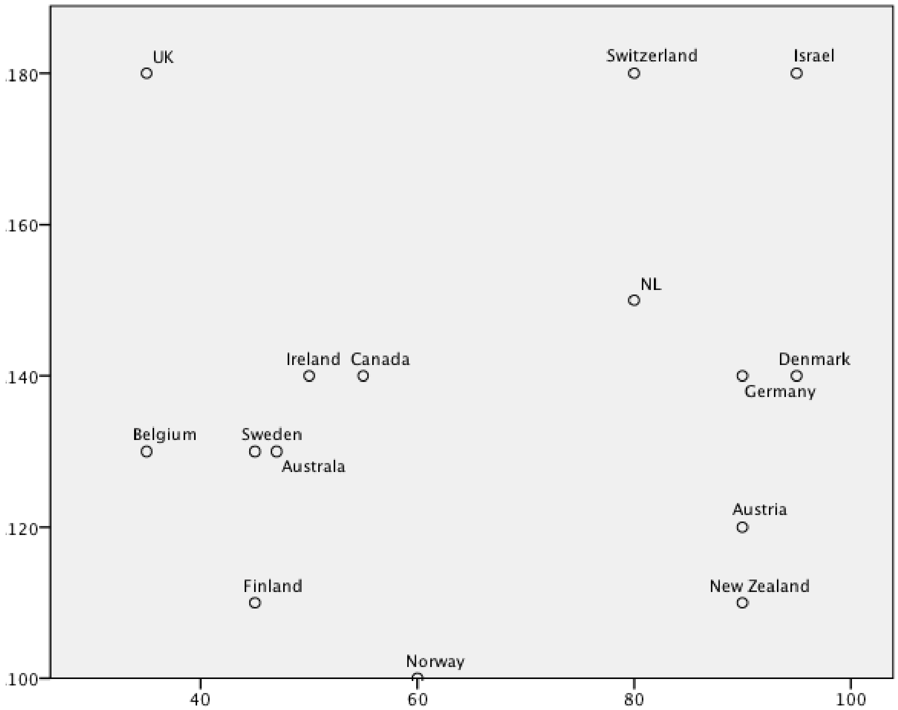


Figure 4. Performance, country's best university (share in top 10% cited papers), by % institutional funding

As Figure 3 shows, there is no indication that the share of institutional funding correlates with differences between universities at the country level. That could either mean that competitiveness plays no role, or that competitiveness is not adequately measured by the share of project/institutional funding. The same holds for the relation between the residual and the level of the top university (measured as its share in the top 10% cited articles) as becomes clear in figure 4.

8. Are performances at Swedish universities equally distributed?[4]

We use data per university per individual researcher's performance for of the eleven main Swedish universities (same type of data was used by the Italian team). For each of the universities the Coefficient of Variance ($Cv_{)}$ of the (field normalized) top 10% cited papers and the MNCS (so two indicators for the differences in performance levels) were calculated for all researchers per university. We then took the weighted average of the two measures to have one measure for the level of homogeneity in performance levels for each university. We also took the average of the ranking based on the top 10% cited papers, and based on the MNCS. Table 3 show the findings.

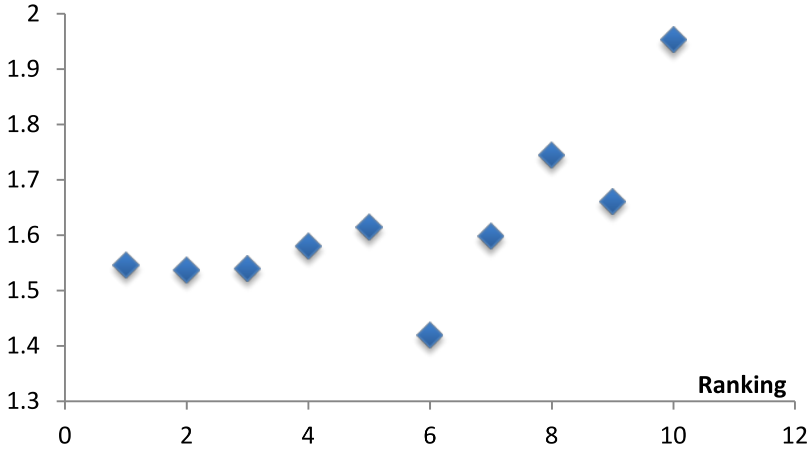
[4] The CWTS ranking can be used for this too: performance indicators per discipline can be used to derive a proxy for performance variety within universities.

Table 3. Indicator variety within universities

Coefficient of Variation	MNCS	Top10%	Weighted average	Ranking
Swedish Univ Agr Sci	0.98	2.30	1.42	6
Karolinska Inst	1.10	2.41	1.54	2
Univ Gothenburg	1.09	2.43	1.54	3
Stockholm Univ	1.21	2.22	1.55	1
Chalmers Univ Technol	1.14	2.46	1.58	4
Lund Univ	1.13	2.54	1.60	7
Uppsala Univ	1.21	2.43	1.61	5
Linköping Univ	1.15	2.69	1.66	9
KTH	1.35	2.54	1.75	8
Umeå Univ	1.64	2.58	1.95	10

There seem to be a relation between the ranking and the level of homogeneity of performance: the lower the Cv, the higher the position on the ranking. This may also be an effect of the skewed distribution of the parameters. Figure 5 shows the association between the Cv and the ranking.

Figure 5. Performance homogeneity* by ranking



* Concentration is the 'weighted average' in table 3. The higher the score, the more the variation in performance. The lower the ranking#, the better the university

9. Conclusion

Although previous contributions have formulated ideas about the level of competitiveness and the performance of research systems, our analysis indicates that the relation between these two variables is less obvious than suggested. First of all, the share of institutional funding does not correlate with competitiveness, overall performance, and top performance. And, more competitive systems do not result in larger differences between performances of universities. Finally, better performing universities seem to have a somewhat more homogeneous performance at the individual level than lower performing universities, but this is also not in line with the hypothesis that the within university variety of performance is related to the competitiveness of the research system.

Obviously, there is a lack of understanding concerning the nature of competition, and how competitive mechanisms manifests themselves at the level of university, in order to establish a relationship between national systems' performance, and national systems' competitiveness.

References

Abramo, G., Cicero, T., & D'Angelo, C.A. (2012). The dispersion of research performance within and between universitites as a potential indicator of the competitive intensity in higher education systems. *Journal of Informetrics*, 6, 155-168.

Allik, J. (2013). Factors affecting bibliometric indicators of scientific quality. *TRAMES*, 17, 199-214.

Auranen, O. & Nieminen, M. (2010). University research funding and publication performance-an international comparison. *Research Policy*, 39, 822-834.

Bornmann, L. Stefaner, M. de Moya Anegon, F. & Mutz, R. (2014). What is the effect of country-specific characteristics on the research performance of scientific institutions? arXiv:1401.2866v2.

Cole, S. & Phelan, T.J. (1999). The scientific productivity of nations. *Minerva*, 37, 1-23.

Granberg, A & Jacobsson, S (2006). Myths or reality-a scrutiny of dominant beliefs in the Swedish science policy debate. *Science and Public Policy*, 33, 321-340.

Grant, J. & Lewison G. Government Funding of Research and Development. *Science*, 278, 878-879.

Jacobsson, S & Rickne, A. (2004). How large is the Swedish 'academic' sector really? A critical analysis of the use of science and technology indicators. *Research Policy*, 33, 1355-1372.

King, D.A. (2004). The scientific impact of nations. *Nature*, 430, 311-316.

Leydesdorff, L. & Wagner, C. (2009). Macro-level indicators of the relations between research funding and research output. *Journal of Informetrics,* 3, 353-362.

Maass, G. (2003). Funding of public research and development: trends and changes. *OECD Journal of Budgeting* 3, 41-69.

May, R.M. (1997). The scientific wealth of nations. *Science*, 275, 793-796.

May, R.M. (1997a). Response (to Grant & Lewison). *Science*, 278, 879-880.

Steen, J. v. (2012), "Modes of Public Funding of Research and Development: Towards Internationally Comparable Indicators", *OECD Science, Technology and Industry Working Papers*, 2012/04

Stephan, P.E. (2012. *How economics shape science*. Harvard U.P.

A textual approach to measure the interdisciplinary character of research proposals for ERC starting grants[1]

Edgar Schiebel**, Ivana Roche*, Dominique Besagni*, Marianne Hörlesberger**,Claire François*

**ivana.roche@inist.fr; dominique.besagni@inist.fr; claire.francois@inist.fr*
CNRS, Institut de l'Information Scientifique et Technique, UPS 76, 2 allée du Parc de Brabois, Vandoeuvre-lès-Nancy, F-54519 Cedex, France

***edgar.schiebel@ait.ac.at; marianne.hoerlesberger@ait.ac.at;*
AIT, Austrian Institute of Technology GmbH, Donau-City-Strasse 1, 1220 Vienna, Austria

Abstract
Interdisciplinary research is discussed to be one key issue for a better transfer of knowledge created in science to innovation. Therefore there is a political interest that the European funding system promotes more interdisciplinary research. The ERC was founded to foster and encourage frontier research in Europe with interdisciplinary research as a key dimension. In the DBF project that was funded by the European Research Council we were interested whether interdisciplinary proposals had a better or a worse chance for selection by the review process.

The indicator interdisciplinary was designed to measure the interdisciplinary character of a proposal. The approach is a proxy to infer self-consistently the presence and proportions of characteristic terminology associated with individual ERC main research fields (panels). It was designed upon an approach that the frequency of occurrence and distribution of research field specific keywords of scientific documents can classify and characterise research fields. The keywords then classified itself as specific for research fields were used to measure the interdisciplinary character of documents. Comparing the distribution of proposals by the degree of interdisciplinarity shows a bias to less successful starting grant proposals with higher interdisciplinarity.

Introduction
The work summarized in this paper was as a part of the DBF project "Development and Verification of a bibliometric model for the Identification of Frontier Research" (ERC 2014). This paper describes the approach how to measure interdisciplinarity of proposals defined as one dimension of the definition of frontier research.

We used the following definition from the high level expert group as a starting point:
"Frontier research pursues questions irrespective of established disciplinary boundaries. It may well involve multi-, inter- or trans-disciplinary research that brings together researchers

[1] This work was supported by the European Research Council (ERC) and is a further development of work performed in the DBF project that was a Coordinated Support Action (CSA) and was carried out from 2009-09-01 to 2013-02-28. It was one of two CSAs that were financed in 2009 (two others having been financed in 2008) as part of a process of building up a comprehensive portfolio of projects and studies to support the ongoing monitoring and evaluation work as well as to the future strategy and policy development at European Research Council (ERC).

from different disciplinary backgrounds, with different theoretical and conceptual approaches, techniques, methodologies and instrumentation, perhaps even different goals and motivations. " (European Commission (2005)).
Wagner et al (2011) published a review on approaches to measuring interdisciplinary research. They found that among the quantitative measures bibliometrics are the most developed and that some newly emerging measures like diversity, entropy and network dynamics are promising. They concluded that these approaches leave gaps in understanding require sophisticated interpretations or carry burdens of expense and lack of reproducibility.

We decided to develop a pragmatic and affordable approach to measure the interdisciplinary character of proposals based on available information for the ERC.
The initial task was to translate this characteristic of frontier research into a textual based indicator. The basic idea was to examine if terminology of different disciplines was used in submitted proposals.

Methodology
Initially there were two different methods chosen to operationalize *interdisciplinarity*. Both methods were based on looking at the occurrence of key words. The idea being that key terms could be assigned to specific disciplines and that a proposal that contained key words from different disciplines were more interdisciplinary. We used the panels and the panel keywords of the ERC as disciplines and defined two indicators.
Indicator 1: The first method was designed to look at whether the proposals were inter-disciplinary according to the number of different ERC Panels that have been allocated in the proposal by the applicant.
Indicator 2: The second method involved a lexical analysis and extracted key words from the summaries of proposals in order to see whether the applicants used key words from different disciplines (ERC keywords).

Calculation of the Indicator 1
The hypothesis we worked with was that the interdisciplinary character of a proposal is higher or lower the more or less other panels have been specified in the proposal.
The calculation of interdisciplinarity indicator 1 (CPI) needed the following steps:

1. Counting the different number of panels assigned by the author of the proposal.
2. Calculation of the indicator by the following formula: $I_i = (N_i-1)/3$, with I_i the indicator value for the i-th proposal, N_i number of different panels. We normalize the indicator by the maximum possible number of different panels without the main panel[2]

Calculation of the indicator 2 (keyword based indicator)
The hypothesis we work with is that the interdisciplinary character of a proposal is higher or lower the more or less keywords from other disciplines than the home discipline occur in the summary of the proposal. We used a similar approach what we called "diffusion model" in a former work, see Schiebel et al (2010).
The calculation of interdisciplinarity indicator 2 needs the following steps:

[2] The applicants could assign one panel as the main panel

1. Extracting all phrasemes (keywords with several single terms such as ”gene expression”) from the summaries of the proposals by automated indexing tagged with the main panel given by the applicant.
2. Creating lists of keywords for each panel using phrasemes from proposals and the tagged main panels
3. Calculating the TFIDF for all phrasemes of a panel
4. Each phraseme was tagged with the panel as a home panel (discipline) with the highest TFIDF.
5. Calculation of the number of home panel keywords and the number of all keywords for each proposal. Calculation of the indicator by the following formula: indicator value = ((number of all keywords) – (number of main panel keywords)) / (number of all keywords) in percent as an integer value.

The assignment of home panels to keywords was performed with the help of a modified TFIDF. Instead of a single document, we took a panel as an artificial document. The corpus of one panel consists of all keywords of proposals that are assigned to this panel. We defined the TFIDF as follows:

$TFIDF_{ij} = (hkw_{ij}/Hj)*\log(N/nkw_i)$,

with i: keyword i

j: panel j,

hkw_{ij}: the number of proposals where a keyword i occurs at least once in a panel j

H_j: The number of proposals assigned to a panel j

N: number of all proposals

nkw_i: the number of proposals where a keyword i occurs at least once

To determine the home panel of a keyword we took the panel with a keyword’s highest TFIDF for all panels.

Data

We used proposal data of starting grants from the year 2009 (SG2009) for the measurement of the indicator. Additionally the definition of panels and related panel keywords in the version for the year 2009 as available.

Proposals for starting grants

The table of proposal abstracts included the following information: proposal ID, successful or not successful, main panel, 4 possible panel keywords, free keyword given by the author, acronym, title, abstract and the summary, The number of successful (SGA2009) and non-successful (NGA2009) starting grant applications was 130 and 628, respectively.

Panels and panel keywords

The panel definition of the ERC in the current version of 2009 had defined 25 panels to cover all the fields of science, engineering and scholarship assigned to three research domains: Social Sciences and Humanities, Physical Sciences and Engineering and Life Sciences. This

analysis focuses on the scientific domains “Physics & Engineering” (PE) and “Life Sciences” (LS). There are ten (nine) main research fields in PE: PE1_x to PE10_y (LS: LS1_x to LS9_x) and about 170 (100) subfields.
Table 1 gives is an example of key words of Life Science Panel LS1. The applicant could allocate the proposal to a total of four different panel keywords on the third level (e.g. LS1_5 “Protein synthesis, modification and turnover”). This information was used to calculate a rough interdisciplinary indicator (indicator 1).

Table1: Panel Keywords in the Life Science’s Panel LS1, Example of Panel LS1 - Molecular, cellular and developmental biology: molecular biology, biochemistry, biophysics, structural biology, cell biology, cell physiology, signal transduction and pattern formation in plants and animals

panel	**panel keyword**
LS1_1	Molecular biology and interactions
LS1_2	General biochemistry and metabolism
LS1_3	Nucleic acid biosynthesis, modification and degradation
LS1_4	RNA processing and modification
LS1_5	Protein synthesis, modification and turnover
…	…

The third domain “Social Sciences & Humanities (SSH)” is excluded as it is expected to differ in terms of publishing, citation behaviour, and other features from those observed in PE and LS (e.g., national/regional orientation, less publications in form of articles, different theoretical ‘development rate’, number of authors, non-scholarly publications), which make it less assessable for approaches developed for natural and the life sciences (Nederhof 2006; Juznic et al. 2010).

Preprocessing of data
The indicator 1 was calculated on the level of panels (ie: PEx). The data was electronically available as shown in table 2. The example of a proposal was assigned to the main Level LS5: “Neurosciences and neural disorders: neurobiology, neuroanatomy, neurophysiology, neurochemistry, neuropharmacology, neuroimaging, systems neuroscience, neurological disorders, psychiatry”. The proposed work was about surface interaction of biobased nanocrystals and addressed the following panel keywords: LS5.10: “Neuroimaging and computational neuroscience”, PE4.5: “Surface science”, PE5.2 “Polymer chemistry” and PE3.4 “ Transport properties of condensed matter”. We extracted all characters up to “.” with the following following results: LS5; PE4, PE5 and PE3. At the end we obtained four different panels including the main panel. The indicator 1 value for this proposal was: I_i = (4-1)/4=0.75.

Table 2. Proposal i with main panel PE5 and four panel keywords.

proposal	**main panel**	**panel keyword 1**	**panel keyword 2**	**panel keyword 3**	**panel keyword 4**
i	LS5	LS5.10	PE4.5	PE5.21	PE3.4

The calculation of indicator 2 started with the generation of disciplinary distinctive keywords for proposals. We used the keyword extraction feature of the software BibTechMon^{TM} to extract phrasems (example: “gene expression”) from the “summary” field. The tool offered the usage of general stop word lists and the possibility to enter stop words individually.

Table 3: individually selected not disciplinary distinctive stop words

keyword	text frequency
project	1005
development	601
study	584
understanding	423
propose	418
aim	397
1	384
research	375
2	350
proposal	350
develop	349
role	333
work	314
g	281
provide	265
field	257
3	240
approach	226

It was used to eliminate often used keywords in the proposals that do not contribute to any meaningful disciplinary distinction. Examples of the first 18 of 3007 most often used keyword that are eliminated are shown in table 3.

The second step to prepare the calculation of indicator 2 was to calculate the TFIDF.

Table 4: $TFIDF_{ij}$ and frequencies for the keyword kw_i “gene expression”

panel j	panel description	$TFIDF_{ij}$	hkwij	H_j
LS2	Genetics, Genomics, Bioinformatics and Systems Biology	0.82	11	43
LS4	Physiology, Pathophysiology and Endocrinology	0.44	3	22
LS1	Molecular and Structural Biology and Biochemistry	0.39	4	33
LS5	Neurosciences and neural disorders	0.30	4	43
LS8	Evolutionary, population and environmental biology	0.29	3	33
LS3	Cellular and Developmental Biology	0.17	2	37
LS6	Immunity and infection	0.10	1	31
PE4	Physical and Analytical Chemical sciences	0.08	1	40
LS7	Diagnostic tools, therapies and public health	0.07	1	48
PE5	Materials and Synthesis	0.05	1	62

Table 4 shows an example for keyword i (kw_i): “gene expression” and the values to calculate the TFIDF for this keyword for different panels. The number N of all proposals was 785, the number of proposals nkw_i where kw_i occurs is 31. The number of proposals in panel j were kw_i occurs at least once (hkw_{ij}) is 11.

In a next step we assign the panel with the highest TFIDF to the keyword kw_i as the home panel. In our case the panel LS2 is allocated as home panel.

The allocation of panels and keywords work quite good as it is exemplary shown for two panels in table 5. The assignment of home panels to keywords results in 967 keywords for panel LS2 “Genetics, Genomics, Bioinformatics and Systems Biology” and 2058 keywords for PE6 “Computer Science and informatics”.

Table 5: Panels LS2 and PE 6 and the allocated “Home panel” keywords (occurring in more than 9 proposals

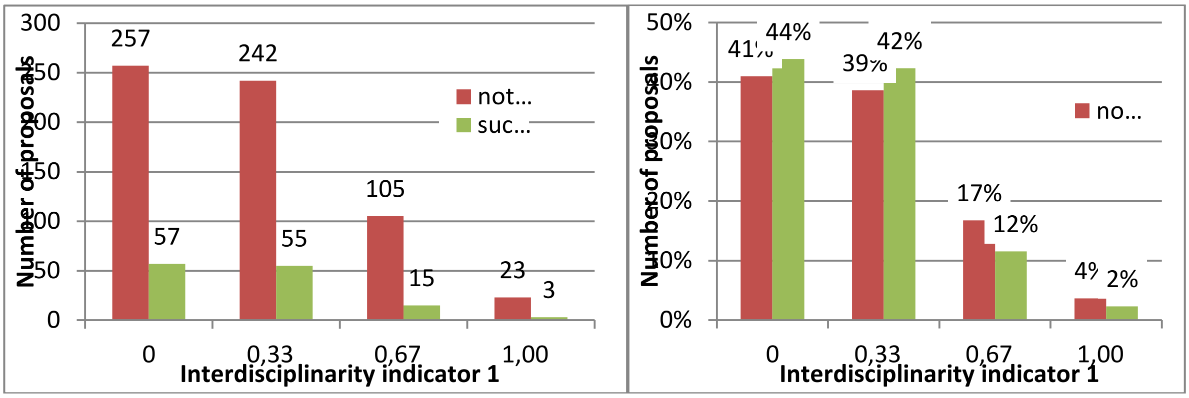
panel	**“home panel” keywords**
LS2 Genetics, Genomics, Bioinformatics and Systems Biology	genes, diseases, gene expression, pathways, genome, DNA, gene, Mutations, Tissue, biochemistry, variation, organism, prediction, metabolism,
PE6 Computer science and informatics	algorithms, simulation, Machine learning, Computer Science, task, platform, computation, internet, software, hardware

In a last step we calculate the indicator as percentage of the number of terms with other home panels than the home panel.

Results

The calculation of indicator 1 was based on the occurrence of the number of different panels normalized to 1. We obtained a number of 314 strict disciplinary (1 panel) proposals, 297 proposals with 2 different panels, 120 proposals with three different panels and just a few (26) high interdisciplinary proposals with four different panels, compare figure 1.

Figure 1 Distribution of successful and not successful proposals - absolute (on the left) and in percentage (on the right) of the number of successful resp. non successful proposals



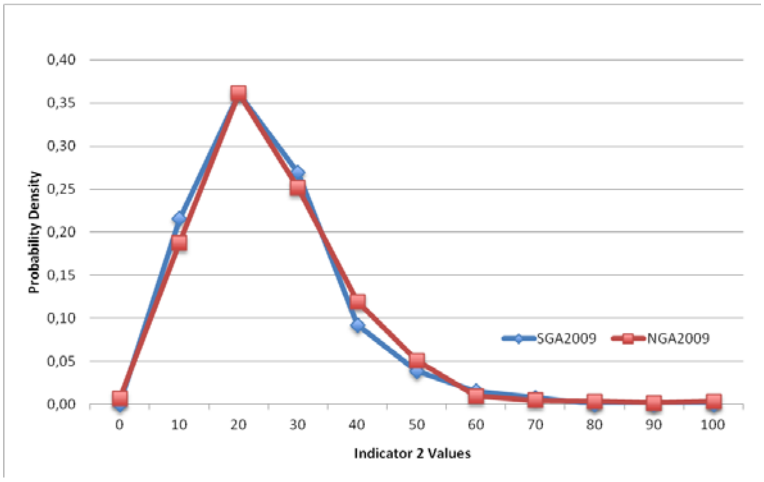
The question whether interdisciplinary proposals were more successful or not can be answered by percentage values for the indicator categories with the number of all successful respective all not successful proposals as 100%. It could be noticed that in both categories for low interdisciplinarity we obtained higher shares (3 percentage points) of successful proposals and of course for the higher categories of interdisciplinarity higher shares of not successful proposals: 5 resp. 2 percentage points.

It can be said that proposals with higher interdisciplinarity measured by the number of panel keywords are to a small extend less successful than proposals with more than two panel keywords different from the main panel.

The results for the calculation of the text based interdisciplinarity indicator 2 are shown in figure 2. The x-axis is defined by the indicator values and the y-axis by the probability density of the 130 successful and 628 not successful proposals.

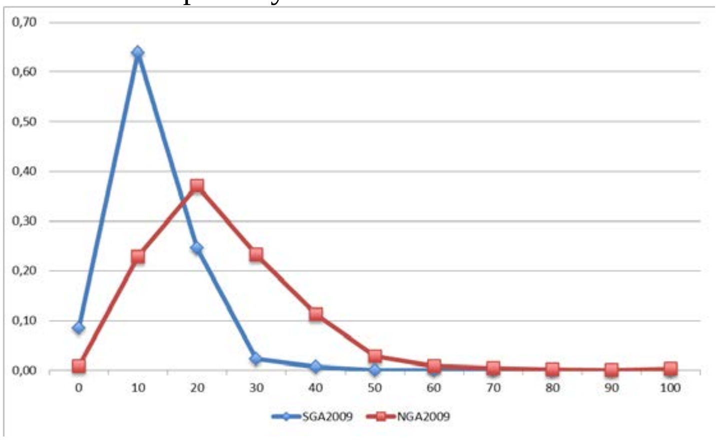
Figure 2: Probability density function for
successful and non-successful proposals for indicator 2

Both distributions indicated that most of the proposals had a range between 0 and 50 for the



indicator with a maximum of 20 %. This means that most proposals included 0-50 percent of keywords from other disciplines and that a very small number of proposals used more than 50 percent of keywords from other disciplines. The distribution of not successful proposals had a very small shift to higher interdisciplinary values in comparison to the distribution of successful proposals.

Figure 3: Probability density function for successful and not-successful proposals for indicator 2 calculated separately for successful and not successful proposals



Due to the results of indicator 1 and the experience of the history in the outcome of the review process we expected to have a higher shift between the two distributions. A reason for the lack of a shift could be that the characteristics of the non-successful proposals that are much more in numbers dominate the whole set. Therefore we calculate the TFIDF and the indicator 2 separately for SGA2009 and NGA2009. We divide the data set in successful and not successful proposals and apply the above described method for the indicator 2 twice. The two curves have a much higher displacement, see figure 3. In a statistical sense not successful proposals have a higher share of keywords from other disciplines than successful proposals.

Conclusions

The two versions of interdisciplinarity indicators pick up the interdisciplinarity aspects of frontier research, and may serve as useful input in an ex post and ex ante evaluation context of grant proposals or peer-review processes.

Our proposed approach builds on titles, abstracts and summaries of the proposals and on the panels as reference to disciplines. Both indicators can be calculated straight forward. The data is electronically available in a machine readable format and no further information was needed from other data sources or concepts. It has the advantage that it uses data of the submitted proposals and not historical information or definitions. The community of scientist deliver actual keywords for their disciplines via their proposals and we just use the probability of occurrence in different panels to measure the interdisciplinary character.

It could be shown that for the indicator interdisciplinarity, the peer review panels had a tendency to select projects that were more of disciplinary than interdisciplinary nature. However, this result is not surprising as it confirms earlier experiences from the ERC. As interdisciplinarity seems to have a negative effect on a proposals selection probability, respective measures may be implemented by the ERC to address this problem. The model could be used in future evaluation studies to indicate any improvement for a higher success rate of interdisciplinary proposals

There are also some weaknesses in the concept. We used the panels for the definition of interdisciplinarity. The definition of panels is not strong disciplinary.

With the exception of predefined “ERC stop words” we used all extracted keywords from proposals for the calculation of indicator 2. The relevant keywords for the assignment of home panels were selected by the TFIDF on the panel level. The assignment of a home panels to keywords that were relevant for different disciplines (like “cell”) do not really indicated interdisciplinary usage and were not be taken into account.

Another point that could affect the indicator 2 values was the number of keywords that were extracted from one proposal. The probability to use more home panel keywords from other panels could be higher if there is a longer text.
The indicator 2 values just indicate interdisciplinarity in a statistical sense. The application for individual proposals needs some verification and further work:

a. Consistent definition of panels and panel keywords
b. Selection of disciplinary specific keywords by improving the ERC stop word list.
c. A test phase with a verification of the interdisciplinary character of single proposals based on assigned keywords and home panels in comparison to the content of the proposal followed by the improvement of the calculation of the indicator.

References

ERC (2014), DBF - Development and Verification of a Bibliometric Model for the Identification of Frontier Research; Synthesis Report from: http://erc.europa.eu/sites/default/files/document/file/DBF_final_report.pdf

EUROPEAN COMMISSION (2005) Frontier research: The European Challenge High Level Expert Group Report

Juznic P., Peclin S., Zaucer M., Mandelj T., Pusnik M., Demsar F., (2010) Scientometric indicators: peer-review, bibliometric methods and conflict of interest, *Scientometrics*, 85, pp. 429-441

Nederhof A.J., (2006) Bibliometric monitoring of research performance in the Social Sciences and the Humanities: A review. *Scientometrics*, 66, 81-100

Schiebel E., Hörlesberger M., Roche I., Francois C., Besagni D. (2010). An advanced diffusion model to identify emergent research issues: the case of optoelectronic devices, *Scientometrics,* 83(3), 765-781

Wagner, C.S., Roessner, J.D., Bobb, K., Klein, J.T., Boyack, K.W., Keyton, J., Rafols, I., and Borner, K. (2011), Approaches to understanding and measuring interdisciplinary scientific research (IDR): A review of the literature, *Journal of Informetrics*, 5(1), 14-26.

What happens when funding is linked to (differentiated) publication counts?
New insights from an evaluation of the Norwegian Publication Indicator

Jesper W. Schneider, Kaare Aagaard and Carter W. Bloch

jws@cfa.au.dk, ka@cfa.au.dk; carter.bloch@cfa.au.dk;
Danish Centre for Studies in Research & Research Policy, Department of Political Science & Government, Aarhus University, Bartholins Allé 7, Aarhus, DK-8000, Denmark

Introduction

The first part of the title is inspired by Butler (2004). In a succession of papers she demonstrated how researchers in Australia responded when funding, at least partially, was linked to productivity measures undifferentiated by any measure of "quality" in the early 1990s (Butler, 2002; 2003a; 2003b; 2004). Australian publication output increased considerably with the highest percentage increase in lower impact journals. For a consecutive number of years, this lead to a general drop in overall citation impact for Australia. The pattern was visible across all research fields but only for sectors covered by the funding model. The patterns, were not, however, uniform across all institutions affected by the model. Since Butler's documentation of the adverse effects, the experience from Australia has stood as a warning for what would most likely happen if funding was linked to publication activity. Nevertheless, in the early 2000s a so-called "quality reform" of the higher education sector in Norway included a performance-based model where publication activity again was linked to funding. The main political purpose with the model was in fact to encourage more research activity (and thereby more publication activity) both at universities and university colleges, and preferably more international publication activity (UHR, 2004). The indicator was first used to distribute funds to universities and colleges in 2006 and in 2008 the system was expanded with a common database and classification system for the university and university college sector, health care organizations and the institute sector. The performance-based indicator redistributes approximately two percent of the annual funding among the institutions in the higher education sector (UH-sector).

Obviously, the designers of the Norwegian indicator were well-aware of the adverse behavioural effects documented in the Australian case. As a consequence, a slightly more sophisticated model was developed (Schneider, 2009; Sivertsen, 2010). A primacy in the model was to reflect the encouragement to publish in international outlets (i.e., international journals and academic book publishers) and at the same time to counter so-called "perverse" publication effects, where researchers seek to publish more but with least effort, i.e. least-publishable units and/or in outlets with high acceptance rates and meagre peer review measures. Hence, a differentiated publication model was constructed where publication channels are classified in two levels. Level one comprises in principle all scholarly eligible publication channels, where eligibility criteria are some basic norms such as a standard external peer review process. Level two, is an exclusive number of publication channels which are deemed to be leading in a field and preferably with an international audience. Level two is exclusive in as much as the number of publication channels designated at any given time to this level should produce roughly one-fifth of the publications produced in a field. Publication channels and nominations to level two are carried out annually in a number of specific subject committees; at the moment there are close to 70 such committees. Publication channels are only treated by one committee and thus have one classification. Correspondingly, three different types of scholarly publications are included in the model: journal publications (articles and reviews), book chapters (contributions to anthologies and conference papers) and books. A point system is then implemented where the different publication types yield different points and the point values also differ according to the level of the publication channel. Hence, the basic idea behind this two-tiered classification system is that publications on level two receive more publication points than publications on level one. Finally, publication points are fractioned 1/n so that an institution eventually receives 1/n points depending on their number of contributing authors. Eventually the annual sum of publication points for an institution is exchanged for funds, where the exchange rate is determined by the total number of publication points in the system in a given year; notice all fields are included, thus a level one journal article with one author is worth the same in physics and literature studies.

It is assumed that this weighted or differentiated point system to some extent will discourage researchers to speculate in "easy publications" resulting in a levelling out effect at the aggregate level, where a situation like the one in Australia is avoided. Consequently, contrary to the Australian model, the Norwegian one has implemented differentiated publication counts. The question we ask is therefore: "what happens when funding is linked to differentiated publication counts?" This question is important as it can further help our understanding of performance-based funding and especially models based on publication activity. Hitherto, we only have evidence from the Australian case. The Norwegian indicator is an interesting case for several reasons. First, as mentioned above, the indicator seeks to avert the deficiencies experienced in Australia by using a differentiated model. Second, the indicator has been in place for almost a decade which enables longitudinal analyses of its potentially lagged effects. Third, the Norwegian model has recently been "adopted" in several European countries and also used here as a national performance-based publication indicator (Hicks, 2012). Notice, this "adoption" has happened basically without empirical knowledge or evidence about the potential effects of the model in Norway as no large-scale evaluation of the model existed at the time.

This paper draws on data and results from a recent evaluation of the Norwegian Publication Indicator, carried out in the autumn of 2013 (Aagaard et al., 2014). The evaluation was commissioned by the Norwegian Association of Higher Education Institutions (UHR) and the main purpose was to examine whether the objective of the indicator has been met: whether the indicator has stimulated more research and research of higher "quality"? A number of different analyses based on large-scale surveys among researchers and leaders, case studies among different institutes, and national and international bibliometric data has been used in a mixed-method design (Bloch et al., 2014) to examine:

- the effects of the indicator on general publication patterns and individual publication behaviour, including the ability of the differentiated model to balance the publication activity between the different levels of publication channels;
- the properties of the indicator such as the perceived neutrality of publication patterns across fields and the quality of the reported data;
- the administrative organization and functioning of the indicator such as the nomination process to level two and its transparency;

- the use of the indicator for purposes other than the budgetary allocation of funds between institutions, including the use of the indicator at lower institutional and individual levels for recruitment, assessment of staff qualifications etc.

The present study addresses the narrow question to what extent the differentiated publication indicator seems to have prevented a situation as the one experienced in Australia where a collective change in publication behaviour resulted in more publication activity but lower national impact. We examine this primarily with aggregated time series of bibliometric data for Norway supplemented with insights from other analyses reported in the evaluation (Aagaard et al., 2014). Given the limited space, we can only present some major findings.

The next section briefly presents the data and main methods and indicators used for the analyses. The subsequent section presents some main results, and the final section contains the conclusions and a brief reflection upon the results.

Data and methods

The results presented below are based primarily on bibliometric data, partly from the publication database supporting the Norwegian indicator and partly from CWTS' *Web of Science* (WoS) citation database. Further, to supplement these time series we include some insights from the surveys in the discussion to shed more light on individual behavioural issues. In order to document the development in overall publication activity we use publication counts and publication points from the Norwegian documentation system. Notice, journal articles and articles in books counts as one publication whereas a book counts as 5 publication equivalents. Publications points reflect the two-tiered classification system[1]. In order to document the international visibility of Norwegian journal publications we use WoS and to contextualize the development we compare Norway to developments in the database and three comparable countries: Finland, Sweden and Denmark.

In order to examine potential changes in overall publication behaviour, we use the mean normalized journal score indicator (MNJS), which measures the impact of the journals in which Norwegian researchers has published (Waltman et al., 2012). This is different from Butler (2002; 2003a; 2003b); she examined changes in publication behaviour by dividing journals into quartiles based on their impact factor and subsequently calculated the development in Australian shares of publications in the different quartiles. The MNJS is basically a field normalized journal impact indicator that examines the impact of the portfolio of journals a unit has published in. The interpretation of the MNJS indicator is the following, if a unit has an MNJS indicator of one, this means that on average the unit has published in journals that are cited equally frequently as would be expected based on their field. An MNJS indicator of, for instance, two means that on average a unit has published in journals that are cited twice as frequently as would be expected based on their field citation activity.

Finally, in order to examine the development in Norwegian impact we use two well-known indicators: mean normalized citation score (MNCS) and the proportion of publications among the 10 percent most highly cited in the database (PPtop10%). The MNCS is a field normalized average citation score and the PPtop10% is a non-parametric percentile indicator

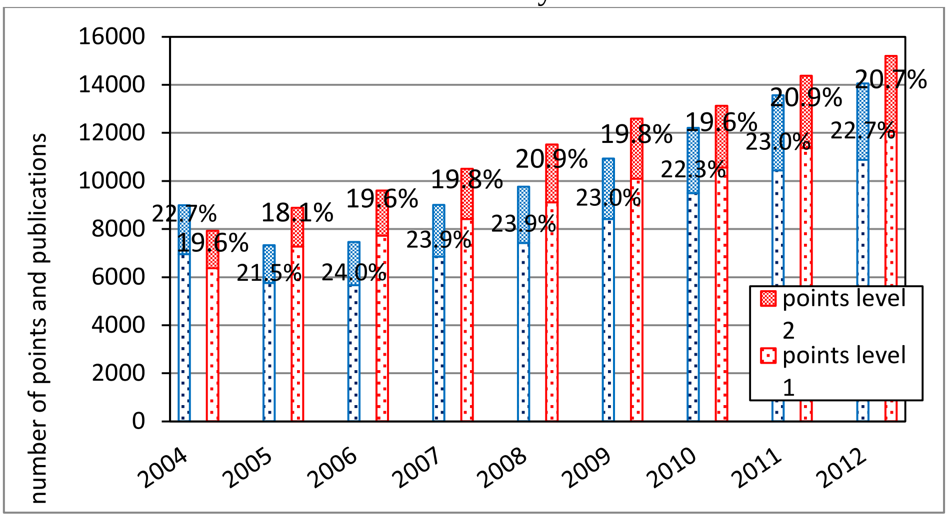
[1] Publications are classified by two levels and three types of publications. Level 1 and 2 journal articles yield 1 and 3 points, respectively, conference papers and contributions to anthologies 0.7 and 1 points, and 5 and 8 points for monographs.

(see Waltman et al., 2014). The next section presents the main results of the macro time series.

Results

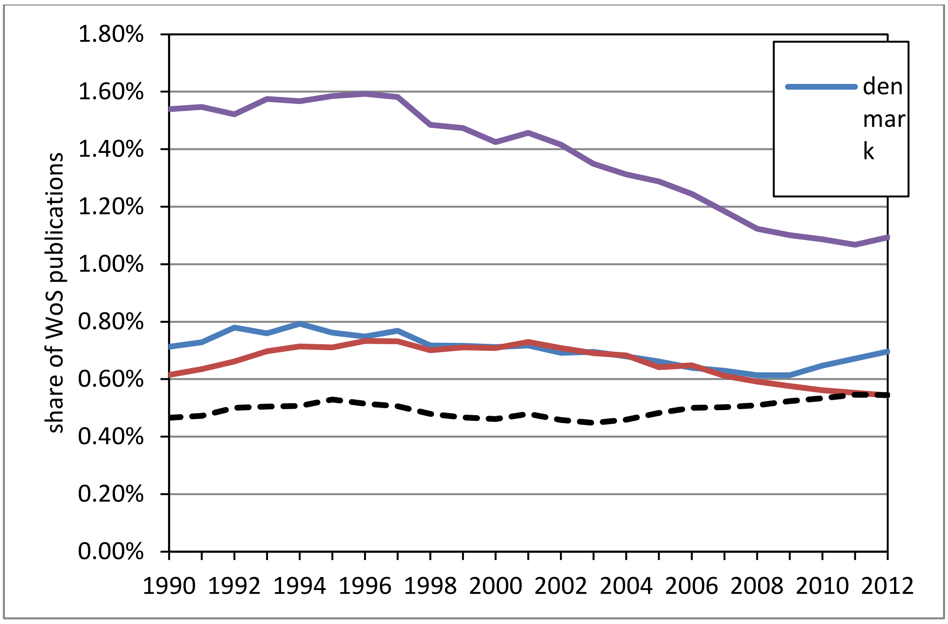
First we present the development in publication counts and points from 2004 to 2012 in Figure 1. Notice, the first publication year used to allocate funds was 2005. Total publication points for the UH-sector have risen from 8,327 in 2004 to 15,189 in 2012, which amounts to 82% increase over the period. The trend in publication counts is very similar. As we document in the evaluation, this dramatic rise in activity is to be explained by several factors including a substantial input of economic and human resources to the UH-sector, but also the actual indicator itself (Aagaard et al., 2014). Also visible from Figure 1 is the very stable distributions of level 1 and level 2 publications and points over time. Publication points in level 2 channels have remained very stable around 20% of total points throughout the period. These figures and their synchronous development are a first indication of stability rather than adverse effects at the aggregate level. Nevertheless, changes in numbers of and nominations to level 2 channels during the period can have influenced the share of level 2 publications, potentially blurring fluctuations in publication behaviour.

Figure 1. Development in publication counts and points; data from the Norwegian indicator's documentation system.



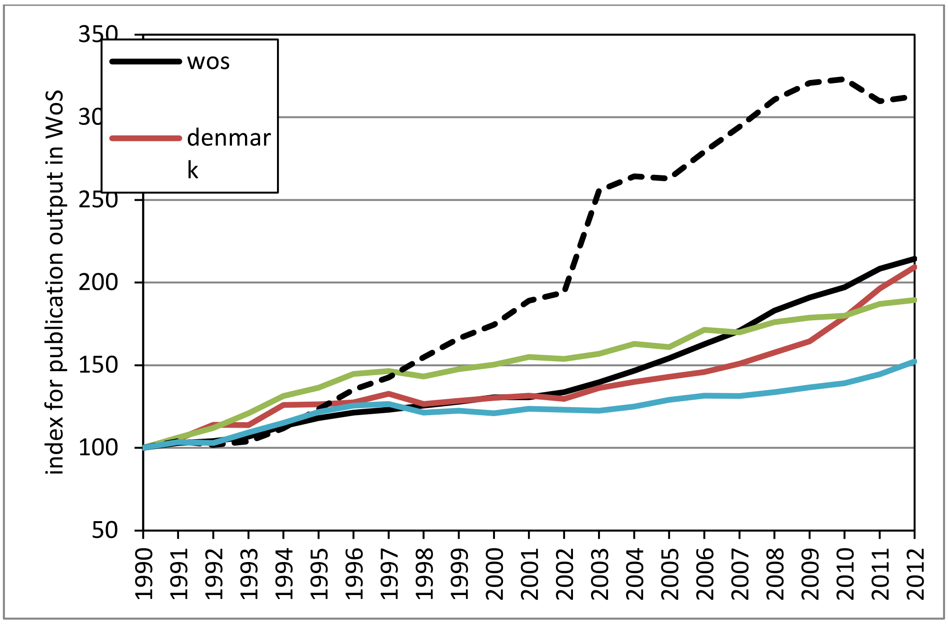
To examine the international visibility in more detail, Figure 2 shows the development in share of WoS publications for Norway and the three benchmark countries. Notice, in 1995 Norway's share peaks for the first time around 0.5%, then we see a drop and a subsequent continuous rise after 2004 and the implementation of the indicator.

Figure 2. Development in the share of total publications in the WoS databse for Norway and benchmark countries.



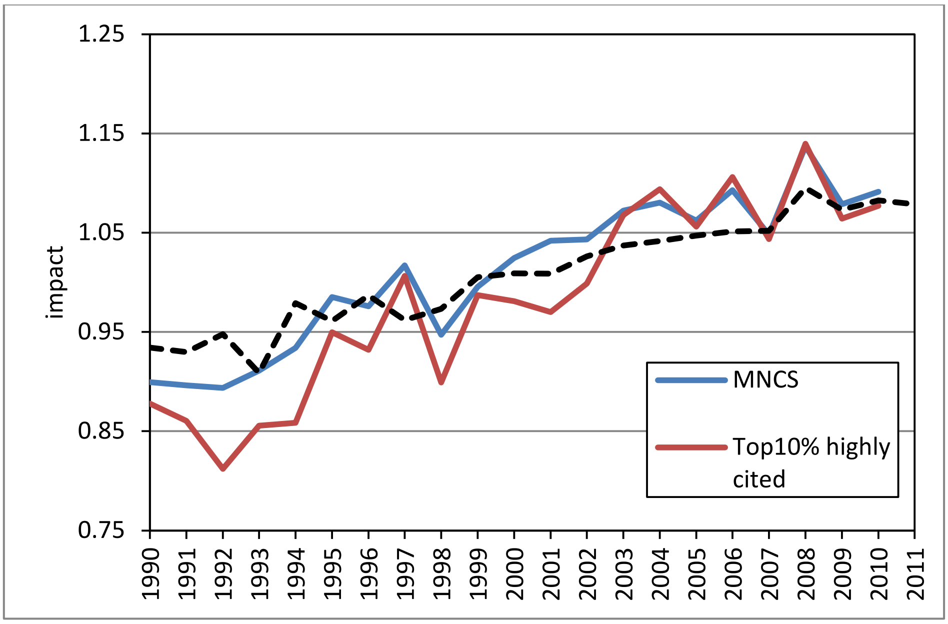
The rise in Norwegian publication activity in international journals indexed by WoS becomes more dramatic if we examine the development as index where index 100 is the publication output in 1990 shown in Figure 3. Whereas the benchmark countries, more or less, relatively follow the general development in the WoS database, the Norwegian development is quite different. From 1990 to 2003 there is a substantial rise above the general rise in the database. However, in 2003 we see a steepening rise bringing Norwegian publication activity beyond index 300 from 2007 onwards.

Figure 3. Developments in indexed publication activities for Norway, the WoS database and benchmark countries.



It is evident that the political intention behind the model of more publication activity and preferably more international publication activity seems to have been achieved. The indicator seems to have stimulated to more research activity, as was the case in Australia; the important question therefore is whether the marked rise in publication activity also has resulted in lower national impact as in the case of Australia.
Figure 4 presents three indicators that reflect this issue: the journal indicator MNJS, and the two article-level indicators MNJS and PPtop10%. Notice, the citation impact of Norway, both MNCS and PPtop10% has seen a more or less continuous rise albeit a slow one since 1990. Nevertheless, the current Norwegian impact level would still be considered meagre in a Nordic and west-European context (Nordforsk, 2010). But the important thing here is that we cannot see any trace of a general drop in Norwegian national impact. If we at the same time examine the development in the MNJS-indicator, which tells us something about the overall publication behaviour of Norwegian researchers, i.e., at what journal impact level are they publishing?, then we can see that the indicator basically follows the same trajectory. In general, there has been a continuous rise in the indicator, meaning that each year (with few exceptions) on the aggregate level, Norwegian researchers have “improved” their publication profile by publishing in journals with higher impact levels. However, the MNJS level both before and after the introduction of the performance-based model must also be considered fairly low.

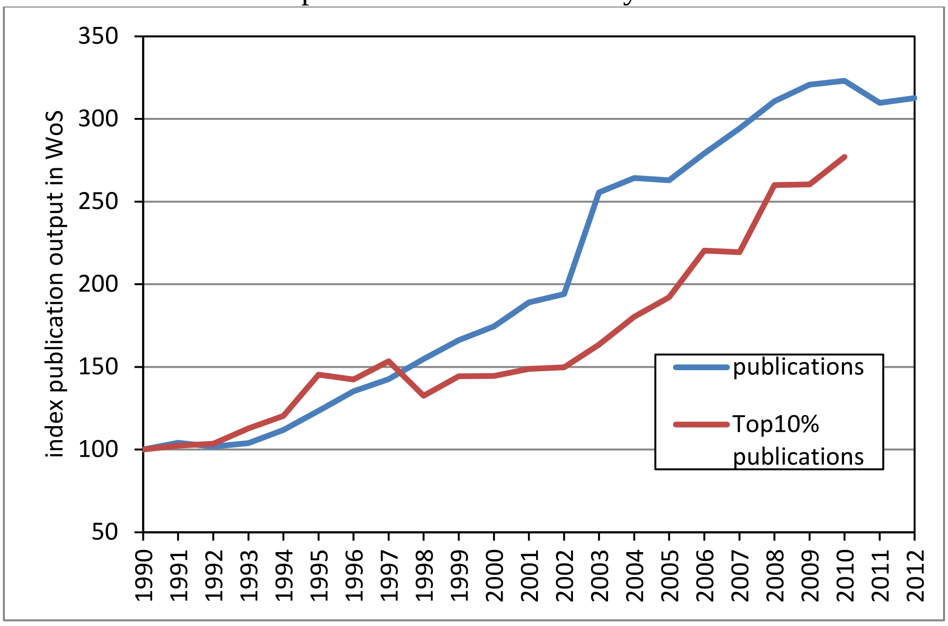
Figure 4. Citation impact for Norway measured by the MNJS, MNCS and PPtop10% indicators.



But again, the important point here is that on the aggregate level there is no indication of a marked drop and thus an indication of more publication activity in lower impact journals. What we cannot see from these figures, however, is to what extent variation in publication activity among low, medium and high impact journals contribute to this rather stable MNJS-indicator. Indeed, we cannot rule out considerable changes in individual publication behaviour, as indeed our survey results may indicate, but for the time being there is no indication of this on the aggregate level.

If we recall the main purpose of the evaluation: to examine whether the indicator has stimulated more research and research of higher "quality" then we can establish that the first element seems to have been achieved. What is referred to as "quality" is in fact (just) impact and here it is not so evident that the Norwegian indicator has stimulated to higher impact. Impact has risen, but the trend was well under way before the implementation of the indicator. Nevertheless, as Figure 5 demonstrates, around 2002 just before the indicator was implemented, a marked rise is seen in publication output, but at the same time we also see a marked rise in the number of Norwegian journal publications among the most highly cited. In other words, Norway manages to both increase its general output and its number of highly cited articles resulting in the stable relative impact levels. It should also be pointed out that the Norwegian indicator was not explicitly designed to increase impact, but instead to avoid deterioration in impact (i.e., publication in lower impact journals).

Figure 5. Indexed development in the number of publications and number of highly cited publications from Norway.



Discussion

What happens when funding is linked to differentiated publication counts? In the Norwegian case, as expected publication activity goes up but impact remains stable. This is a different experience compared to Australia where impact eventually dropped. In this study, we can only determine that on an aggregate level there seems to be no collective adverse effects in publication behaviour. However, as our evaluation suggests, the use of the indicator at lower levels for purposes such as hiring, promotion or salary may indeed lead to changes in individual publication behaviour, but at this time such potential changes are not visible on the aggregate level. Why does the differentiated publication counts in Norway result in overall stability in impact and collective publication behaviour. A tempting answer is the two-tiered classification system and thus the differentiated counts themselves. However, we have no evidence what so ever of such a causal claim for the time being. Also, at the moment we are not able to establish to what extent the indicator contributes to the rise in output in as much as it has encourage researchers already in the system to produce more publications; future analyses of individual level publication data will shed light on this. What we can establish so far is that the experience in Norway with a differentiated publication indicator linked to funding is different from the experience in Australia with undifferentiated indicator. This is an important observation because currently the Norwegian model is adopted in several European countries.

References

Aagaard, K., Bloch, C., Schneider, J.W., Henriksen, D., Lauridsen, P.S. and Ryan, T.K. (2014). *Evaluation of the Norwegian Publication Indicator* (in Danish). Oslo: Norwegian Association of Higher Education Institutions.

Bloch, C. W., Sørensen, M. P., Graversen, E., Krogh; Schneider, J. W., Kalpazidou Schmidt, E., Aagaard, K. & Mejlgaard, N. (2014). Developing a methodology to assess the impact of research grant funding – a mixed methods approach. *Evaluation and Program Planning*, vol. 43, p. 105–117.

Butler, L. (2002). A list of published papers is no measure of value - The present system rewards quantity, not quality - but hasty changes could be as bad. *Nature*, 419(6910), 877.

Butler, L. (2003a). Explaining Australia's increased share of ISI publications - the effects of a funding formula based on publication counts. *Research Policy*, 32(1), 143-155.

Butler, L. (2003b). Modifying publication practices in response to funding formulas. *Research Evaluation*, 12(1), 39-46.

Butler, L. (2004). What happens when funding is linked to publication counts? In: *Handbook of Quantitative Science and Technology Research*. Eds. H. Moed et al. Kluwer Academic Publishers, Netherlands, pp. 389-405.

Hicks, D. (2012). Performance-based university research funding systems. *Research Policy*, Volume 41, Issue 2, March 2012, Pages 251–261.

Nordforsk (2010). *Bibliometric Research* Performance *Indicators for the Nordic Countries*. Editor: J. W. Schneider. Nordforsk report: http://www.nordforsk.org/no/publikasjoner/bibliometric-research-performance-indicators-for-the-nordic-countries.

Norwegian Association of Higher Education Institutions (UHR) (2004). A Bibliometric Model for Performance-based Budgeting of Research Institutions. Oslo. http://www.uhr.no/documents/Rapport_fra_UHR_prosjektet_4_11_engCJS_endelig_versjon_av_hele_oversettelsen.pdf.

Schneider, J.W. (2009). An outline of the bibliometric indicator used for performance-based funding of research institutions in Norway. *European Political Science*, 8 (3), 364-378.

Sivertsen, Gunnar (2010). A performance indicator based on complete data for the scientific publication output at research institutions. ISSI Newsletter (International Society for Scientometrics and Informetrics), vol. 6, no. 1, p. 22-28

Waltman, L., Calero-Medina, C., Noyons, E.C.M, Tijssen, R.J.W., van Eck, N.J., van Leeuwen, T.N., van Raan, A.F.J., Visser, M, and Wouters P. (2012) The Leiden ranking 2011/2012: Data collection, indicators, and interpretation. *Journal of the American Society for Information Science and Technology*, 63, 12, 2419-2432.

Identifying potential ‘breakthrough’ research articles using refined citation analyses: Three explorative approaches

Jesper W. Schneider* and Rodrigo Costas**

*jws@cfa.au.dk
Danish Centre for Studies in Research & Research Policy, Department of Political Science & Government, Aarhus University, Bartholins Allé 7, Aarhus, DK-8000, Denmark

**rcostas@cwts.leidenuniv.nl
Centre for Science and Technology Studies (CWTS), Faculty of Social and Behavioral Sciences, Leiden University, PO Box 905, 2300 AX, Leiden (The Netherlands)

Introduction

‘Breakthrough’ research is a challenging concept that can be coupled with solutions to broad and complex research problems, challenging established theories and scientific paradigms, establishing fundamental new ways of using methods and instruments, or sometimes to the interdisciplinary integration of different research perspectives. The factors influencing breakthrough research have been studied intensively from several perspectives (e.g., Kuhn, 1970; Merton, 1973; Simonton, 1988; Heinze et al., 2007; Heinze, 2008; Öquist & Benner, 2012). There has been some interest in more predictive approaches in order to identify potential breakthrough research in its early stages (e.g., Simonton, 1988, 2004; Chen et al., 2009; Small & Klavans, 2011; Chen, 2012; Ponomarev et al., 2014). Citation data play a noticeable role in some of these approaches. Using citation data to identify or predict ‘excellent’ or ‘breakthrough’ research has been an aim for decades (e.g., Garfield & Welljamsdorof, 1992; Tijssen, Visser, & van Leeuwen, 2002). Common for most of these approaches is that identification and model building are typically retrospective in as much as excellent or breakthrough research is determined by other means than citation analyses and then from the citation patterns of these exemplars, comparisons and predictions are made. In general, very highly cited units are considered good predictors for prizes, awards and peer acknowledgement of excellence.

The aim of the present study is to identify potential breakthrough research by focusing on very highly cited articles. We present three related conceptions and approaches of identifying breakthrough publications. Contrary to previous studies, we perform these analyses on a large-scale clustered network of publications. This network is created through direct citation links, clustering publications according to their linkage with other articles (Waltman & van Eck, 2012).

Our approach for identifying potential breakthrough research starts from the following three assumptions:

1) Breakthrough research tends to become highly cited. However, a limitation here is that a paper can indeed report breakthrough research, but breakthrough research can also be the sum of knowledge claims in a number of papers, where some of them are perhaps not even highly cited. Initially, we assume that at least one publication should become highly cited.

2) A highly cited paper does not *per se* convey breakthrough research. Therefore focusing on highly cited papers solely is not enough, it is important to identify those that most likely convey breakthrough research and separate them from just highly cited articles that do not report true breakthrough results.
3) Finally, breakthrough research should be identified within its discourse community that is among a set of related and self-organized research articles (i.e. the network of publications).

In the present paper we concentrate on just presenting a citation-based methodology for identifying potential breakthrough research. Forthcoming work will discuss the challenging concept of breakthrough research, outline the methodology in more detail, and especially focus upon the evaluation of the citation-based methodology. The following section briefly outlines our operationalization of breakthrough research, introduces the citation-based methodology comprising the three refined citation analyses as well as the data set used for initial exploration. The results section briefly summarizes some of the overall results and we briefly wrap the study in the in the final section.

Data and methods

Firstly, we propose the following definition of a 'breakthrough paper' for this study: *a highly cited paper, with an important spread over its own field(s) and also other fields of science, and it must be a paper that is not a mere follower*[1] *of other highly cited publication(s) but that it has a genuine relevance on its own*. From this definition we focus on highly cited journal articles. Their identification is carried out in the context of the network and clustering of articles worldwide. Review papers are excluded as potential breakthroughs as they mostly condensate and discuss the most recent and important developments in a scientific domain, thus qualifying as 'followers' and not as true breakthroughs.

Three citation-based approaches and the network of clustered papers

We approach the detection of breakthrough papers from three different perspectives, thus also providing three different typologies of breakthrough papers. In all three cases we use a classification of all 16.2 million publications indexed in *Web of Science* (WoS) between 1993 and 2012 developed at CWTS (Waltman & van Eck, 2012). This network of publications is created by direct citation links, thus it is assumed that publications in the same cluster have common research interests. Publications are clustered at three levels: there are 21 macro-fields that represent main scientific disciplines. These macro-fields contain themselves 784 different meso-fields, and finally we have a micro-classification composed by 21,167 micro-fields. All these levels have been used in our methodology for detecting breakthroughs in one way or another.

The three approaches explored are characterized as follows:

- *Approach 1* is very simple but also extremely exclusive. It is based on the idea that the most cited paper of every micro-field can most likely be considered as a breakthrough paper because it has the highest impact in it its micro-domain. This is a very restrictive definition of a breakthrough paper, because only one (or occasionally several) papers pass this filter. In fact, only 21,670 out of the 16.2 million publications pass this filter as breakthroughs (i.e., 0.13% of all publications).

[1] A follower paper is a paper citing another highly cited publication and benefiting form the impact of the first (e.g. a good example is the high impact of publications that have followed the h-index indicator proposal)

- *Approach 2* is based on two citations methods outlined below: 1) the 'Characteristics Scores and Scales' method (CSS) (Schubert, Glänzel & Braun, 1987) and 2) a filtering of 'followers'. This approach is relatively less restrictive compared to Approach 1 in as much as 179,349 out of the 16.2 million publications qualify as potential breakthrough papers (i.e., 1.1% of all publications).
- *Approach 3* is also based on the CSS-method and filtering of 'followers', but this approach is more restrictive than Approach 2 because it introduces a knowledge diffusion criterion where it is also required that a potential breakthrough publication has impact beyond its own macro field. Some 59,617 out of 16.2 million publications qualify as potential breakthroughs according to this approach (i.e., 0.37% of all publications). The knowledge diffusion filter is also outlined below.

Filtering methods: The CSS-method, filtering of 'followers' and knowledge diffusion

The CSS-method. For Approach 2 and 3, we use the CSS-method suggested by Schubert, Glänzel and Braun (1987). The CSS method focuses on the common characteristics of citation distributions across fields and is based on the principle that citation distributions share some fundamental characteristics and similarities. The CSS method basically consists of the reduction of the original citation distribution to 'self-adjusting' classes by iteratively truncating the distribution to conditional mean values from the low end up to the high end. In the present study we end up with four typologies to which we assign publications:

- Typology 1. Lowly cited publications: those that have an impact below the average of the entire field (m1). They are the vast majority of publications in every field representing around 74 percent of all the publications and accounting for approximately 22 percent of all citations.
- Typology 2. Moderately cited publications: those that have an impact above the average of the entire field (m1) but below the second mean (m2). These publications represent approximately 19 percent of all the publications in their fields and receive 32 percent of all the citations in the field.
- Typology 3. Highly cited publications: these are publications that have an impact higher than m2 but below m3. They constitute approximately 5 percent of all publications within each meso-field and receive more than 21 percent of the citations in their respective fields.
- Typology 4. Outstanding publications. These publications represent barely 2 percent of all publications in every meso-field, but they alone receive around 25 percent of all citations in their meso-fields. In other words, these are the 2 percent most cited publications of every field and one in four citations given in their meso-fields goes to them.

There is a remarkable regularity across the fields of science and across the meso-fields in this study. This is very useful for our purposes as it allows us to apply the same approach across fields when we characterize the 'success' and 'dissemination' of the impact of these publications. Based on this methodology, the 16.2 million WoS publications from 1993-2012 and that has a meso-field in the CWTS classification have been classified (in this case both articles and reviews). 314,944 (1.9%) publications belong to type 4 (i.e. outstanding publications), of which 263,148 are of the document type 'article' (1.7% of all articles in the period). All these 263,148 publications have been considered as potential breakthroughs.

Filtering of 'followers'. Being highly cited is in itself not sufficient to be considered a breakthrough, because publications should not be "a mere follower of other highly cited publication(s)" – it must have "a genuine relevance on its own"! To operationalize this filtering of the 'followers' we have performed the following steps:

- Identification of all pairs of potential breakthrough papers. Basically, we have identified potential breakthroughs citing another potential breakthrough(s). If we find such linkages, we label the citing breakthrough as B2 and the cited breakthrough as B1. Thus B2-papers are potential 'followers'.
- We then analyse the papers that cite B2 and check if they also cite B1, if so, we count these papers as double citers of B1 and B2.
- Finally, for B2 publications, we count how many of their citing papers that either simultaneously cites both B1 and B2 or only B2. Subsequently we enforce a threshold to filter 'followers'. For every B2 paper we filter out those that do not have 70% or more of its citations alone (i.e. they are co-cited with B1 in more than 30% of its citations). The main idea behind this strong threshold is that a breakthrough should not benefit too much from the 'spell' of a previous breakthrough and it should have a genuine impact on its own.

We have applied this filter to the 263,148 outstanding articles previously detected and 179,347 passed the followers filter, thus being considered as potential breakthrough candidates (they are the basis for Approach 2).

Knowledge diffusion filter. Based on the 179,349 publications considered for Approach 2, we have included a new filter for determining breakthroughs. Here we introduce a knowledge diffusion criterion in the delineation of breakthroughs enforcing that breakthroughs also must have impact beyond their own macro-domains (i.e. they must have impact across other major fields of Science). To operationalize this we followed a relatively simple approach:

- Taking all the citers of the 179,349 publications previously filtered, we counted the number of different macro-categories (i.e. a total of 21) from which they have received at least one citation.
- We then calculate the average of different external macro-fields where the breakthroughs of every meso-category have had some impact.
- Thus, based on the previous values, we consider a breakthrough those publications within the same meso-category with an impact outside their own macro-field higher than the average of all the potential breakthroughs in the same meso-category.

A potential breakthrough according to this third approach is potential breakthrough papers that have an impact in more macro-categories than an average potential breakthrough within the same meso-category.

Data

The three citation-based approaches are explored retrospectively in a case study based on publications from 66 Centres of Excellence (CoE) funded by the Danish National Research Foundation (DNRF) (Schneider & Costas, 2013). Notice, the foundation was explicitly set up to identify and fund potentially excellent or breakthrough research. It is therefore expected that many of these CoE will eventually produce excellent or breakthrough research. Parallel to, but independent of our approach, the DNRF selected a sample of *eight* CoE considered to have produced breakthrough research in the period. The sample is therefore used as an initial 'golden standard' against which to compare the results or our methodology.

Results

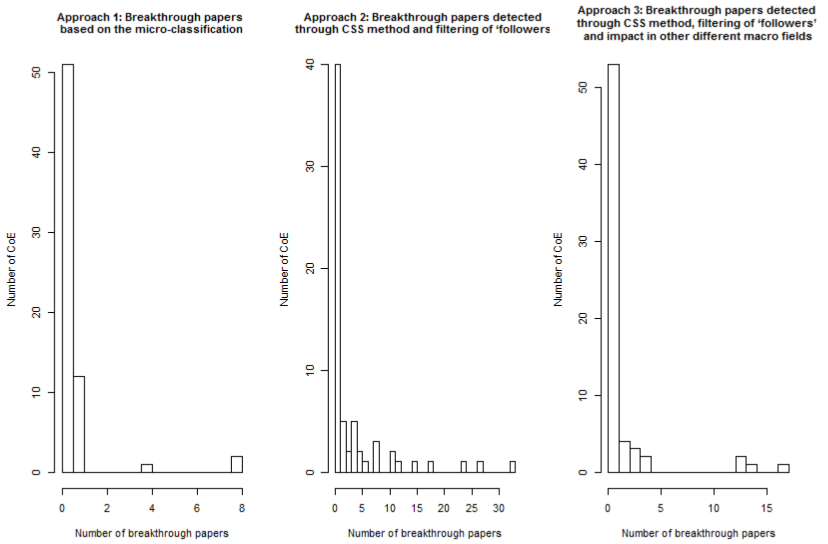
Due to the limited space, we are only able to present some general results. Table 1 presents the overall results for the three approaches. The table focuses upon the overall number of potential breakthrough papers identified and the number of CoE producing these papers.

Table 1. Performance of the 3 approaches with the DNRF CoE.

	No. of potential BRK-papers from all CoE funded by the DNRF	No. of CoE with potential BRK-publications	Percentage of BRK-papers from the total of all CoE papers (%)	No. of CoE from the sample of eight where BRK-papers are detected
Approach 1	32	15	0.30 (*0.13*)	4
Approach 2	241	40	2.23 (*1.1*)	6
Approach 3	97	27	0.90 (*0.37*)	5

Table 1 shows the different restrictiveness of the approaches when looking at the number of total potential breakthrough papers identified. It should be emphasized that the CoEs actually have produced more potential breakthrough papers than expected (in all approaches the share of CoEs breakthroughs is higher than the share of the database, in brackets in the 3rd column). But it is also noticeable that the distribution of potential breakthrough papers is skewed among the CoEs (Figure 1).

Figure 1. Distribution of potential breakthrough papers across CoEs.



We can see that potential breakthrough papers are concentrated on relatively few CoEs and interestingly, but not visible in Figure 1, it is the same four CoEs that produce the majority of them in all three approaches. These CoEs include two working in bioinformatics, one in nanoscience and one with register-based epidemiological research. If we compare the results to the eight breakthrough cases expected by the DNRF, then we see that four CoE of them are identified in all three approaches; five of them are detected in two or more of the approaches, and six of them are detected by any of the approaches. Consequently, for two examples we did not identify any potential breakthrough papers. One of them is still active and still recent and its publication activity for the period under investigation is limited, hence not identifying potential breakthrough papers in this case might be an effect of time. This explanation does not hold for the other case, and simply no papers from this particular CoE qualify as a potential breakthrough papers according to our approaches. This means that either our method fails in identifying potential breakthrough papers from this CoE or the CoE is not a suitable benchmark?

A positive outcome of our methodology however is the fact that four of the expected breakthrough CoEs are detected by the three approaches. Also the fact that other CoEs have been found to have produced some potential breakthrough papers could be a nice opportunity to discuss the consideration of breakthrough publications and research within the DNRF.

Discussion

We have tried empirically to detect potential breakthrough papers assumed to be proxies for breakthrough research. Obviously, the approach has several limitations. Particularly we like to emphasize that we are only able to detect strong signals through citation analysis. As this signal becomes weaker we are not able to detect it and given our definition and operationalization not able to identify potential breakthrough papers. In this line it could be also argued that the approaches are too restrictive and therefore some other signals could go unnoticed. However, if the signal turns out to be strong, we will - other things being equal - detect it and this we have done in this study. We acknowledge however that other thresholds and definitions could have been applied and so forth (and we expect to do it in further research). However, within its limits, we have delineated our current approach with reasoned and sound arguments. The methodology is simple and replicable and there is consistency among the approaches and the results, especially among the strongest signals. The results are also consistent to the examples suggested by the DNRF. As a result, our approach could be incorporated in the bibliometric toolset of research organizations to identify potential breakthrough research by use of refined citation analyses; we consider the methodology an important extension to traditional citation analyses.

References

Chen, C. (2012). Predictive effects of structural variation on citation counts. Journal of the American Society for Information Science and Technology, 63(3), 431-449.

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.

Garfield, E., & Welljamsdorof, A. (1992). Of nobel class - a citation perspective on high-impact research authors. Theoretical Medicine, 13(2), 117-135.

Glänzel, W., Thijs, B., & Debackere, K. (2013). The application of citation-based performance classes in institutional research assessment. Paper presented at the Proceedings of 18th International Conference on Science and Technology Indicators, Berlin, Germany.

Heinze, T. (2008). How to sponsor ground-breaking research: A comparison of funding schemes. Science and Public Policy, 35(5), 302-318.

Heinze, T., Shapira, P., Senker, J., & Kuhlmann, S. (2007). Identifying creative research accomplishments: Methodology and results for nanotechnology and human genetics. Scientometrics, 70(1), 125-152.

Kuhn, T. (1970). *Stucture of* Scientific *Revolutions*. 2nd edition. The University of Chicago: Chicago, IL.

Merton. R.K. (1973). *The* sociology *of science: Theoretical and empirical investigations*. University of Chicago Press, Chicago, IL.

Öquist, G., & Benner, M. (2012). Fostering breakthrough research: A comparative study. In F. All (Ed.): Kungliga Vetenskapsakademien.

Ponomarev, I. V., Williams, D. E., Hackett, C. J., Schnell, J. D., & Haak, L. L. (2014). Predicting highly cited papers: A method for early detection of candidate breakthroughs. Technological Forecasting and Social Change, 81, 49-55.

Schneider, J. W., & Costas, R. (2013). Bibliometric analyses of publications from centres of excellence funded by the danish national research foundation. From http://fivu.dk/en/publications/2013/files-2013/appendiks-5_bibliometrisk_report_03122013.pdf.

Schubert, A., Glänzel, W., & Braun, T. (1987). Subject field characteristic citation scores and scales for assessing research performance. Scientometrics, 12(5), 267-291.

Simonton, D. K. (1988). Scientific genius: A psychology of science. Cambridge, UK: Cambridge University Press.

Simonton, D. K. (2004). Creativity in science: Chance, logic, genius, and zeitgeist. Cambridge, UK: Cambridge University Press.

Small, H., & Klavans, R. (2011). Identifying scientific breakthroughs by combining co-citation analysis and citation context. In E. Noyons, P. Ngulube & J. Leta (Eds.), Proceedings of issi 2011: The 13th conference of the international society for scientometrics and informetrics, vols 1 and 2 (pp. 783-793).

Tijssen, R., Visser, M., & van Leeuwen, T. (2002). Benchmarking international scientific excellence: Are highly cited research papers an appropriate frame of reference? Scientometrics, 54(3), 381-397.

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.

A double shift in researchers' activity profiles: an actor-based analysis of the making of quality in high standing academic departments[1]

Antoine Schoen *, Catherine Paradeise**, Lionel Cauchard***and Marianne Noël ****

*a.schoen@esiee.fr
Université Paris-Est, ESIEE, IFRIS, Cité Descartes
5, boulevard Descartes
Champs-sur-Marne
77454 MARNE-LA-VALLÉE Cedex 02 (FR)

** catherine.paradeise@univ-mlv.fr
Université Paris-Est, UPEM, IFRIS, Cité Descartes
5, boulevard Descartes
Champs-sur-Marne
77454 MARNE-LA-VALLÉE Cedex 02 (FR)

*** lionel.cauchard@gmail.com
Université Paris-Est, UPEM, IFRIS, Cité Descartes
5, boulevard Descartes
Champs-sur-Marne
77454 MARNE-LA-VALLÉE Cedex 02 (FR)

**** noel@ifris.org
Université Paris-Est, CNRS, IFRIS, Cité Descartes
5, boulevard Descartes
Champs-sur-Marne
77454 MARNE-LA-VALLÉE Cedex 02 (FR)

Objectives of the study

This study aims at opening the black box of high standing academic departments for analysing the combination of individual contributions in the production of a shared quality label. By analysing personal CVs, this study investigates three research questions:

- To what extent is academic quality the result of individual performances or a collective achievement?
- Can we characterise a specialisation of activities within the departments between « junior » and « senior » researchers?
- Is there a shift across cohorts in the academic activity profile during the early career phase?

Theoretical or conceptual frameworks

Excellence has become the motto of academic evaluation. The formalized measurement of excellence is largely based on analytical, impersonal and presumably non-contextual performance indicators that concentrate mostly on research. It contrasts sharply with substantial judgments of prestige that do not open the black box of academic production and are based on social circulation of idiosyncratic and synthetic judgments made by social networks.

[1] This work was supported by the PrestEnce project funded by ANR and the Université Paris Est Marne-la-Vallée (UPEM)

This research aims at investigating how do academic collective actors, at their own organizational level, use resources and constraints provided by their various environments (national, regional, academic organization, departmental and individual levels) to enhance their own reputation of quality.

The research work is based on a twofold comparative study that combines disciplinary variety (chemistry, management, history) and national (France, Italy, the USA) variety.

This study, that is part of the PrestEnce research project, is based on the cross-characterization of departments according to excellence as measured by the sum of individual contributions and the available (human, material, social) resources, both in terms of structural capital (positions in various disciplinary or institutional networks) and operational tools (of internal and external coordination) gathered by in-depth investigation.

Modes of inquiry

This study analyses researchers' activities through the networks they are involved in.

Two types of networks are mainly explored, which reflect respectively the department academic outreach (through publications, academic and not, WoS and not WoS) and the institutional outreach, through participation in journals' editorial boards - aka scientific gatekeeping - and involvement in external advisory bodies - aka external involvement.

The scientific networks, stemming from the researchers' publications, link individuals with other researchers (co-authors), academic journals (of publication) and institutions (based on co-authors affiliations). The second type of - hybrid - networks, link the department researchers with various institutions they are connected with either through their scientific gate keeping activities or their external involvement.

The department networks are built as the union of its members' ego network. Their topology will provide quantified information regarding the collective making of a department's quality.

Data sources

Table 1. 6 case studies
Globally: 119 researchers, 2891 publications over the 2003-2012 period.

	nb of persons	cumulated nb of years of publication	nb of all publi	nb of acplus publis	nb of ac publis	nb of nonac publis
Institution 1 Business Juniors (US)	10	85	176	110	55	11
Institution 1 Business Seniors (US)	11	100	154	45	76	33
Institution 1 Business Total (US)	21	185	330	155	131	44
Institution 1 History Juniors (US)	9	56	125	45	62	18
Institution 1 History Seniors (US)	9	90	168	54	91	23
Institution 1 History Total (US)	18	146	293	99	153	41
Institution 1 Chemistry Juniors (US)	4	35	249	155	93	1
Institution 1 Chemistry Seniors (US)	11	110	1094	262	829	3
Institution 1 Chemistry Total (US)	15	145	1343	417	922	4
Institution 2 Business Juniors (FR)	19	156	303	50	187	66
Institution 2 Business Seniors (FR)	9	90	163	17	102	44
Institution 2 Business Total (FR)	28	246	466	67	289	110
Institution 3 Business Juniors (IT)	21	195	289	29	122	138
Institution 3 Business Seniors (IT)	16	160	170	16	93	61
Institution 3 Business Total (IT)	37	355	459	45	215	199

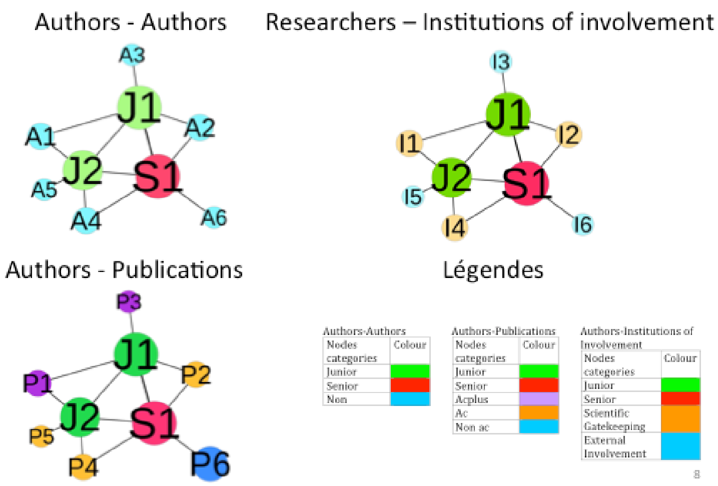
We carry out first an internal comparison, i.e. within an institution (US University), between 3 disciplines: business administration; chemistry; history; and second a comparison within a discipline (business administration) between 3 institutions: US University; French business school, Italian University.

- The information contained in the department members CV has been codified for building individual activity profiles, distinguishing two categories in the subset of department researchers: “juniors”; “seniors” according to their date of PhD.
- Publications have been analysed, distinguishing three tiers: top academic; academic; non academic
- The involvement activity distinguishes “scientific gatekeeping” and “external involvement”

Various original semi-automated strategies have been used for the data treatment. This codification process still required a lot of manual work: disambiguation, harmonisation.
Individual publications have been analysed as count (integer)

The Social Networks Analysis (powered by IFRIS CorText Manager), focused on linkages as potential resources for the department (figure 1):
- Authors – Authors
- Authors – Journals
- Department Researchers – Institutions of involvement

Figure 1: Three types of networks



Two main topology traits are considered for the statistical analysis: overlap of networks components – i.e. intersection (Junior network and Senior network) vs. Union (Junior network and Senior network); the share of edges connected to Juniors vs. share of edges connected to Seniors

Results
First, regarding the contribution of individual performances in the collective achievement in the making of the department academic quality: seniors’ and juniors’ respective network components strongly overlap in the 3 networks above-described. Does this shared space delineate a department backbone or identity?

Figure 2: Overlap of junior and senior components

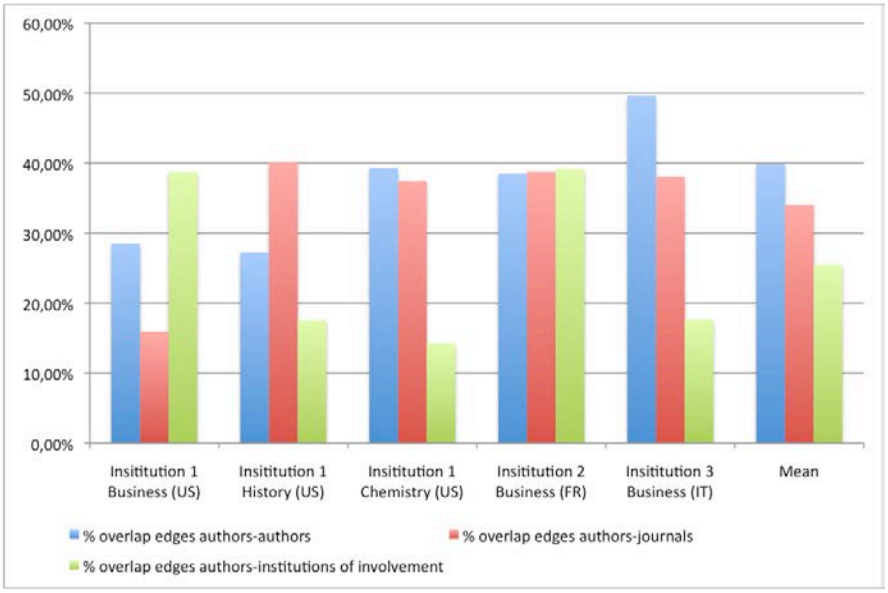
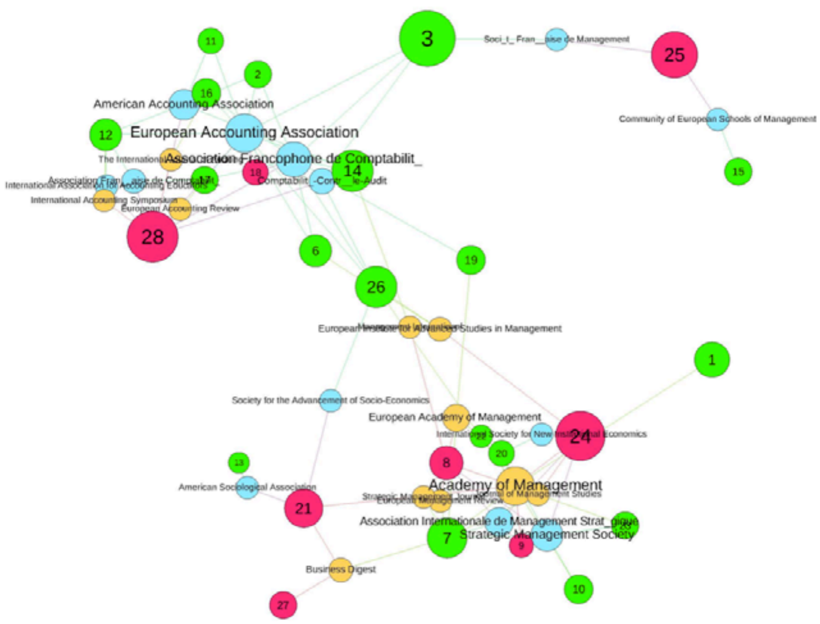


Figure 3: Overlap of junior and senior components

Authors-Institutions of Involvement	
Nodes categories	Colour
Junior	
Senior	
Scientific Gatekeeping	
External Involvement	



Second, regarding the specialisation of activities within the departments between « juniors » and « seniors » researchers: juniors appear specialised in top academic publications; seniors appear specialised in involvement, for both scientific gatekeeping and external involvement. See figures 4 and 5.

Figure 4: Junior specialisation index in top academic publications

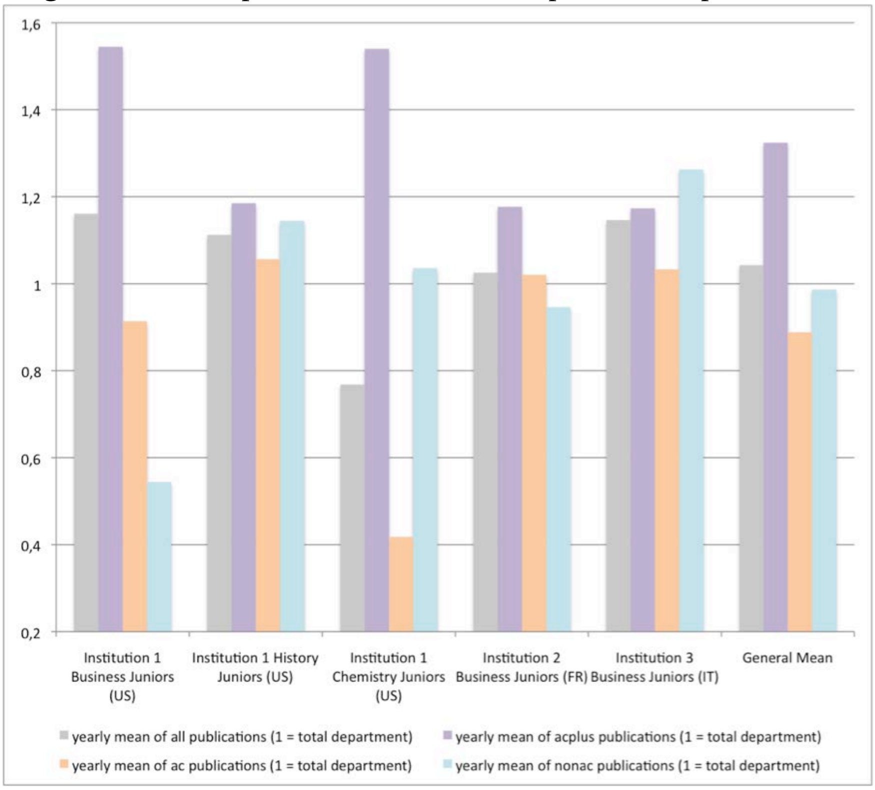
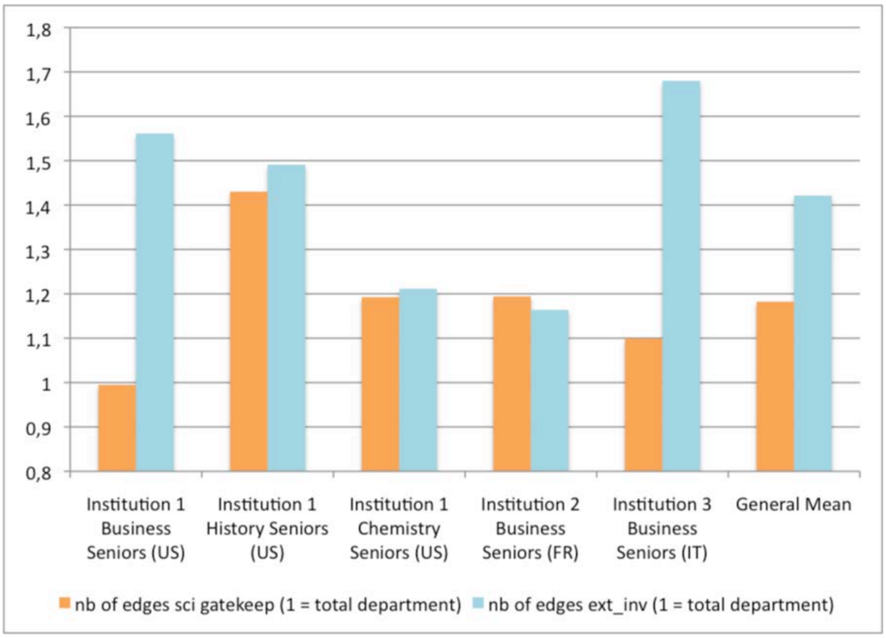


Figure 5: Senior specialisation index in networks of involvement



Finally, with a cohort-oriented perspective, we have identified a shift in the activity profiles during early career phase between junior and senior researchers: during the first 10 years period following the PhD, the former appear to present an activity profile more oriented towards top academic publications. See table 2.

Table 2: Average yearly academic production during early careers

Business administration in US University	Total yearly production of publications	Yearly production of acplus publications	Yearly production of ac publications	Yearly production of nonac publications
junior	2,13	1,38	0,59	0,17
senior	2,18	1,25	0,75	0,19

Business School (FR)	Total yearly production of publications	Yearly production of acplus publications	Yearly production of ac publications	Yearly production of nonac publications
junior	2,01	0,30	1,26	0,44
senior	4,33	0,21	2,62	1,50

Selected References

Bleiklie I. (2008). Excellence and the diversity of higher education systems, Paper presented at the *21th CHER Annual Conference*, Excellence and diversity in Higher education. Meanings, goals and instruments, Università degli studi di Pavia, Sept. 11-13 2008.

Musselin Ch. & Paradeise C. (2009). French universities in transition. In Paradeise C. , Ferlie E., Bleiklie I., Reale E. *University governance: Western European comparative perspectives*, Dortrecht, Springer.

Schoen A. et al (2006). *Observatory of the European University Methodological Guide, Strategic management of University research activities*, PRIME project, 237 p. http://www.prime-noe.org/Local/prime/dir/Projects/OEU/OEU%20methodological%20guide%20v2007.pdf

Braam R & van den Besselaar P. (2010). Life cycles of research groups: the case of CWTS, *Research Evaluation*, 19(3), September 2010

van der Weijden I et al. (2008). Implications of managerial control on performance of Dutch academic (bio)medical and health research groups, *Research Policy* 37 1616–1629

Morris N., (2002) The developing role of departments, *Research Policy* 31 817–833

Schmoch U. et al. (2010). How to use indicators to measure scientific performance: a balanced approach, *Research Evaluation*, 19(1), March 2010, pages 2–18

A methodological study to structure liquid biofuel science and put its publication dynamics into context[1]

Mirjam CB Schomaker* and Ed CM Noyons**

* *mirjam.schomaker@wanadoo.fr*
CWTS, Leiden University, Wassenaarseweg 62A, Leiden (the Netherlands)
African Studies Centre Leiden, Wassenaarseweg 52, Leiden (the Netherlands)

** *noyons@cwts.leidenuniv.nl*
CWTS, Leiden University, Wassenaarseweg 62A, Leiden (the Netherlands)

Abstract
Climate change is unequivocal, human influence on the process is clear, and substantial and sustained reductions of greenhouse gas emissions will be required to mitigate and adapt to it. An early global response has been to use liquid biofuel for transport, a development that has triggered a vivid public debate. We try to provide insight in interactions, and the timing thereof, in this debate between science, policy and society. We look at the entire biofuel knowledge base, which is a heterogeneous body of scientific publications and applied 'grey' literature. As a first step we present this paper on the methodology applied to structure the *scientific part* of biofuel knowledge and to put its time dynamics in perspective. We combine quantitative bibliometric data and contextual, qualitative expert interpretation. We used the Web of Science and the CWTS global science classification system to create a body of publications, structured into biofuel related clusters that are based on citation relations and patterns. Each cluster contains both 'pure' liquid biofuel publications *and* surrounding articles that are relevant but not specific enough to be part of the biofuels core. We annotated the selected clusters, extracted data on key players, calculated cluster ages, and presented the number of publications per year for each cluster graphically, always distinguishing between core biofuel- and related surrounding articles. In our analysis, and the main reason for this paper, we then illustrate that it is the actual *integration* of these different types of detailed data that provides *additional* information to arrive at an understanding of publication dynamics.

Keywords
research structuring, expert interpretation, publication dynamics, biofuels

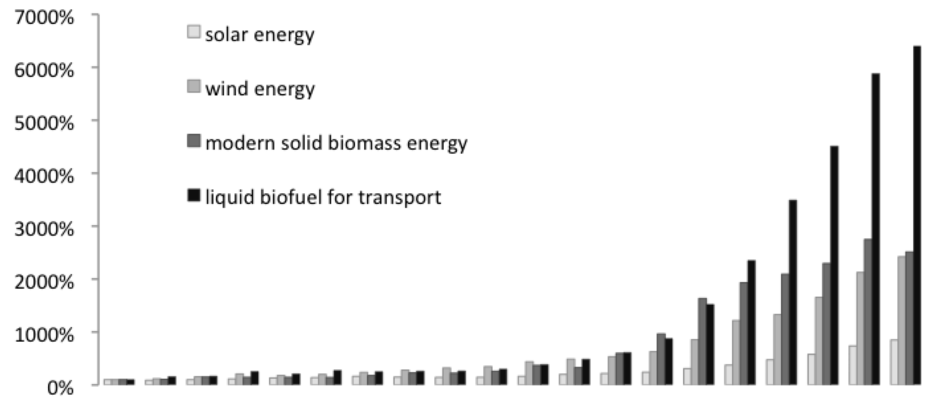
Introduction

Climate change is unequivocal; human influence is clear, mostly through fossil fuel combustion for heat, electricity and transport; limiting it will require substantial and sustained reductions of greenhouse gas emissions (IPCC 2013). After 25 years the question no longer is 'whether climate change is human induced or not', but 'how to prepare for and adapt to its impacts' and 'how to mitigate it' (IPCC 2014a and b). An early global response has been to use modern forms of renewable energy (its global share of final energy consumption is already nearing10%); liquid biofuel for transport is one of them (it meanwhile provides 3.4% of global road transport fuels); and global new investment and public spending on renewable energy research and development increases (REN21 2013, FS 2013). Also the number of

[1] All supporting data referred to but not included in this paper is available at http://www.cwts.nl/research/schomaker2014/data.zip

scientific publications on renewable energy grows, on liquid biofuels for transport in particular (Figure 1). In this paper we focus on the latter (mostly referred to as 'biofuel').

Figure 1: Relative increase in publications on some forms of renewable energy, 1993 – 2012 (1993 = 100%, WoS 2013).



Biofuel publications in the Web of Science cover fundamental, mostly chemical and physical sciences, while also more applied, mostly environmental, social, economic and developmental aspects of biofuels are well represented. The number of publications increases rapidly since 2005. In particular for liquid biofuel for transport a wide variety of arguments for and against are documented, and this scientific debate also takes place vividly in the public domain.

Several researchers have published on the interaction in the biofuel debate between science and society, often based on case studies with semi-structured interviews (such as Dunlop 2010, Sharman and Holmes 2010, Pilgrim and Harvey 2010). Major institutional reports on biofuels present comprehensive global or regional reviews and integrated assessments (such as REN21 2013 to give only one example). Clearly the biofuel knowledge base is a heterogeneous body of both fundamental scientific publications and 'applied studies in grey literature', the latter adding value by integrating and presenting science in a societal context. We are trying to frame this broad biofuel knowledge base, considering both scientific and applied ('grey') literature.

Our first objective is to only study the *scientific part* of the biofuel knowledge base. In this initial paper we give a methodological description of how we delineated and structured scientific publications on liquid biofuels for transport, and how we put its time dynamics in perspective, by integrating quantitative bibliometric data and contextual, qualitative expert interpretation.

Data sources, tools and approach

CWTS has obtained a licence from Thomson Reuters to create and use a dedicated Web of Science (WoS) database for bibliometric analysis, research and services. CWTS updates its WoS database on a regular basis to make it specifically suited for advanced bibliometric research. Where we mention WoS we mean the CWTS version.

To structure the biofuel publications collected from the WoS we use the CWTS global classification system of science. It is based on automated clustering of citation relations and patterns at the individual publication level (Waltman and Van Eck 2012). The classification clusters over 18 million articles, published between 1993 and 2012. It provides a hierarchical, three level structure; we used the detailed third level.

In this paper we describe how we collected and structured scientific biofuel publications and then present the actual cluster data and analysis on publication dynamics.

Data collection and structuring

Collecting subject-specific WoS publications

We searched the WoS for publications related to liquid biofuel for transport, written in English, and published between 1993 – 2012. To cover the entire range of research fields that are relevant, we searched in titles, abstracts and keywords. We used an iterative process to define optimal lists of search terms to properly delineate our subject. Delineation problems encountered:

- Definitions: the word 'biofuel', for example, is usually reserved for liquid biofuel for transport. However, in some Nordic and Central European countries 'biofuel' is specifically used for modern solid biomass energy for heating and electricity. And in Asia and Africa some authors these days also use 'biofuel' for fuelwood, charcoal or dung burned for comfort-heat and cooking (so-called traditional biomass energy). And while 'biodiesel' is common, 'bioethanol' is not. Authors rather use 'fuel ethanol', 'ethanol for bioenergy', or 'ethanol' linked to feedstock (corn, soy, waste).
- Habits: 'biofuel' is often not used in fundamental research. Authors rather give their own descriptions (such as "sustainable, drop-in replacements for petroleum based fuels"), or are so focussed on and caught up in their own research community that they do not see the need nor the opportunity to link up to other communities.
- Different contexts: 'ethanol', for example, is common both in medical or pharmaceutical research and in biorefinery. 'Combustion' is used for vehicle engines and for stationary engines in power plants.
- Terminology changes over time: in experiments on novel green fuels, for example, other language is used than for first generation commercially available biofuels. Or in new biorefinery technologies the same biomass can be pre-treated for both liquid and solid energy conversions.

Keeping the above restrictions in mind, our final search script collects two sets of publications that, combined, result in our final corpus:
set_0 : a temporary set of publications that have 'biofuel', 'biodiesel' and/or several fuel related words linked to 'ethanol' in their title, abstract or keywords.
not_set_0 : a temporary set of publications that we do *not* want to include in our corpus. We singled out traditional bio(mass)fuel issues, modern solid biomass (co-)generation, and *non*-liquid biofuel for transport.
final_set : a set of publications that do occur in *set_0* and do not occur in *not_set_0* (in other words: excluding the overlap between *set_0* and *not_set_0*, Figure 2).
In this *final_set* we collected 21,422 WoS publications.

Figure 2: Different sets of publications produced to come to a final corpus.

Structuring WoS publications into CWTS classification clusters

We joined the collected WoS publications (final_set) with the CWTS science classification database to structure the publications into biofuel related clusters. This resulted in 18,545 core biofuel WoS publications spread over 1,689 CWTS clusters. The number of publications is lower than the 21,422 in our final_set because WoS publications without citations, such as editorials, are not included in the citation-relations based CWTS classification. Two other classification characteristics to note: (1) the number of publications per cluster varies considerably (roughly between 50 and 5000), where small clusters are often less informative, but could also point to a specific niche; and (2) apart from the subject-specific *core* publications (our 18,545 biofuel core) each cluster also contains *surrounding* publications that do not match our search rules, but that are in most cases relevant and informative in a biofuel context (this dual property is important in the next step of selecting clusters for further analysis, and in the publication dynamics analysis itself).

Deciding on a cut-off point to limit the 1,689 clusters for further analysis is a matter of finding a convenient balance between ‘not too many and not too small clusters’ and ‘not missing out on possibly important small clusters’. We checked three aspects. First we calculated the subject coverage for each cluster (the percentage of subject-specific core biofuel publications over the total number of publications). The higher this percentage, the more relevant a cluster is for analysis. Next we did a quick scan on relevance of titles in low coverage clusters (we did not go below 3%). And finally we searched for an acceptable minimum number of biofuel publications per cluster. Based on these three checks we set our limits at a coverage of at least 5% and at least 10 biofuel titles in a cluster, which resulted in 46 clusters.

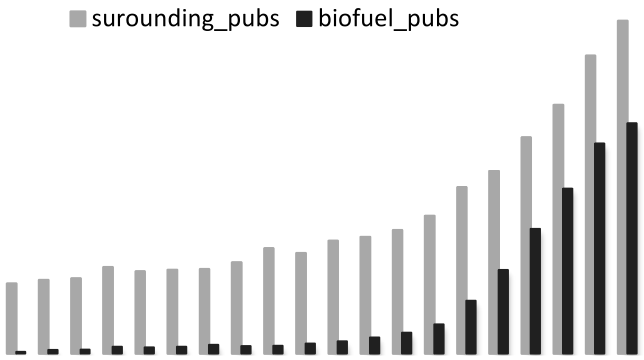
We then reviewed titles, abstracts and keywords in these 46 clusters, and named and annotated each cluster for both core biofuel and for surrounding publications. In six small clusters the geographical bias turned out to be too obvious and titles too inconsistent or irrelevant. This seems due to too incessant peer citation behavior; a problem that unfortunately is inherent to clustering techniques based on citation patterns. We deleted these clusters, so arriving at 40 clusters for further analysis that together contain 44,373 publications: 11,754 core biofuel and 32,619 surrounding publications. The biofuel core in the 40 clusters comprises 62% of the total number of biofuel publications collected in the 1,689 clusters. Descriptions and statistical details on the 40 clusters, for example on coverage, can be accessed at the web address given in footnote 1. For the ten clusters analyzed for this paper some data are given in Table 2.

Presenting bibliometric and contextual data for further analysis

In this section we first present general statistics on our collected publications, which already illustrate some unexpected and some unsurprising patterns, followed by specific calculations for 10 clusters that represent the most important patterns occurring in the 40 selected clusters (only 10 due to space restrictions).

Figure 3 shows the total number of publications per year between 1993 and 2012 in the 40 selected clusters, distinguishing between core biofuel (dark) and surrounding publications (light); the surrounding publications increased nearly 5-fold in 20 years, while the biofuel core increased over 125-fold.

Figure 3: Total number of publications per year in the 40 selected clusters, 1993 – 2012.



Looking at the top 10 ranking for countries that are affiliated with *all* articles in the 40 clusters (institute of first authors only), we note large differences compared to the ranking for *all* WoS publications in the same period (Table 1). Countries like China, India and Brazil have a very high 'biofuel position' compared to their overall WoS ranking, while the USA, Great Britain and Germany, for example, have significantly lower scores. When looking at the core biofuel articles only the list changes even more; Turkey and South Korea push France and Germany out. The high position of Brazil is obvious as Brazil has had national policies and programmes on sugarcane ethanol for their vehicle fleet in place since the mid 1970-s. It requires a detailed analysis of publication abstracts to provide further explanations on the observed shifts.

Table 1: The top 10 scientific publishing countries in 1993 – 2012: (a) for all 18 million publications in the WoS that are also included in the CWTS classification; (b) for all 44,373 publications in our 40 clusters; and (c) for our 11,754 biofuel core publications.

(a) All pubs in WoS		**(b) All pubs in 40 clusters**		**(c) Biofuel core in 40 clusters**	
USA	28%	USA	21.2%	USA	24.1%
Japan	7%	Peoples R. of China	9.2%	Peoples R. of China	9.8%
Great Britain	7%	Japan	5.6%	India	6.3%
Peoples R. of China	6%	India	4.5%	Brazil	5.9%
Germany	6%	Great Britain	4.1%	Japan	3.6%
France	4%	Brazil	4.1%	Spain	3.6%
Canada	4%	Canada	3.8%	Great Britain	3.3%
Italy	3%	Spain	3.7%	Canada	2.9%
Spain	3%	Germany	3.6%	South Korea	2.7%
India	2%	France	3.5%	Turkey	2.6%

The top 20 scientific journals in which articles in the 40 clusters were published show no surprises: 8 journals on bioenergy and energy policy; 5 on microbiology, biotechnology, biochemistry, bioengineering, 4 on "not-bio-specific" energy and 3 on "not-bio-specific" chemistry.

To get a feel of the publication dynamics we calculated the cluster age (average publication year for each selected cluster) along the lines that Kajikawa and Takeda (2008) applied to describe their research on structuring biomass and biofuel to obtain insight in emerging research fields. We calculated cluster ages both for all publications per cluster (the 'overall' cluster age), and for core biofuel publications only (Table 2). The overall cluster ages lie between 2003.1 and 2009.2; the biofuel cluster age range starts higher, going from 2006.8 to 2010.7, and has more young clusters. In the Analysis below we discuss ins and outs of overall cluster ages and specific biofuel cluster ages.

Table 2: Data for the 10 clusters analysed on publication dynamics (from highest (a) to lowest (j) number of biofuel publications per cluster).

#	cluster code			N all pubs	N BF pubs	cluster impression	cover-age	overall cluster age	biofuel cluster age
(a)	10	22	1	4916	4267	biodiesel blends - production and performance	86.8%	2008.58	2009.01
(b)	10	8	1	3636	1788	lignocellulosic feedstock decomposition for bioethanol	49.2%	2007.87	2008.92
(c)	6	7	8	1393	853	the biofuel debate	61.2%	2009.22	2009.86
(d)	10	22	2	2816	783	microalgae for large-scale production of biomass	27.8%	2006.16	2010.74
(e)	6	7	3	1869	421	non-food (energy) crop yield improvement	22.5%	2005.37	2008.16
(f)	20	4	1	3146	286	thermo-chemical biomass conversion into bioenergy (pyrolysis)	9.1%	2006.71	2008.74
(g)	10	8	4	2486	257	lignocellulosic feedstock decomposition	10.3%	2003.78	2009.50
(h)	1	55	7	919	253	potential of biobutanol	27.5%	2006.18	2010.31
(i)	10	16	4	1380	245	fermentation; beer, wine, bioethanol, …	17.8%	2003.05	2006.75
(j)	10	22	4	771	153	re-use of glycerol, waste from biodiesel	19.8%	2009.19	2009.64

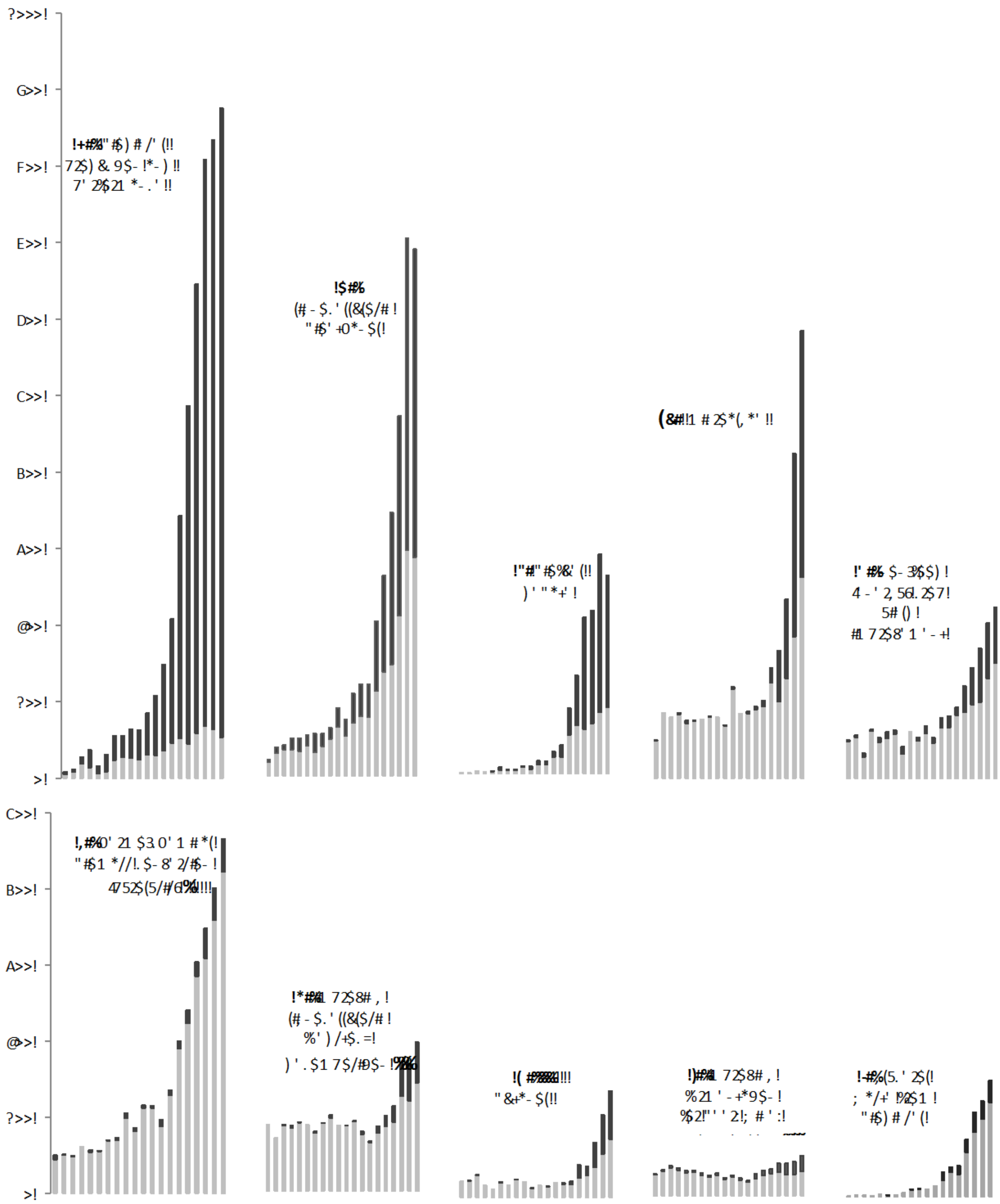
To obtain optimal insight in the data we looked at the cluster ages averaged over 20 years, listing the number of publications *per year* within each cluster (also done by Kajikawa and Takeda 2008). We then went one step further by differentiating publications per year for core biofuel publications (dark grey) and surrounding publications (light grey) (Figure 4 below). The resulting graphs nicely illustrate the coverage percentages (more dark colours equal higher percentages), and the spreading of biofuel specific publications over the 20 years period. As we will demonstrate in the Analysis below, these specific cluster patterns provide interesting additional information to arrive at an understanding of the publication dynamics.

Analysis and discussion

In this section we will analyse the data for the ten clusters briefly presented above.

First: the *overall* cluster ages point to only two young clusters (2009+), cluster (c) on the biofuel debate and (j) on glycerol waste; this is understandable from a contextual perspective. The next overall youngest (2008+) cluster, (a) on biodiesel, is somewhat surprising knowing from annotations that it also covers articles on 'old' forms of commercially available biodiesel. The overall middle aged (2006+) clusters (d) and (h) seem strange as they cover recent subjects (biodiesel from microalgae and butanol fuel respectively); the same is true for overall old (2003+) cluster (g) that covers new experiments to improve woody biomass decomposition.

Figure 4: Number of publications per year per cluster, 1993 - 2012; dark = biofuel core; light = surrounding publications.



Second: looking at *core biofuel* cluster ages gives a different picture. Our *overall* youngest clusters (c) and (j) are also among the five youngest biofuel clusters (2009+); this matches the context of a recent subject. The *overall* younger cluster (a) also has a young biofuel age (± 2009), which, as noted above, is surprising considering its context. Our overall middle aged clusters (d) and (h) have the two youngest biofuel ages (2010+); this fits much better with our contextual impression. And finally, our overall old cluster (g) turns out to have a relatively young biofuel age (2009.5), again more in line with its context.

In other words, even the two types of cluster ages do not give enough handles to come to satisfactory conclusions and in some cases even contradict each other. We therefore combined the two types of cluster ages with, for each cluster, graphs showing core biofuel and surrounding publications per year, and with annotations obtained from titles (and if needed from abstracts).

Below we present our analysis for individual clusters. But first a note up front, related to delineation problems while separating surrounding publications (light grey) from core biofuel publications (dark grey). We distinguished three situations: (1) surrounding publications really do not cover liquid biofuel issues, or may be potentially biofuel relevant, but do not (yet) discuss it; (2) surrounding publications often do deal with liquid biofuels, but due to different definitions, habits, contexts, or terminology dynamics authors do not always use the 'right' language to end up in the (dark grey) biofuel core selection; and (3) a cluster has a bit of both. Often only expert impressions from abstracts can give a final clue on which of the three situations is most applicable.

Clusters (a) and (b)
(a) covers biodiesel production, blending and performance; (b) covers degradation processes of (mainly) non-edible dry plant matter for bioethanol production. The clusters mostly fit situation 2 described above, so in fact the dark bars should be even longer (more so for (b) than for (a)). Both the overall and the biofuel cluster ages are young, in principle hinting to a new or even emerging topic. The annotations and graphs, however, show that biofuel has been on the agenda since 1993. The graphs also show gradually increasing small numbers in the first decade and very rapid growth in the second, particularly so for the (dark) biofuel publications. Integrating all this, we may conclude that the young cluster ages do not so much seem to point to recent of even emerging topics, but rather to increasing popularity of long existing ones?

Cluster (c)
Covers the biofuel debate, looking at potential and perceived impacts of biofuels on societal aspects (environment, economy, development, ethics). The cluster type is comparable to (a) and (b), a situation 2 cluster (even longer dark bars than shown): both types of cluster ages are young; annotations note a younger topic. The graph shows an almost zero start in 1993, with a gradual incerease including biofuel publications from the start. However, the numbers in the early years are much lower than for (a) and (b), the period of gradual increase is slower and lasts 5 years longer, and the very rapid increase starts only in 2007 (fastest for the biofuel publications). Still, the debate is not emerging as it starts increasing around 2002; it rather becomes extremely popular in the last five years?

Clusters (d), (g) and (h)
(d) provides an overview of microalgae development and its potential for cheap, large-scale production of biomass for food and lipid industries (cosmetics, medicine, food additives, animal nutrition, aquaculture, and biorefineries). The cluster fits situation 1 (a clear division between light and dark grey). The cluster is overall middle aged and has a very young biofuel age. Since 1993 the surroundings (with around 100 publications per year) hardly increase, and the graph shows very few biofuel publications until 2008. From 2008 onwards all publications increase at a very high rate, those on biofuels clearly faster than the surroundings. We could conclude that microalgae publications in general (light grey) have been existing from the start and are becoming much more popular, while its application for biofuels is emerging?

(g) on improving lignocellulosic feedstock decomposition for enhanced bioethanol production and (h) on butanol production experiments, are comparable clusters: both in situation 1; old overall cluster age for (g) and middle aged for (h); biofuel cluster ages both young. Graphs show slower increases than for (d), but the same pattern, with very few biofuel publications until 2008 and rapid increases in the last five years, so indicating emerging biofuel topics?

Clusters (f) and (e)
(f) covers pyrolyse for development of enhanced bioenergy using modern feedstock (waste, lignocellulosic biomass, microalgae). It is a situation 3 cluster, and both overall and biofuel cluster ages are ‘seriously middle aged’. Surroundings (light grey) gradually increase from about 60 publications per year in 1993 to around 100 by 2005 and start increasing very rapidly in 2006. A few biofuel publications (dark grey) were published every year from the beginning, also rapidly increasing from 2006 onwards. More detailed abstract reviews are required to pinpoint the cause of the strong increase for both surrounding and biofuel publications in the last six years, but it is already clear that the recent increase could not have been detected from average cluster ages only.

(e) on yield improvement of non-food energy crops for liquid biofuel for transport and solid biomass energy resembles cluster (f): a situation 3 cluster (the dark bars should probably be a bit shorter) and both overall and biofuel middle aged. The only difference with (f) is that the graph shows growth from 2006 onwards at a slower rate, and that the ratio between biofuel and surrounding publications seems to favour biofuels.

Cluster (i)
Covers fermentation through more stress-tolerant yeasts to increase ethanol yields, used in wine making, beer brewing, specific food processing, and bioethanol production. The cluster has a different pattern from all others. It is a situation 1 cluster (clear division). Both types of cluster ages are the oldest of the ten, and indeed some of the subjects originate from ancient times. The graph pattern shows a rather stable volume of publications ranging between 40 and 60 per year for the entire 20 year period, with a slow dip from 1995 to 2005. The biofuel publications are present from year one, and slightly increase in the last 4 years. We conclude that this is a stable cluster, currently with no new developments. The gentle dip needs further study of abstracts.

Cluster (j)
Covers re-use of glycerol, a waste output of the biodiesel industry: a situation 1 cluster; young cluster ages, both overall and for biofuel. The graph shows an almost zero start; very few publications in the first decade followed by a few years with a slight increase, almost non on

biofuels (but this may be due to definition problems); then a rapid growth since 2005, when biofuels start coming up. Abstracts indicate that growing amounts of glycerol waste are becoming problematic. This is the only cluster for which all indicators point to relatively new topics, but the rapid increase since 2009 is most obvious for the surrounding publications. More abstracts need to be studies to pinpoint which subjects are emerging.

Conclusions and next steps

In this analysis of *scientific* biofuel publication dynamics we compared four sub-sets of data: (1) overall cluster age, (2) core biofuel cluster age, (3) contextual annotations, and (4) cluster graphs on annual trends in core biofuel and surrounding publications. Our overall conclusion is that it is the integration of these different types of data that provides new insights and additional information to arrive at an understanding of publication dynamics. Looking at the first ten clusters it seems possible to differentiate between emerging issues and 'mere' popular topics; such knowledge could give clues on policy influence (or not) in science development and vice versa.

Next steps in our effort to obtain more insight in the interaction between science and policy are:

- *Complete mapping of the biofuel science base*, analysing all selected 40 clusters using bibliometric tools (term maps, citation networks, geographical coverage, publication dynamics, biofuel position in the overall global science classification map, etc.);
- In parallel *collect bibliometric details on applied grey biofuels literature* that is not included in the WoS
- Then integrate those two large data sets and *analyse this new dedicated set, zooming in on societal aspects of biofuels*: linkages and impacts (or not) between science and policy, with a focus on Europe and Africa.

Acknowledgements

Many thanks to Prof. Dr. Ton Dietz and three anonymous reviewers for their very useful comments and suggestions.

References

Dunlop, C.A. (2010). The temporal dimension of knowledge and the limits of policy appraisal: biofuels in the UK. Policy Science, 43, 343-363.

FS (2013). Global Trends in Renewable Ene*rgy Investment*. Frankfurt am Main: Frankfurt School of Finance & Management, Nairobi: United Nations Environment Programme, and Bloomberg New Energy Finance.

IPCC (2013). *Climate Change 2013: The Physical Science Basis. Working Group I contribution to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change (IPCC)*. Cambridge UK and New York NY: Cambridge University Press.

IPCC (2014a). *Climate Change 2014: Impacts, Adaptation, and Vulnerability.Working Group II contribution to the* Fifth *Assessment Report of the Intergovernmental Panel on Climate Change*. Final Draft Report. Cambridge UK and New York NY: Cambridge University Press.

IPCC (2014b): *Climate Change 2014: Mitigation of Climate Change. Working Group III contribution to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge UK and New York NY: Cambridge University Press.

Kajikawa, Y. and Takeda, Y. (2008). Structure of research on biomass and bio-fuels: a citation approach. *Technological Forecasting and Social Change*, 75, 1349-1359.

Pilgrim, S. & Harvey, M. (2010). Battles over Biofuels in Europe: NGOs and the Politics of Markets. *Sociological Research Online*, 15, august 31. London: SAGE publications ltd. DOI 10.5153/sro.2192

REN21 (2013). *Renewables 2013 Global Status Report.* Paris: Renewable Energy Policy Network for the 21 Century Secretariat.

Sharman, A. & Holmes, J. (2010). Evidence-Based Policy or Policy-Based Evidence Gathering? Biofuels, the EU and the 10% Target. *Environmental Policy and Governance,* 20, 309–321.

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, 2378–2392.

Invisible College Ranking: An Empirical Study of How Chinese Graduate Student Mobility Underpins Research Universities[1]

Nobuuyki Shirakawa* and Takao Furukawa**

* *nshiraka@sfc.keio.ac.jp*
School of Media and Governance, Keio University, 5322 Endo, Fujisawa, 252-0882 (Japan)

** *furukawa@nistep.go.jp*
Science and Technology Foresight Centre, National Institute of Science and Technology Policy, 3-2-2 Kasumigaseki Chiyoda-ku, Tokyo, 103-0013 (Japan)

Introduction

International student mobility in higher education has often been discussed from the aspect of the migration of potential highly skilled workers, which is relevant to brain drain and brain circulation in the context of global workforce issues (OECD 2009, 2010). These data enable a macro analysis of international student mobility among countries. The issue of international student mobility has had a profound effect on the policy decision making of the higher-education system of every country; however, the statistical data on this subject are insufficient, especially in the case of graduate students.

Sources and data

This article focuses on international graduate student mobility between institutions, although public statistical data has been insufficient to assess worldwide student mobility between institutions. We therefore collect original data for micro analysis.

Biographical notes and author trajectory

To comprehend international student mobility between institutions of higher education, we investigate the authors' biographical notes that are attached to journal articles. Scholarly articles published by the Institute of Electrical and Electronics Engineers (IEEE), which is the largest professional association for electronics-related disciplines, normally contain brief biographical notes as follows:

> *The author received the BTech degree in electrical engineering from Indian Institute of Technology, Bombay in 1997 and the MS degree from the University of Illinois, Urbana-Champaign in 1999. Since August of 1999, he has been working as a system engineer at QUALCOMM Inc., where he is working on design and development of the cdma2000 reverse link.*

By analysing such biographical notes, we can obtain the following information about the institutions the author attended, in chronological order:

- Institution awarding the bachelor's degree: Indian Institute of Technology, Bombay
- Institution awarding the master's degree: University of Illinois, Urbana-Champaign
- Institution awarding the doctoral degree: Not available in this case

[1] This work was supported by JSPS KAKENHI Grant Numbers 20556071, 25285236, 22500875

Data

Table 1 summarizes the source data: journal title; the period of publication and the numbers of articles, including both authors and unique authors. The listed journals have high impact factors in their respective research domains. Although the number of authors includes the duplication of scientists who have published several articles, the number of unique authors does not contain this duplication. Consequently, we obtained the trajectories of more than 7,000 unique authors in total and more than 2,200 unique authors for each research domain. We extracted the institutions awarding the bachelor's, master's and doctoral degrees from the biographical notes of the authors, as mentioned above. However, the institution names showed variations owing to the different languages in which the descriptions were written, the abbreviations etc. We therefore checked and unified these variations to maintain consistency. Consequently, we obtained 1,647 academic institutions from 93 countries.

Table 1. Selected Research Domains, Journal and Data Sets

Research domain	Journal title	Period of publication	Number of papers	Number of authors	Number of unique authors
Computer vision	IEEE Transactions on Pattern Analysis and Machine Intelligence	1997–2009	1,294	3,437	2,361
Robotics	IEEE Transactions on Robotics	2004– 2009	493	1,487	1,157
	Robotics and Autonomous Systems	2004–2009	540	1,672	1,382
	Subtotal*		1,033	3,159	2,441
Electron devices	IEEE Transactions on Electron Devices	Aug. 2008–Dec. 2009	584	2,919	2,251

* Subtotal does not equal the sum of journals in robotics, because some authors published different articles in both journals.

Results

Characteristic institutions

- Institutions providing researchers to other institutions abroad

Universities in China typify large outflow institutions that provide numerous researchers to other institutions abroad in all three domains. In the robotics domain, universities in South Korea, Greece and Iran also provide many researchers to institutions of other nations. In the computer vision domain, a research institute in China and a university in Israel are also typical of such institutions. In the electron devices domain, many researchers from the universities in South Korea and Taiwan go abroad, indicating that institutions in East Asia tend to provide researchers to various institutions abroad.

- Institutions receiving researchers from abroad

Universities in the United States and Singapore receive many researchers from abroad, typify large inflow institutions in all three research domains. It is noteworthy that both the National University of Singapore and the Nanyang Technological University in Singapore receive as many researchers from abroad as do the top-ranked research universities in the United States.

Characteristic institutions in China

Concentrating especially on computer vision, we digitally illustrate the international/inter-institutional mobility of graduate students relating to five major Chinese national research universities. All of the following are representative universities sending their holders of bachelor's degrees to other universities for graduate studies: Tsinghua University, the University of Science and Technology of China, Zhejiang University, Peking University, and Xi'an Jiaotong University.

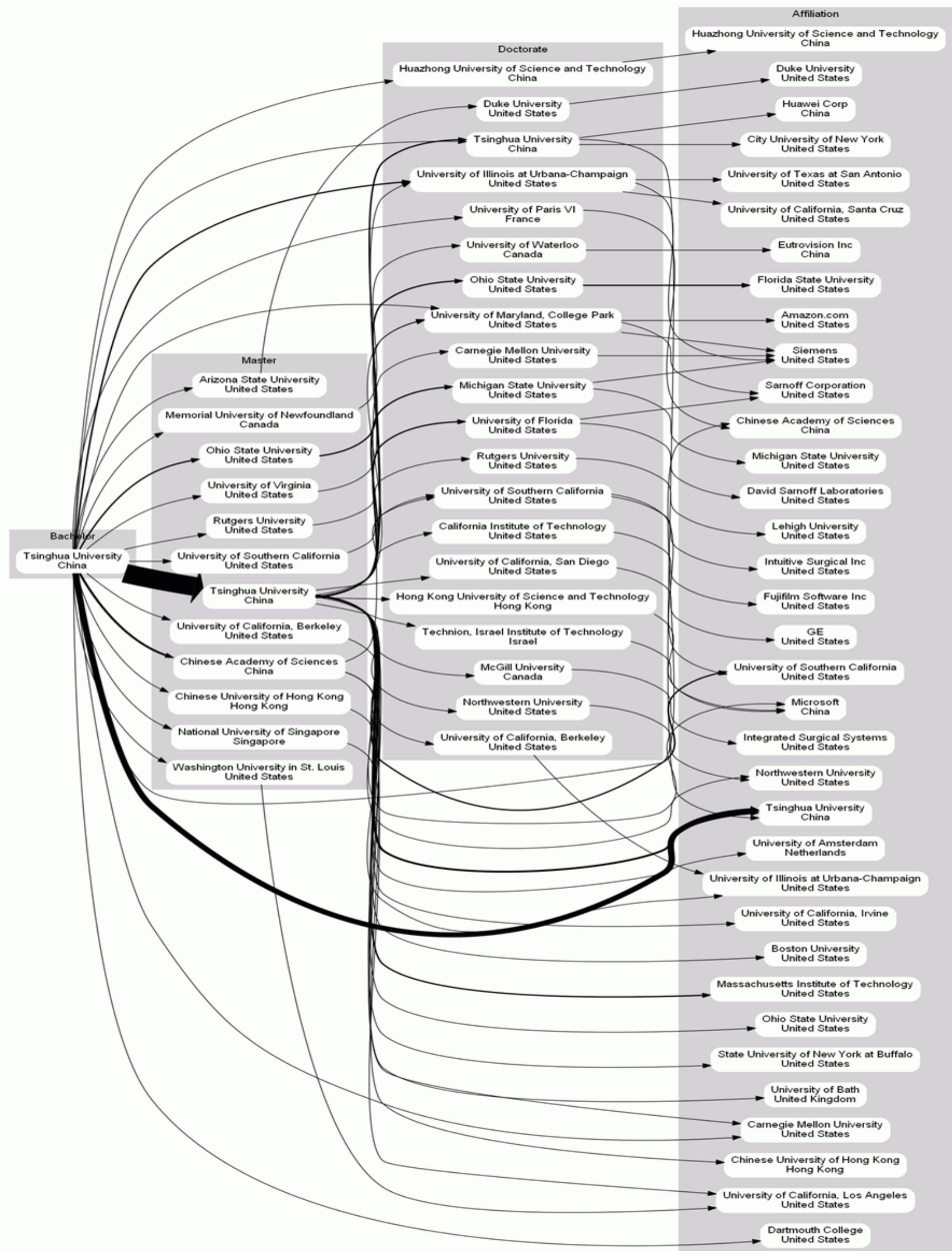


Figure. 1: Trajectories of 61 authors who earned bachelor's degrees from Tsinghua University, as obtained from biographical notes with their journal articles on computer vision. The arrow thickness indicates the number of students who moved between institutions.

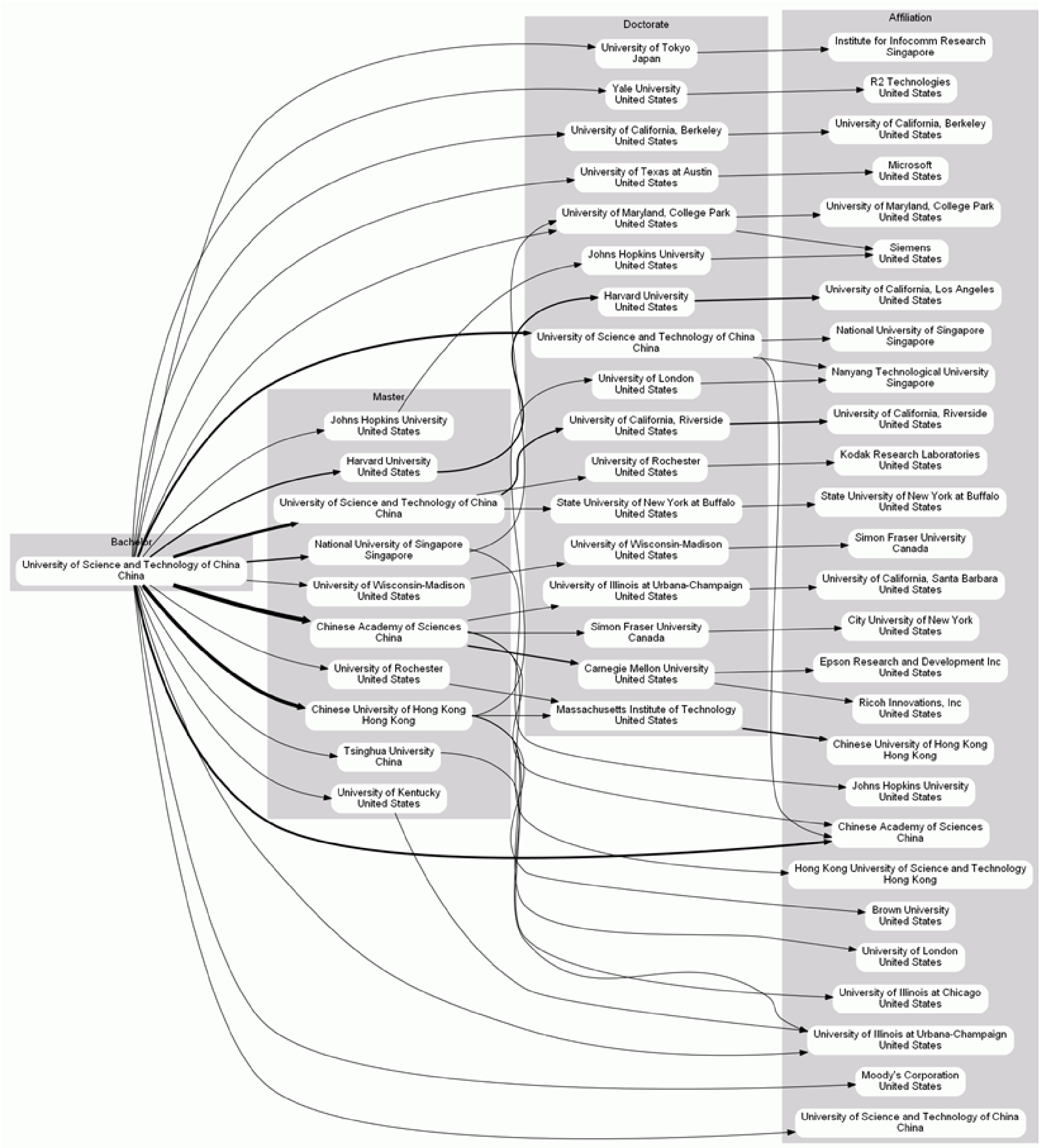


Figure. 2: Trajectories of 38 authors who earned bachelor's degrees from the University of Science and Technology of China, as obtained from the biographical notes with their journal articles on computer vision. The arrow thickness indicates the number of students who moved between institutions.

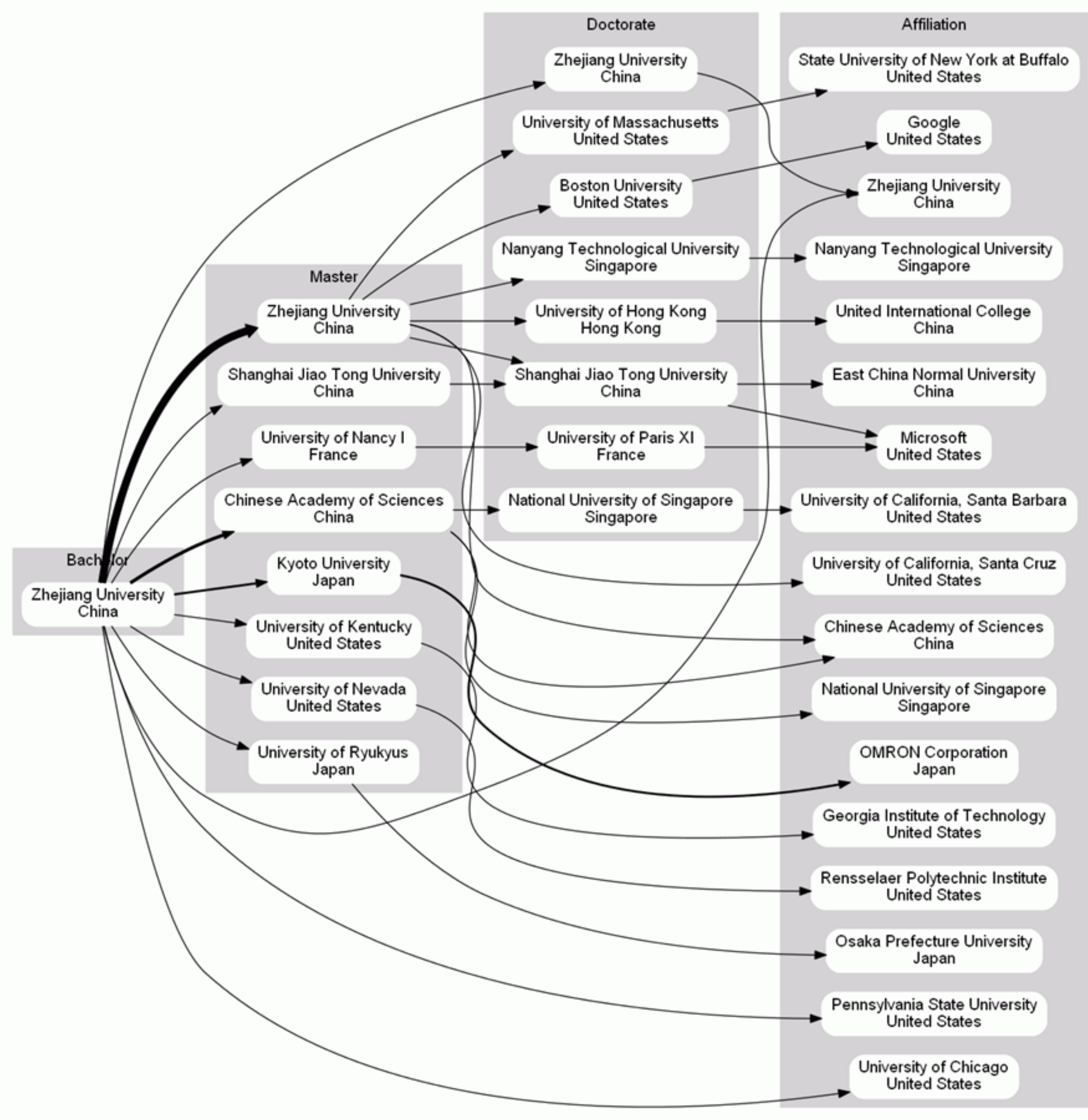


Figure. 3: Trajectories of 21 authors who earned bachelor's degrees from the University of Science and Technology of China, as obtained from biographical notes with their published journal articles on computer vision. The arrow thickness indicates the number of students who moved between institutions.

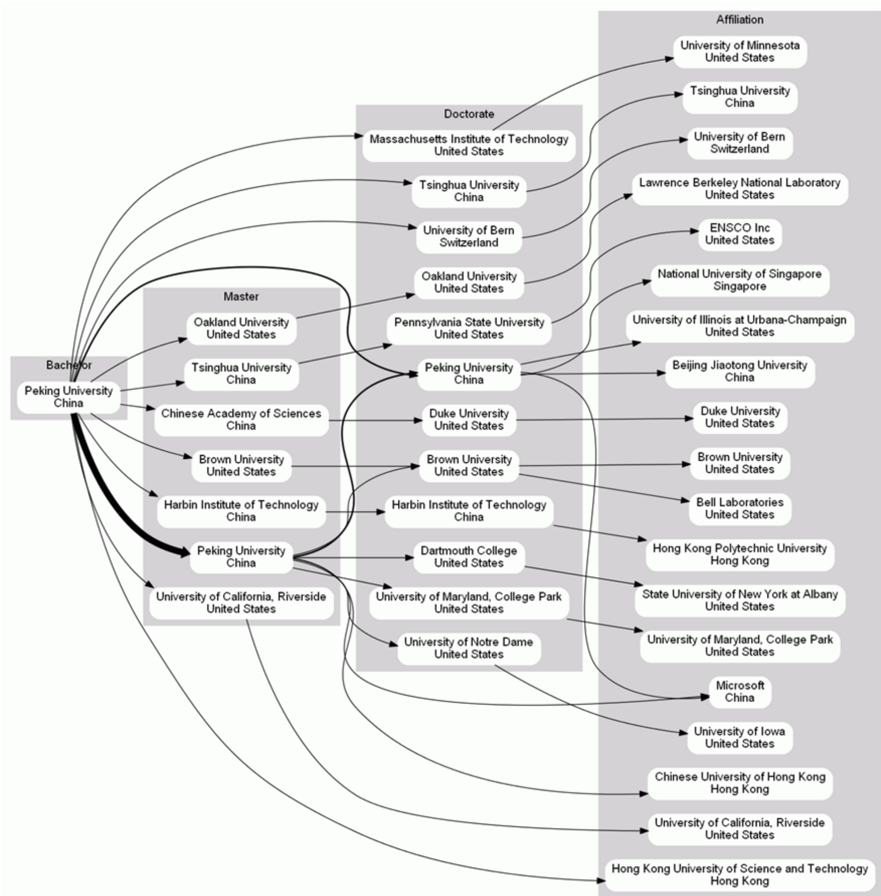


Figure. 4: Trajectories of 20 authors who earned bachelor's degrees from Peking University, as obtained from biographical notes with their published journal articles on computer vision. The arrow thickness indicates the number of students who moved between institutions.

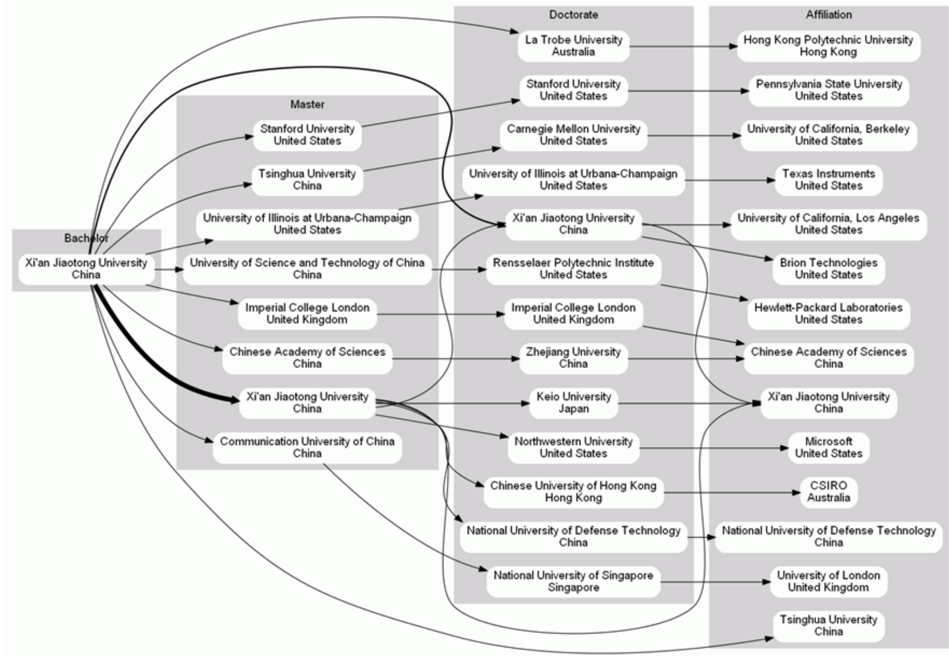


Figure. 5: Trajectories of 17 authors who earned bachelor's degrees from Xi'an Jiaotong University, as obtained from biographical notes with their journal articles on computer vision. The arrow thickness indicates the number of students who moved between institutions.

Conclusion

Our results roughly coincide with these categories and also indicate additional issues concerning graduate student mobility. This result suggests that there are complementary relationships between the top-ranked research universities and the exporting national research universities such as Tsinghua University etc. Consequently, exporting national research universities underpin top research universities in terms of providing excellent undergraduate students to the top institutions. China and India have exported professionals as well as students to Western countries since the 1970s, and brain drain has been considered a serious problem in those two countries. In China, the situation has been changing since 1990, following the improvement of economic and academic conditions under the policies in the 985 and 211 projects (Altbach 2009), which foster the inbound mobility of academic staff. Statistics on the international mobility of academic staff are now being noted in European countries, but there have been difficulties in classifying the types of academic staff, the disciplines etc. (Teichler et al. 2011).

References

Altbach, P. G. (2009). One-third of the globe: The future of higher education in China and India. *Prospects*, 39(1), 11–31.

Brooks, R., & Waters, J. (2009). International higher education and the mobility of UK students. *Journal of Research in International Education*, 8(2), 191–209.

Brooks, R., & Waters, J. (2011). *Student mobilities, migration and the internationalization of higher education*. London: Palgrave Macmillan.

Furukawa, T., Shirakawa, N., & Okuwada, K. (2011). Quantitative analysis of collaborative and mobility networks. *Scientometrics,* 87(3), 451–466.

Furukawa, T., Shirakawa, N., Okuwada, K., & Sasaki, K. (2012). International mobility of researchers in robotics, computer vision and electron devices: A quantitative and comparative analysis. *Scientometrics*, 91(1), 185-202.

Furukawa, T., Shirakawa, N., & Okuwada, K. (2013). An empirical study of graduate student mobility underpinning research universities. *Higher Education*, 66(1), 17–37

Laudel, G. (2003). Studying the brain drain: Can bibliometric methods help? *Scientometrics*, 57 (2), 215–237.

Laudel, G. (2005). Migration currents among the scientific elite. *Minerva*, 43(4), 377–395.

Organization for Economic Co-operation and Development (OECD) (2009). Higher education to 2030: Vol. 2. *Globalisation*. Paris: OECD.

Organization for Economic Co-operation and Development (OECD) (2010). *Education at a glance 2013: OECD indicators*. Paris: OECD.

Stephan, P. E., & Levin, S. G. (2001). Exceptional contribution to US science by the foreign-born and foreign-educated. *Population Research and Policy Review*, 20(1–2), 59–79.

Teichler, U., Ferencz, I., & Wächter, B. (Eds.) (2011). *Mapping mobility in European higher education: Vol. 1. Overview and trends*. Brussels: European Commission.

Appendix

Table A-1. Detected researchers as sorted by affiliation on computer vision.

Ranking	Insitute	Country	Number of detected reserchers	Total number of authors
1	Microsoft	United States	50	85
2	Carnegie Mellon University	United States	35	64
3	Chinese Academy of Sciences	China	33	43
4	Siemens	United States	30	43
5	University of Maryland, College Park	United States	29	56
6	Massachusetts Institute of Technology	United States	28	45
	INRIA	France	28	44
8	University of Illinois at Urbana-Champaign	United States	27	47
9	Technion, Israel Institute of Technology	Israel	26	42
10	Georgia Institute of Technology	United States	23	34
11	Chinese University of Hong Kong	Hong Kong	21	34
	Swiss Federal Institute of Technology Lausanne	Switzerland	21	26
13	IBM	United States	19	25
	National University of Singapore	Singapore	19	23
	Tsinghua University	China	19	21
16	Michigan State University	United States	18	52
17	University of Toronto	Canada	17	24
18	Rutgers University	United States	16	33
	University of Amsterdam	Netherlands	16	29
	Brown University	United States	16	24
	Johns Hopkins University	United States	16	23
	University of California, Santa Barbara	United States	16	21
23	University of South Florida	United States	15	26
24	University of California, Los Angeles	United States	14	33
	University of Oxford	United Kingdom	14	26
	University of Groningen	Netherlands	14	25
	Rensselaer Polytechnic Institute	United States	14	23
	Pennsylvania State University	United States	14	22
	University of California, Berkeley	United States	14	22
30	University of California, San Diego	United States	13	26
31	Hong Kong University of Science and Technology	Hong Kong	12	31
	Catholic University of Leuven	Belgium	12	18
	Tel Aviv University	Israel	12	17
	University of Washington	United States	12	14
	Imperial College London	United Kingdom	12	12
36	University of Southern California	United States	11	32
	Weizmann Institute of Science	Israel	11	30
	University of Cambridge	United Kingdom	11	24
	Boston University	United States	11	23
	Yale University	United States	11	23
	State University of New York at Buffalo	United States	11	21
	University of Florida	United States	11	18
	University of Florence	Italy	11	17
	Purdue University	United States	11	15
	National Technical University of Athens	Greece	11	14
	Nanyang Technological University	Singapore	11	13
	National Cheng Kung University	Taiwan	11	12
48	Heriot-Watt University	United Kingdom	10	13
	Autonomous University of Barcelona	Spain	10	12
	GE	United States	10	12
	University of Minnesota	United States	10	12
52	Columbia University	United States	9	33
	Hebrew University of Jerusalem	Israel	9	33
	Duke University	United States	9	20
	IDIAP Research Institute	Switzerland	9	18
	University of Central Florida	United States	9	15
	Osaka University	Japan	9	10
	Stanford University	United States	9	10
	University of Chicago	United States	9	10

Table A-2. Detected researchers as sorted by doctoral degree holders on computer vision.

Ranking	Insitute	Country	Number of detected reserchers	Total number of authors
1	Massachusetts Institute of Technology	United States	76	124
2	University of Maryland, College Park	United States	38	77
3	Carnegie Mellon University	United States	35	60
	University of Illinois at Urbana-Champaign	United States	35	51
5	University of Oxford	United Kingdom	29	56
	University of Tokyo	Japan	29	48
7	Stanford University	United States	26	55
8	University of California, Berkeley	United States	25	35
9	University of Southern California	United States	22	56
10	Michigan State University	United States	19	21
	University of Paris VI	France	19	21
12	Cornell University	United States	18	36
13	University of Cambridge	United Kingdom	17	35
	Purdue University	United States	17	28
	Brown University	United States	17	21
16	Harvard University	United States	16	36
	University of Paris XI	France	16	26
	California Institute of Technology	United States	16	23
	University of Amsterdam	Netherlands	16	21
20	Hebrew University of Jerusalem	Israel	15	35
	University of Washington	United States	15	24
22	University of Toronto	Canada	14	34
	Technion, Israel Institute of Technology	Israel	14	32
	University of Pennsylvania	United States	14	26
	Princeton University	United States	14	17
	Swiss Federal Institute of Technology Lausanne	Switzerland	14	16
	Yale University	United States	14	15
28	State University of New York at Buffalo	United States	13	23
	University of Edinburgh	United Kingdom	13	21
	Georgia Institute of Technology	United States	13	14
31	Ohio State University	United States	12	46
	University of Massachusetts	United States	12	26
	Columbia University	United States	12	22
	Institut National Polytechnique de Grenoble	France	12	19
	Washington University in St. Louis	United States	12	18
	University of London	United Kingdom	12	17
	Imperial College London	United Kingdom	12	16
38	University of Texas at Austin	United States	11	19
	University of California, Los Angeles	United States	11	15
	University of Rennes I	France	11	14
	Rensselaer Polytechnic Institute	United States	11	11
42	Catholic University of Leuven	Belgium	10	19
	Weizmann Institute of Science	Israel	10	16
	Rutgers University	United States	10	14
	INRIA	France	10	13
46	Delft University of Technology	Netherlands	9	16
	Duke University	United States	9	16
	University of Valencia	Spain	9	15
	University of Wisconsin-Madison	United States	9	13
	Chinese Academy of Sciences	China	9	12
	National Technical University of Athens	Greece	9	11
	Osaka University	Japan	9	11
	Tsinghua University	China	9	11
	University of Florida	United States	9	11

Table A-3. Detected researchers as sorted by master's degree holders on computer vision.

Ranking	Insitute	Country	Number of detected reserchers	Total number of authors
1	Massachusetts Institute of Technology	United States	48	79
2	Tsinghua University	China	40	56
3	Technion, Israel Institute of Technology	Israel	24	37
	Chinese Academy of Sciences	China	24	29
5	University of Tokyo	Japan	23	31
6	Stanford University	United States	21	42
7	University of Maryland, College Park	United States	19	53
	University of Illinois at Urbana-Champaign	United States	19	23
9	Hebrew University of Jerusalem	Israel	14	38
	Weizmann Institute of Science	Israel	14	26
	Georgia Institute of Technology	United States	14	15
12	Delft University of Technology	Netherlands	13	24
	McGill University	Canada	13	20
	Imperial College London	United Kingdom	13	18
15	University of Southern California	United States	12	37
	Korea Advanced Institute of Science and Technology	South Korea	12	18
17	Indian Institute of Science Bangalore	India	11	24
	Cornell University	United States	11	21
	Catholic University of Leuven	Belgium	11	19
	University of Toronto	Canada	11	16
	Xi'an Jiaotong University	China	11	13
	National Cheng Kung University	Taiwan	11	12
	Rensselaer Polytechnic Institute	United States	11	11
24	Tel Aviv University	Israel	10	19
	University of Cambridge	United Kingdom	10	13
	Chinese University of Hong Kong	Hong Kong	10	12
	Peking University	China	10	11
	University of California, Santa Barbara	United States	10	11
29	Washington University in St. Louis	United States	9	15
	University of California, Berkeley	United States	9	14
	Johns Hopkins University	United States	9	13
	Carnegie Mellon University	United States	9	12
	Brown University	United States	9	11
	Kyoto University	Japan	9	11
	Michigan State University	United States	9	10
	Zhejiang University	China	9	10
37	University of Pennsylvania	United States	8	21
	Columbia University	United States	8	13
	University of Wisconsin-Madison	United States	8	13
	National University of Singapore	Singapore	8	11
	University of Paris VI	France	8	11
	Osaka University	Japan	8	10
43	Politehnica University of Bucharest	Romania	7	14
	Purdue University	United States	7	11
	Shanghai Jiao Tong University	China	7	10
	University of Massachusetts	United States	7	10
	University of Washington	United States	7	10
	Swiss Federal Institute of Technology Lausanne	Switzerland	7	9
	Bogazici University	Turkey	7	8
	University of Science and Technology of China	China	7	8
	Middle East Technical University	Turkey	7	7

Table A-4. Detected researchers as sorted by bachelor's degree holders on computer vision.

Ranking	Insitute	Country	Number of detected reserchers	Total number of authors
1	Tsinghua University	China	61	77
2	University of Science and Technology of China	China	38	66
3	Technion, Israel Institute of Technology	Israel	30	46
4	Hebrew University of Jerusalem	Israel	26	64
5	Massachusetts Institute of Technology	United States	24	41
6	University of Cambridge	United Kingdom	21	49
	University of Tokyo	Japan	21	30
	Zhejiang University	China	21	29
9	National Technical University of Athens	Greece	20	32
	Peking University	China	20	25
11	Xi'an Jiaotong University	China	17	24
12	Seoul National University	South Korea	16	24
13	McGill University	Canada	15	22
14	Tel Aviv University	Israel	14	32
	University of Oxford	United Kingdom	14	21
16	University of Padua	Italy	13	28
	Kyoto University	Japan	13	25
	Ecole Polytechnique	France	13	16
	National Taiwan University	Taiwan	13	16
20	Indian Institute of Technology Kanpur	India	12	57
21	Chinese University of Hong Kong	Hong Kong	11	26
	Indian Institute of Technology Kharagpur	India	11	26
	California Institute of Technology	United States	11	16
	Nanjing University	China	11	15
25	Princeton University	United States	10	20
	Harvard University	United States	10	11
	Indian Institute of Technology Bombay	India	10	11
28	TELECOM ParisTech	France	9	25
	University of Illinois at Urbana-Champaign	United States	9	12
	Shanghai Jiao Tong University	China	9	10
	University of Karlsruhe	Germany	9	10
32	University of Bonn	Germany	8	15
	Indian Institute of Technology Madras	India	8	14
	Birla Institute of Technology and Science	India	8	13
	University of Manchester	United Kingdom	8	13
	National Chiao Tung University	Taiwan	8	9
	Indian Institute of Technology Delhi	India	8	8
38	Aristotle University of Thessaloniki	Greece	7	14
	University of Florence	Italy	7	14
	Cornell University	United States	7	12
	Carnegie Mellon University	United States	7	10
	University of California, Berkeley	United States	7	10
	Middle East Technical University	Turkey	7	9
	Osaka University	Japan	7	9
	Polytechnic University of Catalonia	Spain	7	9
	Bogazici University	Turkey	7	8
	Yonsei University	South Korea	7	8
	Nankai University	China	7	7
	Tianjin University	China	7	7
	Tohoku University	Japan	7	7

Table B-1. Detected researchers as sorted by affiliation on robotics.

Ranking	Insitute	Country	Number of detected reserchers	Total number of authors
1	University of Tokyo	Japan	54	76
2	Massachusetts Institute of Technology	United States	37	47
3	University of Karlsruhe	Germany	32	37
4	National Institute of Advanced Industrial Science and	Japan	30	34
5	Carnegie Mellon University	United States	27	31
6	Stanford University	United States	24	31
7	National University of Singapore	Singapore	23	29
8	ATR	Japan	22	36
	Scuola Superiore Sant'Anna	Italy	22	33
10	University of Malaga	Spain	21	36
	Osaka University	Japan	21	34
	Swiss Federal Institute of Technology Lausanne	Switzerland	21	32
	Johns Hopkins University	United States	21	27
	Nanyang Technological University	Singapore	21	27
15	University of British Columbia	Canada	20	25
	Vanderbilt University	United States	20	25
17	University of Zaragoza	Spain	19	42
	INRIA	France	19	26
	University of Toronto	Canada	19	20
	Orebro University	Sweden	18	35
	University of Illinois at Urbana-Champaign	United States	18	26
22	Catholic University of Leuven	Belgium	17	25
	Korea Advanced Institute of Science and Technolog	South Korea	17	18
24	Northwestern University	United States	16	21
	University of Oxford	United Kingdom	16	21
	University of Sydney	Australia	16	21
	Technical University of Munich	Germany	16	19
28	University of Pennsylvania	United States	15	22
29	Nara Institute of Science and Technology	Japan	14	23
	University of Rome	Italy	14	19
	University of Delaware	United States	14	18
	CNRS	France	14	14
33	Polytechnic University of Catalonia	Spain	13	28
34	Swiss Federal Institute of Technology Zurich	Switzerland	12	20
	Cornell University	United States	12	19
	Free University of Brussels	Belgium	12	18
	McGill University	Canada	12	17
	University Carlos III of Madrid	Spain	12	16
	Polytechnic University of Madrid	Spain	12	14
	Tsinghua University	China	12	13
	Peking University	China	12	12
	University of Alcala	Spain	12	12
	University of Zurich	Switzerland	12	12
44	Royal Institute of Technology	Sweden	11	20
	Purdue University	United States	11	13
	University of Paris VI	France	11	12
	Sakarya University	Turkey	11	11
48	University of Michigan	United States	10	19
	University of Freiburg	Germany	10	18
	Technion, Israel Institute of Technology	Israel	10	17
	University of California, Berkeley	United States	10	15
	University of Southern California	United States	10	15
	City University of Hong Kong	Hong Kong	10	13
	CSIRO	Australia	10	12
	Imperial College London	United Kingdom	10	11
	University of New South Wales	Australia	10	10

Table B-2. Detected researchers as sorted by doctoral degree holders on robotics.

Ranking	Insitute	Country	Number of detected reserchers	Total number of authors
1	University of Tokyo	Japan	73	98
2	Massachusetts Institute of Technology	United States	40	49
3	Stanford University	United States	33	49
4	Carnegie Mellon University	United States	29	44
5	Tokyo Institute of Technology	Japan	26	31
6	Polytechnic University of Madrid	Spain	21	28
7	University of California, Berkeley	United States	19	31
	University of Pennsylvania	United States	19	29
9	California Institute of Technology	United States	17	27
	University of Toronto	Canada	17	21
11	Osaka University	Japan	16	28
	University of Oxford	United Kingdom	16	22
	Institut National Polytechnique de Grenoble	France	16	20
14	Catholic University of Leuven	Belgium	15	22
15	Polytechnic University of Catalonia	Spain	14	35
	Nagoya University	Japan	14	18
	University of Sydney	Australia	14	17
18	McGill University	Canada	13	23
	University of Michigan	United States	13	19
20	University of Zaragoza	Spain	12	31
	Kyoto University	Japan	12	29
	University of Bonn	Germany	12	21
	University of Illinois at Urbana-Champaign	United States	12	16
	University of Paris VI	France	12	16
	University of Maryland, College Park	United States	12	12
26	University of Southern California	United States	11	24
	University of Malaga	Spain	11	21
	University of Bologna	Italy	11	17
	Technical University of Munich	Germany	11	16
	University of Karlsruhe	Germany	11	13
31	Harvard University	United States	10	14
	Scuola Superiore Sant'Anna	Italy	10	12
33	University of Rennes I	France	9	21
	Purdue University	United States	9	16
	Swiss Federal Institute of Technology Zurich	Switzerland	9	14
	Johns Hopkins University	United States	9	13
	Georgia Institute of Technology	United States	9	10
	Tsinghua University	China	9	10
	University of Nantes	France	9	10
40	University of Edinburgh	United Kingdom	8	29
	Royal Institute of Technology	Sweden	8	17
	Yale University	United States	8	14
	Korea Advanced Institute of Science and Technolog	South Korea	8	13
	University of Texas at Austin	United States	8	12
	Rensselaer Polytechnic Institute	United States	8	10
	Complutense University of Madrid	Spain	8	9
	University of British Columbia	Canada	8	9

Table B-3. Detected researchers as sorted by master's degree holders on robotics.

Ranking	Insitute	Country	Number of detected reserchers	Total number of authors
1	University of Tokyo	Japan	58	74
2	Massachusetts Institute of Technology	United States	38	53
3	Osaka University	Japan	24	30
4	Tokyo Institute of Technology	Japan	22	27
5	Stanford University	United States	21	28
6	Korea Advanced Institute of Science and Technology	South Korea	16	24
7	Carnegie Mellon University	United States	15	21
8	Seoul National University	South Korea	14	15
9	University of Pennsylvania	United States	13	19
	Sharif University of Technology	Iran	13	17
	University of Illinois at Urbana-Champaign	United States	13	16
12	Kyoto University	Japan	12	28
	University of Michigan	United States	12	18
	University of Pisa	Italy	12	18
	Johns Hopkins University	United States	12	17
	University of California, Berkeley	United States	12	17
	Tsinghua University	China	12	15
18	Swiss Federal Institute of Technology Lausanne	Switzerland	11	18
	Catholic University of Leuven	Belgium	11	17
	Nagoya University	Japan	11	15
21	University of Malaga	Spain	10	23
	University of Zaragoza	Spain	10	19
	Technion, Israel Institute of Technology	Israel	10	13
	Nara Institute of Science and Technology	Japan	10	12
	Waseda University	Japan	10	11
26	Royal Institute of Technology	Sweden	9	16
	Polytechnic University of Madrid	Spain	9	10
28	University of Southern California	United States	8	21
	Rensselaer Polytechnic Institute	United States	8	17
	Ohio State University	United States	8	15
	Tohoku University	Japan	8	10
	Polytechnic University of Valencia	Spain	8	9
	Brigham Young University	United States	8	8
	Center of Research and Advanced Studies of the Nati	Mexico	8	8
	Zhejiang University	China	8	8
36	Harvard University	United States	7	10
	Pohang University of Science and Technology	South Korea	7	10
	Swiss Federal Institute of Technology Zurich	Switzerland	7	10
	Hiroshima University	Japan	7	9
	Georgia Institute of Technology	United States	7	8
	Peking University	China	7	8
	University of British Columbia	Canada	7	7
	University of Toronto	Canada	7	7
44	University of California, Los Angeles	United States	6	11
	Indian Institute of Science Bangalore	India	6	8
	Linkoping University	Sweden	6	8
	Middle East Technical University	Turkey	6	8
	University of Karlsruhe	Germany	6	8
	University of Maryland, College Park	United States	6	7
	University of Rome	Italy	6	7
	Imperial College London	United Kingdom	6	6
	Kyushu University	Japan	6	6
	Lund University	Sweden	6	6
	University of Cambridge	United Kingdom	6	6
	University of Utah	United States	6	6
	Vanderbilt University	United States	6	6

Table B-4. Detected researchers as sorted by bachelor's degree holders on robotics.

Ranking	Insitute	Country	Number of detected reserchers	Total number of authors
1	University of Tokyo	Japan	53	70
2	Seoul National University	South Korea	29	37
3	Osaka University	Japan	20	27
	Massachusetts Institute of Technology	United States	20	22
5	Sharif University of Technology	Iran	19	25
	Tsinghua University	China	19	22
7	Polytechnic University of Madrid	Spain	18	22
	University of Pisa	Italy	18	21
9	Kyoto University	Japan	17	35
	Middle East Technical University	Turkey	15	19
11	Technion, Israel Institute of Technology	Israel	14	21
	University of Karlsruhe	Germany	14	17
13	National Technical University of Athens	Greece	13	22
	University of Toronto	Canada	13	19
	Tokyo Institute of Technology	Japan	13	15
16	Catholic University of Leuven	Belgium	12	21
	Complutense University of Madrid	Spain	12	19
	University of Rome	Italy	12	13
19	Technical University of Munich	Germany	11	26
	Nagoya University	Japan	11	14
	National University of San Juan	Argentina	11	11
	Zhejiang University	China	11	11
23	University of Padua	Italy	10	24
	University of Bologna	Italy	10	15
	University of Michigan	United States	10	13
	Harbin Institute of Technology	China	10	12
	Carnegie Mellon University	United States	10	11
	Waseda University	Japan	10	11
29	Polytechnic University of Catalonia	Spain	9	27
	University of California, Los Angeles	United States	9	14
	Peking University	China	9	11
	Brigham Young University	United States	9	9
33	University of Oxford	United Kingdom	8	13
	Polytechnic University of Milan	Italy	8	12
	Shanghai Jiao Tong University	China	8	10
	Swiss Federal Institute of Technology Zurich	Switzerland	8	10
	Johns Hopkins University	United States	8	9
	Technical University of Istanbul	Turkey	8	8
	University of Cambridge	United Kingdom	8	8
	University of Science and Technology of China	China	8	8
41	University of Bonn	Germany	7	12
	Korea Advanced Institute of Science and Technolog	South Korea	7	10
	University of British Columbia	Canada	7	9
	University of Malaga	Spain	7	9
	National University of Singapore	Singapore	7	8
	Cornell University	United States	7	7
50	University of Seville	Spain	6	13
	University of Zaragoza	Spain	6	11
	Pohang University of Science and Technology	South Korea	6	9
	Bogazici University	Turkey	6	8
	National Taiwan University	Taiwan	6	8
	Tianjin University	China	6	8
	Huazhong University of Science and Technology	China	6	7
	Princeton University	United States	6	7
	University of California, Berkeley	United States	6	7
	University of Ljubljana	Slovenia	6	7
	University of Queensland	Australia	6	6
	University of Waterloo	Canada	6	6
	Yonsei University	South Korea	6	6

Table C-1. Detected researchers as sorted by affiliation on electron devices.

Ranking	Insitute	Country	Number of detected reserchers	Total number of authors
1	IMEC	Belgium	54	101
2	Toshiba	Japan	51	52
3	National Chiao Tung University	Taiwan	48	71
4	IBM	United States	43	47
5	Arizona State University	United States	39	47
	Massachusetts Institute of Technology	United States	39	43
7	Peking University	China	32	51
	Chinese Academy of Sciences	China	32	39
9	Stanford University	United States	30	42
10	University of Electronic Science and Technology of	China	29	52
	Taiwan Semiconductor Manufacturing Company Lim	Taiwan	29	36
12	MINATEC	France	28	50
13	Nanyang Technological University	Singapore	27	45
	National Cheng Kung University	Taiwan	27	35
	Seoul National University	South Korea	27	33
16	University of Cambridge	United Kingdom	25	31
	National Tsing Hua University	Taiwan	25	28
18	NXP	Netherlands	23	30
19	Samsung	South Korea	22	22
20	National University of Singapore	Singapore	21	42
	Renesas Technology Corporation	Japan	21	21
22	Korea Advanced Institute of Science and Technology	South Korea	20	52
	NEC	Japan	20	25
	Hitachi	Japan	20	21
	National Taiwan University	Taiwan	20	21
26	Purdue University	United States	19	31
27	STMicroelectronics	Italy	18	33
28	University of California, Berkeley	United States	16	29
	Indian Institute of Technology Bombay	India	16	21
	National Institute of Advanced Industrial Science and	Japan	16	16
31	Chang Gung University	Taiwan	15	16
	French Atomic Energy Commission	France	15	15
33	US Navy	United States	14	23
	Agency for Science, Technology and Research	Singapore	14	21
35	Institute of Microelectronics, Singapore	Singapore	13	24
	Sony	Japan	13	22
	University of Tokyo	Japan	13	16
	United Microelectronic Corporation	Taiwan	13	15
39	University of Glasgow	United Kingdom	12	17
	Intel	India	12	13
	DALSA	United States	12	12
42	Ferdinand-Braun-Institut	Germany	11	14
	Gwangju Institute of Science and Technology	South Korea	11	13
	Polytechnic University of Milan	Italy	11	13
	Pennsylvania State University	United States	11	11
46	Texas Instruments	United States	10	14
	Georgia Institute of Technology	United States	10	13
	University of Texas at Austin	United States	10	11
	Applied Materials Inc	United States	10	10
	Chungnam National University	South Korea	10	10
	National Institute of Standards and Technology	United States	10	10
	Semiconductor Leading Edge Technologies, Inc.	Japan	10	10

Table C-2. Detected researchers as sorted by doctoral degree holders on electron devices.

Ranking	Insitute	Country	Number of detected reserchers	Total number of authors
1	Stanford University	United States	39	57
2	University of Tokyo	Japan	34	40
3	National Chiao Tung University	Taiwan	33	43
4	University of California, Berkeley	United States	30	69
5	Catholic University of Leuven	Belgium	24	50
6	Purdue University	United States	20	32
	Arizona State University	United States	20	24
	Seoul National University	South Korea	20	24
	University of Michigan	United States	20	24
10	University of Cambridge	United Kingdom	19	22
11	National Cheng Kung University	Taiwan	18	21
	University of Texas at Austin	United States	18	20
13	Osaka University	Japan	17	18
14	Institut National Polytechnique de Grenoble	France	15	25
	Cornell University	United States	15	19
	University of Paris XI	France	15	18
17	Tohoku University	Japan	14	21
	University of Illinois at Urbana-Champaign	United States	14	19
	Massachusetts Institute of Technology	United States	14	17
	North Carolina State University	United States	14	17
	Tokyo Institute of Technology	Japan	14	17
	University of Maryland, College Park	United States	14	17
23	National University of Singapore	Singapore	12	19
	Chinese Academy of Sciences	China	12	15
25	University of Electronic Science and Technology of	China	11	26
	National Taiwan University	Taiwan	11	14
	Waseda University	Japan	11	11
28	INSA	France	10	16
	University of Lille	France	10	14
	University of Stuttgart	Germany	10	11
31	Korea Advanced Institute of Science and Technology	South Korea	9	14
	University of Florida	United States	9	14
33	Polytechnic University of Milan	Italy	8	20
	Kyoto University	Japan	8	12
	Rensselaer Polytechnic Institute	United States	8	11
	National Tsing Hua University	Taiwan	8	10
	University of California, Los Angeles	United States	8	10
	Harvard University	United States	8	8
39	Tsinghua University	China	7	10
	University of Arizona	United States	7	7
41	University of Bologna	Italy	6	15
	Lehigh University	United States	6	10
	Nanyang Technological University	Singapore	6	10
44	Eindhoven University of Technology	Netherlands	6	8
	Lund University	Sweden	6	8
	Georgia Institute of Technology	United States	6	7
	RWTH Aachen University	Germany	6	7
	Technical University of Munich	Germany	6	6

Table C-3. Detected researchers as sorted by master's degree holders on electron devices.

Ranking	Insitute	Country	Number of detected reserchers	Total number of authors
1	National Chiao Tung University	Taiwan	51	66
2	University of Tokyo	Japan	43	51
3	Seoul National University	South Korea	39	56
4	Stanford University	United States	34	49
5	National Tsing Hua University	Taiwan	33	40
6	National Cheng Kung University	Taiwan	27	32
7	Tokyo Institute of Technology	Japan	24	31
8	University of California, Berkeley	United States	23	34
9	Korea Advanced Institute of Science and Technology	South Korea	20	45
	Tohoku University	Japan	20	24
	Osaka University	Japan	20	22
	Waseda University	Japan	20	20
13	Catholic University of Leuven	Belgium	19	40
14	Kyoto University	Japan	17	21
15	INSA	France	16	20
16	National Taiwan University	Taiwan	15	18
17	Arizona State University	United States	14	14
18	University of Michigan	United States	13	15
19	Institut National Polytechnique de Grenoble	France	12	19
20	Technical University of Munich	Germany	11	11
21	National University of Singapore	Singapore	10	19
	Yonsei University	South Korea	10	16
	Massachusetts Institute of Technology	United States	10	11
24	University of Florida	United States	9	13
	Indian Institute of Technology Bombay	India	9	12
	Chinese Academy of Sciences	China	9	11
	Shanghai Jiao Tong University	China	9	11
	University of Naples Federico II	Italy	9	11
	National Central University	Taiwan	9	9
30	Polytechnic University of Milan	Italy	8	21
	University of Catania	Italy	8	15
	Peking University	China	8	12
	Tsinghua University	China	8	10
	University of Delhi	India	8	9
	Chang Gung University	Taiwan	8	8
	Chemnitz University of Technology	Germany	8	8
	University of Texas at Austin	United States	8	8
38	University of Electronic Science and Technology of	China	7	14
	Cornell University	United States	7	9
	Nagoya University	Japan	7	9
	University of Calcutta	India	7	9
	University of Stuttgart	Germany	7	8
	Delft University of Technology	Netherlands	7	7
	University of Cambridge	United Kingdom	7	7
45	Indian Institute of Technology Madras	United States	6	8
	Keio University	Japan	6	8
	Lund University	Sweden	6	8
	Chungbuk National University	South Korea	6	7
	Fudan University	China	6	7
	Georgia Institute of Technology	United States	6	7
	Gwangju Institute of Science and Technology	South Korea	6	7
	RWTH Aachen University	Germany	6	7
	Xi'an Jiaotong University	China	6	7

Table C-4. Detected researchers as sorted by bachelor's degree holders on electron devices.

Ranking	Insitute	Country	Number of detected reserchers	Total number of authors
1	Seoul National University	South Korea	63	84
2	University of Tokyo	Japan	47	56
3	National Cheng Kung University	Taiwan	41	50
4	National Chiao Tung University	Taiwan	40	56
5	National Taiwan University	Taiwan	39	52
6	National Tsing Hua University	Taiwan	32	39
7	Tsinghua University	China	27	38
8	Peking University	China	26	38
9	National University of Singapore	Singapore	22	44
10	Kyoto University	Japan	20	24
11	Waseda University	Japan	19	20
12	Osaka University	Japan	17	19
13	Tohoku University	Japan	16	17
14	Yonsei University	South Korea	15	23
15	Polytechnic University of Milan	Italy	14	21
	Kyungpook National University	South Korea	14	16
17	Korea University	South Korea	13	19
	University of Electronic Science and Technology of	China	13	19
	University of California, Berkeley	United States	13	15
20	Nanjing University	China	12	22
21	Nanyang Technological University	Singapore	11	18
	Korea Advanced Institute of Science and Technology	South Korea	11	17
	Tokyo Institute of Technology	Japan	11	12
24	Massachusetts Institute of Technology	United States	9	13
	Shanghai Jiao Tong University	China	9	9
26	Catholic University of Leuven	Belgium	8	17
	University of Bologna	Italy	8	16
	University of Science and Technology of China	China	8	12
	Chungnam National University	South Korea	8	10
	Keio University	Japan	8	10
	National Central University	Taiwan	8	10
	Indian Institute of Technology Madras	United States	8	9
	North Carolina State University	United States	8	9
	University of Delhi	India	8	9
	Chung Yuan Christian University	Taiwan	8	8
	Indian Institute of Technology	India	8	8
	University of Naples Federico II	Italy	8	8
38	Jilin University	China	7	11
	University of Calcutta	India	7	11
	Xi'an Jiaotong University	China	7	9
	Feng Chia University	Taiwan	7	8
	NIT India	India	7	8
	University of California, Los Angeles	United States	7	7
	University of Wisconsin-Madison	United States	7	7
45	Nagoya University	Japan	6	8
	Southeast University	China	6	8
	University of Michigan	United States	6	8
	Chungbuk National University	South Korea	6	7
	Middle East Technical University	Turkey	6	7
	Technion, Israel Institute of Technology	Israel	6	6
	University of Tehran	Iran	6	6
	Zhejiang University	China	6	6

A Fully Automated Method for the Unification of Funding Organizations in the Web of Knowledge[1]

Daniel Sirtes* and Mathias Riechert*

*sirtes@forschungsinfo.de; riechert@forschungsinfo.de
iFQ – Institute for Research Information and Quality Assurance, Schützenstraße 6a, D-10177 Berlin (Germany)

Introduction

As of August 2008 Thomson Reuters includes funding acknowledgements in their Web of Knowledge (WoK) database. One of the major obstacles in using this resource for assessing the output of funding organizations (FO) is the vast amount of aliases included in the funding organizations list (over 10000 entries for the German Research Foundation (DFG)). Numerous further problems with FO data have been discussed previously (see Costas & Yegros-Yegros, 2013; Rigby, 2011a, 2011b; Sirtes, 2013; Yegros-Yegros & Costas, 2013). This proof-of-concept paper presents a highly efficient, precise and fully automated method with minimal manual configuration to unify many of these aliases and almost all of the publications associated with a funding organization.

One of the things that we have learned from our semi-automated method developed for cleaning the DFG data was that many aliases include only the sub-programme of the DFG instead of the German Research Foundation itself (see (Sirtes, 2013)). However, we have also realized that in many cases these aliases appear together with the DFG acronym (e.g. DFG cluster of excellence). This circumstance led to the idea of fishing for the different names and sub-programmes of a funding organization with the help of its acronym(s).

Method

The approach incorporates findings from previous studies on funding acknowledgement data in the WoK with the help of the in-house database developed by the Competence Center for Bibliometrics. In order to further stimulate the debate on which steps should be included in automated FA data cleaning, they are described explicitly:

As previously proposed (Sirtes, 2013; Wang & Shapira, 2011), we first create a thesaurus of funding organization aliases.

(1) Get all WoK funding organizations aliases that were used in more than 80 publication items.
(2) Extract abbreviations out of the funding organization text by using a regular expression that selects strings with at least two capital letters. The resulting list was manually reviewed, and 23 combinations were blacklisted as they are not funding organizations (e.g. USA).
(3) Extract a list of unified short funding organization texts (USFO) from the original funding organization field by removing commas, brackets, hyphens and the abbreviations themselves. Furthermore, we only consider terms that include at least one blank space, as we are searching for the long form of the abbreviations or full

[1] This work was supported by the German Federal Ministry of Education and Research (BMBF, project number: 01PQ08004A).

names of sub-programmes. Additionally, for each of the USFO, the dominant country of all items stating that FO is computed, to ensure that the term is not used by multiple organizations in different countries[2]. Out of all 549 unified funding organization texts, two had differing dominant countries: “Chinese Academy of Sciences (LAMOST)” (with China and USA) and “Cancer Institute” (with USA and Australia).

(4) Assign the 549 USFO to the 668 abbreviations by identifying terms containing both the funding organization text and the abbreviation (for example: “Deutsche Forschungsgemeinschaft (DFG)”. This results in a thesaurus with abbreviation-USFO mappings ranging from 16 different variants (for example NIH) to only abbreviations (for example IDRIS), where no corresponding text was found. As in the previous step, dominant countries are controlled for to address synonyms. Out of the 668 abbreviation-USFO mappings, 12 differ concerning the dominant country of the mapped USFO (most prominently the America NSF and NSF of China).

Building on the thesaurus the search for funding organizations and publication items stating them is implemented as a VB.net program combining two search strategies. The first strategy searches for the abbreviation with regular expressions. In order to include misspellings in funding organizations (which cumulate to about 40% of all funding organization texts) we additionally compute Levenshtein distances of possible variants of the WoK funding organization text and the USFO texts in the second search strategy.

(5) Search in the WoK funding organization full text (FFO) for the extracted abbreviations using regular expression search patterns. The abbreviation can be surrounded by a non-character letter or build the start or the end of the term. This is computed for all 668 abbreviations. The found FFOs are then inserted into a mapping table, if the abbreviation has a common dominant country (see last step). For those abbreviations with multiple dominant countries, only the second search strategy can be applied in order to prevent semantic mismatching.

(6) Search in the WoK FFO for the extracted USFOs: Each USFO text is split and each part-term’s first two letters are used as the basis for the regular expression search. Searching for “Deutsche Forschunsgemeinschaft” for example, uses the following regular expression string: “(^|\s|[[:punct:]])(De\w*.Fo\w*)”, meaning that the funding organization can be at the start of the term, after a blank space or punctuation character. This results in a wide range of different terms, which are possibly misspelled variants of the USFO. Then we compute the Levenshtein-wordlength-ratio (LWLR) to assess the closeness to the USFO. Out of the found funding organizations texts, only funding organizations with a LWLR<0.4 or 0.5 (different variants were tested) are added to a matching table. Again, multiple dominant countries are controlled for.

Finally, the results of both strategies are combined by matching those USFO according to the three possible dominant country combinations:

(7) If the mapping of the USFO to the abbreviations has only one dominant country (276 of 668 cases): In this case, all FOs of a unified short form have the same

[2] A dominant country is the country, which is most often associated with publications of a funding organization. We have also flagged the FOs with low share of the dominant country and whitelisted FOs that have a multinational nature, like the ERC.

dominant country[3]. Both search strategies can therefore be combined. Consequently, the mappings from the funding organization to abbreviations and the mappings from the funding organizations to the USFO texts are unified for each abbreviation.

(8) If the mapping of the USFO to the abbreviations has multiple dominant countries (12 of 668 cases): In this case, combining both search strategies would confound different organizations from different countries. Therefore, the USFO mappings from the same dominant countries are combined. Consequently, each FO/dominant country combination gets unified into an abbreviation with a country index (e.g. NSF_CHN for the National Science Foundation of China)[4].

(9) If there is only an abbreviation and no USFO text (381 of 668 cases): Only the abbreviation matches are used to assign funding organizations from the WoK.

Results and Outlook

DFG

We are in the fortunate position of having a complete manually cleaned list of all publications in the WoK associated with the DFG, aided by the semi-automated method described in Sirtes (2013). We have restricted our comparison with the data generated in our current fully automated method to the 21,963 publications from the year 2010. Our new fully automated method has associated 21,072 items with the DFG. Of these 21,072 we found 21,002 again in our manual set, which amounts to a recall of 95.6% and an astounding precision of 99.7%, with the University of Georgia Research Foundation as the top culprit with 14 false positive publications. However, if one compares the success of the method on the basis of FO aliases instead of publications, then the picture looks considerably grimmer, as most items are concentrated in a few aliases. Out of the 3,061 aliases for 2010, only 1,834 have been found, which amounts to a recall of 59.9%. The precision however, is again very high at 98.3% with 31 false positives. The high recall in publications compared to aliases is explained by two factors: First, the 1,227 missing aliases amount only to 2,201 publications, and second, 1,240 of these have a second funding organization alias associated with it, that is included in our set.

NIH

As the German Research Foundation is probably the funding organization with the most diverse list of aliases and is therefore extremely hard to unify, we compared our method to the largest funding organization in the database, the NIH. We used one external bit of knowledge to enhance our search, which is the list of national institutes with other acronyms than NIH itself. We used 26 of the 28 acronyms listed on the front page of www.nih.gov (we left out the two letter acronyms CC and OD). Thus, we have used the 27 acronyms (including NIH) and their associated texts in the most common occurrences to search for aliases in the WoK. We compared this dataset with 200 publications from 2010 that we have randomly picked from the NIH's own database of their publications: the NIH RePORTER (http://projectreporter.nih.gov/reporter.cfm). We have found 188 of these publications in the WoK (WoK recall: 94%). Out of these publications, 163 had a funding organization associated with it (Funding acknowledgement to NIH RePORTER recall of 81.5%, share ofWoK items with FA 86.7%). Our method was successful in finding 155 out of these 163, which amounts to a recall of 95.1%. However, 7 out of the false negatives did not credit the NIH at all (including one "funding organization" called 'Public Service Grant', which might

[3] Or the USFO is whitelisted as multinational.

[4] The FO aliases with less than 80 articles (or in future versions less than 30) cannot be used for this kind of homonymous FO names as of now there is no way to determine their dominant country.

or might not allude to the public health service grants of the NIH). Thus, our method caught 155 out of 156 publications that can be associated manually to the NIH i.e. a recall of 99.3%. A single publication has evaded our method due to the fact that some researchers are rather sloppy with the prepositions in the names of funding organizations, like ‘of’, ‘for’, etc.

To develop this very promising method further, we plan to leave these kind of words out of our regular expression search. Furthermore, we plan to lower our initial starting set to FOs with 30 instead of 80 publications per alias and possibly include grant number patterns in our query.

References

Costas, R. & Yegros-Yegros, A. (2013). Possibilities of funding acknowledgement analysis for the bibliometric study of research funding organizations: case study of the Austrian Science Fund (FWF). In *Proceedings of ISSI 2013* (pp. 1401–1408). Vienna, Austria.

Rigby, J. (2011a). Systematic Grant and Funding Body Acknowledgement Data for Publications: New Dimensions and New Controversies for Research Policy and Evaluation. *Research Evaluation, 20*(5), 365–375. doi:10.3152/095820211X13164389670392

Rigby, J. (2011b). Systematic Grant and Funding Body Acknowledgment Data for Publications: An Examination of New Dimensions and New Controversies for Bibliometrics. *SSRN eLibrary*. Retrieved from http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1742845

Sirtes, D. (2013). Funding Acknowledgements for the German Research Foundation (DFG). The Dirty Data of the Web of Science Database and How to Clean It Up. In J. Gorraiz, E. Schiebel, C. Gumpenberger, M. Hörlesberger, & H. Moed (Eds.), *Proceedings of the 14th International Society of Scientometrics and Informetrics Conference* (Vol. 1, pp. 784–795). Vienna, Austria: AIT GmbH.

Wang, J. & Shapira, P. (2011). Funding Acknowledgement Analysis: An Enhanced Tool to Investigate Research Sponsorship Impacts: The Case of Nanotechnology. *Scientometrics, 87*(3), 563–586. doi:10.1007/s11192-011-0362-5

Yegros-Yegros, A. & Costas, R. (2013). Analysis of the web of science funding acknowledgement information for the design of indicators on “external funding attraction.” In *Proceedings of the 14th International Society of Scientometrics and Informetrics Conference, 15th-19th July 2013* (pp. 84–95). Vienna, Austria.

Scholarly publication patterns in the social sciences and humanities and their coverage in Scopus and Web of Science

Gunnar Sivertsen

gunnar.sivertsen@nifu.no
Nordic Institute for Studies in Innovation, Research and Education
P.O. Box 5183 Majorstuen, N-0302 Oslo, Norway.

Introduction

The social sciences and humanities (SSH) are known to have a more limited coverage in Scopus and Web of Science than science, technology and medicine (STM) (Bourke & Butler, 1990; Hicks, 2004; Nederhof, 2006; Martin et al, 2010; van Leeuwen, 2013). 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 (Archambault et al, 2006; Engels, Ossenblok & Spruyt, 2012; Sivertsen & Larsen, 2012).

In recent years, there has been a positive trend towards broader coverage of the SSH in the international bibliographic databases. There has been an expansion in the number of journals covered in Scopus and Web of Science (WoS), especially in Scopus, where the coverage is now clearly beyond that of WoS. At the same time, Thomson Reuters has supplemented WoS with the Book Citation Index (BCI), while Elsevier has included scholarly books and book series from a selected set of publishers in Scopus.

In this study, we analyse the differences in coverage between the three mentioned data sources (Scopus, WoS, BCI). We also estimate the degree to which they deviate from a full coverage of all peer-reviewed scholarly publications. The focus is on what types journals and book publishers that make up for these differences. The purpose of the study is to discuss the options for an improved coverage of the SSH in international bibliographic data sources in the future.

Methods

Several European countries – among them Belgium (Flanders), Croatia, Czech Republic, Denmark, Estonia, Finland, Hungary, Norway, Portugal, Slovenia, Spain and Sweden – have established, or are in the advent of establishing, national current research information systems with complete, quality-assured bibliographic metadata for the country's scholarly publication output in the public sector. Such national information systems represent a potential for a more comprehensive coverage of the scholarly literature of the social sciences and humanities (Hicks & Wang, 2009) if connected to a scheme for institutional funding (Hicks, 2012).

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). A dataset of more than 70,000 scholarly publications from the eight years 2005-2012, 44 per cent of which are in the SSH, will be used here to study the publication patterns in the SSH and their coverage in Scopus and Web of Science.

In an earlier study (Ossenblok, Engels & Sivertsen, 2012), we compared the publication patterns in the SSH in two countries, Flanders (Belgium) and Norway, on the basis of data from similarly structured and defined comprehensive national data. We could confirm the observation in an earlier study (van Leeuwen, 2006) that publication patterns differ between the disciplines of the SSH while they are similar across countries. Even in disciplines with a nationally oriented publication pattern, the pattern itself is international. As an example, the publication pattern in sociology (degree of international publishing; percentage book publishing versus journal publishing; coverage of publications in the WoS) was much the same in the two countries and it also differed from that of economics in a similar way.

In this study, we assume that the disciplines of the SSH have specific publication patterns that are similar across countries. Although we will use data from only one country, the purpose of our study is more general. On the level of individual disciplines, we want to enlighten the situation in general with regard to the coverage of the scholarly publication patterns of the SSH. We also make use of the data to give special attention to the types of journals and book publishers that are not covered.

The methodology of the bibliographic data collection in the Norwegian Cristin database (www.cristin.no) has been published earlier (Sivertsen, 2010; Sivertsen & Larsen, 2012). 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. Three main types of peer-reviewed publications are registered: articles in ISSN-titles: journals, series, and yearbooks; articles in edited volumes (individual ISBN-titles), and monographs (individual ISBN-titles). All publication channels (journals, series, book publishers) are standardized in the database. This is the basis for the following independent variables used in this study:

- Publication type (article in journal or series; article in book; book)
- Publication channel (journal, series, book publisher)
- Field and subfield (based on a general classification of all journals and series, and a classification of each individual book)
- Language (“Norwegian” or “International”), based on the main language in the publication channel
- Scopus, WoS, or BCI coverage (publications from the specific publication channel was actively indexed in 2013).

Both law and psychology are included in the major field of social sciences in our study.

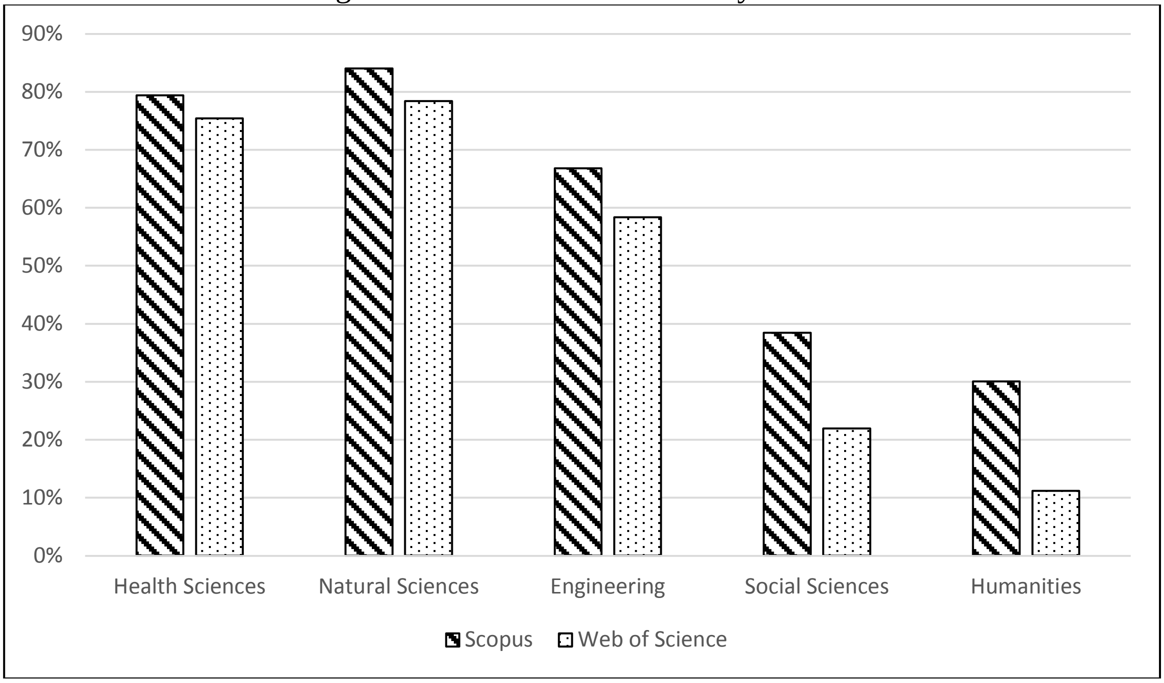
The last variable, data source coverage, is constructed in this way: A list of all publication channels (journals, series, book publishers) with at least one publication in the Norwegian database 2005-2012 was matched with the lists of publication channels that were actively indexed for Scopus, WoS or BCI in 2013. Our study therefore demonstrates to what extent the publishing patterns in the Norwegian higher education sector would be covered in the three data sources in 2013. We can thereby use a large dataset covering eight years to cover the more recent situation.

Publication counts are fractionalized between contributing institutions according to the number of authors from each of them. The three publication types mentioned above are given the same weight. An alternative could have been to give monographs more weight than articles, as in productivity studies, but we find equal weights more appropriate for studying the database coverage of scholarly publication patterns.

Results

Since the Norwegian data include all fields of research, we start by giving and overall comparison of SSH fields with STM fields in *figure 1*. Both book and journal publishing is included.

Figure 1: Coverage in *Scopus* and *Web of Science* of 70,500 scholarly publications from the higher education sector in Norway 2005-2012.



The general picture is that although Scopus has a broader coverage, the two data sources follow the same pattern in the representation of major fields. We can see from our data that the deficiencies in the SSH are mainly due to incomplete coverage of the international journals, limited or no coverage of national disciplinary journals and very limited coverage of peer-reviewed scholarly books.

In general, articles in journals and series represent only 48 % of the publications in the humanities in our data. This result can be compared to the fact that some 45% of all documents in the *Arts & Humanities Citation Index* (A&HCI) are book reviews (Zuccala & van Leeuwen, 2011). Books are also important in the social sciences, where 58 % of the publications are in journals or series. These shares can be compared to 89 % in the health sciences, 88 % in the natural sciences and 71 % in engineering. These differences are stable throughout the eight years 2005-2012.

Of all publications in the humanities, 48 per cent are in international languages. The corresponding shares are 53 per cent in the social sciences, 83 per cent in the health sciences, and 97 per cent in the natural sciences and engineering. These percentages are only slightly increasing during the period.

Tables 1 and 2 give a more detailed picture on the level of disciplines with regard to publications in *journals*. Note the large disciplinary variations in coverage. As explained in the introduction, we maintain that the differences in coverage of e.g. economics versus law is not only related to Norwegian scholarly publication practices or the coverage of Norway in Scopus and WoS, but rooted in differences in publication patterns (and the missions, subject areas and methods in the field) that are similar across all countries.

Table 1. Scopus and WoS coverage of disciplines in the humanities with special focus on journal publishing. Based on complete data for peer-reviewed publications in the Norwegian HE sector 2005-2012.

	All publications		Journal publications		
Subfield	Total	% in journals	% in int. language	% in WoS	% in Scopus
Classical Studies	259	66 %	48 %	23 %	26 %
Theatre and Drama	129	66 %	61 %	14 %	29 %
Linguistics	1057	61 %	93 %	39 %	41 %
Ethnology	392	57 %	47 %	12 %	16 %
Literature	764	57 %	31 %	16 %	18 %
Archaeology and Conservation	765	56 %	52 %	26 %	30 %
Slavonic Studies	231	56 %	84 %	17 %	43 %
Architecture and Design	424	54 %	38 %	11 %	24 %
Philosophy and History of Ideas	1121	54 %	45 %	28 %	33 %
Art History	278	54 %	45 %	21 %	25 %
Musicology	403	50 %	43 %	28 %	26 %
Theology and Religion	2126	50 %	42 %	16 %	34 %
History	1645	45 %	44 %	40 %	44 %
Media and Communication	1073	39 %	73 %	19 %	47 %
Asian and African Studies	237	39 %	99 %	42 %	49 %
Germanic Studies	238	38 %	100 %	39 %	37 %
Romance Studies	304	35 %	100 %	47 %	55 %
Scandinavian Studies	1777	35 %	17 %	2 %	2 %
English Studies	329	32 %	100 %	39 %	60 %
Total	**13551**	**49 %**	**52 %**	**23 %**	**32 %**

Table 2. Scopus and WoS coverage of disciplines in the social sciences (including law and psychology) with special focus on journal publishing. Based on complete data for peer-reviewed publications in the Norwegian HE sector 2005-2012.

	All publications		Journal publications		
Subfield	Total	% in journals	% in int. language	% in WoS	% in Scopus
Library and Information Science	389	83 %	98 %	56 %	80 %
Psychology	1940	79 %	79 %	66 %	72 %
Geography	853	78 %	86 %	72 %	78 %
Economics	1081	75 %	83 %	73 %	77 %
Business & Administration	2904	63 %	76 %	39 %	57 %
Law	2108	61 %	31 %	6 %	13 %
Anthropology	597	53 %	65 %	32 %	82 %
Gender Studies	358	48 %	38 %	19 %	37 %
Sociology	1157	46 %	60 %	40 %	48 %
Political Science	1655	45 %	76 %	64 %	73 %
Education & Educational Research	4861	43 %	51 %	22 %	35 %
Total	**17903**	**58 %**	**66 %**	**42 %**	**54 %**

We see that Scopus covers **32 per cent,** while Web of Science covers **23 per cent**, of all peer-reviewed scholarly articles in journals and series in the humanities from Norway's higher education institutions. The corresponding figures for the social sciences (including law and psychology) are **54 per cent** in Scopus versus **42 per cent** in Web of Science.

The difference between Scopus and WoS is not due to journals published in the Norwegian or Scandinavian languages. With very few exceptions, such journals are not covered in any of the data sources. The number of journals thereby not covered is small, since at the national level, many articles are concentrated in only a few journals (Sivertsen & Larsen, 2012). The difference between Scopus and WoS is that Scopus has a wider coverage of international journals in the SSH. There is no journal covered by WoS in the Norwegian data that is not also covered by Scopus.

While Scopus has a broader coverage of journals, *table 3* shows that the situation is the opposite with regard to scholarly publishing in books. Thomson Reuter's *Book Citation Index* covers 17 per cent of the peer-reviewed monographs and articles in edited volumes in the humanities, while Scopus covers only 5 per cent. The corresponding figures for the social sciences (including law and psychology) are 28 per cent in the Book Citation Index versus 7 per cent in Scopus.

Table 3. Scopus and Book Citation Index coverage of disciplines in the humanities with special focus on book publishing. Based on complete data for peer-reviewed publications in the Norwegian HE sector 2005-2012.

Major field	**Subfield**	**Publications in books**	**BCI**	**Scopus**
Humanities	Classical Studies	278	9 %	2 %
Humanities	Theatre and Drama	163	23 %	6 %
Humanities	Linguistics	129	33 %	19 %
Humanities	Ethnology	175	45 %	9 %
Humanities	Literature	441	23 %	8 %
Humanities	Archaeology and Conservation	80	32 %	12 %
Humanities	Slavonic Studies	799	13 %	3 %
Humanities	Architecture and Design	146	5 %	1 %
Humanities	Philosophy and History of Ideas	112	8 %	0 %
Humanities	Art History	324	41 %	19 %
Humanities	Musicology	266	7 %	1 %
Humanities	Theology and Religion	538	16 %	3 %
Humanities	History	163	21 %	2 %
Humanities	Media and Communication	1015	4 %	3 %
Humanities	Asian and African Studies	168	26 %	3 %
Humanities	Germanic Studies	93	26 %	1 %
Humanities	Romance Studies	44	0 %	0 %
Humanities	Scandinavian Studies	915	16 %	5 %
Humanities	English Studies	136	24 %	19 %
Humanities	**Total**	**5977**	**17 %**	**5 %**
Social Sciences	Library and Information Science	58	14 %	5 %
Social Sciences	Psychology	158	33 %	8 %
Social Sciences	Geography	161	14 %	4 %
Social Sciences	Economics	2275	8 %	3 %
Social Sciences	Business & Administration	720	20 %	9 %
Social Sciences	Law	228	48 %	22 %
Social Sciences	Anthropology	244	29 %	9 %
Social Sciences	Gender Studies	511	23 %	5 %
Social Sciences	Sociology	773	40 %	7 %
Social Sciences	Political Science	890	35 %	12 %
Social Sciences	Education & Educational Research	345	28 %	8 %
Social Sciences	**Total**	**6363**	**22 %**	**7 %**

Interestingly, we find that *Scopus* and Thomson Reuter's *Book Citation Index* has a broader coverage of Norwegian scientists's book publishing in the STM fields that they do in the

SSH, where the coverage is quite narrow, even of publications from prestigious international book publishers in these fields. Scopus has a very narrow selection of publishers that mainly operate in the STM market. The BCI has a better representation of the SSH, but the coverage of publishers still seems to in a starting phase, as mainly English language publishers have been selected.

Discussion

Throughout eight years, there is no sign that the use neither of national language nor of book publishing is decreasing significantly in the SSH. Using the Norwegian database, which registers publications on the level of individuals, we could also demonstrate that the normal publication pattern for an individual researcher in the SSH is to publish professionally in *both monographs, edited books, and journals*, and in a *minimum of two languages*, one of which is the national language.

The stability of the publication patterns and their differences in the SSH 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.

Since the publication patterns are relatively stable and rooted in the missions and methodologies of SSH, improved coverage in Scopus, WoS and BCI cannot to a large degree come from a change in the publication practices, but rather as changes in the databases themselves.

So far, the available bibliographic data sources that are used in library information systems for literature search have not been able to provide comprehensive metadata on an international level for the scholarly publication patterns of the SSH. Neither have the digital repositories for institutional archiving “green” open access to publications been able to provide this comprehensiveness yet. Most promising so far in this respect are the current research information systems (CRIS) with a quality assured production of metadata at the level of institutions (most common) or nations (still very few, but several countries are in a development phase).

CRIS systems on the institutional level have become widespread recently. They provide not only publication metadata, but also other types of information that can be used in the public interface and for management, statistics and assessment – as a kind of dynamic “annual report” which also serves as a communication channel and working space. Institutional CRIS systems both come in individual non-commercial solutions and through commercial products that are specifically designed for the purpose. The two dominant solutions on the European market has recently been acquired by Elsevier (PURE) and Thomson Reuters (Converis). These solutions are now integrated with Scopus and the management tool SciVal by Elsevier, and with Web of Science and InCites by Thomson Reuters. The effect is that the commercial providers are already working in the space between limited and comprehensive coverage of the humanities and social sciences. A probable consequence is that “disinterest” or “caution” in the treatment of the SSH in bibliometric studies will be replaced by an increasing interest, not only in performing new bibliometric studies of SSH, but also in the type validating studies demonstrated here.

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, 329–342.

Bourke, P., & Butler, L. (1996). Publication types, citation rates and evaluation. *Scientometrics*, 37(3), 473-494.

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.

Hicks, D. and Wang, J. (2009). Towards *a Bibliometric Database for the Social Sciences and Humanities. A European Scoping Project* (Appendix 1 to Martin et al., 2010). Arizona: School of Public Policy, Georgia University of Technology.

Hicks, D. (2012). Performance-based university research funding systems. *Research Policy*, 41, 251–261.

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: A report produced for DFG, ESRC, AHRC, NWO, ANR and ESF*. Sussex: Science and Technology Policy Research Unit.

Nederhof, A. J. (2006). Bibliometric monitoring of research performance in the social sciences and the humanities: A review. *Scientometrics*, 66(1), 81-100.

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.

van Leeuwen, T.N. (2006). The application of bibliometric analyses in the evaluation of social science research. Who benefits from it, and why it is still feasible, *Scientometrics*, 66, 133-154.

Zuccala, A and T.N. van Leeuwen (2011), Book reviews in humanities research evaluations, *JASIST*, 62, 1979-1991

van Leeuwen, T.N. (2013). Bibliometric research evaluations, Web of Science and the Social Sciences and Humanities: a problematic relationship? *Bibliometrie - Praxis & Forschung*, 1-18. http://www.bibliometrie-pf.de/article/viewFile/173/218.

The contribution of the postdoctoral fellows to the advancement of knowledge[1]

Held Barbosa de Souza* and Vincent Larivière*

* *held.bsouza@gmail.com; vincent.lariviere@umontreal.ca*
École de bibliothéconomie et des sciences de l'information, Université de Montréal,
C.P. 6128, Succ. Centre-ville, Montréal, QC, H3C 3J7 (Canada) and
Observatoire des Sciences et des Technologies (OST), Centre Interuniversitaire de Recherche sur la Science et la Technologie (CIRST), Université du Québec à Montréal, CP 8888, Succ. Centre-Ville, Montréal, QC, H3C 3P8 (Canada)

Introduction

In most fields, postdoctoral training is a necessary step after the PhD for those who seek an academic career. However, given the increasing number of PhD graduates and the relatively stable number professor/researcher positions available, the number of postdoctoral fellows and the length of this period in the life of recent PhD graduates have been increasing over the recent years. Hence, several studies have documented an overpopulation of postdoctoral fellows in some fields (Nerad and Cerny 1999; Jones 2013). In Canada only, the Canadian Association of Postdoctoral Scholars (CAPS) estimated that there were around 9,000 postdoctoral scholars in 2012 (Mitchell 2013). Moreover, in most North American universities, little is known on the characteristics of the population of postdocs, mainly because their administrative status is often unclear — are they students or researchers? — and not systematically managed (AAU 1998, 2005). Also, given their increasing number and the longer time during which they remain postdoctoral fellows, many of them have started to regroup into associations and are pleading for better salary levels and benefits in agreement with their professional status (Mitchell 2013; Åkerlind 2005; NIH 2012).

Along these lines, CAPS has been encouraging academic institutions to adopt better policies. To address these recommendations it is crucial to draw a clear picture of the situation of Canadians postdocs, which means, among other things, to analyze their demographic data as well as their research output. Due to the lack of proper administrative record, these tasks are rather difficult to accomplish. Hence, despite the fact that the contributing to research — and increasing one's research output — is the main purpose of a postdoctoral fellow, the research output of these high-qualified Canadian researchers stays unknown, and few small studies were undertaken to bring up bibliometric indicators about these researchers. Using data on postdoctoral fellowships award by the Canadian and the Quebec governments, this paper aims at assessing the global contribution of Canadian postdoctoral fellows to the advancement of science, in the different fields and countries of destination.

Methodology

This paper studies postdocs who have received competitive fellowships from the Canadian and the Quebec provincial government between 2004 and 2008 (N = 3,454). These fellowships provide an annual salary of about 40,000$Can (Provincial scholarships) and 30,000$Can (Provincial fellowships) to awardees for a period of two years. All scientific

[1] This work was supported by the *Observatoire des sciences et des technologies* and the Canada Research Chair program.

articles authored by these researchers between the last year preceding their postdoc and the 5 years following it were retrieved from Web of Science (WoS), by matching their names and respective institution of fellowship address with the authors' names and institutional affiliations. The names of the postdocs as well as basic information about their research project were obtained from each one of the 6 research councils: IRSC and FRQS for the health field, CRSNG and FRQNT for life sciences and engineering, and CRSH and FRQSC for social sciences and humanities.

Given that the WoS did not index the given names of authors before 2008, this first match of articles and postdocs generated a high number of false positives — papers authored by other researchers having the same name as a postdoc fellow (homonyms). These false positives were detected and removed by manual validation, using the information on the topic of the grant, the topic of the papers, as well as information on each researcher available on the Internet. We also looked for false negatives, especially for postdocs with compound surnames, by confirming the names of postdocs through web searches Personal author validation was not used in this study as it is very time consuming and often has low response rate..

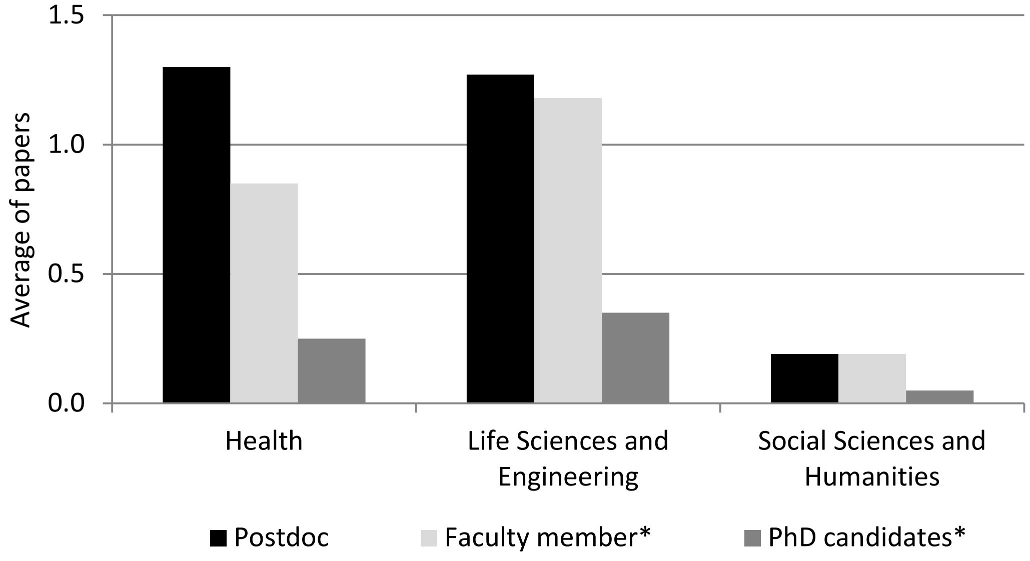
The final set included 11,327 papers authored by 3,014 postdocs. Hence, 440 (12.7%) did not author any paper indexed in the WoS. Publication counts does not include editorials, book reviews, letters to the editor or meeting abstracts, as they are generally not considered as original contributions to scholarly knowledge (Moed 1995). Thus, results are limited to articles, notes, and review. Documents authored during the competition year and the two subsequent years are considered as "postdoc" papers. We also measure the evolution of these indicators from the year before the competition until four years after. However, data for these 6 years is not always complete, as it was often impossible to know the exact affiliation of researchers after their postdoc, as they often move from one country to another.

To measure the contribution of the postdocs we used full counts of authorship, so each paper signed by the postdocs are considered as full contributions. As a measure of the scientific impact of their publications, we present the average of relative citation (ARC), which takes into account the specialty in which the paper is published, and normalizes its impact value according to the average value of this specialty. When its value is above 1, it means that the impact of the studied group of researchers is above average. When it is below 1, it means the opposite. Self-citations are excluded from the ARC counts. Given the source of data used (WoS), only journal articles are included in the analysis which, of course, underestimates the output of researchers in the social sciences and humanities.

Results and Discussion

For all fields combined, postdoc fellows published an average of 3.27 papers during the three year period including the competition year and the two years following. Figure 1 shows the average number of papers to which postdocs contributed, compared with that of faculty members and PhD students from the province of Québec (Larivière 2010), by domain. We can see that postdocs have published a similar or higher — depending of the domain — number of papers than faculty members and, unsurprisingly, publish also more than PhD students. Also, the proportion of postdocs who have published at least one paper is much higher than that of faculty and students, with 95% in health and life sciences and engineering, and 70% in social sciences and humanities.

Figure 1. Average number of papers of postdocs during their fellowship (competition year plus two years), and by faculty members and PhD students from Québec, 2000-2007



The scientific impact of these publications is shown in the Figure 2. We can see a very important difference between the postdocs and the two other groups. In Health and Natural Sciences and Engineering, the difference is quite striking, with postdocs having ARC values that are 60% above the world average, and about 40 percentage points above that of Faculty. The difference is smaller in social sciences and humanities, but postdoc are still 20% above Faculty members.

Figure 2. Average of relative citations (ARC) of papers published by postdocs during their fellowship program (competition year plus two years), and by faculty members and PhD students from Québec, 2000-2007

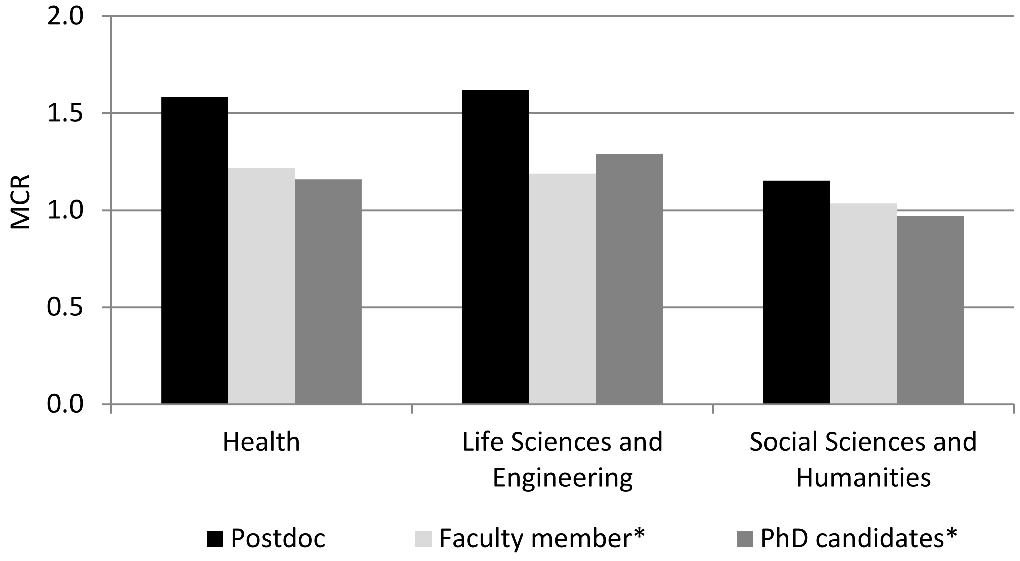
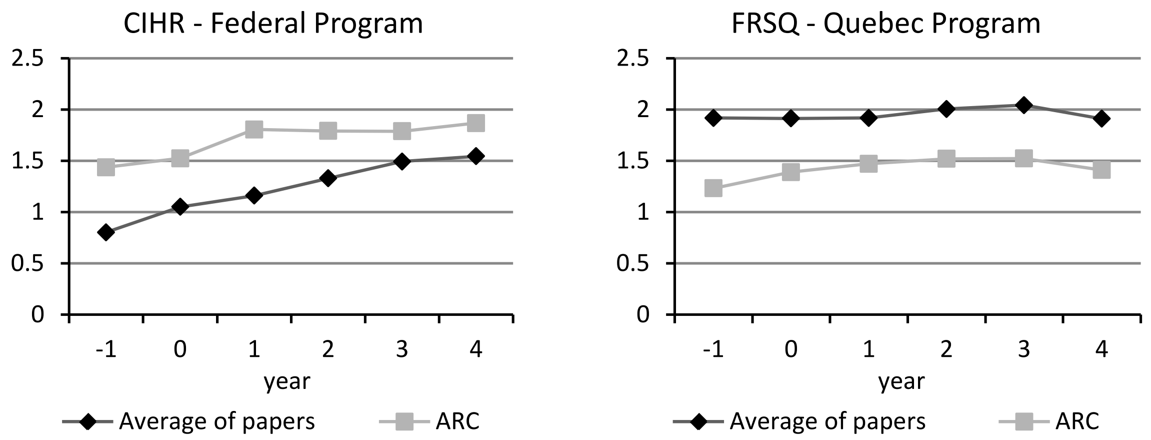


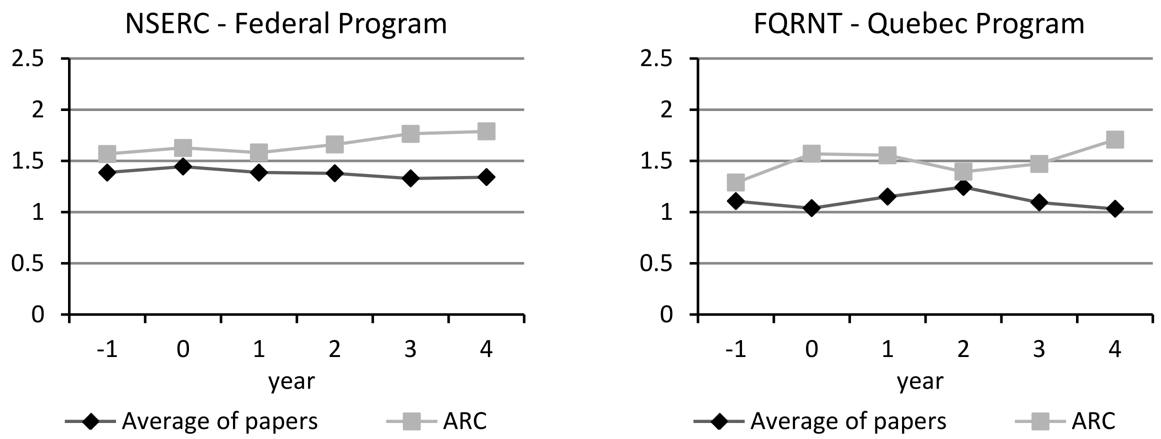
Figure 3 provides the evolution of scientific productivity and impact of the postdocs studied, separated by research council and by field. The year -1 indicates the year before the competition, 0 the competition year, and 1-4 are the years following the competition. Globally, we can notice that there is no major difference in the number of papers published, with the exception of the postdocs funded by the CIHR — and to a certain extent, those funded by NSERC — who have been significantly increasing throughout the 6 years studied. Fellows funded by CIHR and SSHRC also increased their impact throughout the period. For FRQSC, however, we see the opposite — their scientific impact drops — while for other research councils, ARC values are relatively stable.

Figure 3. Average number of papers and Average of relative citations (ARC) of postdocs, by research council and domain, by year before and after the competition

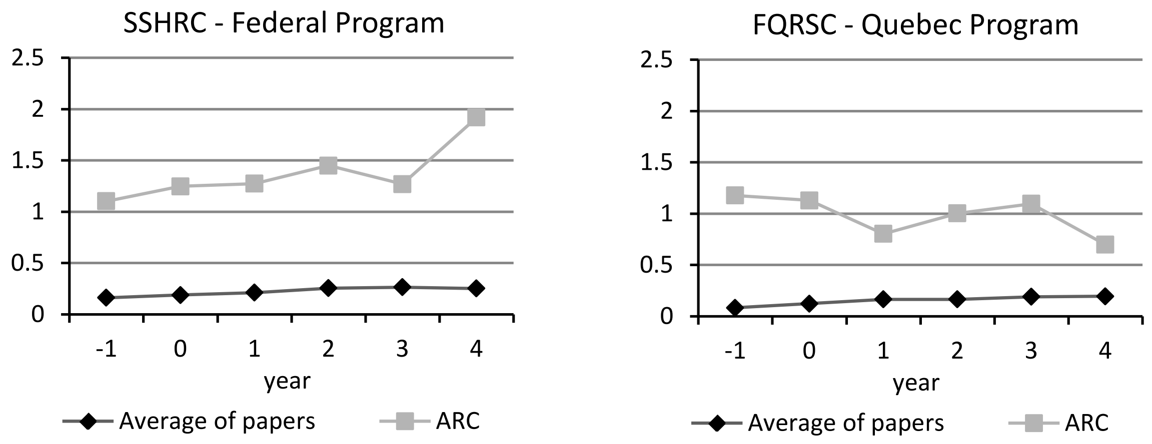
Health



Natural Sciences and Engineering



Social Sciences and Humanities



Conclusion

As the first bibliometric study about Canadian postdocs — and one of the largest at the international level — this paper constitutes a major contribution to the understanding of this group of researchers, especially on their research output compared to that of Faculty members and PhD students. More specifically, we note that, in general, postdocs have a higher productivity and scientific impact than established researchers, as one could expect from previous literature on the relationship between age and scientific output (Gingras et al. 2008). The evolution of the indicators during the 6-year period studied shows little changes, except mainly for CIHR, whose funded postdocs increase both their scientific impact and productivity throughout the period.

One of the critical problems mentioned in some studies is the lack of information about the postdocs all around the world, and this was not different in Canada. Thus, we had to work out with the available data, which represent about 8% of the whole population of postdocs in Canada according to the survey by Mitchell (2013). Moreover, this list comprises only postdocs that have won their own grant through peer review — only 13.7% of grant applications submitted to NSERC program on 2013, for example, got funded[2] — and, hence, can be considered as the elite of postdocs. This could explain the very high scores obtained. This suggests that we cannot generalize the findings presented here to all postdoctoral fellows in Canada, but, rather, should consider this as a first step towards the understanding of postdoctoral fellows' contribution to the advancement of knowledge in the country.

References

Akerlind GS. (2005). Postdoctoral researchers: roles, functions and career prospects. Higher Education, 24 : 1, 21-40.

Association of American Universities (AAU). (1998). Postdoctoral education committee report. Retrieved 5 july 2013 from : www.aau.edu/WorkArea/DownloadAsset.aspx?id=6834.

Association of American Universities (AAU). (2005). Postdoc education survey: summary of results. Retrieved 5 july 2013 from : www.aau.edu/WorkArea/DownloadAsset.aspx?id=1944.

[2] Data available at: http://www.nserc-crsng.gc.ca/NSERC-CRSNG/FundingDecisions-DecisionsFinancement/ScholarshipsAndFellowships-ConcoursDeBourses/index_eng.asp

Gingras Y, Larivière V, Macaluso B, Robitaille J-P. (2008). The effects of aging on researchers publication and citation patterns. Plos One, 3: 12, e4048.

Jones A. (2013). The explosive growth of postdocs in Computer science. Comunications of the ACM, 56 : 2, 37-39.

Larivière V. (2010). A bibliometric analysis of Quebec's PhD students' contribution to the advancement of knowledge. PhD Thesis, McGill University.

Mitchell JS, Walker VE, Annan RB, Corkery TC, Goel N, Harvey L, Kent DG, Peters J, Vilches SL. (2013). The 2013 Canadian Postdoc Survey: Painting a Picture of Canadian Postdoctoral Scholars. Canadian Association of Postdoctoral Scholars and Mitacs.

Moed, HF, De Bruin RE, Van Leeuwen TN. (1995). New bibliometric tools for the assessment of national research performance: database description, overview of indicators and first applications. Scientometrics, 33 : 3, 381-422.

National Institutes of Health (NIH), Working Group of the Advisory Committee to the Director. (2012). Biomedical research workforce working group report. Retrieved 15 september from : http://report.nih.gov/investigators_and_trainees/ACD_BWF/.

Nerad M, Cerny J. (1999). Postdoctoral patterns, career advancement, and problems. Science, 285 : 5433, 1533-1535.

Innovative women: an analysis of global gender disparities in patenting[1]

Cassidy R. Sugimoto*, Chaoqun Ni*, Jevin D. West,** and Vincent Larivière***

* *{sugimoto; chni}@indiana.edu*
School of Informatics and Computing, Indiana University Bloomington, 1320 East 10th St., Bloomington, IN, 47405 (USA)

** *jevinw@uw.edu*
University of Washington Information School, Box 352840, Seattle, WA, 98195 (USA)

*** *vincent.lariviere@umontreal.ca*
École de bibliothéconomie et des sciences de l'information, Université de Montréal, Pavillon Lionel-Groulx C.P. 6128, Succ. Centre-ville, Montréal, QC, H3C 3J7 (Canada) and
Observatoire des Sciences et des Technologies (OST), Centre Interuniversitaire de Recherche sur la Science et la Technologie (CIRST), Université du Québec à Montréal, CP 8888, Succ. Centre-Ville, Montréal, QC, H3C 3P8 (Canada)

Introduction

Innovation is critical to economic development (Schumpeter, 1934) and depends upon the full participation of the scientific workforce (Hunt, Garant, Herman, & Munroe, 2013). Yet, the field of "innovation studies" (Fagerberg, Fosaas, & Sapprasert, 2012) demonstrates that there are many disparities in the exploitation of human capacity for innovation. Foremost among these is the dearth of female inventors (Ding, Murray, & Stuart, 2006; Thursby & Thursby, 2005; Whittington & Smith-Doeer, 2005). The first patent granted to a woman was in 1637; however, female contribution failed to exceed more than 2% through the first half of the 20th century (Jaffe, 2003). Contemporary studies have shown that fewer women patent and when they do, they produce fewer patents per person than men (Ding, Murray, & Stuart, 2006). A number of correlates have been noted: women with higher degrees are more likely to patent than those without (Hunt, Garant, Herman, & Munroe, 2013), and when women inventors are involved, patents tend to have higher diversity in terms of the number of IPC codes assigned (Meng & Shapira, 2011).

The need to understand inventor diversity in patenting was stressed in the America Invents Act (2010), which mandated that the USPTO "establish methods for studying the diversity of patent applicants" (Pub.L. 112-29). The Federal Register (Focarino, 2013) disclosed the first analysis of the 2005-2006 USPTO data, discussed the poor matching with Census data, and called for others to study the diversity of patent applicants. Previous work in this area has relied on purposive sampling of specific populations (e.g., all college graduates, doctoral degree recipients) and single disciplines (e.g., nanotechnology, biochemistry). This paper answers the USPTO call and fills a gap in the literature by providing a global analysis of women in patents from 1976 to 2013.

[1] The primary support for this work came from the Science of Science, Innovation and Policy Program within the National Science Foundation (Grant No. 1158670) and from the the Canada Research Chairs program

Methods

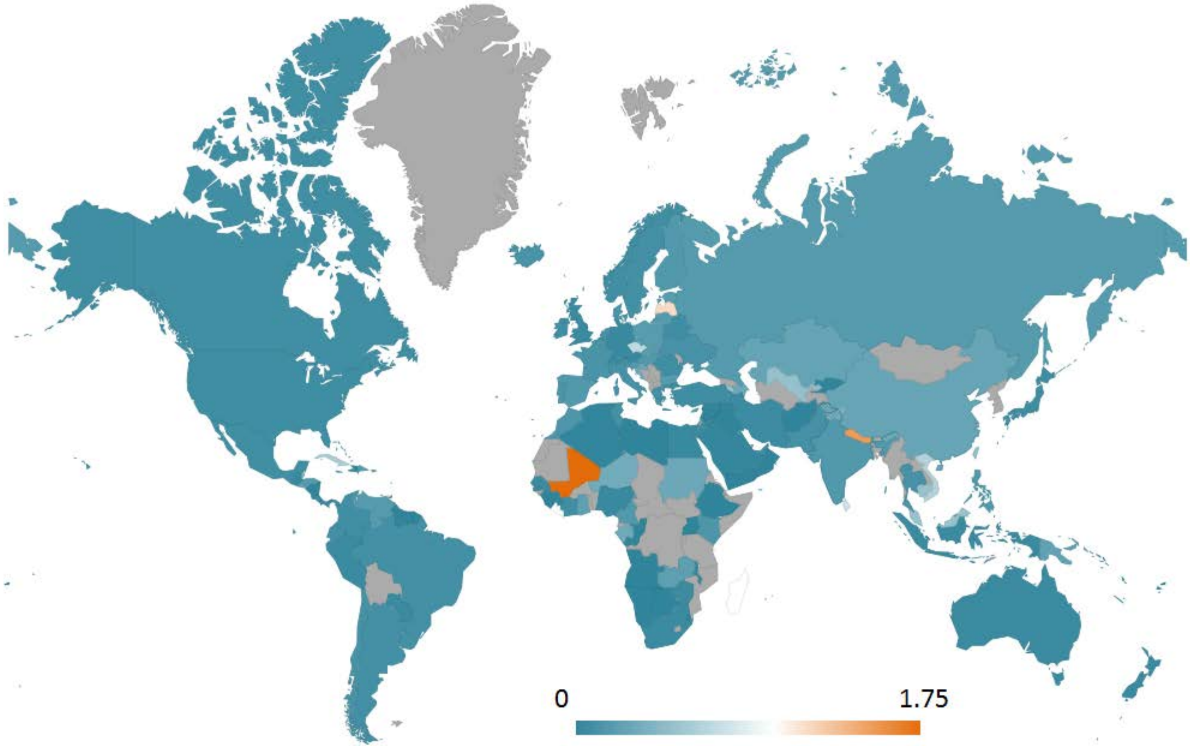
Data was downloaded from the USPTO database and transformed into an SQL relational database. The gender of inventors was categorized using first names, which was matched with worldwide and country-specific name lists, as developed in previous work (Larivière et al., 2013). 87% of 11.7 million inventorships analysed were assigned to a gender. Nationality of the assignees was listed in the patent and was used to identify fractionalized counts of patents per country.

The number of patents for female and male inventors was calculated based on fractionalized counting of patents (see Larivière et al., 2013). That is, each inventor is given 1/x count of the inventorship where *x* represents the number of inventors for which a gender could be assigned on the given patent. Therefore, if there are 5 inventors listed in a patent, 2 of them were identified as female and 3 of them as male, then the female fractionalized count is 2/5, and the male fractionalized count is 3/5.

Findings

We first sought to examine the proportion of female inventorships by country. Women contributed less than 8% of all patent authorships for the entire period (1976-2013). In 2013, women contributed to slightly more than 10% of patents. Figure 1 displays the ratio between female and male productivity in terms of patenting (with fractionalized counts). As demonstrated, men dominate production in nearly every country (in 42 countries, there are no female inventors). Five countries are female dominated; however, these all have fewer than 35 fractionalized patents (Mali, Nepal, Latvia, Madagascar, and Liberia).

Figure 1: Female to Male Productivity Ratio by Country



Ten countries make up more than 90% of the world share of patents. These countries, and associated female-male ratios and fractionalized inventorship counts are provided in Table 1.

Table 1. FMRatio in top 10 countries by number of patents (93.6% of the world total)

Country	FMRatio	Fractionalized count
United States	0.07	2,349,090.00
Japan	0.07	850,786.10
Germany	0.04	311,242.40
United Kingdom	0.05	114,264.80
France	0.12	106,867.80
Republic of Korea	0.16	97,578.94
Taiwan	0.47	95,741.60
Canada	0.08	90,578.42
Italy	0.11	47,412.98
Switzerland	0.04	46,708.73

As shown, Germany, the United Kingdom and Switzerland have the lowest levels of parity; whereas Taiwan is closest to parity (followed by Korea). We further investigated male dominance in terms of those countries producing more than 1,000 patents (Table 2), with Austria, Germany, Switzerland and the UK having the most extreme male dominance.

Table 2. Countries with highest male dominance (more than 1,000 patents)

Country	F	M	FMRatio	TotalN
Austria	3.14%	96.86%	0.03	15,924.24
Germany	3.91%	96.09%	0.04	311,242.4
Switzerland	3.96%	96.04%	0.04	46,708.73
United Kingdom	4.50%	95.50%	0.05	114,264.8
Australia	4.97%	95.03%	0.05	25,616.45
South Africa	4.47%	95.53%	0.05	3426.379
New Zealand	5.17%	94.83%	0.05	3197.525
United States	6.57%	93.43%	0.07	2,349,090
Japan	6.74%	93.26%	0.07	850,786.1
Canada	7.16%	92.84%	0.08	90,578.42

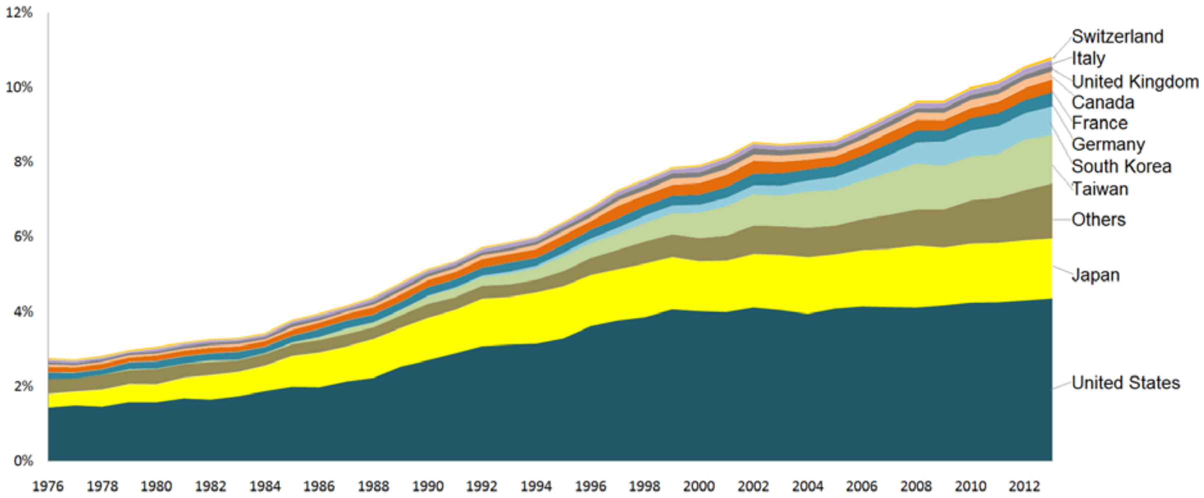
Few countries come close to parity. Table 3 ranks, by gender partiy, countries that have more than 1,000 patents. As shown, Asian countries and the USSR/Russia come closest to parity, though men are still dominant.

Table 3. Countries coming closest to achieving parity

Country	F	M	FMRatio	TotalN
Malaysia	33.91%	66.09%	0.51	1544.403
Taiwan	32.13%	67.87%	0.47	95,741.6
Singapore	21.42%	78.58%	0.27	6401.447
China	20.94%	79.06%	0.26	23,713.66
Poland	15.63%	84.37%	0.19	1044.417
Union of Soviet Socialist Republics	15.56%	84.44%	0.18	4219.982
Russian Federation	15.56%	84.44%	0.18	4198.689
Korea, Republic of	13.60%	86.40%	0.16	97,578.94
Israel	13.77%	86.23%	0.16	24,789.74
Finland	14.02%	85.98%	0.16	16,999.29

Figure 2 displays the ten countries contributing most to the total share of female patents over time. As is shown, Taiwan and Korea have seen large increases in their overall contribution to female patenting since the mid-1990s. The proportional contribution to female patenting from the United States and Japan has remained fairly stable since 2000.

Figure 2: Top 10 countries' (by number of patents) contribution to global female patenting



Work-in-progress

While we have provided an overview of the global statistics, we also need to analyse the contribution of women to different areas of patenting and the contexts in which this patenting occurs. For example, it has been suggested that women are more risk averse and the lack the social networks necessary to effectively commercialize their work (e.g., Abreau & Grinevich, 2013). However, others have suggested that institutional setting and resource allocation, rather than personal proclivities, are better predictors of potential patenting (e.g., Colyvas, Snellman, Bercovitz, & Feldman, 2012). It may also be that women are concentrated in fields or countries where patenting is either discouraged or less incentivized.

One thing that remains constant is women's patenting remains lower than would be predicted given their representation in science, technology, engineering, and mathematics fields and professions (Mauleon & Bordons, 2010) and their relative productivity in publishing (Larivière et al., 2013). More work needs to be done to understand why this valuable human resource is not being captured in the innovation process and mechanisms that can be used to support full participation of the scientific workforce in patenting activities and how this relates to other types of gender disparities. However, for richer analyses, triangulating data from qualitative and quantitative sources may be necessary. This may be particularly useful in understanding why, for instance, women's names are included on publications related to a patent, by disappear between the articles about the patents and the patents themselves (Lisson, Montobbio, & Zirulia, 2013). Using country-level data is only an initial step in investigating the types of environments and policies that are more conducive to gender parity.

References

Abreau, M. & Grinevich, V. (2013). The nature of academic entrepreneurship in the UK: Widening the focus on entrepreneurial activities. *Research Policy, 42*, 408-422.

Colyvas, J.A., Snellman, K., Bercovitz, J., & Feldman, M. (2012). Disentangling effort and performance: a renewed look at gender differences in commercializing medical school research. *Journal of Technology Transfer, 37*, 478-489.

Ding, W.W., Murray, F., & Stuart, T.E. (2006). Gender differences in patenting in the academic life sciences. *Science*, *313*, 665-667.

Fagerberg, J., Fosaas, M., & Sapprasert, K. (2012). Innovation: Exploring the knowledge base. *Research Policy, 41*, 1132-1153.

Focarino, M.A. (2013). Department of Commerce: United States Patent and Trademark Office: Request for comments on methods for studying the diversity of patent applicants. *Federal Register,* 78(231), 72064-72065.

Hunt, J., Garant, J.-P., Herman, H., & Munroe, D.J. (2013). Why are women underrepresented amongst patentees? *Research Policy*, *42*, 831-843.

Larivière,V., Ni, C., Gingras, Y., Cronin, B., & Sugimoto, C.R. (2013). Gender Disparities in Science, *Nature*, 504, 211–213

Lissoni, F., Montobbio, F., & Zirulia, L. (2013). Inventorship and authorship as attribution rights: An enquiry into the economics of scientific credit. *Journal of Economic Behavior & Organization*, 95, 49-69.

Mauleon, E., & Bordons, M. (2010). Male and female involvement in patenting activity in Spain. *Scientometrics*, *83,* 605-621.

Meng, Y., & Shapira, P. Women and patenting in nanotechnology: Scale, scope and equity. In Nanotechnology and the challenges of equity, equality and development. Susan E. Cozzens and Jameson Wetmore (Eds.). (2011). New York, NY: Springer

Schumpeter, J.A. (1934/2008). The theory of economic development: An inquiry into profits, capital, credit, interest, and the business cycle. Revers Opie (trans.). New Brunswick, NJ: Transaction Publishers.

Thursby, J.G., Thursby, M.C., 2005. Gender patterns of research and licensing activity of science and engineering faculty. *Journal of Technology Transfer*, 30, 343–353.

Whittington, K. B., & Smith-Doerr, L. (2005). Gender and commercial science: Women's patenting in the Life Sciences. *Journal of Technology Transfer*, *30*, 355–370.

Improved author profiling through the use of citation classes[1]

Bart Thijs*, Koenraad Debackere**, Wolfgang Glänzel***

[*] *Bart.Thijs@kuleuven.be*
Centre for R&D Monitoring (ECOOM) and Dept. MSI, KU Leuven, Leuven (Belgium)

[**] *Koenraad.Debackere@kuleuven.be*
Centre for R&D Monitoring (ECOOM) and Dept. MSI, KU Leuven, Leuven (Belgium)

[***] *Wolfgang.Glänzel@kuleuven.be*
Centre for R&D Monitoring (ECOOM) and Dept. MSI, KU Leuven, Leuven (Belgium)
Department of Science Policy & Scientometrics, Library of the Hungarian Academy of Sciences, Budapest (Hungary)

Introduction

At the level of research teams and individual scientists bibliometric standard indicators are easily distorted by the citation impact of papers that considerable differ from their expected value. In previous studies by the authors (Glänzel, 2013; Glänzel et al., 2014), a parameter-free solution providing four performance classes has been proposed to replace mean-value based indicators by performance classes. This has been found useful at the level of national and institutional assessment of citation impact. The approach, which is based on the method of Characteristic Scores and Scales (CSS) that has originally been introduced by Glänzel & Schubert (1988) in the context of journal analysis, has several striking advantages over traditional methods. The application of these performance classes completely eliminates the effect of the heavy tails of citation distributions including their typical outlier-based biases at lower levels of aggregation. Unlike in methods based on percentiles (e.g., Leydesdorff et al. (2011), this approach is not sensitive to ties and ensures seamless integration of measures of outstanding and even extreme performance into the standard tools of scientometric performance assessment. The method can be interpreted as a reduction of the original citation distribution to a distribution over a given number of performance classes with self-adjusting thresholds without requiring arbitrarily pre-set values. The method proved thereby to be insensitive to both subject-specific peculiarities and the particular choice of publication years and citation windows. The application to the micro-level, that is, to the level of individual scientists and research teams, however, still remained a challenge. Career evolution of individual scientists and changing team composition along with the typically small paper sets underlying the citation distribution at this level require extremely stable and robust solutions. In the present study we will analyse in how far the CSS-based method meets these requirements, and will give examples for its application to the level of individual scientists. We will also show that the method can be applied using reference standards defined by any appropriate base-line distribution forming a superset of the publication set under study.
An additional advantage is that the classes can, because of the high robustness of the distribution of papers across fields (see Glänzel et al., 2014), be calculated at any level of field aggregation and that multiple field assignments do not hinder the calculation of the specific thresholds and the performance scores for individual authors. Furthermore, only the calculated

[1] All data presented in the tables and figures of this study are based on data sourced from Thomson Reuters Web of Science.

threshold values are needed in a real world application of the method in an evaluative exercise without the underlying citation distributions.

Data

The data set was built on downloaded Researcher-ID data of 4.271 registered researchers from eight selected countries. For reasons of statistical reliability only authors with at least 20 publications were taken into consideration for the retrieval. This selection of data creates a build-in bias towards more prolific and excellent authors. Heeffer et al., (2013) showed that authors registered for a Researcher ID are indeed more productive than others. Also the applied threshold of 20 publications implies the exclusion of many occasional authors whose impact is rather limited. However, as the goal of this paper is not to define any particular reference standard but to investigate the applicability of the CSS-method, we can proceed from these data sets without loss of generality. Moreover, it can be expected that in a real-world evaluative exercise or selection procedure, the studied authors are indeed the more active ones.

Publication data were matched with data retrieved from Thomson Reuters Web of Science (WoS). Only journal publications indexed as article, letter, note or review between 1991 and 2010 are taken into consideration. Papers are assigned on the basis of the journals in which they appeared to subfields according to the Leuven-Budapest scheme (Glänzel & Schubert, 2003). Also papers only assigned to Social Sciences or Humanities have been excluded in the analysis. Moving three-year citation windows were used throughout the analysis.

Methods and results

First we briefly recall the outlines of the model (cf. Glänzel, 2013). Characteristic scores are obtained by iteratively calculating the mean value of a citation distribution and subsequently truncating this distribution by removing all papers with less citations than the conditional mean. As described earlier, the process is stopped after three iterations. This results in three scores b_k with (k=1, 2, 3). By adding b_0=0 and b_4=$\propto$, four distinct performance classes can be created each defined by a pair of threshold values [b_{k-1},b_k) with (k=1, 2, 3, 4). This definition solves the problem of ties that otherwise might occur in ranking approaches as each paper can be uniquely assign into one of these four half-open intervals.

These scores are now calculated at the field level and for each publication year. As papers can be assigned to multiple fields, this requires some special attention both in the calculation of the scores as with the attribution to the performance classes. The applied methodology is similar as the calculation of the subfield-expected citation rate. The contribution of a paper assigned to multiple subfields is fractionated based on the number of assigned fields. This means, for instance, that a paper classified in two subfields counts only as a half in the numerator of the mean calculation. And only the half of its citations contributes to the denominator. A detailed description of this method can be found in a recent study by Glänzel et al., (2014).

The results of the CSS-based methods are also gauged against traditional measures. For this purpose we use normalized citation rates, particularly Relative Citation Rate (RCR) and the Mean Normalized Citation Rate (NMCR) (Glänzel et al., 2009). The first indicator compares the observed citation rate to a journal based expected citation rate while the second one uses a subfield expected citation rate. It is important to mention here that score b_1 is in fact an expected citation rate that depends on the given reference distribution.

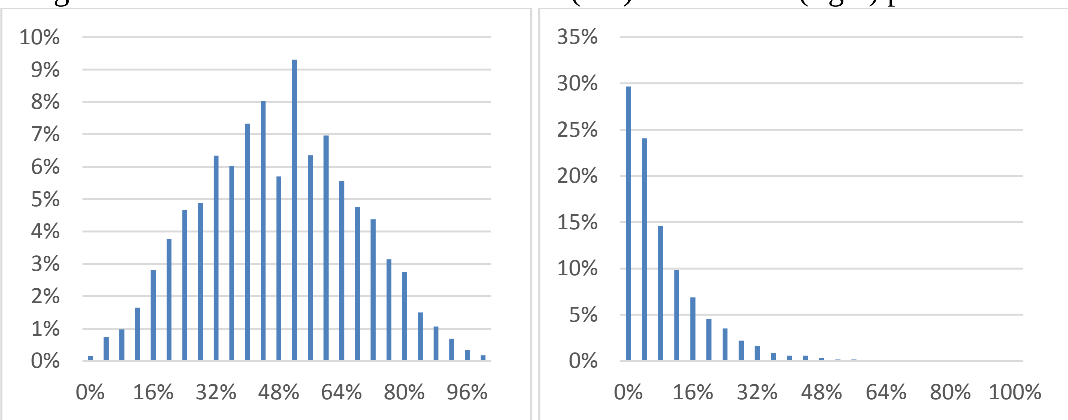
Performance Classes

The first result we present is the share of publications in each performance class in Table 1. The first row contains the average over all the distinct authors of the distribution of papers across classes. And the second row presents the share of the distinct classes within the total set of 116.467 unique papers. This data is extended in this table with the results for the distribution over classes of all the publication data indexed in 2010 in the WoS database. Because of the robustness of the method this could considered a suitable reference standard (cf. Glänzel, 2007 and 2013). The shares in the last row are in line with earlier reported shares (see Glänzel et al., 2014). Also Albarran & Ruiz-Castillo (2010) found the same 70-21-9 rule where they merged the two upper classes for publications indexed between 1998 and 2002 with a 5-year citation window and classified into 22 fields according to Thomson Reuters *Essential Science Indicators* (ESI). As mentioned above, only prolific authors are included in the analysis as the calculation of shares within small publication sets is otherwise subject to possible fluctuations, which might distort the resulting statistics and thus reduce their statistical reliability. As a consequence of this reliability-related selection criterion shares are shifted towards the upper classes as compared with the reference standard of the complete population. This "bias" can also be observed in the citation indicators for the set of papers of the selected authors (RCR=1.42 and NMCR=2.10). The implications of this discrepancy will be discusses later with respect to the choice of benchmarks for comparison in particular applications.

Table 1. Distribution of papers over performance classes.

Data Set	**Class 1**	**Class 2**	**Class 3**	**Class 4**
Average over all authors with R-ID	42.8%	33.1%	15.0%	9.1%
All papers of authors with R-ID	44.3%	32.9%	14.5%	8.4%
All papers indexed in 2010	69.8%	21.5%	6.2%	2.5%

Figure 1. Distribution of shares in Class 1 (left) and Class 4 (right) publications



The distribution of shares in the first two classes proved to be normal. The third class deviates slightly but the last class deviates strongly from a normal distribution. In order to illustrate this effect we show the patterns for Class 1 and 4 in Figure 1. The strong deviation from normality in Class 4 reflects once again the problematic behaviour of extreme values and their particular distribution as already reported by Glänzel (2013).

Citation Indicators

In a second analysis we compare the distribution across classes with the traditional citation indicators. A Spearman rank correlation is calculated among the classes. The results are presented in Table 2. The first class is negatively correlated with the three other classes. The negative correlation of the first class with the other ones *and* the traditional relative indicators is in line with our expectations since the first class relates to poorly cited papers. It is striking that Class 2 is not correlated at all with the upper two classes and with the RCR and that it has a moderate correlation with the other citation indicators. Both classes 3 and 4 have higher correlations with the citation indicators but have an inter-correlation of 0.48. These observations substantiates that citation behaviour cannot sufficiently be represented by one class or any individual indicator alone. This is a strong argument for the choice of this method with four performance classes at this aggregation level too.

Table 2. Rank correlation between performance classes and citation indicators for data set 1. (values marked with * do not statistically deviate from 0)

	Class 1	**Class 2**	**Class 3**	**Class 4**	**RCR**	**NMCR**
Class 2	-0.459					
Class 3	-0.743	0.058*				
Class 4	-0.685	-0.055*	0.485			
RCR	-0.630	-0.001*	0.547	0.734		
NMCR	-0.839	0.125	0.693	0.863	0.786	
NMCR/RCR	-0.632	0.202	0.482	0.547	0.163	0.701

Author profiling

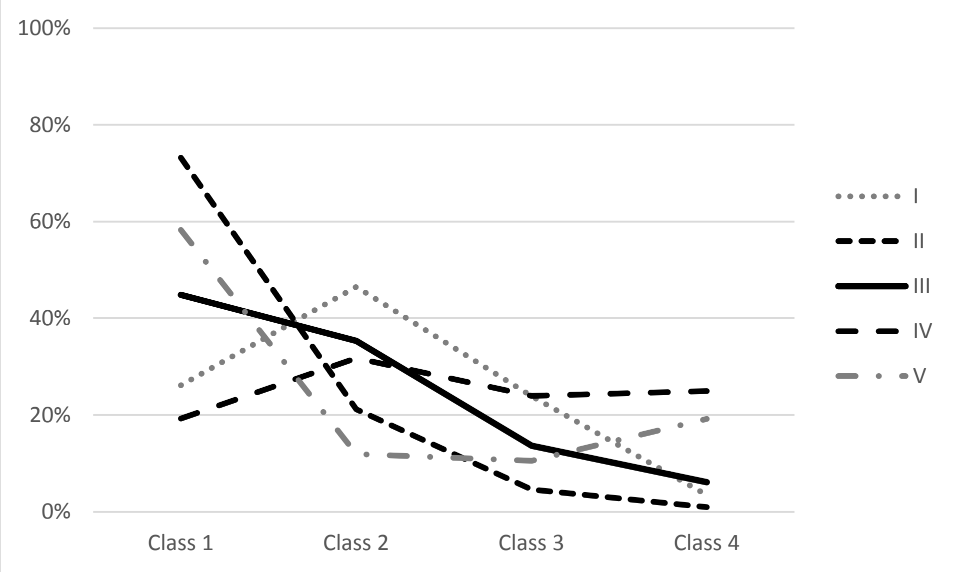
As already described in an early study (Glänzel et al., 2014), different profiles according to the deviation from the base-line distribution in each class are possible. Unlike in the case of traditional indicators, where just higher/lower than the expectation can occur, here a variety of deviations are possible. Each author can thus be characterized by an individual profile indicating the distribution of his papers over the four performance classes. The advantage of these profiles is that they enable a direct comparison between distinct authors but it may also be applied for comparison with an appropriate reference distribution.

First, a χ^2-test indicates whether or not an individual profile deviates from the reference (at a confidence level of $p = 0.05$). The null hypothesis assumes that the distribution of an author's papers across these four classes is consistent with the chosen reference. In our case the number of publications instead of the shares is used to take the size of the publication set into account. Only if H_0 is rejected, the deviation in each of the four classes from the reference is used to identify performance types to be distinguished.

I. The χ^2-test indicates that the profile deviates from the reference and the shares of both classes 1 and 4 are lower than the reference.
II. The χ^2-test indicates that the profile deviates from the reference and the share of class 1 is higher and class 4 is lower than the reference.
III. The χ^2-test indicates that the profile does not deviate from the reference.
IV. The χ^2-test indicates that the profile deviates from the reference and the share of class 1 is lower and class 4 is higher than the reference.
V. The χ^2-test indicates that the profile deviates from the reference and the shares of both classes 1 and 4 are higher than the reference.

We present the classification of each of the 4271 authors according to these profile types as an example. The choice of an appropriate reference is crucial here. It is clear that the total set of publications as presented in the last row of table 1 is not suitable as our sample consists of authors with a more prolific and excellent profile. Therefore, the distribution of the total set of unique publications published by at least one of the authors was taken as reference standard here (see second row in Table 1). The χ^2-test indicates whether or not the individual author profile deviates from this reference standard. Six out of ten authors have a profile that does not significantly deviate from the reference and are classified as type III. Type II and IV contain respectively 13.6% and 19.8% of all authors. The two remaining types I and V comprise a much smaller number of authors. Figure 2 presents the average share within each class for the five types. The solid line of type III shows a neat decline.

Figure 2. Average share of four performance classes across five types of authors



These types do not only have different shares among the performance classes nor on other indicators – as Table 3 shows. Type III authors have, on an average, less publications than the other types. This is reasonable as the sensitivity of the χ^2-test increases with higher number of publications. This effect thus confirms the reliability of the applied classification rule. Another striking observation is the high values for the citation indicators for type I. The lack of highly cited publications in the fourth class is here compensated by the low share in Class 1 publications and high share in Class 2. These publications in Class 2 have already received citation rates higher than the expected citation impact used for the calculation of the citation indicators in Table 3.

Table 3. Standard indicators for the five profile types

	I	**II**	**III**	**IV**	**V**
Number of authors	211	579	2603	845	33
Average Publication	37.6	38.3	30.7	37.8	31.6
Average RCR	1.33	0.88	1.28	2.22	3.39
Average NMCR	2.06	0.85	1.80	4.36	5.87
Average NMCR/RCR	1.62	0.99	1.44	2.01	1.49

To conclude the section on author profiling, Table 4 presents a sample of 20 authors taken from data set 1. The total set of unique publications published by at least one of the authors from data set 1 are taken as benchmarks and its distribution over the four classes is taken as reference point.

Table 4. Average distribution of papers over performance classes.

Author	Class 1	Class 2	Class 3	Class 4	RCR	NMCR	NMCR/RCR	Type
1	31.8%	27.3%	36.4%	4.5%	1.32	2.01	1.53	I
2	39.3%	60.7%	0.0%	0.0%	1.13	1.24	1.09	I
3	32.5%	45.0%	22.5%	0.0%	0.98	1.59	1.61	I
4	40.0%	24.4%	33.3%	2.2%	1.36	1.65	1.21	I
5	67.5%	28.6%	1.3%	2.6%	1.30	1.01	0.77	II
6	58.3%	32.5%	7.3%	2.0%	1.00	1.18	1.19	II
7	74.2%	16.1%	9.7%	0.0%	1.21	0.75	0.62	II
8	72.0%	20.0%	8.0%	0.0%	0.65	0.82	1.27	II
9	68.2%	27.3%	0.0%	4.5%	0.82	0.79	0.96	III
10	33.3%	52.4%	9.5%	4.8%	1.56	1.82	1.17	III
11	61.9%	38.1%	0.0%	0.0%	0.70	0.80	1.15	III
12	33.3%	28.9%	20.0%	17.8%	1.89	3.35	1.77	III
13	21.4%	21.4%	14.3%	42.9%	2.65	5.85	2.21	IV
14	9.4%	50.0%	21.9%	18.8%	1.68	2.72	1.62	IV
15	17.9%	14.3%	28.6%	39.3%	3.21	6.34	1.97	IV
16	30.0%	20.0%	40.0%	10.0%	1.80	2.29	1.27	IV
17	56.8%	22.4%	11.2%	9.6%	1.47	1.83	1.24	V
18	62.5%	16.1%	12.5%	8.9%	1.38	1.92	1.39	V
19	47.6%	9.5%	9.5%	33.3%	14.58	35.46	2.43	V
20	50.0%	10.7%	14.3%	25.0%	3.75	3.69	0.98	V
Reference	**44.3%**	**32.9%**	**14.5%**	**8.4%**				

Conclusions

In the present study we showed the general applicability of the CSS-method to the individual level and to author profiling of candidates with scientific excellence, in particular. Publications from authors with a Researcher ID and at least 20 registered publications were classified according to the field and year specific thresholds. The distribution of shares in the first two classes came close to normal and in the third class is deviated only slightly. But once again, it is the tail of citation distributions represented by Class 4 that showed strong deviation. When correlated with traditional citation indicators it became clear that the distribution over classes is only partially correlated with these, especially with the journal based citation indicator (RCR).

We also compared the individual authors with a chosen reference standard and could define five different profile types. In our example the reference was calculated based on all the publications of the selected authors. In this case, the set is a result from within selection procedure but the reference set could also be defined prior to the start of this procedure, e.g., publications form a certain country or institute. Two types, I and V have shares in the two outer classes that are both above or below the reference share. In both cases the presences/absence of publication in both classes compensates for the citation scores. Finally we would like to stress that the reduction of this method to two instead of four performance classes would bring us back to system of traditional indicators. This and the above observations confirm the seamless integration of the CSS method into the standard toolset of scientometric research evaluation.

References

Albarrán, P., & Ruiz-Castillo, J. (2011). References made and citations received by scientific articles. *JASIST*, 62(1), 40–49.

Glänzel, W., & Schubert, A. (1988). Characteristic scores and scales in assessing citation impact. *Journal of Information Science*, 14(2), 123–127.

Glänzel, W., Schubert, A. (2003), A new classification scheme of science fields and subfields designed for scientometric evaluation purposes. *Scientometrics*, 56 (3), 357-367.

Glänzel, W. (2007), Characteristic scores and scales. A bibliometric analysis of subject characteristics based on long-term citation observation. *Journal of Informetrics*, 1(1), 92-102.

Glänzel, W., Schubert, A., Thijs, B., & Debackere, K. (2009). Subfield-specific normalized relative indicators and a new generation of relational charts: Methodological foundations illustrated on the assessment of Institutional Research Performance. *Scientometrics*, 78(1), 165–188.

Glänzel, W., Thijs, B., & Debackere, K. (2014). The application of citation-based performance classes to the disciplinary and multidisciplinary assessment in national comparison and Institutional Research Assessment. *Scientometrics*, DOI: 10.1007/s11192-014-1247-1, in press.

Heeffer, S., Thijs, B., & Glänzel, W. (2013). Are registered authors more productive? *ISSI Newsletter*, 9(2), 29-32.

Leydesdorff, L., Bornmann, L., Mutz, R., & Opthof, T. (2011). Turning the tables on citation analysis one more time: principles for comparing sets of documents. *JASIST*, 62(7), 1370-1381.

How many citations are there in the Data Citation Index?[1]

Daniel Torres-Salinas*, Evaristo Jiménez-Contreras** and Nicolas Robinson-García**

*torressalinas@gmail.com
EC3Metrics, EC3 Evaluación de la Ciencia y de la Documentación Científica, Universidad de Navarra, C/ Alhóndiga, 6 2ºA, Granada, 18002 (Spain)

** evaristo@ugr.es; elrobin@ugr.es
EC3 Evaluación de la Ciencia y de la Documentación Científica, Universidad de Granada, Colegio Máximo de Cartuja, s/n, Granada, 18071 (Spain)

Introduction

Lately we have witnessed a renewed interest for data sharing and the development of reproducible research (Anon, 2008). Although the claim for transparency in research is not new (King, 1995), in the last few years researchers have been challenged with the management and processing of huge amounts of datasets for conducting large-scale studies in what is known as the 'Big Data' phenomenon (Lynch, 2008). But data sharing practices are relatively common in some fields such as Genomics or Astronomy (Borgman, 2012). Their experience has allowed the development of infrastructure and a slow expansion towards the rest of fields, but still these practices are far from common. In order to promote data sharing practices, journals and evaluation agencies have started to introduce policies that encourage and in some cases, demand authors to share their datasets (an overview of such policies is offered by Borgman, 2012; Torres-Salinas, Robinson-Garcia & Cabezas Clavijo, 2013).

One of the main concerns researchers have for sharing data has to do with the idea that such practices are not 'worth it' as they are time-consuming and are not acknowledged by colleagues and funding bodies. In order to surpass such fear, some authors have analyzed the citation effects of publications sharing data concluding that there is a positive relation between them (Piwowar, Day & Fridsma, 2007; Piwowar & Chapman, 2010). In this context, many tools are being developed in order to track 'impact' of data such as DataCite, CrossRef or Thomson Reuters Data Citation Index (Costas, Meijer, Zahedi & Wouters, 2012). Here we will focus on the latter, a multidisciplinary database launched in 2012 which indexes major data repositories from all areas of the scientific knowledge along with citation data associated to them (Thomson Reuters, 2012).

Torres-Salinas, Martín-Martín and Fuente-Gutiérrez (2014) recently studied the coverage of the Data Citation Index (DCI). From their analysis they concluded that the DCI is heavily biased towards the Hard Sciences, the most common document type is datasets (94% of the total share) and four repositories represent 75% of the database. This paper builds up on their work focusing on the citation distribution of the DCI by areas and by repositories, offering the first citation analysis so far of the DCI.

[1] Nicolas Robinson-Garcia is currently supported by a FPU Grant of the Spanish Ministerio de Economía y Competitividad

Material and methods

In this paper we conduct an analysis of the citation distribution of the Data Citation Index by areas and repositories. Between May and June, 2013, we retrieved all records indexed in the DCI and created a relational database for data processing. Subject categories to which repositories were assigned were aggregated into four broad areas (Science, Engineering & Technology, Social Sciences and Arts & Humanities). The DCI includes three different document types: datasets, data studies and repositories. However, the distribution of each of them varies by repository. While some repositories include both datasets and data studies, other only include one of them. Also, not all fields in records seem to be common to all repositories. This can be seen especially in the case of the fields dedicated to assigning keywords to each record.

Results

In table 1 we show the main figures by document type. There are a total of 2,626,528 records in the DCI. Most of these are datasets, representing, 94% of the database. Regarding the total number of citation received, 88% of all records remain uncited. Data studies receive more citation in average (0.69) than datasets (0.12), but again, datasets accumulate most of the citation included in the DCI (73%).

Table 1. Indicators for all records, datasets and data studies

	All Document Types	Datasets	Data studies
Total Citations	404,211	294,051	106,895
Total Records	2,623,528	2,468,736	154,674
Uncited Records	2,311,553	2,185,062	126,428
% Uncited	88.11	88.51	81.74
Citation Average	0.15	0.12	0.69
Standard Desviation	3.06	0.36	9.56

When focusing on the analysis by areas, 81% of the records belong to the area of Science, followed by far by Social Sciences (18%). On the other hand, Engineering & Technology is the most underrepresented area with 0.1% of the whole share. This pattern is also seen when focusing on datasets where Science, was again represents 81% of the database followed by Social Sciences with a share of 17%. However, this picture changes slightly when focusing on datasets. Although the distribution is still severely biased towards Science (74%), Social Sciences has a higher presence (24%). Regarding the citation distribution, only in the area of Engineering & Technology we see a citation average above 0.5, highlighting the high degree of uncitedness. Science accumulates most citations (79%) followed by the Social Sciences (18%), Arts & Humanities (5%) and finally, Engineering & Technology (0.2%). But there are significant differences when analyzing each document type. While in the fields of Engineering & Technology and Science, researchers tend to cite datasets (97% of all citation received in Engineering & Technology and 92% in Science are directed to datasets), the opposite occurs in Social Sciences and Arts & Humanities, where most of the citations were directed to data studies (96% in the case of the former and all except one citation in the case of the latter).

Table 2. Indicators for all records, datasets and data studies by area

A. All document types

	Total Records	Total Citations	Citation Average	Standard Deviation
Engineering & Technology	1,786	916	0.51	0.90
Humanities & Arts	51,444	20,460	0.40	7.99
Science	2,118,855	319,458	0.15	0.59
Social Sciences	462,826	72,855	0.16	6.84

B. Datasets

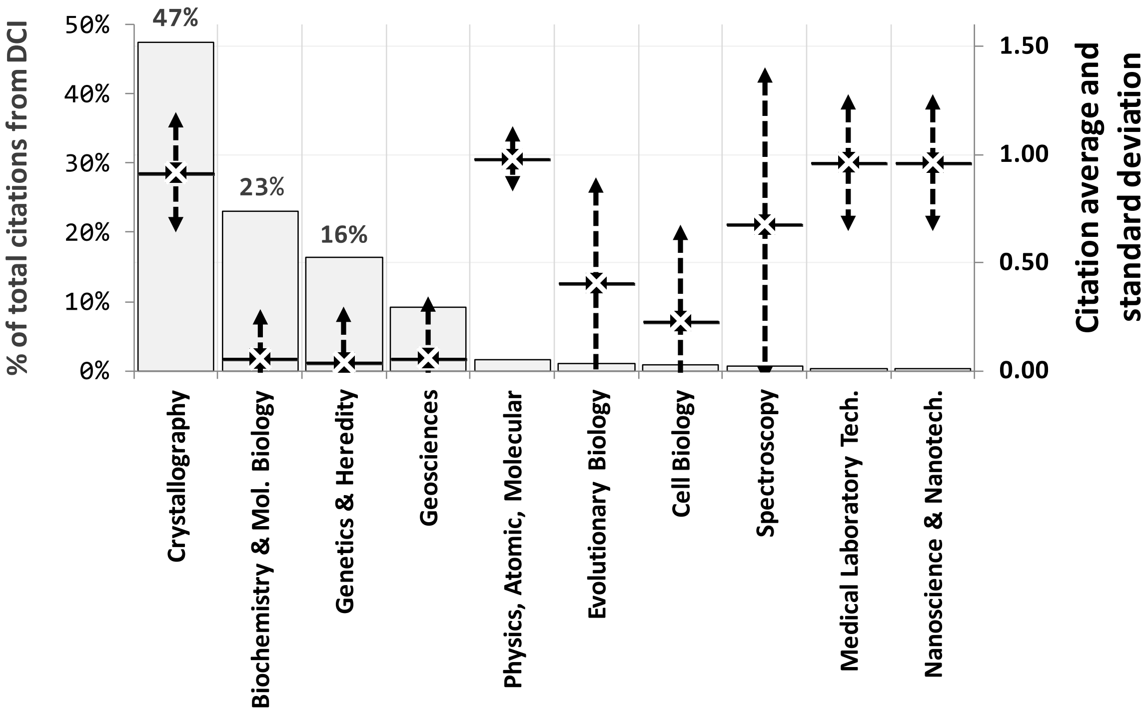
	Total Records	Total Citations	Citation Average	Standard Deviation
Engineering & Technology	1,545	890	0.58	0.94
Humanities & Arts	44,588	1	0.00	0.00
Science	2,004,449	293,193	0.15	0.40
Social Sciences	424,952	7	0.00	0.01

C. Data studies

	Total Records	Total Citations	Citation Average	Standard Deviation
Engineering & Technology	240	26	0.11	0.50
Humanities & Arts	6,847	20,459	2.99	21.72
Science	114,338	26,189	0.23	1.91
Social Sciences	37,855	69,659	1.84	17.34

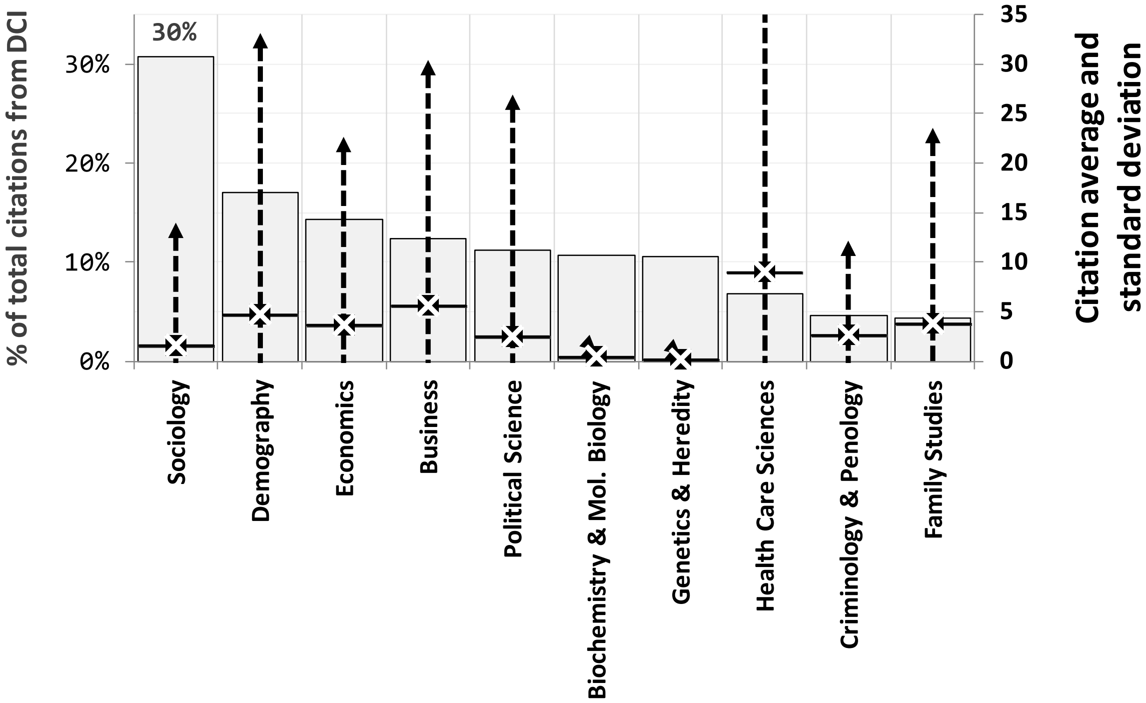
This phenomenon is later confirmed when analyzing the citation distribution by subject categories. In figures 1 and 2 we show the top 10 subject categories according to the DCI with a higher number of citations received. Hence, we see that all top ten subject categories for datasets receiving citations belong to the area of Science (Figure 1). Also, we observe that a single subject category, Crystallography, accumulates nearly half of all citations to datasets. Indeed, this category along with Biochemistry & Molecular Biology and Genetics & Heredity represent 86% of all citations.

Figure 1. Top 10 subject categories with a higher number citations received, citation average and standard deviation for datasets indexed in the Data Citation Index.



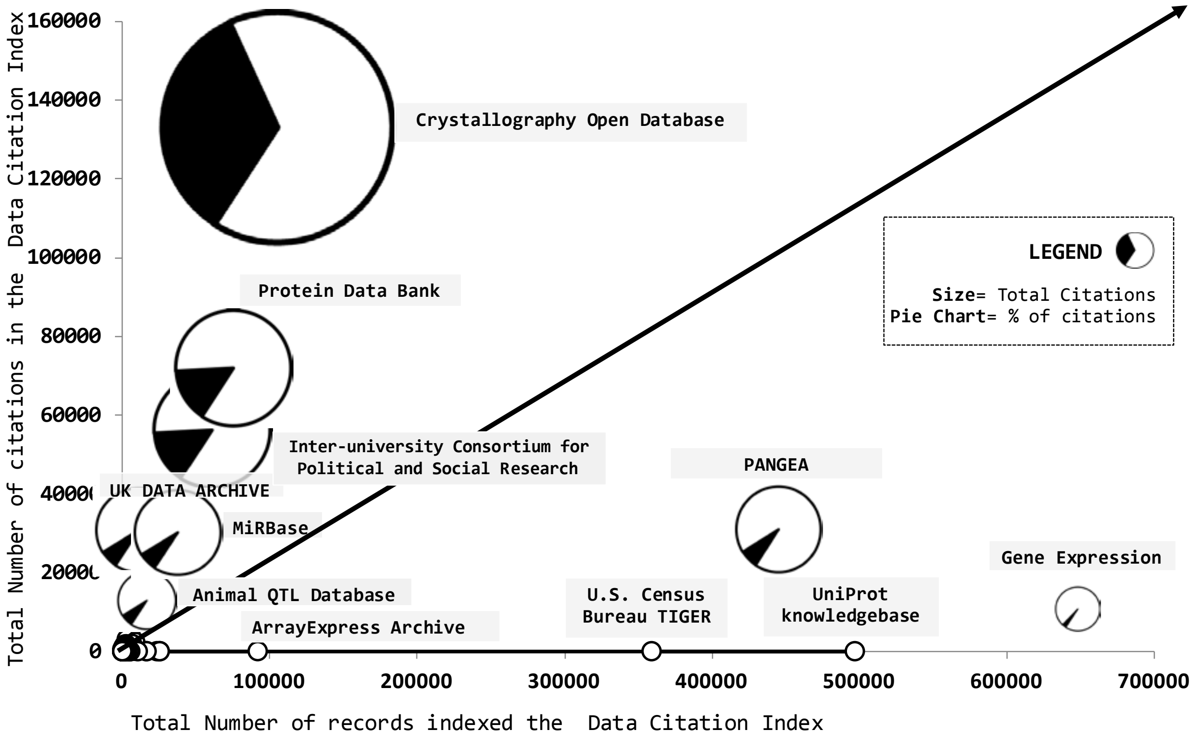
The pictures changes radically in the case of data studies (figure 2). Here, seven of the top ten categories belong to the area of Social Sciences. However, Biochemistry & Molecular Biology and Genetics & Heredity also make it to the top ten along with Health Care Sciences.

Figure 2. Top 10 subject categories with a higher number citations received, citation average and standard deviation for data studies indexed in the Data Citation Index.



In order to explore if such accumulation of citations in specific categories is due to the number of records in these categories, in figure 3 we relate the number of records with the number of citations received for the largest repositories indexed in the DCI. Here, we see that the largest repository is specialized on Crystallography (Crystallography Open Database), followed by the Protein Data Bank and the Inter-university Consortium for Political and Social Research. Also, these three repositories are the ones containing a higher number of citations.

Figure 3. Main repositories in the DCI, citations received and total number of records



Discussion and concluding remarks

In this paper we conduct the first analysis on the citation distribution of the Thomson Reuters' Data Citation Index, a new database launched in 2012 which includes a large number of data repositories associated with citation information. As observed, most of its records have no citations related with them, showing a high rate of uncitedness (88%). This demonstrates that data citation practices are far from common within the scientific community. Also, the DCI is heavily biased towards certain fields from the Hard Sciences as shown by Torres-Salinas, Martín-Martín & Fuente-Gutiérrez (2014), with almost no representation for Engineering & Technology which influences heavily the citation distribution. The reasons for this may not only be attributed to the criteria followed by Thomson Reuters, but to the expansion of data sharing practices within the research community. As indicated before, data sharing practices are not common to all areas of scientific knowledge and only certain fields have developed an infrastructure that allows to use and share data.

Even so, we observe different citation patterns depending on the area of study. While in Science and Engineering & Technology citations are concentrated among datasets, in the Social Sciences and Arts & Humanities, citations are normally referred to data studies. This fact is of extreme importance when conducting a citation analysis on data sharing practices as the chosen field will determine the suitability of focusing on one document type or the other. Similarly to what we see in scholarly publication.

The DCI seems a promising tool which may play an important role as data sharing expands among research fields. Citation analysis may encourage researchers to make their data publicly available as they will be able to analyze the impact of their contribution and the use of their work as well as developing a more open and transparent research process. In this sense, other repositories of a multidisciplinary nature have been launched in the recent years such as Figshare (http://figshare.com) which also seek at including metrics that will indicate the use and discussion awakened by the data displayed. Although data citation analyses do not seem yet appropriate as data sharing practices have not been fully expanded, if they do so, the DCI may well be a tool with a high potential in a near future.

References

Anon. (2008). Community cleverness required. *Nature*, 455, 1.

Borgman, C.L. (2012). The conundrum of sharing research data. *Journal of the American Society for Information Science and Technology*, 63, 1059-1078.

Costas, R., Meijer, I., Zahedi, Z. & Wouters, P. (2012). The value of research data - Metrics for data sets from a cultural and technical point of view. A Knowledge Exchange Report. Retrieved March 3, 2014 from: www.knowledge-exchange.info/datametrics

King, G. (1995). Replication, replication. *PS: Political Science and Politics*, 28, 444-452.

Lynch, C. (2008). Big Data: How do your data grow? *Nature*, 455, 28-29.

Piwowar, H.A., Day, R.S. & Fridsma, D.B. (2007). Sharing detailed research data is associated with an increased citation rate. *PLOS One*, 2, e308.

Piwowar, H.A. & Chapman, W.W. (2010). Public sharing of research datasets: A pilot study of associations. *Journal of Informetrics*, 4, 148-156.

Thomson Reuters (2012). Repository evaluation, selection, and coverage policies for the Data Citation Index within Thomson Reuters Web of Knowledge. Retrieved March 3, 2014 from: http://wokinfo.com/products_tools/multidisciplinary/dci/selection_essay/

Torres-Salinas, D., Martín-Martín, A. & Fuente-Gutiérrez, E. (2014). Analysis and coverage of the Data Citation Index - Thomson Reuters: Disciplines, document types and repositories. *Revista Española de Documentación Científica*, 37, e036.

Torres-Salinas, D., Robinson-Garcia, N. & Cabezas-Clavijo, Á. (2013). Compartir los datos de investigación: Una introducción al 'Data Sharing'. *El profesional de la información*, 21, 173-184.

The impact of academic patenting on publication performance: The multi-dimensional framework[1]

Tung-Fei Tsai-Lin*, Yuan-Chieh Chang**, Bernhard R. Katzy***

*s9873801@m98.nthu.edu.tw
Institute of Technology Management, National Tsing Hua University, 101, Sec.2, Kuang-Fu Road, Hsinchu, 30013 (Taiwan, R.O.C)
Center for Technology and Innovation Management (CeTIM), University of BW Munich, Werner Heisenberg-Weg 39, Neubiberg, 88559 (Germany)

**yucchang@mx.nthu.edu.tw
Institute of Technology Management, National Tsing Hua University, 101, Sec.2, Kuang-Fu Road, Hsinchu, 30013 (Taiwan, R.O.C)

***prof.katzy@cetim.org
Center for Technology and Innovation Management (CeTIM), University of BW Munich, Werner Heisenberg-Weg 39, Neubiberg, 88559 (Germany)
CeTIM, Leiden University, Plantsoen 97, Leiden, 2311 KL (the Netherlands)

Introduction

In the recent decades, commercial knowledge production (e.g. patenting) is emerging new mode of research outcome disclosure in university. However, such proprietary knowledge production obviously is in contradiction with the mandate of public knowledge sharing (e.g. paper publishing). In consequence, there are increasingly patent-paper studies examined the contradiction which employed diversified models. But, those results are not convergent (**Table 1**), and it is lack of a comprehensive framework. Therefore, the aims of this paper are try to frame a multi-dimensional framework (**Figure 1**), and show a new case from National Tsing Hua University in Taiwan.

Patenting behavior	Publication performance			
	Quantity	Quality	Orientations (Journal)	
			Applied	Basic
Inventorship				
Inventors vs. Non-inventors	[1] [2][3](+)	[2](+) [3](x)	[1](x)	[1](+)
Timeframes				
Pre-patenting years vs. Post-patenting years	[1] [2] [4] [5](+)	[4](+)		[4](+)
First patenting year vs. Non-first patenting year	[4](x)	[4](+)		[4] (x)

[1] The authors would thank the China Foundation for the Promotion of Education and Culture, Ministry of Science and Technology (Taiwan), and Research Center for Humanities and Social Sciences (NTHU) for travel support of STI 2014.

Patenting years vs. Non-patenting years	[5](+) [10](x)	[5](+)	[5](+)
Commitment			
Annual filed patents	[7][2][3][8](+) [6](x) [9](-)	[7][3][8][9](+) [6] (x) [2](-)	
Accumulation of patenting years	[7](+) [5](x)	[7](+)	
Accumulation of filed patents	[10](x)	[2](-)	

[1] Van Looy et al. (2006), [2] Fabrizio and Di Minin (2008), [3] Buenstorf (2009), [4] Breschi et al. (2008), [5] Azoulay et al. (2009), [6] Agrawal and Henderson (2002), [7] Czarnitzki et al. (2007), [8] Wang and Guan (2010), [9] Kelchtermans and Veugelers (2011),and[10] Crespi et al. (2011).
Parentheses indicate patent-paper relationship is "+" significantly positive, "-" significantly negative, and "x" not significant.

Table 1 Impact of patenting behavior: inventorship, timeframes, and commitment on publication performance [a]

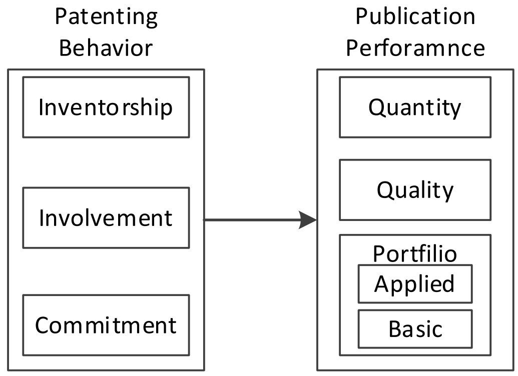


Figure 1 The multi-dimensional framework of patent-paper analysis

An overall view of National Tsing Hua University, Taiwan

Sample

The study constructs a panel dataset, Triple-P (TP), during 2001 to 2010, includes personnel data from faculty's biography, and bibliometrics data from patent filings (USPTO, TWP) and journal publications (SCI). In 2010, there are 615 full-time faculty members. After checking those records, TP has 377 faculty members are shared 61.31% of faculty.

In the final, there are 1,909 patent filings. And, there are 207 faculty members have patenting records, and, the size of control group is 170 faculty members who are non-inventors. In publication data, the total of counts is 8,401, and we identify the nature of publications based on journal title (Hamilton, 2003). In the final, there are 8,169 journal publication counts are classified, and non-classified counts are 232 and shared 2.76% of total counts.

Results

We follows prior research on the design of independent and dependent variables, and we control the career years, junior faculty (joining NTHU after 2001), college (e.g. science, engineering). And, we employ negative binomial regression with fix effect. Table 2 summaries all estimations.

Table 2 Results of regressions: general patenting behavior on publication performance

Independent variables	Relationship to faculty's annual journal publication counts[a]			
	Quantity	Quality	Orientations	
			Applied	Basic
	$Paper.Num_{it}$	$Paper.JIF.Num_{it}$	$Paper.Applied.Num_{it}$	$Paper.Basic.Num_{it}$
Inventorship				
$Inventor_{i}$	.578***	.464***	.209	.547**
Timeframes				
Post $Patent.Year_{it}$	.276***	.058	.392***	.206***
First $Patent.Year_{it}$	.030	.079	.140	-.027
$Patent.Year_{it}$	.013***	.174***	.117**	.157***
Commitment				
$Patent.Num_{it}$	.033***	.039***	.038**	.029*
$Patent.Cum.Year_{it}$	-.010	.032***	-.050**	.008
$Patent.Cum.Num_{it}$	-.003	.003	-.014*	.004

p<0.1, ** p<0.05, *** p<0.01.

Findings

In this paper, we also confirmed academic patenting has significantly positive influence on publication performance from the multi-dimensional results. And, we find that the inventors would increase applied-oriented publications when they are engaging in patenting, but it would not influence the orientations of following publications when they have more experience.

References

Agrawal, A., & Henderson, R. (2002). Putting Patents in Context: Exploring Knowledge Transfer from MIT. *Management Science, 48*(1), 44-60.

Azoulay, P., Ding, W., & Stuart, T. (2009). The Impact of Academic Patenting on the Rate, Quality, and Direction of (Public) Research Output. *The Journal of Industrial Economics, 57*(4), 637-676. doi: 10.1111/j.1467-6451.2009.00395.x

Breschi, S., Lissoni, F., & Montobbio, F. (2008). University patenting and scientific productivity: a quantitative study of Italian academic inventors. *European Management Review, 5*(2), 91-109.

Buenstorf, G. (2009). Is commercialization good or bad for science? Individual-level evidence from the Max Planck Society. *Research Policy, 38*(2), 281-292.

Crespi, G., D'Este, P., Fontana, R., & Geuna, A. (2011). The impact of academic patenting on university research and its transfer. *Research Policy, 40*(1), 55-68.

Czarnitzki, D., Glänzel, W., & Hussinger, K. (2007). Patent and publication activities of German professors: an empirical assessment of their co-activity. *Research Evaluation, 16*(4), 311-319.

Fabrizio, K. R., & Di Minin, A. (2008). Commercializing the laboratory: Faculty patenting and the open science environment. *Research Policy, 37*(5), 914-931.

Funtowicz, S., Shepherd, I., Wilkinson, D., & Ravetz, J. (2000). Science and governance in the European Union: a contribution to the debate. *Science and Public Policy, 27*(5), 327-336.

Hamilton, K. (2003). Subfield and Level Classification of Journals. Document regarding the CHI classification from CHI Research. *Inc. CHI*(2012).

Kelchtermans, S., & Veugelers, R. (2011). The great divide in scientific productivity: why the average scientist does not exist. *Industrial and Corporate Change, 20*(1), 295-336.

Van Looy, B., Callaert, J., & Debackere, K. (2006). Publication and patent behavior of academic researchers: Conflicting, reinforcing or merely co-existing? *Research Policy, 35*(4), 596-608.

Wang, G., & Guan, J. (2010). The role of patenting activity for scientific research: A study of academic inventors from China's nanotechnology. *Journal of Informetrics, 4*(3), 338-350.

The influence of career perspectives on the job choice of recent PhD graduates: a survey of five Dutch universities[1]

Cathelijn J.F. Waaijer*

*c.j.f.waaijer@cwts.leidenuniv.nl
Centre for Science and Technology Studies, Faculty of Social and Behavioural Sciences, Leiden University, P.O. Box 905, Leiden, 2300 AX (The Netherlands)

Introduction

The academic labor market is strongly competitive for its employees. There is a considerable disparity between the aspirations of young scientists and the reality that, depending on field, at most 25% will get a (permanent) faculty position (Stephan 2012). Most studies on the career trajectories of PhD graduates focus on the graduates staying in academic R&D, although a higher percentage will leave academic R&D and start working in other sectors. In the Netherlands we lack reliable information on the sector recent PhD graduates work in and which factors influence their job choice. In this study, we collected such information, determine their perception of career perspectives in academic R&D, non-academic R&D and outside R&D, and assess to what extent these career perspectives influence their job choice. Here, we present preliminary results of this study.

Data and methods

Survey sample

Two sources of PhD graduates were used as a sample for our survey. The first was the survey sample of the 2008 Netherlands Survey of Doctorate Recipients (Sonneveld, Yerkes & Van de Schoot 2010). This survey was sent to 1,096 individuals who had registered for a PhD defense taking place between April 2008 and March 2009 at Utrecht University (a large, broad research university), Delft University of Technology (a technical university), Wageningen University (a university historically focused on agriculture but now broadening its scope to life sciences and environmental research), and Erasmus University Rotterdam (focused on medicine and social sciences, especially economics and management). The second source of PhD graduates in the present study were all individuals who obtained their PhD at Leiden University (a large, broad research university) between January 2008 and May 2012: a total of 1,351 PhD graduates. Combined with the sample from the other four universities, this amounted to a total of 2,447 PhD's.

For these 2,447 PhD graduates on-line searches were conducted using standard search engines, Web of Science and Pubmed to find their current email address or LinkedIn profile. The email addresses were used to send an invitation to our survey. If only a LinkedIn profile was found, the study's researchers sent an invitation to "link" on this website. If a PhD graduate accepted this invitation, the researcher sent the actual survey invitation. For a total of 1,847 PhD's an email was delivered and 378 were invited through LinkedIn and other similar channels, which resulted in a survey sample of 2,225 PhD graduates. The survey

[1] I would like to thank Cornelis van Bochove, Rosalie Belder, Inge van der Weijden and Hans Sonneveld for discussion on the survey questionnaire. Furthermore, I am very grateful to Moniek de Boer, Danique van den Hanenberg, Malu Kuhlmann, Lisa van Leeuwen, Lisette van Leeuwen, Suze van der Luijt-Jansen, and Laura de Ruiter for their work in finding the respondents' current email addresses.

remained open for 91 days. During this period a total of three reminders was sent to potential respondents who had not visited the survey or not completed it.

Results

Description respondents

In total 1,156 PhD graduates started our survey, resulting in a (partial) response rate of 52.0%. Of those, 976 progressed to the final question, resulting in completed response rate of 43.9%. Results are described for all answers given (regardless of completion status). The ratio of male and female respondents was 55-45. Most respondents received their PhD in 2008 (560), 247 did so in 2009, 153 in 2010, 142 in 2011, and 54 in 2012.

Most PhD graduates did their PhD in the medical or health sciences, followed by the natural sciences (Table 1). The majority of respondents obtained their PhD from Leiden University (Table 1).

Table 1. Scientific field by university

	Leiden	*Utrecht*	*Delft*	*Rotterdam*	*Wageningen*	*Total*
Medical and health sciences	230	65	0	64	7	366
Natural sciences	131	58	18	5	34	246
Agricultural sciences	6	3	0	0	22	31
Social sciences	111	29	7	32	7	186
Humanities	114	16	4	3	1	138
Engineering and technology	18	12	77	1	9	117
Unknown	43	13	3	10	3	72
Total	653	196	109	115	83	1156

Employment status and sector of employment

Unemployment is very low among the respondents, with only 2.4% not doing paid work and looking for a job (Table 2). The vast majority of working PhD graduates are employed as employees. We assessed whether respondents were still involved in R&D and if so, if they were working in academic R&D (working at a university, university medical centre or research institute) or in non-academic R&D (performing R&D at other public institutions or at private companies). Over sixty per cent of all respondents are working in academic R&D, a quarter is working in non-academic R&D and just over one tenth is not performing R&D in their job (Table 3). Employment in academic R&D is especially high for the social sciences, whereas it is lower than average for engineering and technology PhD graduates. Instead, a high percentage of the latter is involved in non-academic R&D. Finally, humanities PhD’s are more often working outside R&D than PhD’s from other scientific fields.

Table 2. Respondents by employment status (multiple answers possible)

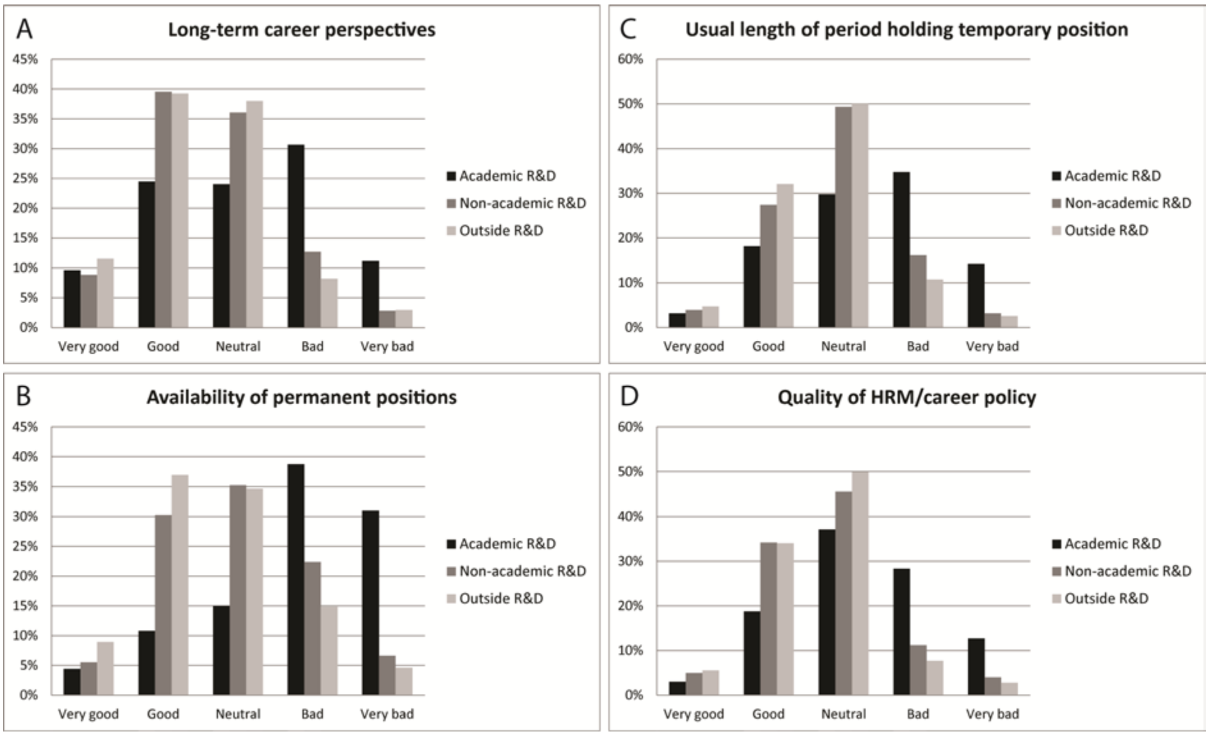
	%
Working	**96.3**
Employee	89.4
Self-employed	8.7
Working paid or unpaid for own or family's business	1.7
Away from work ill, on maternity leave or temporarily laid off	2.0
Doing any other kind of paid work	0.9
Unemployed	**2.4**
Looking for job	2.4
Waiting to start job	0.3
Inactive	**1.3**
Retired	0.4
Student	-
Looking after home or family	0.4
Long-term sick or disabled	0.1
Other	0.5
n	*1118*

Table 3. Sector of employment by scientific field (%)

	n	Academic R&D	Non-academic R&D	Outside R&D
		%		
Medical and health sciences	*328*	65	22	13
Natural sciences	*230*	64	28	8
Agricultural sciences	*27*	67	33	0
Social sciences	*55*	74	15	11
Humanities	*122*	64	16	20
Engineering and technology	*108*	46	41	13
Total	*970*	64	24	12

Perception career-related aspects in different sectors
Our first question is how recent PhD graduates perceive several career-related aspects in academic R&D, non-academic R&D and outside R&D. Respondents were asked to rate (1) general long-term career perspectives, (2) the availability of permanent positions, (3) the usual length of holding a temporary position, and (4) the quality of HRM and career policy on a five-point Likert scale. All aspects were rated more negatively for academic R&D than for non-academic R&D and outside R&D (Fig. 1, $p < 0.001$ between groups, ANOVA using five-point scale). The difference is especially pronounced for the perceived availability of permanent positions (Fig. 1B).

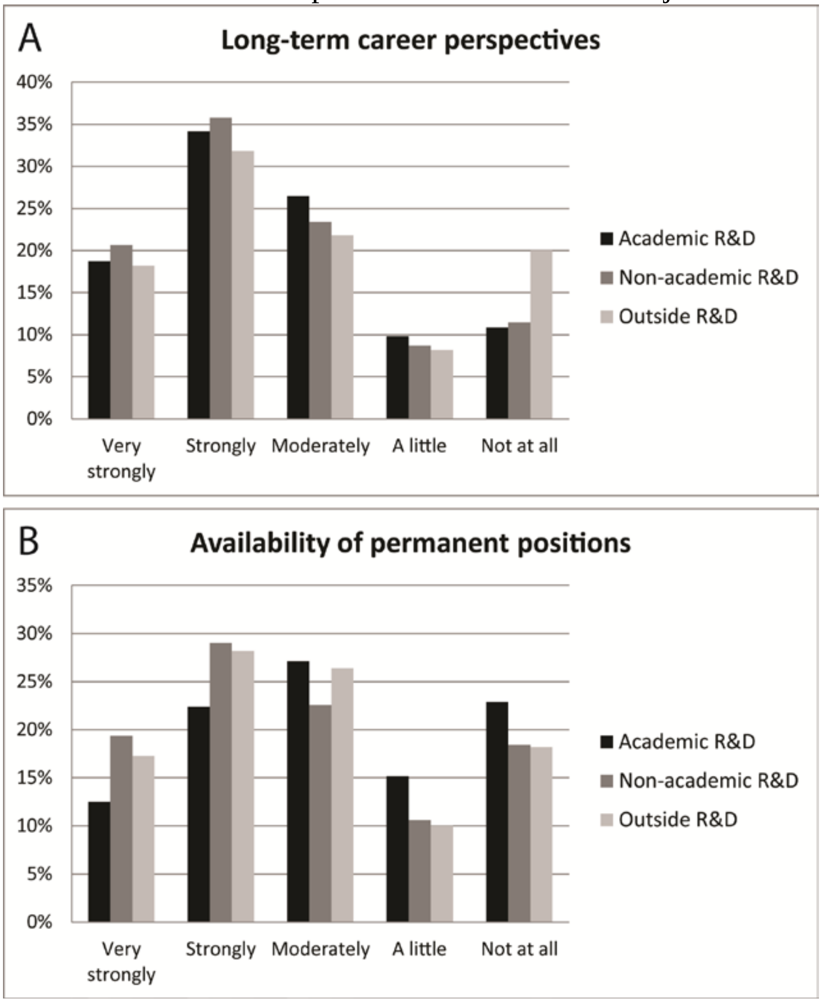
Figure 1. Perception of career-related aspects in academic R&D, non-academic R&D and outside R&D



Self-reported influence of career-related aspects on job choice
The next step was to determine whether the more negative perception of career-related aspects in academic R&D influences the job choice of PhD graduates, by asking the question whether these aspects have made them hesitate to choose a job in academic R&D. Indeed, over eighty per cent of respondents indicate that long-term career perspectives have made them hesitate moderately, strongly or very strongly to choose an academic research job. The same holds true for the lack of available permanent positions: about 65% say it has made them hesitate moderately, strongly or very strongly to pursue a job in academic R&D. Conversely, PhD graduates do not report a large influence on career choice of the usual length of the period in temporary posts nor of the quality of HRM/career policy in academic R&D.

When breaking down the reported effects by sector of employment, hardly any differences can be observed (Fig. 2). PhD graduates working in non-academic R&D do report a significantly larger influence of the availability of permanent positions on their job choice than their counterparts working in academic R&D, but the difference is not very large (Fig. 2B).

Figure 2. Influence of career-related aspects in academic R&D on job choice



Discussion and conclusion

In this paper we present the first results of a survey among 1,156 recent PhD graduates (response rate: 52%). The main finding of our study is that several career-related aspects, such as long-term career perspectives and the availability of permanent positions are judged much more negatively for the sector of academic R&D than for non-academic R&D and non-R&D. Furthermore, a majority of PhD graduates indicate that these two aspects made them hesitate to choose a job in academic R&D.

Future work

In the future we will further elaborate the effect of career-related aspects on job choice and will look at their influence on job satisfaction. Furthermore, we will assess the effects of temporary job contracts on the quality of R&D and on the personal lives of PhD graduates.

References

Sonneveld, H., Yerkes, M. & van de Schoot, R. (2010). Ph.D. trajectories and labour market mobility: a survey of recent doctoral recipients at four universities in the Netherlands.

Stephan, P. (2012). *How* economics *shapes science*. Cambridge: Harvard University Press.

Characterizing the research landscape of Influenza A: new tools to inform policymaking

Matthew L. Wallace* and Ismael Rafols**

* *matwal@ingenio.upv.es;*
INGENIO (CSIC-UPV), Ciudad Politécnica de la Innovación, Valencia, 46005 (Spain)

** *i.rafols@ingenio.upv.es*
INGENIO (CSIC-UPV), Ciudad Politécnica de la Innovación, Valencia, 46005 (Spain)
and
Science Policy Research Unit, University of Sussex, Falmer, Brighton, BN1 9SL (United Kingdom)

Abstract

In this paper, we present a novel application of mapping and clustering tools for the management and deliberation on research portfolios. It is based on a need for new ways to make decision on how public funding is allocated according to entire portfolios of an organisation, given a specific societal problem such as avian flu or climate change. We present empirical evidence for 1) identifying and characterizing different areas of a complex multi-faceted research landscape; 2) understanding the *de-facto* prioritization of research or setting of research portfolios by some of the main research funders. As an example, we apply this methodology to the case of Influenza A as an instance of public, outcome-driven research, where science policy can benefit from a holistic, portfolio-oriented approach to funding allocation decisions.

Introduction

Tackling complex global challenges – climate change, food security, the risk of global pandemics – not only requires increases in R&D expenditure and coordination, but also requires a willingness to consider a broad range of research options. More generally, recent years have seen greater effort to account for public funds spent on R&D and an effort to expand the evidence base for measuring impacts of science, but also for designing scientific programmes and allocating resources (Feeling et al., 2011). However, accounting for public research investment in purely monetary terms is not necessarily desirable or practical, leading scholars and policymakers to inquire as to other measures of “value” of public research, namely in terms of desired social outcomes (Pielke and Sarewitz, 2007; Sarewitz and Bozeman, 2011). We contend that these efforts towards a systematic understanding of a “research portfolio” and its societal value are a key step forward in this direction.

A “portfolio” approach to research is becoming increasingly popular as science funders and performers strive not only to maximize the quality of individual research projects, but also to maximize the overall performance of a given set of projects in terms of a set of desired outcomes spanning multiple dimensions such academic excellence, economic impact and social benefits (e.g., Srivastava, 2007; Bozeman and Rogers, 2001). Concretely, portfolio or system-based approaches have been used managing multi-faceted research on large-scale problems such as global Malaria eradication (Alonso et al., 2011) or selecting new public energy R&D portfolios (Stirling, 2010). Many large funders of public research are increasingly focusing on enhanced cooperation and coordination to address major societal challenges, which often require multidisciplinary and multisectoral approaches.

In order to make use of such a "portfolio" perspective in the context of social impacts-driven research, we propose to examine the broader research landscape and focus on linkages between research avenues or projects within portfolios, providing the evidence base for decisions which reflect a variety of options for resource allocation. The landscape can be based on cognitive information such as bibliometric data, and can be designed to reflect social, institutional and geographical characteristics, all critical information for policymakers.

We illustrate some of these ideas through statistics and maps derived from bibliometric data related to research on Influenza A. This category comprises a variety of more or less pathogenic strains of the virus, which, in some cases, can be transmitted to humans via carriers such as pigs and birds. Most recently, certain strains (e.g., H1N1, H5N1 and H7N9) of the disease have been the centre of increased attention due to the possibility of worldwide pandemics, reminiscent of the 1918 so-called "Spanish" flu. While each strain presents distinct issues, there are common research challenges related to zoonosis (transmission from animals to humans), virus mutations, mechanisms of transmission, the development of vaccines and other treatments. Most importantly, the global nature of the challenge has revealed distinct, and often competing, narratives related to the control of avian strains, following the series of pandemics beginning in 2003. Specifically, the international response has been dominated by a focus on (i) agriculture and livelihoods, (ii) on public health and (iii) pandemic preparedness, accompanied by other strong narratives relating to risk (Abeysinghe and White, 2011; Scoones, 2010). In this paper, we propose a method for gaining insight into research landscape of Influenza A, as well as into the constitution of *de facto* research portfolios, which reflect a diversity of priorities among funding organisations.

Data and methods

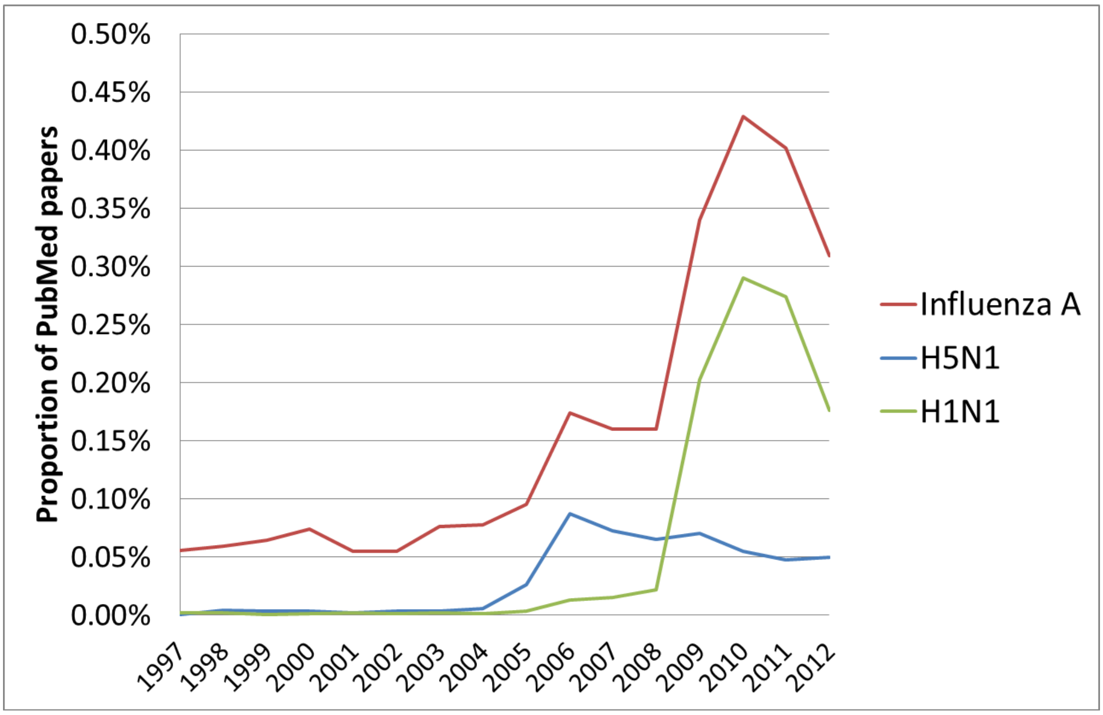
Our approach is driven by a need to examine the research landscapes centred on a specific challenge, which can encompass a series of outcomes. Within a research landscape, an organisation or agency covers certain areas – hence we distinguish between the overall research landscape of Influenza A, and the research portfolio of a given organisation such as the Wellcome Trust.

Recent developments in techniques for overlaying maps has allowed for a new glimpse into the structure of science in terms of cognitive distances (Rafols, Porter and Leydesdorff, 2010). In the case of specific research topics such as Influenza, it appears that co-word and co-citation analyses provide insights and content associated with different parts of the research field. We use here a co-term analysis, a means to visualize and analyze domains of research (Ding, 2001), in this case through clustering. Using a textual analysis allows us to more easily make connections to the narratives that underlie the expected social outcomes of research. We make use of the VOSViewer software in order to identify the relevant terms from a set of bibliometric articles, then perform a clustering and mapping analysis (Eck and Waltman, 2010; Waltman, Eck and Noyons, 2010).

We rely on the Medical Subject Headings (MeSH), provided through the Public Library of Medicine (PubMed), which, unlike titles or abstracts, provide a means to accurately capture bodies of research relating to a specific medical disease or condition (Lu, 2011; Leydesdorff, Rotolo and Rafols, 2012). PubMed provides a range of information on the nature of research (e.g., through "qualifiers") and the branch structure of MeSH provides a means to characterize the "Influenza A" category in terms of a set of dominant Influenza strains. The MeSH method allows us to identify over 20,000 papers related to Influenza A between 2000 and 2012. Like many other cases of multi-faceted research, “avian influenza” or “swine flu” are terms widely used within the policy arena, which should be analyzed within the parent category of “Influenza A” in order to appreciate existing and potential connections within the landscape. Finally, these papers are linked to Web of Science (WoS) entries using the method described in (Leydesdorff & Opthof, 2013). While this only captures on the order of 80% of PubMed entries, it allows us to use the bibliometric data from both sources.

Finally, we can produce overlays of the maps by using the funding acknowledgment data from WoS. While this data only exists for more recent WoS entries, for which the coverage is not complete, and does not quantify the levels of funding, we can nevertheless clean the data manually in order to identify at the most basic level in which areas of the research landscape a given funder is most likely to be involved. In order to produce the overlay itself, we have developed a simple template which uses the VOSviewer data mapping each term to the documents where it is contained. This data is then combined with the “local” (i.e., restricted to the subset of documents from the given funder) term network, in order to identify the links between individual elements of the portfolio and the occurrence of the elements. The overlay is then developed by fixing the positions of each term in the global network, but adjusting their size and showing the links between them based on the local network.

Figure 15: Number of papers with “major topic” MeSH terms as “Influenza A”, to sub-types (H5N1 and H1N1).



Results: Understanding the research landscape of Influenza A

Figure 1 shows the evolution of research on different strains, which reflects the several (relatively) small H5N1 outbreaks occurring as of 2003, and a large H1N1 outbreak in 2009. Bibliometric data can provide insight as to how the scientific community responds to these drivers. Figure 2 shows the overall evolution of research on “Influenza A”, keeping only terms (from abstracts) which have more than 10 occurrences and which are among the 60%

most "relevant" (i.e., least common). The evolution reflects new themes (e.g., H5N1), as well as the rise or decline of approaches such as developing vaccines and understanding outbreaks. After 2009, for instance, there is less work related to understanding the spread of disease and monitoring it from a public health standpoint. The red cluster, which is primarily related to clinical medicine and public health, becomes more focused on clinical trials and diagnosis, distancing itself from epidemiological research and molecular biology.

Figure 16: Research landscapes of Influenza A prior to 2004 (top), in 2004-2009 (middle), and in 2010-2012 (bottom).

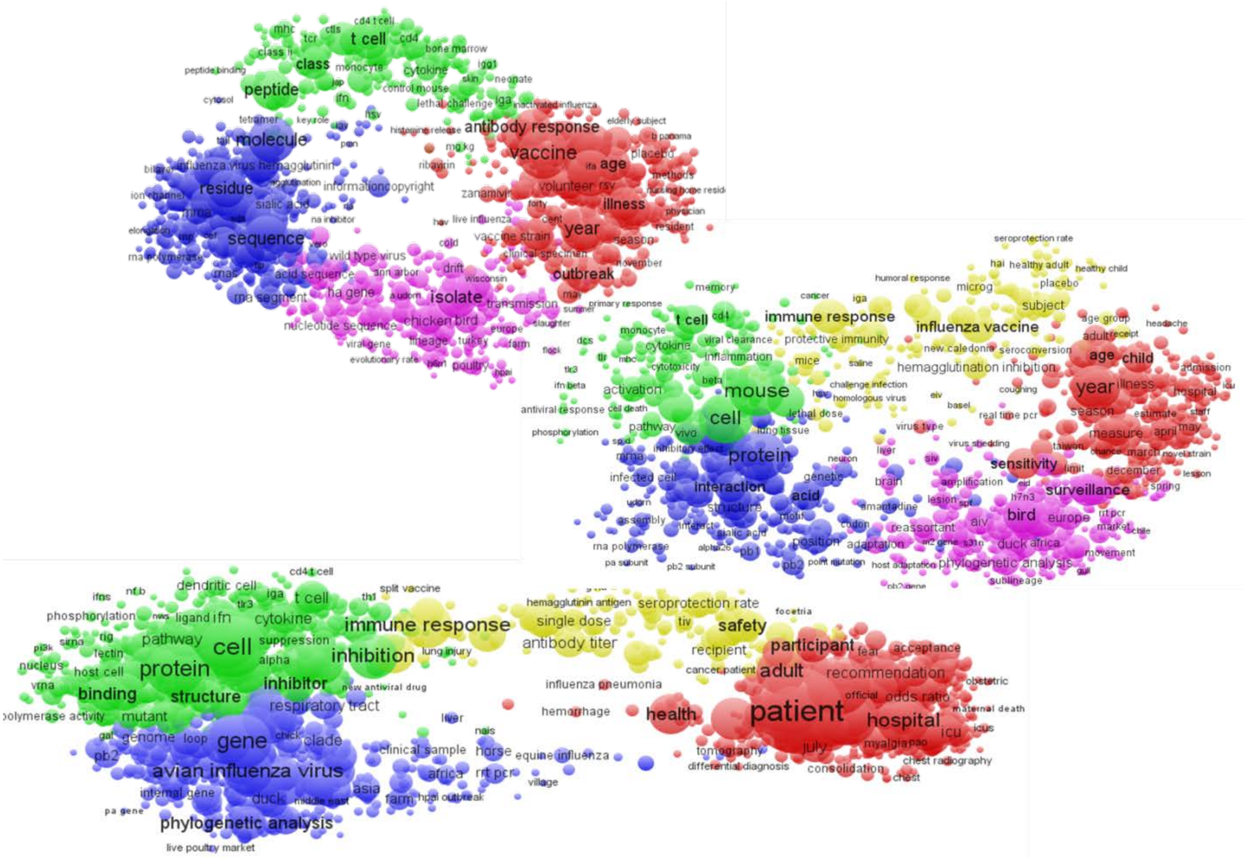
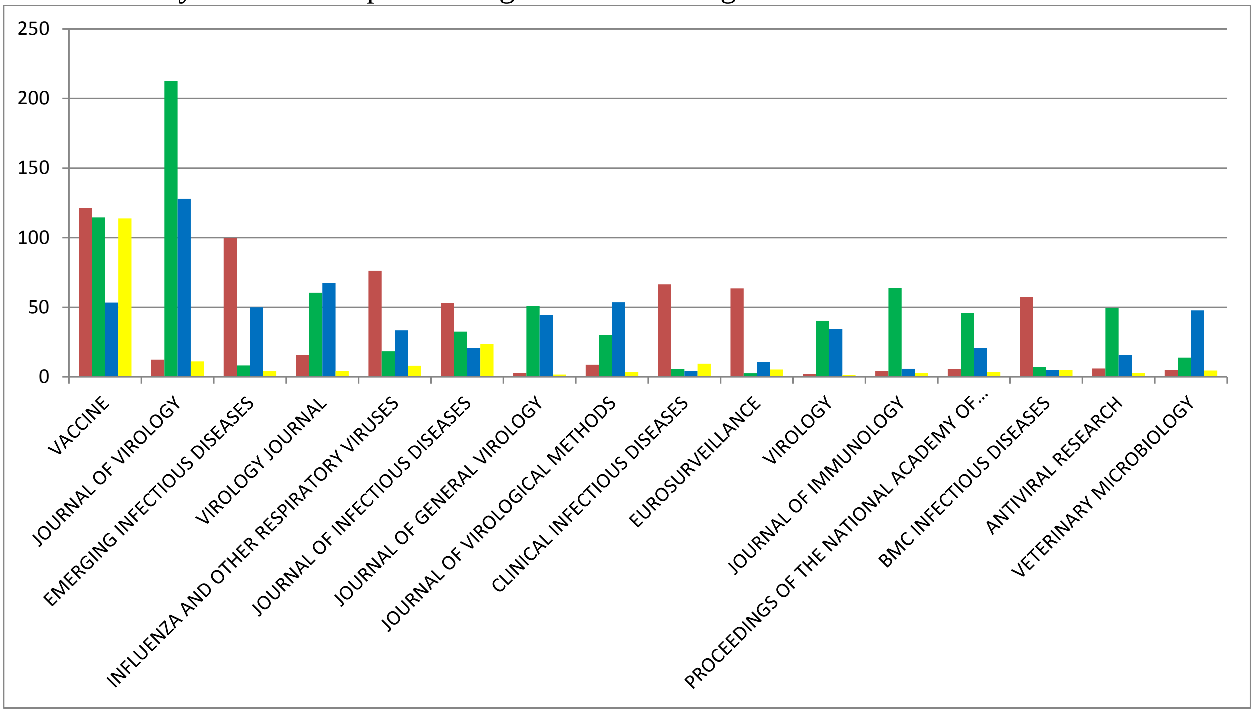


Figure 17: Top journals, according to each of the clusters. Red: Clinical Research. Green: Immunology. Blue: Virology (with a focus on zoonosis). Yellow: Vaccine development. The scores on the y-axis are computed using the relative weights of each cluster



In order to interpret the results from the co-term analysis, we can examine what specifically characterizes the four clusters identified, using the frequency of publications in a given journal (Figure 3), combined with data on MeSH term qualifiers and other bibliometric information. Indeed, the research landscape is subtle in that the different topics are not dominated by a given journal (or scientific specialization). Specifically, In this case, we are not considering the size of the journal, but simply linking the clusters, via the occurrence of each individual term, to each journal. This only allows us to compare the among clusters in a given journal, but does not provide a means to compare between journals in a meaningful way. One can nevertheless look for journals where there is a dominance of one or two main clusters, which allows us to infer a certain similarity in audience between the green and blue clusters, as well as the yellow and red clusters.

Based on comparisons with co-citation maps and other complementary bibliometric information, we can characterize the dominant Influenza A research avenues in 2010-12 as: “clinical medicine” (red), immunology (green), virology (blue) and vaccine development (yellow). The overall description of the research landscape through the maps, as well as the characterization of each cluster, is assessed as part of a series of 15 interviews with leading researchers and stakeholders. While the interviewees found the Influenza A maps to be representative and revealing (to various degrees), responses also point to the existence of other ways of characterizing the landscape (e.g., focusing on applied and basic research, or on lab-based and field-based research).

The main question is then how the characterization of research avenues can be used to (very partially) inform how future research can be organized across research avenues (clusters). First, the maps provide a sense of the cognitive distance between different parts of the research landscape. In the case of Influenza A, Figure 2 points to a progressively greater distance between more applied clinical research, and “basic” virology or immunology research. This cognitive distance can be both a barrier and an opportunity to setting up a research portfolio with complementary pathways, either facilitating or hindering new collaborations or translational research (Hörig, Marincola and Marincola, 2005).

In addition to cognitive considerations, the social and institutional characteristics of research are critical pieces of the landscape. Concretely, expanding upon the visualizations to provide a richer analysis of the landscape can be accomplished through descriptive statistics or by overlaying the data of subset of the landscape. We can also present bibliometric examples of some of the social characteristics of the work done within each cluster or research avenue. For example Figure 4 shows the percentage of a given number authors collaborating in a paper for each research avenue and Figure 5 the percentage of different countries.

Figure 18: Distribution of collaboration sizes (as inferred from number of authors) by cluster, showing similar practices within virology and immunology research. Vaccine development, on the other hand, is dominated by larger. The x-axis represents the number of co-authors, while the y-axis is the normalized probability density.

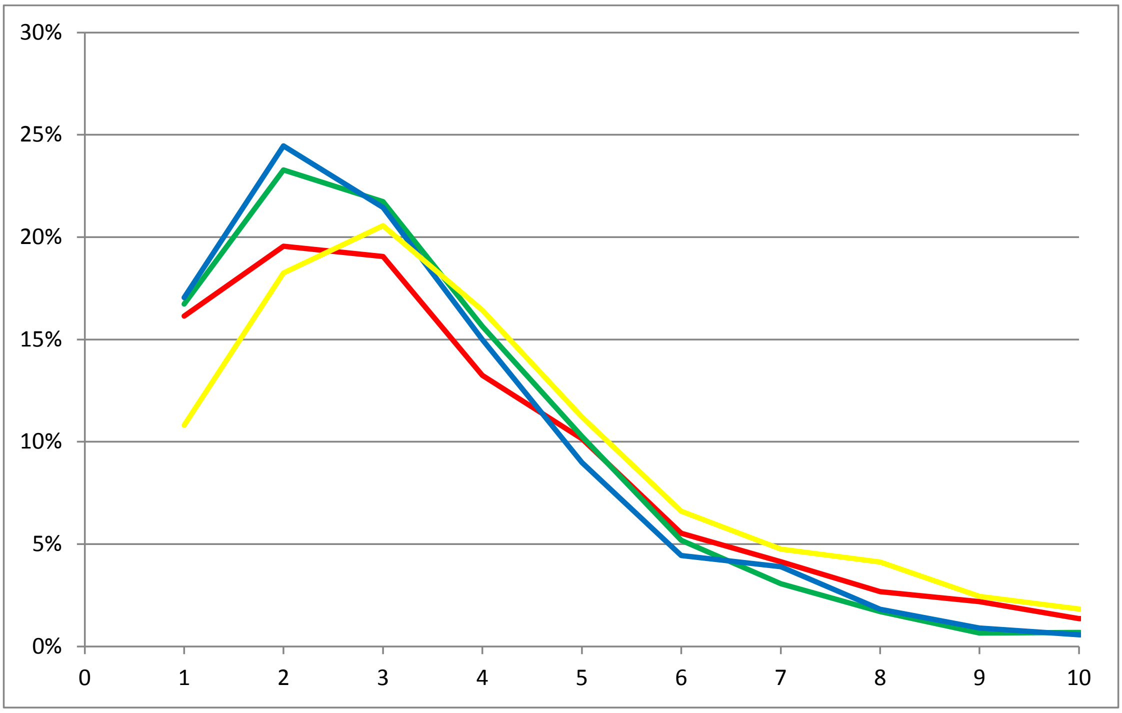
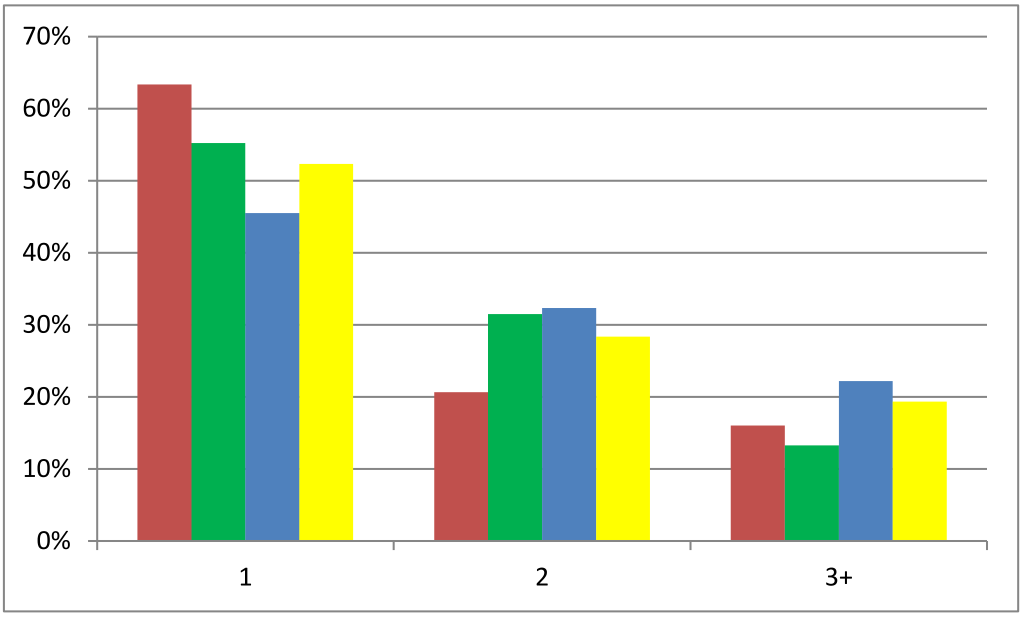


Figure 19: Relative proportion of the average number of countries involved in co-authoring each research paper, an important indicator for understanding how "global" biomedical problems are tackled. The x-axis represents the number of different countries represented among the list of co-authors, while the y-axis is the normalized probability density (by cluster).



Discussion: using research landscapes and research portfolios in decision-making

Here we describe three policy considerations supported by the visualisation of research landscapes and portfolios: (i) helping to find links between research avenues and societal outcomes; (ii) thinking about interactions and trade-offs between research avenues; (iii) locating the position of the portfolio of an organisation in the research landscape. As mentioned above, the generally positive reception of the maps as a heuristic by policymakers and researchers is encouraging. However, we must caution against taking the bibliometric maps as representing a unique "reality" of the research landscape.

First, this analysis can help connect research avenues to social outcomes and values favoured by stakeholders. While research is usually thought in terms of content, social outcomes can refer to mitigating the current and potential effects of Influenza A, but also to supporting capacity development in terms of human resources. The landscape is useful to think one can begin to explore concrete risks, costs and benefits associated with various research avenues, as well as options for portfolios which combine the various approaches. For example, some stakeholders associate certain topics (e.g., manipulation of genetic material) with high levels of return and risk, not only in terms of achieving an objective of developing solutions to future mutations, but also in terms of a perceived risk of use for military or terrorist purposes, the so called "dual-use" research (Kuehn, 2012).

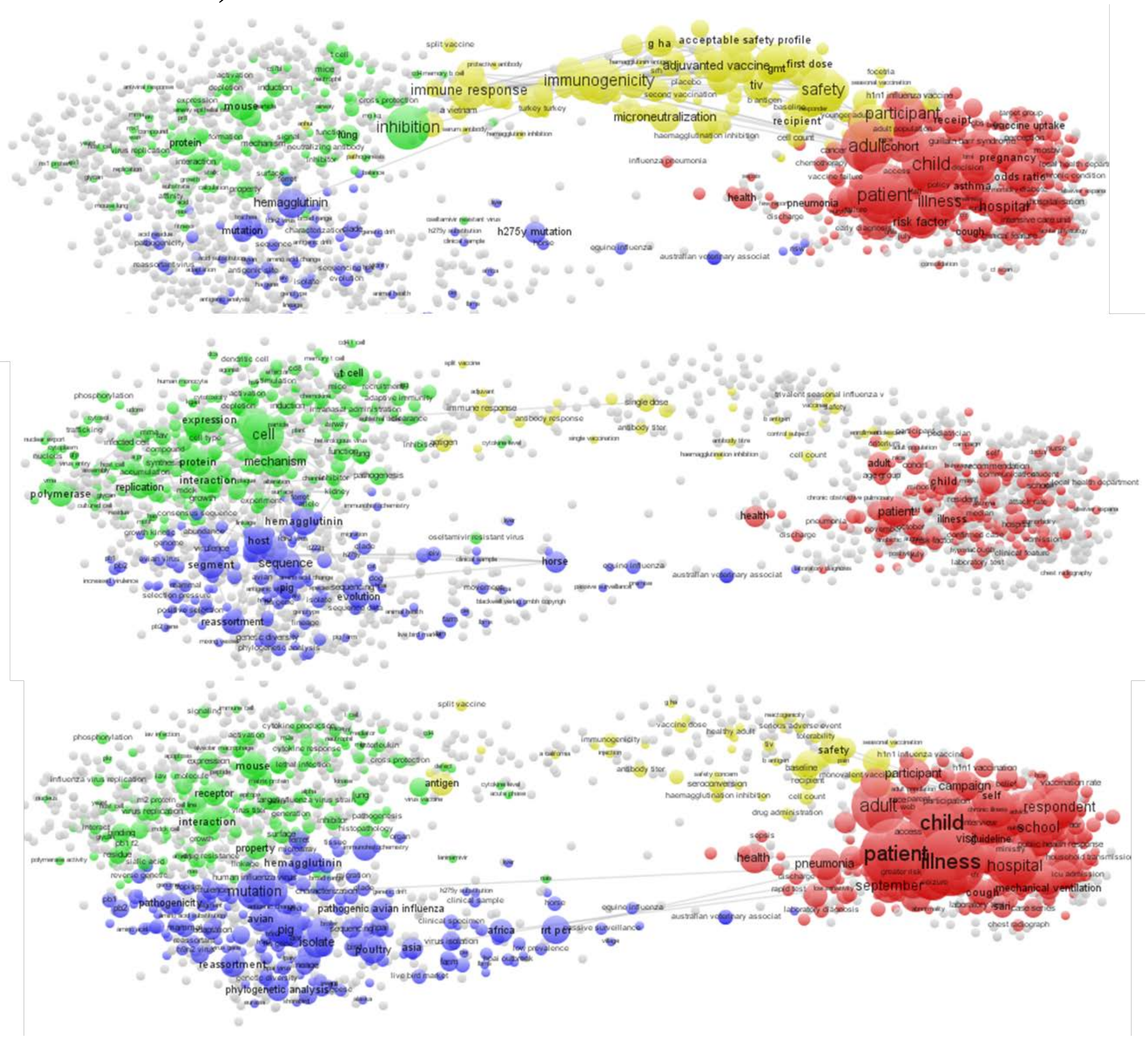
More generally, one can conceive of trade-offs between research which aims to operate within a short timeframe and with a relatively high chance of "success" (e.g., public health or clinical trials), as opposed to that which could be expected to provide larger, global solutions, but with a lower likelihood of "success" and on a longer timeframe (e.g., understanding viruses). Furthermore, one can begin to conceive of potential synergies or complementarities between specific topics or general research avenues. In this regard, it is possible to overlay a wide variety of data onto the research landscape to identify countries or research groups which have been found to make links across topics.

Descriptions of the social dynamics associated with different types of research also inform how research portfolios are designed and how these trade-offs are viewed. For example, Figures 4 and 5 show that while larger collaborations occur in clinical medicine and vaccine development, there is more international collaboration in the field of epidemiology. This type of information can help foster specific types of collaborations or seek out synergies between research areas.

Finally, an overlay technique we have developed for term-based maps allows us to visualize existing research portfolios by combining the data from PubMed with funding data from Web of Science from the 2010-12 period, retaining only original research and review articles. Three of the main funders of Influenza A research are shown in Figure 6. This approach provides indications of possible gaps in public research and can be used not only to take a critical look at existing portfolio configurations, but also at possible complementary portfolios among funding organizations. As shown in Figure 6, by also adding in the overlay maps information on the links between areas (based on only the elements contained within a given portfolio), we can get a sense of the degree to which an individual portfolio is explicitly creating relations between the research topics it covers.

These types of overlays are a starting point for decision-makers to scrutinize portfolios and potential mixes of different research avenues. They can also serve to raise specific and well-defined normative questions around how different portfolios should interact. In the Influenza A example, to what degree should vaccine development be left to the private sector? How much overlap should there be in basic research done by two public agencies within the same country? How much coordination is required between dominant public funders worldwide? And, more generally, does the current mix of research themes correspond to the range of outcomes sought by key stakeholders?

Figure 20: Overlays maps based publications associated with funding sources: GlaxoSmithKline (top), Wellcome Trust (middle) and the U.S. Centers for Disease Control (bottom). Also shown are the dominant lines of the local networks (most of which are hidden within the clusters).



Conclusion

This paper provides an example of how conventional scientometrics tools for mapping and clustering tools can provide a concrete means to support policy-makers' thinking about a given grand challenge. This is done by characterizing the evolution of the overall research landscape of a given topic in terms of interrelated research avenues. This mapping can be complemented by other bibliometric data and overlay techniques, which can help provide information on the existing research portfolios of specific organizations. Indeed, this type of bibliometric data can both provide evidence to support decisions and can be very useful in eliciting or clarifying the variety of perspectives that are at the core of many "grand challenges" that dominate public research investments on national and international scales.

We recognize that such maps are very partial tools, subject to various interpretations, and possibly missing some part of the actual research landscape, particularly research more associated with practitioners and not always published in international literature. This is more problematic in fields such agriculture, where practices are often developed locally without publication than in the medical area, where codification in publications in more institutionalised.

References

Abeysinghe, S., and White, K. (2011). The avian influenza pandemic: Discourses of risk, contagion and preparation in Australia. *Health, Risk & Society*, *13*(4), 311–326.

Alonso, P. L., et al. (2011). A research agenda to underpin malaria eradication. *PLoS Medicine*, *8*(1), e1000406.

An, X.Y., and Wu, Q.Q. (2011), Co-word analysis of the trends in stem cells field based on subject heading weighting, *Scientometrics*, 88, 133-144.

B. Bozeman and D. Sarewitz (2011), "Public Value Mapping and Science Policy Evaluation", *Minerva* 49, 1-23.

Ding, Y., Chowdhury, G. G., & Foo, S. (2001). Bibliometric cartography of information retrieval research by using co-word analysis. *Information Processing & Management*, *37*(6), 817–842.

Feeling, K.H. *et al.* (ed.), *The Science of Science Policy: A Handbook*, Stanford: Stanford University Press, 2011.

Horig, H., Marincola, E., & Marincola, F. M. (2005). Obstacles and opportunities in translational research. *Nat Med*, *11*(7), 705–708.

Kuehn, B. (2010). US Reviews High-Risk Research Portfolio, *JAMA 307*, 1682.

Leydesdorff, L. Rotolo, D and Rafols, I. (2012). Bibliometric perspectives on Medical Innovation Using the Medical Subject Headings of PubMed, *Journal of the American Society for Information Science and Technology*, 63, 2239-2253.

Leydesdorff, L and Opthof, T. (2013) Citation analysis with Medical Subject Headings (MeSH) Using the Web of Knowledge: A New Routine, *Journal of the American Society for Information Science and Technology*, 64, 1076-1080.

Lu, Z. (2011). PubMed and beyond: a survey of web tools for searching biomedical literature. *Database : The Journal of Biological Databases and Curation*, *2011*, baq036.

Rafols, I., Porter, A. L. and Leydesdorff, L. (2010). Science Overlay Maps : A New Tool for Research Policy and Library Management, *Journal of the American Society for Information Management and Technology*, 61(9), 1871–1887.

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.

Scoones, I. (2010), Unpacking the International Response to Avian Influenza: Actors, Networks and Narratives, in I. Scoones (ed.), *Avian Influenza: Science, Policy and Politics*, London: Routledge, p. 19-64.

Srivastava, C. V., Towery, N. D., & Zuckerman, B. (2007). Challenges and opportunities for research portfolio analysis, management, and evaluation. *Research Evaluation*, *16*(3), 152–156.

Stirling, A. (2010), Multicriteria diversity analysis: a novel heuristic framework for appraising energy portfolios, *Energy Policy 38*, 1622-1634.

Van Eck, N. J., & Waltman, L. (2010). Software survey: VOSviewer, a computer program for bibliometric mapping. *Scientometrics*, *84*(2), 523–538.

Waltman, L., Eck, N. J. Van, & Noyons, E. C. M. (2009). A unified approach to mapping and clustering of bibliometric networks, *Journal of Informetrics 4,* 629-635.

The organization of science: teams and networks

Jian Wang and Diana Hicks***

**jian.wang@kuleuven.be*
Centre for R&D Monitoring (ECOOM) and Department of Managerial Economics, Strategy and Innovation, K.U. Leuven, Belgium
Institute for Research Information and Quality Assurance (iFQ), Berlin, Germany

***dhicks@gatech.edu*
School of Public Policy, Georgia Institute of Technology, Atlanta, GA, USA

Abstract
Science is increasingly produced by collaborative teams, however, the language of teams, and the concepts used borrow from the study of teamwork in situations where teams are fairly well defined, closed and work together over long periods of time. These concepts do not accurately represent collaboration in science where teams are fluid and interdependent. Taking this reality into account, we argue here than an egocentric network approach is a more useful framework for the study of research than either the individual or team as unit of analysis. Using publication records of 1,310 American scientists in five disciplines, this paper demonstrates (1) the instability of scientific teams and (2) significant knowledge spillovers from new collaborations to repeated collaborations in the same egocentric network. Finally, this paper discussed a *multiple-stream model* to illustrate a network form of science organization.

Introduction
Science is increasingly produced collaboratively, not only in *big sciences* but across all scientific disciplines (de Solla Price, 1986; Hicks & Katz, 1996; Wuchty, Jones, & Uzzi, 2007), and team has become an important unit of analysis for science studies. By bringing in insights from the psychology literature on small groups, scholars have portrayed themselves as establishing a new *science of team science*. In contrast to this approach, this paper proposes an egocentric network perspective. We justify collaborative networks as legitimate organizations of science by demonstrating how the context of scientific work challenges the traditional team approach imported from other work contexts and suggesting a network approach can overcome some of these challenges. Specifically, we demonstrate (1) the fluidness of and (2) the interdependence between collaborative teams in science. Since team assembly and operation are embedded in open networks, the external activities responsible for team outcomes should be taken into account for a better understanding of science production. We also discuss a *multiple-stream model* to illustrate a network form of science organization. In summary, this paper defends the collaborative network as the best unit for analysis of scientific work.

Fluidness of teams
One important characteristic distinguishing modern science from other systems of work organization is its autonomy and self-governance (Whitley, 2000). Scientific teams are largely voluntary, and scientists have substantial autonomy to create, maintain, restructure, and dissolve their teams. This autonomy enables fluid teams in science characterized by ill-defined boundaries and unstable memberships.

We demonstrate this using publication data in 2005-2007 (7,678 Thomson Reuters Web of Science (WoS) journal articles) of 1,148 American scientists (i.e., egos) in five disciplines (biology, chemistry, computer science, earth and atmospheric sciences, electrical engineering), we recover collaborative teams associated with these egos. We view a publication as a product of a project and the group of authors as the project team. Teams are classified based on team size: doublet, triplet, quartet, and so on and we examine the stability of these teams. As shown in Table 1, teams are mostly one-time phenomena and do not repeat. In addition, as team size increases, the percentage of repeated teams decreases. About 20% of the doublets have repeated, that is, have more than one paper, and only about 9% of the triplets and 4% of the quartets have repeated. Although teams do not repeat in the exact form, they do frequently "repeat" in slightly different forms. Take triplets as an example, 18% of them have papers authored by the whole team and scientists outside the team. In addition, more than 63% of them have papers authored by a subset of the team and scientists outside the team. To be more specific, a set of authors, S_1, forming a focal team have a 63% chance of appearing on other papers including another set of authors, S_2, such that (a) S_2 contains the ego, (b) at least one member of S_1 is absent from S_2, and (c) at least one member of S_1, other than the ego, is present in S_2. Please refer to (NETWISE, 2007; Wang, 2014; Wang et al., 2012) for more information about the data.

Table 1. Fluidness of collaborative teams

	Doublet	**Triplet**	**Quartet**	**Quintet**	**Sextet**
Total number of teams	1169	1460	1404	891	601
Number of teams having more than one paper (repeated teams)	228 (20%)	133 (9%)	57 (4%)	31 (3%)	13 (2%)
Number of teams having paper(s) with someone outside of the team	517 (44%)	256 (18%)	139 (10%)	56 (6%)	30 (5%)
Number of teams having paper(s) by the ego, a subset of the team, and maybe also someone outside of the team		917 (63%)	991 (71%)	684 (77%)	462 (77%)
Maximum number of papers of the team	16	6	4	3	3

Percentages in parentheses.

Team fluidity has important implications. Teams investigated in traditional team studies have well-defined boundaries and relatively stable members, so one can attribute outcomes to clearly delineated teams. In contrast, because of the autonomy in organizing the scientific team, people constantly come and go. For example, scientific teams have constant turnover of graduate students and postdocs. In addition, teams may acquire new members when new expertise is needed, and a team member may leave the team as he no longer shares the common interest with other teammates. To some extent, a collaborative team is co-evolving with the project, and there is rarely a stable team seeing through the whole creative process. This instability makes it difficult to identify the actual team responsible for a scientific output, which is also reflected in the difficulty in determining authorships for scientific publications (Haeussler & Sauermann, 2013; Laudel, 2002).

Interdependence between teams

Instability and ill-defined boundaries have other consequence. In particular, team production of science is not isolated from but constantly exchanges with the external network of other scientists and team. This means that external activities in the open network but outside of the team provide an input to team performance, in addition to within-team factors.

In addition to porous boundaries, scientists also simultaneously participate in multiple teams which may share members in common and similar research agendas. Therefore, there might be considerable knowledge spillovers across teams. In other words, teams are interdependent. Intellectual and other capital carried by individual members are important inputs to teams, and team members can import lessons learned from previous team experiences into new team situations. For example, by tracking the turnover of keywords, Tang and Hu (2013) showed that scholars pick up new research streams from their international collaborations and further pursue these new streams in their domestic collaborations. Since teams in science are not independent or isolated from the external environment of other teams, such independence should be taken into account in order to better understand the team production of science.

To illustrate the interdependency between teams in the same egocentric collaborative network, we look for knowledge spillover. Specifically, we hypothesize that if teams are truly interdependent, a scientist who has more new collaborators/collaborations will carry more new skills and knowledge to his other teams not involving these new collaborators, and these teams, even though they are repeat collaborations, will see a slight boost in performance that comes with the fresh ideas of newcomers to a collaboration. The *organizational learning* literature suggests that newcomers are important sources of innovation for an organization because they are more likely to bring in different knowledge and perspectives that are not yet shared in the organization (Gupta, Smith, & Shalley, 2006; March, 1991). We further argue that a scientist can transfer the new knowledge and perspectives that he learned from his new collaborators to his other teams not involving these new collaborators. Furthermore, a larger number of new collaborators will provide more diverse knowledge and perspectives and therefore will have a larger positive effect on the performance of repeated collaborations. Therefore, we hypothesize that: *when a scientist has more new collaborators/collaborations, his other collaborative teams not involving these new collaborators will produce papers with higher impact*.

For hypotheses testing, we use the same sample of American scientists, but use the unbalanced panel data of their life-time publication histories. The data have 11,850 observations (i.e., ego-year) of 1,310 egos. We test, for each ego in each year, the relationship between (1) the number of new collaborators/collaboration and (2) the number of citations received by papers from his repeated collaborations. The first step is to classify collaborators of each ego in each year into two categories: *new* (not collaborated in the preceding three years) and *repeated* (collaborated at least once in the preceding three years). Correspondingly, papers are classified into four categories: *solo* (single-authored paper), *new* (involving only *new* collaborators), *repeated* (involving only *repeated* collaborators), and *mixed* (involving both *new* and *repeated* collaborators). The dependent variables are the average and maximum numbers of citations for repeated-collaboration papers (*Cite.AVG.REP* and *Cite.MAX.REP*). For the explanatory variable, the number of new collaborators or collaborations is counted in three different ways: (1) the number of new collaborators (*Collaborator.NEW*), (2) the number of new collaborations, that is, papers involving only new collaborations (*Pub.NEW*), and (3) the number of collaborations involving new collaborators

(*Pub.NEW&MIX*). We also control for the number of repeated-collaboration papers (*Pub.REP*) and lagged citation performance (*Cite.AVG.LAG* and *Cite.MAX.LAG*). We control for standard demographic variables, including *age*, *gender*, and *race*. Given significant differences across scientific disciplines, we also control for *field* in our analysis. Quasi-Poisson random effects models are adopted. Quasi-Poisson models are adopted because the dependent variables are over-dispersed count variables, and random ego effects are incorporated to account for unobserved ego characteristics. To address the endogeneity issue (i.e., high-performing scientists are more attractive partners), we control for lagged citation performance and incorporate ego random effects. A summary of all variables are reported in Table 2.

Table 2. Variable descriptions

Variable	Description
$\text{Cite.AVG.REP}_{i,t}$	The average number of citations per paper received by repeated-collaboration papers of ego *i* published in year *t*.
$\text{Cite.MAX.REP}_{i,t}$	The maximum number of citations received by repeated-collaboration papers of ego *i* published in year *t*.
$\text{Cite.AVG.NEW}_{i,t}$	The average number of citations per paper received by new-collaboration papers of ego *i* published in year *t*.
$\text{Cite.MAX.NEW}_{i,t}$	The maximum number of citations received by new-collaboration papers of ego *i* published in year *t*.
$\text{Cite.AVG.LAG}_{i,t}$	The average number of citations per paper received by all papers of ego *i* published in year *t*-1.
$\text{Cite.MAX.LAG}_{i,t}$	The maximum number of citations received by all papers of ego *i* published in year *t*-1.
$\text{Pub.NEW}_{i,t}$	The number of new-collaboration papers of ego *i* published in year *t*.
$\text{Pub.REP}_{i,t}$	The number of repeated-collaboration papers of ego *i* published in year *t*.
$\text{Pub.NEW\&MIX}_{i,t}$	The number of new- or mixed-collaboration papers of ego *i* published in year *t*.
$\text{Pub.REP\&MIX}_{i,t}$	The number of repeated- or mixed-collaboration papers of ego *i* published in year *t*.
$\text{Collaborator.NEW}_{i,t}$	The number of new collaborators of ego *i* published in year *t*.
$\text{Collaborator.REP}_{i,t}$	The number of repeated collaborators of ego *i* published in year *t*.
$\text{Age}_{i,t}$	Physical age of ego *i* in year *t*.
Field_{i}	Research field of ego *i*, constant over time. Categorical variable: biology (BIOL), chemistry (CHEM), computer science (CS), earth and atmospheric sciences (EAS), electrical engineering (EE).
Gender_{i}	Gender of ego *i*, constant over time. Dummy variable: 1 if female 0 if male.
Race_{i}	Race of ego *i*, constant over time. Dummy variable: 1 if non-Hispanic White and 0 if minority.

Note: a five-year time window is used for counting citations. According to Wang's (2013) calculation on the whole WoS database, the Spearman correlations between five-year citation counts and 31-year citation counts are: 0.810, 0.906, 0.852, 0.888, and 0.792 in fields of biology, biomedical research, chemistry, earth and space, and engineering, respectively. The correlations are sufficiently high for this study.

Table 3. Quasi-Poisson random effects models

	Cite.AVG.REP			Cite.MAX.REP		
	1	2	3	4	5	6
Collaborator.NEW (ln)	0.080***			0.061***		
	(0.016)			(0.017)		
Pub.NEW&MIX (ln)		0.114***			0.078***	
		(0.023)			(0.024)	
Pub.NEW (ln)			0.045*			-0.001
			(0.027)			(0.028)
Pub.REP (ln)	-0.352***	-0.358***	-0.342***	0.580***	0.576***	0.585***
	(0.047)	(0.047)	(0.047)	(0.041)	(0.041)	(0.042)
Cite.AVG.LAG (ln)	0.148***	0.148***	0.155***			
	(0.013)	(0.013)	(0.013)			
Cite.MAX.LAG (ln)				0.181***	0.182***	0.188***
				(0.012)	(0.012)	(0.012)
Age	-0.010***	-0.010***	-0.009***	-0.013***	-0.013***	-0.012***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Age^2	0.000***	0.000***	0.000***	0.001***	0.001***	0.001***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Field-CHEM	-0.072	-0.080	-0.059	-0.094*	-0.098*	-0.084
	(0.054)	(0.055)	(0.055)	(0.057)	(0.057)	(0.057)
Field-CS	-0.472***	-0.475***	-0.489***	-0.468***	-0.473***	-0.481***
	(0.080)	(0.080)	(0.080)	(0.086)	(0.086)	(0.086)
Field-EAS	-0.115*	-0.114*	-0.115*	-0.114*	-0.114*	-0.115*
	(0.059)	(0.059)	(0.059)	(0.062)	(0.062)	(0.062)
Field-EE	-0.388***	-0.399***	-0.391***	-0.396***	-0.405***	-0.397***
	(0.069)	(0.070)	(0.070)	(0.072)	(0.072)	(0.072)
Gender-Female	0.010	0.011	0.006	0.003	0.003	-0.002
	(0.041)	(0.041)	(0.041)	(0.043)	(0.043)	(0.043)
Race-White	-0.119**	-0.119**	-0.125**	-0.117**	-0.118**	-0.123**
	(0.056)	(0.056)	(0.056)	(0.058)	(0.058)	(0.058)
χ^2	16490***	16360***	16096***	17458***	17361***	17340***

Number of observations (ego-year): 4503
Number of ego: 948
Ego random effect not reported
Standard error in parentheses
* p < .10, ** p < .05, *** p < .01

Regression results are reported in Table 3, columns report different combinations of the dependent and explanatory variables. Results show that the number of new collaborators/collaborations has a significant positive effect on the citation impact of papers not involving these new collaborators, except in column 6. When the dependent variable is the maximum number of citations received by repeated-collaboration papers, and the independent variable is the number of new-collaborations (*Pub.NEW*), the coefficient is insignificant. However, *Pub.NEW* is a very restrictive measure since some new collaborators are also involved in mixed-collaborations. While previous studies have explored the benefits of new collaborators for team innovation and creativity, we further demonstrate that such benefits can be transferred by a scientist to his other collaborative teams not involving these new collaborators. Given the significant knowledge spillovers across teams, it would seem misleading to draw conclusions about scientific performance based on assumptions of small groups of people working together in stable, longstanding, well defined, isolated teams.

A *multiple-stream model* of collaborative networks

The preceding two sections have demonstrated that collaborative teams in science are fluid and interdependent, and this section proposes a multiple-stream model to describe collaborative networks as organizations of scientific production. Latourian science studies focus on science in the making rather than ready-made science, tracing controversies in the process before a claim is accepted as a unproblematic scientific fact (Latour, 1987; Latour & Woolgar, 1986). In this context, publications should not be treated as stable end-products but as a part of an evolving stream of fact construction. Situated in this unstable and continual process of science production, the assembly of collaborative teams is also an evolving process. Collaboration is ultimately between people (i.e., scientists). However, a scientist is not a one-dimensional entity but embodies multiple-streams: phenomena, expertise, and resources. Phenomena are looking to be known (discovered), and each scientist is attracted to a set of phenomena. Examples of phenomena include: carbon sinks in forests, inactivating mutations in NOTCH1, green fluorescent protein, origin of retrovirus XMRV, influence of race on peer review of NIH grants. Expertise represents cognitive and epistemological perspectives to approach the phenomena, such as deciding what aspects to observe and how to observe. Expertise depends on personal experiences and disciplinary background. Expertise is seeking to be applied in various situations, much like a hammer seeing everything as a nail. Resources are information, materials, and equipment that are essential inputs for scientific research, for example, cell lines, mutant strains of mice, the BICEP2 instrument at the South Pole Telescope facility, Large Hadron Collider (LHC), survey and archive data. Furthermore, phenomena, expertise, and resources are not stable entities but evolving streams. In the process of science production, the meaning of studied phenomena is shifting and controversial, until science in the making is transferred into ready-made science (Latour, 1987; Latour & Woolgar, 1986). In terms of expertise, some components of the expertise are more resistant to change, such as paradigm, while other components are more flexible, such as methodology. For a successful collaboration, collaborators need to combine heterogeneous cognitive perspectives and reconcile epistemological conflicts, such issues are important topics in the studies of interdisciplinary collaboration (National Academy of Sciences, 2004). Furthermore, resources are also an evolving stream, as new resources are continually produced in current studies and in turn serve as inputs for future research (Hicks, 1995; Hilgartner & Brandt-Rauf, 1994).

Therefore, collaboration networks can be viewed as a collection of these three streams, and a collaborative team is assembled when these three streams are coupled. The team is a group of scientists with complementary expertise and resources working together to reveal a new phenomenon of shared interest that adds to knowledge by solving a puzzle. In this view of collaborative networks, the creative process in science production is not bounded within closed teams but takes place in a fluid, evolving network. Idea generation (finding and defining the problem) precedes team assembly. Even after a team is assembled, the team keeps interacting with other scientists and teams in the network to tap into external sources of knowledge. Furthermore, because of the uncertainty in scientific research, the initial idea behind the team is constantly subject to change, so is the team, adding new expertise and resources or removing exiting ones. This dynamic process leads to the observed instability of teams as described in the preceding sections. Furthermore, in this dynamic process, teams have intensive exchange with the environment, that is, the network of other scientists and teams. Therefore, the boundaries of teams are fuzzy, and teams are interdependent.

In addition, scaling up from teams to networks does not make individuals less relevant. On the contrary, it allows us to study egocentric networks from a more direct perspective. From the team perspective, the outcome of interest is a result of collective action, and therefore individual performance is irrelevant. However, scaling up from teams to networks allows us to see individuals' cross-team activities. Individual performance depends on a scientist's access to diverse streams in the network and his (dis)advantages in managing his multiple collaborative teams. Furthermore, this multiple-stream model of collaborative networks allows us to study egocentric networks from a perspective different from the *social capital* one. Instead of viewing a current egocentric network as social capital of potential value to future science production, a current egocentric network can be treated as an organization of scientific production that is directly responsible for current scientific outputs produced from this network. Different egocentric networks have different streams of phenomena, expertise, and resources, and these differences lead to different scientific outputs. Furthermore, the network structure affects the coupling process between these streams and therefore affects the productivity and creativity at the individual level across teams.

Conclusion

This paper proposes a network approach for the organization of science. Science is increasingly produced collaboratively, which calls for an organizational theory of science. Collaborative teams are the actual "factories" of science production, but teams in science are fluid and interdependent, so a network approach can complement the team approach for a better understanding of the organization of science. First, teams in science are unstable and have fuzzy boundaries. Teams have intense exchanges with their external networks of other scientists and teams, and a considerable amount of the activities responsible for the final team product take place in the open network outside of the team. Therefore, these external activities should be taken into account when explaining team performance, in addition to internal team activities. Second, teams sharing common members are interdependent. There are significant knowledge spillovers across teams in the same egocentric network. Therefore, the egocentric network can be studied as a loosely coupled system to account for not only within-team processes but also cross-team processes. To some extent, collaborative networks can be viewed as a collection of three evolving streams: phenomenon, expertise, and resources. A collaborative team emerges when these streams are joined together. These three streams are dynamic and flowing through the network, so teams are also dynamic and co-evolve with them.

References

de Solla Price, D. J. (1986). *Little Science, Big Science--and Beyond*. New York, NY: Columbia University Press.

Gupta, A. K., Smith, K. G., & Shalley, C. E. (2006). The interplay between exploration and exploitation. *Academy of Management Journal, 49*(4), 693-706. doi: 10.2307/20159793

Haeussler, C., & Sauermann, H. (2013). Credit where credit is due? The impact of project contributions and social factors on authorship and inventorship. *Research Policy, 42*(3), 688-703. doi: http://dx.doi.org/10.1016/j.respol.2012.09.009

Hicks, D. (1995). Published Papers, Tacit Competencies and Corporate Management of the Public/Private Character of Knowledge. *Industrial and Corporate Change*, *4*(2), 401-424. doi: 10.1093/icc/4.2.401

Hicks, D., & Katz, J. S. (1996). Where is science going? *Science, Technology, & Human Values*, *21*(4), 379-406. doi: 10.2307/690087

Hilgartner, S., & Brandt-Rauf, S. I. (1994). Data Access, Ownership, and Control: Toward Empirical Studies of Access Practices. *Science Communication, 15*(4), 355-372. doi: 10.1177/107554709401500401

Latour, B. (1987). Science in action : how to follow scientists and engineers through society. Cambridge, MA: Harvard University Press.

Latour, B., & Woolgar, S. (1986). *Laboratory life: The construction of scientific facts*. Princeton, NJ: Princeton University Press.

Laudel, G. (2002). What do we measure by co-authorships? *Research Evaluation, 11*(1), 3-15.

March, J. G. (1991). Exploration and exploitation in organizational learning. *Organization Science*, *2*(1), 71-87. doi: 10.1287/orsc.2.1.71

National Academy of Sciences. (2004). *Facilitating interdisciplinary research*. Washington DC: The National Academies Press.

NETWISE. (2007). NETWISE Project. from http://netwise.gatech.edu/index.php

Tang, L., & Hu, G. (2013). Tracing the footprint of knowledge spillover: Evidence from U.S.–China collaboration in nanotechnology. *Journal of the American Society for Information Science and Technology, 64*(9), 1791-1801. doi: 10.1002/asi.22873

Wang, J. (2013). Citation time window choice for research impact evaluation. *Scientometrics, 94*(3), 851-872. doi: 10.1007/s11192-012-0775-9

Wang, J. (2014). Unpacking the Matthew effect in citations. *Journal of Informetrics*, *8*(2), 329-339. doi: http://dx.doi.org/10.1016/j.joi.2014.01.006

Wang, J., Berzins, K., Hicks, D., Melkers, J., Xiao, F., & Pinheiro, D. (2012). A boosted-trees method for name disambiguation. *Scientometrics*, *93*(2), 391-411. doi: 10.1007/s11192-012-0681-1

Whitley, R. (2000). *The Intellectual* and *Social Organization of the Sciences* (2nd ed.). Oxford, UK; New York, NY: Oxford University Press.

Wuchty, S., Jones, B. F., & Uzzi, B. (2007). The increasing dominance of teams in production of knowledge. *Science*, *316*(5827), 1036-1039. doi: 10.1126/science.1136099

Funding for some, spills for others: Explaining the emergence of nanotechnology in Chinese regions

Lili Wang*, Jojo Jacob** and Zibiao Li***

*wang@merit.unu.edu
UNU-MERIT, Keizer Karelplein 19, Maastricht, 6211 TC (The Netherlands)

*Jojo.JACOB@grenoble-em.com
Grenoble Ecole de Management, 12 rue Pierre Sémard (F307), Grenoble, 38000 (France)

*** lizibiao@hebut.edu.cn
School of Management, Hebei University of Technology, Tianjin, 300401 (P.R.China)

Introduction and background

For less developed countries, who typically occupy follower positions in mature technologies that have long lost their dynamism, a 'real' catching-up process requires acquiring the capability to develop a new technology system (Perez and Soete 1988). Such a system provides enormous opportunities for successive improvements across a range of technologies that can generate economy-wide technological externalities lasting several decades. An early entry into a new technology system therefore can trigger faster catching up and long run success.

In this respect, nanotechnology represents a set of science-based enabling technologies that are still in the early stages of their technological life cycles and that promise significant long-term pay offs to countries engaging in their development and commercialization. Studies have shown that nanotechnology can serve as a general purpose technology that has applications across a broad spectrum of economic activities spanning almost all fields of manufacturing (Shapira and Youtie, 2008; Wang et. al. 2013). In other words, countries that occupy frontier positions in nanotechnology are likely to lead in many fields of innovation in the years to come.

Large less developed countries with a strong scientific-research tradition, such as China, have been expected to provide global leadership in emerging science-based technologies such as biotechnology and nanotechnology (Niosi and Reid, 2007). In this respect, over the last decade or so, as China began undergoing its transformation from an investment-driven to an innovation-driven economy, the country experienced dramatic progress in the development of nanotechnology. The scientific output in nanotechnology from China, as measured by nanotechnology-related publications with a Chinese address, has been increasing exponentially. Whereas in 2000 the number of nanotechnology-related publications from China stood at a paltry 30% of the US level, by 2010 it rose to 92%[1]. The same period also witnessed a remarkable increase in the number of annual nanotechnology-related patent applications filed locally, from 275 to 6,333.

Financial support from the state is generally viewed as a vital ingredient to the emergence of a new technology system. Private sector investment in such a system, especially in the early

[1] Data source is UNU-MERIT nano-database.

stages, will be less than optimum because of the high levels of uncertainty about not just the technological outcomes but also the commercial potentials of the newly-developed technologies. In China, nanoscience and nanotechnology drew favourable policy interest already in the 1980s when these concepts first emerged. However, it was not until 1990 that serious efforts to promote nanotechnology began, with the Ministry of Science and Technology launching the ten-year "Climbing-Up" project (Bai 2001, Tang, Wang & Shapira 2010). Soon after, the concept began trickling through the scientific ranks and the Chinese Academy of Sciences (CAS), the National Natural Science Foundation of China (NSFC), and the State Science and Technology Commission (SSTC) began funding nanoscience-related work and activities (Chunli Bai, 2005). Today, according to the China Association for Science and Technology, the three most widely used high-tech words in China are "computer", "gene" and "nanometer".

In this paper, we examine the growth of nanotechnology in China with a particular focus on whether the dynamics of this growth vary across Chinese regions with different scientific capabilities. We argue that the large-scale governmental aid for nanotechnology development would have made a notable impact only in regions possessing high scientific capabilities. Regions lagging behind in scientific capabilities would not have the necessary complementary resources either to be major beneficiaries of government support in the first place or to make an efficient use of the support received from the state. However, we suggest, drawing on the economic geography literature, that lagging Chinese regions can leverage their scientists' formal collaboration links to bring in spillovers of nanotechnology from other regions. The collaboration network of scientists acts as an important resource for lagging regions, partly compensating for their weak scientific capabilities. Our focus on the differential sources of nanotechnology development contributes to the economic geography literature on knowledge spillovers and to the catch up literature that stresses the development of a new technology system for faster catching up.

The paper is organized as follows. The following section provides a theoretical and empirical background to the study and raises the specific questions for empirical scrutiny. The third section presents the data and explains the methods. The results of the empirical analysis are discussed in the fourth section, and the final section concludes.

Background and research questions

A technology system perspective of nanotechnology, and nanotechnology's emergence in China

Both the traditional catch up literature (e.g. Gerschenkron, 1962) and the new-growth theories (Grossman and Helpman 1991; Rivera-Batiz and Romer 1991) stress the role of international technology diffusion for the catching up of less developed countries to the income levels of developed countries. In both these perspectives, mature technologies developed in advanced countries represent a major opportunity that less developed countries might exploit so they can avoid the costly, time consuming, and challenging task of developing new technologies from scratch. However, another perspective, whose spirit we embrace in this paper, emphasizes the importance of less developed countries taking a leadership role in the development of a new technology system (Perez and Soete 1988). In this view, a new technology system impacts growth in a broad range of sectors and generates economy-wide knowledge spillovers, thereby accelerating the catching-up process. Well-known examples of this process are South Korea and Taiwan which focused early on in developing the electronics industry, at a time when this industry was fast emerging and when both countries had little prior experience in this or related industries. In this context, given that nanotechnology has applications in a wide spectrum of activities, a leadership position in nanotechnology implies a significant 'window of opportunity' for a large less developed country like China to accelerate its catch up to the global techno-economic frontier.

In developing a science-based technology like nanotechnology, less developed countries are not particularly at a comparative disadvantage vis-a-vis developed countries. This because many in the former category of countries, and in particular China, have universities and research institutes that boast of a rich heritage in scientific research. Realizing the tremendous potential of nanotechnology, China has been adopting an ambitious nanotechnology development strategy. Key to this has been the extensive financing for nanotechnology research under the National Natural Science Foundation program. Following the substantial progress made by China in information and communication technologies over the past decades (Lazonick and Li, 2012; Lazonick, 2004; Lu, 2000), the government's efforts to promote nanotechnology are aimed at setting off another technological wave in China.

The geography of knowledge development

It is widely acknowledged that technological activities tend to be unevenly distributed across regions (or countries), with high-technology activities in particular concentrated in geographic clusters (Verspagen and Schoenmakers, 2004; Henderson, 2003; Niosi, 2001; Antonelli, 2001; Niosi and Queenton, 2010). In China, given the wide regional inequality in scientific capabilities, the emergence of nanotechnology unavoidably started in a few leading regions. Few studies have explored the question of the extent to which other regions are involving in nanotechnology development. Motoyama, et al. (2014) was one of the first attempts to address the question of regional convergence or divergence of nanotechnology development in China. They, adopting a spatial correlation technique, found very little diffusion of knowledge and predicted that the divergence trend would continue. However, we argue that for a fuller understanding of regional dimensions of knowledge development in a large country like China, it is important to go beyond the traditional spatial proximity framework and take into account knowledge flows through the collaboration network of scientists. This is because, as we discuss below conceptually and in section 4.2 empirically in the Chinese context, diffusion of knowledge from other regions can compensate for the initially weak innovation systems of less developed regions.

Channels of knowledge flows

There is a vast literature that examines spillovers of knowledge across regions, nations, firms or industries (for reviews see, Frenken et al. 2010; Wang et al. 2013; Jacob & Meister, 2005). A dominant strand of this literature emphasizes that knowledge externalities occur locally, rather than globally (Jaffe 1989; Antonelli 2001; Abramovsky and Simpson, 2008; Arundel and Guena, 2004). The localized character of knowledge spillovers, the argument goes, stems from the tacit nature of knowledge. This renders the acquisition of knowledge simply from technology blueprints difficult, and therefore calls for close, often informal, people to people interaction. A few authors, however, are more explicit about the specific mechanism of knowledge flows, arguing that it may not necessarily be the 'knowledge in the air', but the locally-bound scientific networks that generate localized knowledge flows (Zucker, et al, 1998; Breschi & Lissoni 2009).

However, there is increasing evidence that geographic distance is not a limiting factor for knowledge spillovers. Formal linkages, such as co-authorship ties, can facilitate knowledge flows over long distances (Cockburn and Henderson, 1997; Ponds et al, 2009). These linkages provide an important means for regions or countries to tap into the resources and knowledge of more advanced regions or countries. Several studies have documented the fast growth of collaboration in science, with some highlighting that international collaborations generate higher quality research (higher citation rates) than domestic collaborations (Frenken et al. 2010; Tang and Shapira, 2011), or facilitate entry into new research fields (Tang and Shapira, 2011).

Empirical framework and Research questions
Drawing on the preceding discussion, we propose an empirical framework for understanding the development of nanotechnology in Chinese regions. Two factors are integral to explaining the growth in nanotechnology across Chinese regions in our framework: the sizeable governmental financial support and inter-regional and international knowledge spillovers. We focus on collaboration networks as the main conduits for knowledge spillovers, while also taking on board the effect of geographic proximity between regions. Given that collaboration networks evolve over time, we treat collaboration as a dynamic construct; existing literature has paid only scant attention to the dynamic aspect of collaboration due primarily to the use of cross sectional data.

A particular novelty of our study is that we carryout separate analysis for leading and lagging regions in scientific capabilities. The dynamics of knowledge development in these two sets of regions are likely to be different. Even if advanced and lagging regions received the same level of funding, they would likely generate differential returns just because advanced regions can leverage their superior capabilities to generate greater *bang for the buck* compared to lagging regions wherein funds would be less efficiently utilized. Nevertheless, lagging regions can benefit from collaborations between their scientists and those from advanced regions. The benefits for advanced regions through these collaborations are likely minimal (aside from the goodwill they have gained).

The following are the specific question we explore in this paper.

- To what extent has funding for nano technology research by the Chinese government succeeded in stimulating development of nanotechnology in Chinese regions?
- To what extent have collaboration networks and geographic proximity generated inter-regional spillovers of nano technology knowledge in general, and funding-induced knowledge spillovers in particular?
- Do differences in the scientific and technological capabilities of regions affect the extent to which regions benefit from state funding and from knowledge spillovers? Specifically, do lagging regions benefit more from regional spillovers than from state funding, and vice versa?

Data and variables
For the econometric analysis, we use a panel data set of 30 Chinese regions[2] spanning 11 years (2000-2010). The dependent variable is patent applications filed from a Chinese region at China's State Intellectual Property Office (SIPO), capturing the region's technological output. We employ over 30,000 nano patent applications gathered from the China Patents Full-text Database[3].

[2] There are in total 31 provincial regions in China. Tibet is not included in the analysis due to lack of data.
[3] Nano patent is defined as a patent with a "nano" word in the title.

The key independent variables are nano funding that a region received from the National Natural Science Foundation; inter-regional spillovers; and international spillovers. Inter-regional spillovers in our framework stem from two sources: one is the patents of a region, and the other is the nano funding received by a region. We identify two carriers of spillovers: the region-spanning collaboration network of scientists and the geographic proximity between regions. The first of these carriers is defined in terms of a dynamic collaboration matrix as follows:

$$\begin{bmatrix} P_{1,1,t} & P_{1,2,t} & \cdots & \cdots & \cdots & P_{1,31,t} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \vdots \\ \cdots & \cdots & \cdots & \cdots & \cdots & \vdots \\ P_{i,1,t} & \cdots & \cdots & P_{i,j,t} & \cdots & \vdots \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ P_{31,1,t} & \cdots & \cdots & \cdots & \cdots & P_{31,31,t} \end{bmatrix}$$

In the matrix, an element P_{ijt} is the number of co-authored nano publications involving the regions *i* and *j* in year *t*. The spillovers from patented technologies ($\mathrm{TECHSPILL_{ijt}}$) and nano funding ($\mathrm{FUNDSPILL_{ijt}}$) that region *i* receives from region *j* are defined respectively as:

$$\mathrm{TECHSPILL_{it}} = \frac{\mathrm{PUB}_{ijt}}{\mathrm{PUB_{jt}}} * PAT_{jt} \text{ (j=region1, region 2, ..., region 30, i} \neq \text{j)}$$

$$\mathrm{FUNDSPILL_{it}} = \frac{\mathrm{PUB}_{ijt}}{\mathrm{PUB_{jt}}} * F_{jt} \text{ (j=region1, region 2, ..., region 30, i} \neq \text{j)}$$

in which PAT_{jt} is the number of nanotechnology-related patents in region *j* in year *t*, and F_{jt} is the nano-funding received by region *j* in year *t*[4]. To construct the publication weights in the above two equations we collected 164,000 Nano-publications from Thomson Reuters' Web of Science (WoS). The database is constructed based on an evolutionary lexical query searching and defining strategy developed by the Georgia Institute of Technology (see more details in Wang and Notten, 2010).

In addition to the collaboration weight above, we also use the geographical proximity between regions to construct a second set of spillover variables. If d_{ij} is the geographical distance between regions *i* and *j*, the spatial spillover weight from *j* to *i* can be expressed as (see also Vinciguerra, et al. 2011; Ertur et al., 2006; Wang, et al. 2013): [5]

$$w_{ij} = w_{ij}^{*} / \sum_{j} w_{ij}^{*}$$

in which the distance weight w_{ij}^{*} is the inverse of squared distance[6] between region *i* and *j* ($1/d_{ij}^{2}$). Using this weight to replace the collaboration weight in the spillover equations above, we derive two additional spillover variables that capture the effect of proximity in generating spillovers.

Next, we construct an international collaboration intensity variable for capturing the effect of knowledge spillovers resulting from collaboration with foreign countries:

[4] The nano-patent collaboration data is not available, hence we use the collaboration extracted from nano-publication to create the interregional, as well as the international collaboration variable that is defined later.

[5] In this model, spillover weight has been standardized by the row total, assuming that the amount of spillovers from *j* to *i* is subject to the spillovers i receives from other regions.

[6] Distance of provinces is measured by their capital cities, considering that a capital city is usually the central business and technology center of each province.

$$CI_{it\_international} = \frac{\sum PUB_{ikt}}{PUB_{it}} \text{ (k= country 1, country 2, ..., country27)}$$ [7]

where $CI_{it\_international}$ represents the international collaboration intensity in nanotechnology-related publications of region *i* in year *t*, PUB_{ikt} is the number of co-authored nanotechnology-related publications involving region *i* and the foreign country *k* in year *t*, and PUB_{it} the total number of nanotechnology-related publications stemming from region *i*. Each of the 27 foreign countries had at least 10 papers co-authored with an author based in China during the period of analysis[8]. These countries, in the order of the number of collaborative nano publications with Chinese regions are U.S.A., Hong Kong, Japan, Germany, Australia, Singapore, England, South Korea, Canada, France, Sweden, Taiwan, Switzerland, Spain, the Netherlands, Belgium, India, Russia, Ireland, Scotland, Pakistan, Norway, Portugal, Austria, Malaysia, Brazil, and Macao. Finally, as control variables we include regional R&D intensity (ratio of total R&D to GDP), non-nano patenting productivity (ratio of non-nano patents to R&D), and per capita income. These variables take into account regional differences in, respectively, general scientific capability, general patenting propensity, and general economic prosperity.

Empirical analysis and findings

In order to further set the stage for the econometric analysis, we first discuss some key aspects concerning the growth of nanotechnology across Chinese regions.

China's position in nano-science and technology

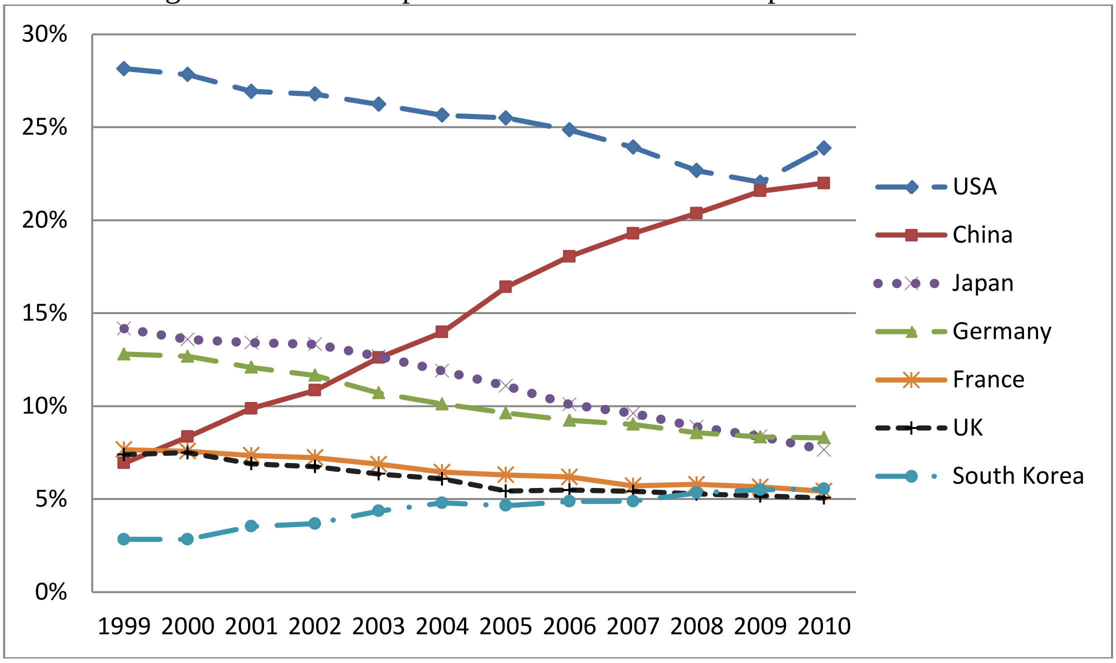
The period 1999-2010 witnessed the number of nano publications with Chinese addresses growing from 2,487, at an annual rate of 23 per cent, to 23,686. While US occupied a leading position in the early years of the emergence of nanoscience and nanotechnology, China has been able to catch-up in an impressive way over the last decade (Figure 1). Between 1999 and 2010 as the share of China increased from 6.9% to 22%, that of most other leading players dropped—from 28% to 24% for the U.S., from 13% to 8% for Germany, from 14.2% to 7.7% for Japan, from 8% to 5% for France, and from 7% to 5% for the UK. Nanotechnogy (as measured by nanotechnology patents) too has been skyrocketing in China. According to the patent records at China's State Intellectual Property Office (SIPO), the annual nano patent filing reached over 6,000 in 2010 from a meagre 98 in 1999. China's position in global nano patenting is difficult to assess, however. This is because Chinese inventors file for patents mainly locally in the Chinese patent office, with only fewer than 2 per cent of patent applications filed outside of China (Harvey, 2011).[9]

[7] This index is a sum of the collaboration intensity between region i and each foreign country. For instance, if region i collaborates with foreign country 1 and 2, this will be counted twice. Thus this calculation takes into consideration the number of foreign countries involved in one collaborated paper. Nevertheless, one should keep in mind that this intensity value is supposed to be slightly higher than the one calculated by directly using the number of internationally collaborated papers with region i divided by the total publications of this region.

[8] Hong Kong and Macao have different S&T systems from mainland of China and don't receive R&D funding from Chinese government. Hence these two regions are counted as "foreign" countries.

[9] However, Chinese inventors file for patents mainly locally in the Chinese patent office, with only fewer than 2 per cent of patent applications filed outside of China (Harvey, 2011); this makes it difficult to compare China's global position with other advanced countries.

Figure 1. Share of top six countries in total nano-publications world-wide

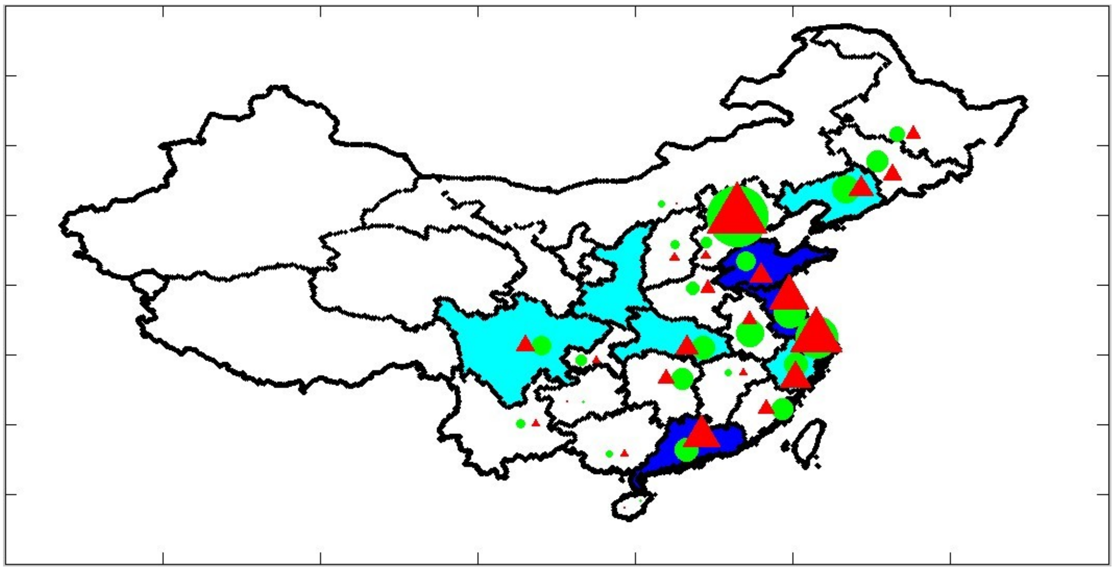


Source: Authors' own calculation based on UNU-MERIT nano-database.

Regional disparity and changes

Any discussion of overall growth of nanotechnology in China masks wide differences in scientific capabilities across Chinese regions. Figure 2 illustrates the strong regional disparities in nano funding, nano patenting, and technological capabilities in China over the 2000-2009 period. With their very high R&D expenditures, coastal regions in Eastern China, and a few inland regions close to them, stand out compared to the rest of China. It is worthwhile to note that the regional disparity of nano funding is more pronounced than that of the general R&D expenditure. As shown in Figure 2, the level of R&D expenditure in some middle regions is reasonably high (see the light blue areas in the map). However, nano funding (green circles) – and consequently patent applications (red triangles) – has been concentrated in coastal regions. In nano patent applications, four regions (Beijing, Shanghai Jiangsu and Guangdong) accounted for more than 50 per cent of the national total.

Figure 2. Distribution of nano patent application, nano funding and general R&D expenditure, 2000-2009



Note: 1) The presented value is the sum of 2000-09. 2) Blue shades represent the general R&D expenditure (the darker the higher level); Green circle is nano funding (the bigger size the greater value); Red triangle represents nano patent applications (the bigger size the greater value).

To further explore this we divide Chinese regions into two categories: leading regions and lagging regions—the former category of regions are defined as those that fall into the top 25% in total scientific publications during period of study; the rest of the regions fall into the lagging category. A look at the trend in patent applications in the two categories of regions (Table 1) suggests an increasing dynamism in lagging regions. While leading regions witnessed a higher growth in nano patent applications during the first half of the period under study (1999-2004), the opposite happened during the later period (2005-2010).

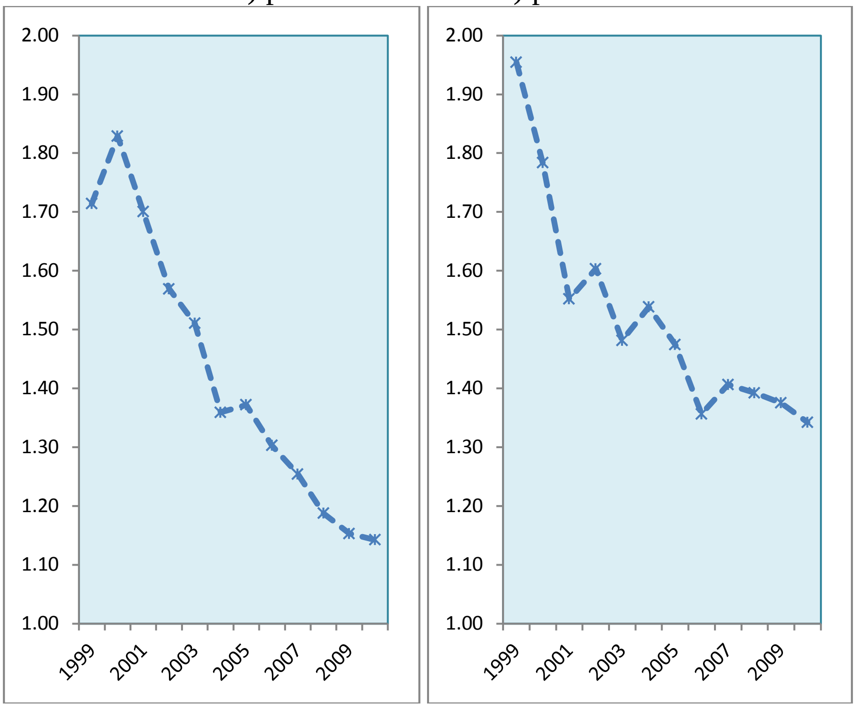
Table 1. Number of patent applications and growth rates, by regional groups

	number of patent applications			exponential growth rate	
Year	**2000**	**2005**	**2010**	**1999-04**	**2005-10**
Leading regions (Top 8)	189	1704	4419	55%	21%
Lagging regions	86	692	1914	52%	23%

Source: patent data from China's State Intellectual Property Office (SIPO).
Note: Leading regions are defined as those that belonged to the top 25% in total scientific publications.

Furthermore, we notice a sharp decline in the coefficient of variation in nanotechnology-related publications and patents between 1999 and 2010: respectively from 1.71 to 1.14 and from 1.95 to 1.34 (Figure 3). These evidences indicate that scientifically lagging regions have increasingly become active in nanotechnology research. The increasing dynamism shown by lagging regions in nanotechnology development requires an explanation. We focus here on the contributions of the linkages that lagging regions have with other regions within China, but also with international partners.

Figure 3. Coefficient variation of regional nano-publications and patents
a) publications b) patents



Source: Authors' own calculation.
Note: 1)Tibet is not included. 2) We removed one extreme outlier: 911 patent applications in 2001 filed by a single person from Beijing. This caused Beijing to account for 85% of the national total that year.

Collaboration patterns in China

In table 2, we explore the intensity of scientific collaborations (1) among Chinese regions, and (2) between Chinese regions and the rest of the world. The top part of table 2 reveals that international collaboration intensity in scientific publications for an average Chinese region was about 19% during 1999-2004, and about 17% during 2005-2009; leading regions, understandably, exhibited a slightly higher international collaboration intensity compared to lagging regions.

Table 2: Collaboration intensity in nano-science

	1999-2004	2005-2010	comparison
	(1)	(2)	(3)=(2)-(1)
	international collaboration		
all regions	18.6	17.3	-1.3
leading regions	21.2	20.2	-1.0
lagging regions	17.6	16.2	-1.4
	national collaboration		
all regions	47.7	56.8	9.2
leading regions	37.2	39.0	1.8
lagging regions	53.6	64.8	11.1

Source: Scientific collaboration data are collected from Web of Science.
Note: Leading regions are defined as those that belonged to the top 25% in total scientific publications.

On the other hand, as the bottom part of table 2 reveals, inter-regional collaboration intensity in scientific publications was much higher for both regional categories: it was close to 50% during the first period, before increasing by about nine percentage points during the second period. Even more interestingly, leading regions on average had a much lower inter-regional collaboration intensity compared to lagging regions. The collaboration intensity in lagging regions furthermore registered an 11 percentage point increase between the two periods (as against just a two percentage point increase in leading regions) – during 2005-2009, approximately 65% of the scientific publications in an average lagging region were written with scientists based in another region. These observations lend credence to our suggestion earlier on that collaboration networks may be an important source of catching up in lagging regions; forging links with the scientific communities in other Chinese regions could help lagging regions compensate for their weak scientific capabilities.

Results and discussion

As our dependent variable is the number of nanotechnology patents, a count data model such as Negative Binomial or Poisson is more appropriate than OLS. Chinese regions exhibit wide variations in patenting so the critical assumption of the equality of mean and variance of the Poisson model does not hold. Therefore we employ Negative Binomial Regression model as our preferred model. Given especially that regional patenting can be shaped by a host of other factors that we cannot fully account for, we employ a fixed effect model. We also include a full set of year dummies to account for unobserved annual events that may affect patenting in all regions.

Table 3. Regression results of negative binomial estimations on nano patent application

	All regions			Leading regions			Lagging regions		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Log of Nano funding		0.065*	0.059		0.292**	0.258*		0.027	0.028
		(0.036)	(0.037)		(0.131)	(0.137)		(0.037)	(0.038)
Nanotech spillovers –Collaboration	0.230***	0.201***		-0.249	-0.245		0.393***	0.378***	
	(0.074)	(0.075)		(0.171)	(0.167)		(0.091)	(0.093)	
Nanotech spillovers - Proximity	0.026	0.004		-0.016	-0.056		0.126	0.120	
	(0.121)	(0.119)		(0.176)	(0.163)		(0.175)	(0.174)	
Funding spillovers -Collaboration			0.128**			-0.304*			0.184**
			(0.065)			(0.169)			(0.077)
Funding spillovers –Proximity			0.127			0.078			0.265
			(0.113)			(0.161)			(0.168)
International collaboration intensity	0.003	0.004	0.003	-0.003	-0.003	-0.003	0.004	0.004	0.003
	(0.004)	(0.004)	(0.004)	(0.009)	(0.009)	(0.009)	(0.004)	(0.004)	(0.004)
Control variables									
R&D/GDP	0.153***	0.140***	0.160***	0.246***	0.179***	0.164***	0.010	0.010	0.101
	(0.044)	(0.044)	(0.043)	(0.056)	(0.062)	(0.061)	(0.114)	(0.113)	(0.110)
Non-nano patent/R&D	0.076	0.089	0.105*	0.174**	0.140*	0.139*	-0.069	-0.052	0.017
	(0.062)	(0.061)	(0.060)	(0.079)	(0.080)	(0.080)	(0.105)	(0.107)	(0.110)
GDP per capita	0.009	0.013	0.027	0.135**	0.145***	0.128**	-0.037	-0.034	-0.005
	(0.034)	(0.034)	(0.033)	(0.053)	(0.051)	(0.052)	(0.069)	(0.069)	(0.064)
Constant	2.320***	2.047***	1.225	4.152***	2.093	2.405	1.759*	1.632	0.521
	(0.664)	(0.671)	(0.794)	(1.033)	(1.371)	(1.487)	(1.033)	(1.041)	(1.258)
Observations	330	330	330	88	88	88	242	242	242
Number of regions	30	30	30	8	8	8	22	22	22

Note: 1) Dependent variable is nano patent applications. 2) Explanatory variables are lagged by one year; 3) Year dummies are not reported. 4) *** at 1% significance level; ** at 5% significance level; and * at 10% significance level. 5) Leading regions are defined as those that belonged to the top 25% in total scientific publications.

Regression results are documented in table 3[1]. All models include the full set of controls and year dummies.

Total sample
Models 1 to 3 present results based on the complete sample with different combinations of the key explanatory variables. In Model 1 we include the two nanotechnology spillover variables: in one spillovers stem from formal collaboration linkages, and in the other, from regional proximity. The results confirm that formal collaborations generate knowledge spillovers; however, proximity has no significant effect. This is not surprising given the vast distances separating Chinese regions. In model 2 we add the nano-funding variable. This variable has a positive and significant coefficient, suggesting that direct financial support for nanotechnology research has had a positive impact on the development of this technology in Chinese regions. In model 3, we replace the nanotechnology spillover variables (proximity-induced and collaboration-induced) with those based on funding. Note that the two sets of variables could not be simultaneously included in a single model due to the high collinearity between them. The results are similar – funding generates inter-regional spillovers through collaboration networks, but not through proximity.

Sub-samples
Next, we carried out separate analysis for leading and lagging regions[2]. Results for leading regions are presented in models 4 to 6, and those for lagging regions in models 7 to 9. Comparison of the results for the two categories reveals interesting insights. First, direct funding has a significant positive effect on patenting only in leading regions (model 5 and 6), not in lagging regions (model 8 and 9). This is consistent with our earlier discussion in section 4 that advanced regions led in both nano funding and nano patenting.

Contrary to the effects of nano funding, spillovers from other regions through collaborations exerted a significant positive impact in lagging regions (model 7, 8 and 9), but not in leading ones (model 4, 5 and 6). This applies for both nano-technology and nano-funding spillovers. These results too are in line with our earlier discussion, demonstrating that collaboration linkages with other regions compensate for the weak capabilities of lagging regions and the low degree of government support they receive. Advanced regions, being the front runners of nanotechnology development, are able to capitalize on governmental support, leveraging their own capabilities.

The international collaboration intensity variable shows little noticeable influence, with negative, though non-significant, coefficients for leading regions. This supports the standard view that the surge of nano patent applications in China was driven by China's indigenous capability, in particular in its leading regions. More broadly, the results are in agreement with the notion that in the development of new technologies, national linkages are likely to be more effective than international ones (Metcalfe and Ranlogan, 2008).

---

[1] Regression results stay similar if we separate collaboration and proximity variables into different models. Results are available upon request.

[2] As noted before, leading regions are defined as those that belonged to the top 25% in total scientific publications. Different definitions of scientific capabilities such as nano publications, nano patents, and total patents yielded similar results as in table 3.

Conclusions

Over the past two decades, China has been attempting to make a giant leap in nanotechnology development. Given China's strong scientific capabilities as reflected in the presence of a number of world class universities and research institutes, already in the late 1990s China was projected to be a leader in emerging science-based technologies such as nanotechnology (Porter et al. 2002). True to these predictions, China has fast emerged as a leading global player in nanotechnology. The evidence presented in this paper suggests that China's success in nanotechnology development in general owes in large part to the fostering of indigenous scientific capabilities through strong financial support from the state.

Our analysis also revealed that the dynamics of nanotechnology development were quite different in regions leading in versus those lagging in scientific capabilities. It is indeed well known that economic development and scientific capabilities are highly uneven across Chinese regions. Sure enough, a few regions with superior scientific capabilities spearheaded the early growth of nanotechnology in China. However, regional inequalities in nanotechnology development are diminishing. In this regard, our study has found that the key source of growth in nanotechnology patenting in lagging regions was the collaborative ties that scientists from these regions forged with those from other regions. These collaborative ties generated significant inter-regional spillovers of nanotechnology knowledge. In leading regions, on the other hand, R&D support received from the government for nanotechnology development was the principal factor behind the rapid growth of nanotechnology output. Spillovers from other regions, or from abroad, played no significant role in the growth of nanotechnology in these regions.

Our study contributes to the catch up literature by highlighting on the one hand how targeted governmental support can help leading regions spearhead the growth of a new technology system, and on the other the role of region-spanning scientific collaborations in helping lagging regions partake in the development of these technologies. The study furthermore contributes to the economic geography literature on knowledge spillovers in that future studies may place greater emphasis on the differences in growth dynamics in leading and lagging regions.

References

Antonelli, C. (2001). *The Microeconomics of* Technological *Systems*, Oxford University Press, Oxford.

Abramovsky, L. & Helen, S. (2008). Geographic Proximity and Firm-University Innovation Linkages: evidence from Great Britain, CMPO *Working Paper Series*, No. 08/200, University of Bristol.

Arundel, A & Guena, A. (2004). Proximity and the use of public science by innovative European firms. *Economics of Innovation and New Technology*, 13, 559- 580.

Breschi, S. & Lissoni, F. (2009). Mobility of skilled workers and co-invention networks: an anatomy of localized knowledge flows. *Journal of Economic Geography*, 9, 439-468.

Broekel, T. (2012). Collaboration intensity and regional innovation efficiency in Germany – A conditional efficiency approach, *Industry and Innovation*, 19, 155-179.

Cockburn, I, M & Henderson, R. (1997). Public-Private Interaction and the productivity of academic research, *NBER Working paper* 6018.

Ernst, D. (2002). Global production networks and the changing geography of innovation systems- Implications for developing countries. *Economics of Innovation and New Technology*, 11, 497-523.

Frenken, K., Ponds, R & Van Oort, F.G. (2010). The citation impact of research collaboration in science-based industries: A spatial-institutional analysis. *Papers in Regional Science*, 89, 351-371.

Gerschenkron, A. (1962). *Economic Backwardness in Historical Perspective: A Book of Essays*, Cambridge, Massachusetts: Harvard University Press.

Grossman, G. & Helpman, E. (1991). Quality Ladders in the Theory of Economic Growth. *Review of Economic Studies*, 58, 43–61.

Harvey, I. (2011). Intellectual property: China in the global economy – Myth and reality, *INTIPSA Briefing Note*.

Henderson, J.V. (2003). Marshall's scale economies, *Journal of Urban Economics*. 53, 1-28.

Hoekman, J., Scherngell, T., Frenken, K. & Tijssen, R. (2013). Acquisition of European research funds and its effect on international scientific collaboration. *Journal of Economic Geography*, 13, 23-52.

Jacob, J. & Meister, C. (2005). Productivity gains, technology spillovers and trade: Indonesian manufacturing, 1980-96. *Bulletin of* Indonesian *Economic Studies*, 41, 37-56.

Jaffe, A. (1989). Real effects of academic research, *American Economic Review*, 79, 957-970.

Jin, B. & Rousseau, R. (2004). Evaluation of research performance and scientometric indicators in China. In: Moed, H.F., Gl¨anzel, W., Schmoch, U. (Eds.). *Handbook of Quantitative Science and Technology Research* (pp. 497–514). Dordrecht: Kluwer Academic Publishers.

Kostoff, R.N. (2012). China/USA nanotechnology research output comparison-2011 update. *Technological Forecasting & Social Change*, 79, 986-990.

Lazonick, W. (2004). Indigenous innovation and economic development: Lessons from China's leap into the information age. *Industry and Innovation*, 11, 273-297.

Lazonick, W. & Lin, Y. (2012). China's path to indigenous innovation, *Paper presented at the annual meeting of the SASE Annual Conference 2012*, M.I.T., Cambridge, MA, June 28 2012.

Li, X. (2009). China's regional innovation capacity in transition: An empirical approach, *Research Policy*, 38, 338-357.

Lu, Q. (2000). *China's Leap into the* Information *Age: Innovation and Organization in the Computer Industry*, Oxford: Oxford University Press.

Marshall, A. (1920). *Principles of* Economics, London: Macmillan.

Metcalfe, S. & Ramlogan, R. (2008). Innovation systems and the competitive process in developing economies. *The Quarterly Review of Economics and Finance,* 48, 433-446.

Motohashi, K. & Yun, X. (2007). China's innovation system reform and growing industry and science linkages. *Research Policy*, 36, 1251-1260.

Motoyama, Y., Cao, C. & Appelbaum, R. (2014). Observing regional divergence of Chinese nanotechnology centers. *Technological Forecasting & Social Change*, 81, 11-21.

Newman, M. (2001). The structure of scientific collaboration networks, *Proceedings of the National Academy of Science of the United States of America (PNAS*). 98, 404-409.

Niosi, J. (2001). The competencies of regions-Canada's clusters in biotechnology, *Small Business Economics*, 17, 31-42.

Niosi, J. and J. Queenton, (2010). Knowledge capital in biotechnology industry: Impacts on Canadian firm performance. *International Journal of Knowledge-based Development*, 1, 136-151.

Perez, C. & Soete, L. (1988). Catching up in technology: entry barriers and windows of opportunity. In G. Dosi, et al. (Eds.). *Technical change and economic theory* (pp. 458–479). London UK: Pinter.

Ponds, R, van Oort, F. & Frenken, K. (2010). Innovation, spillovers and university–industry collaboration: an extended knowledge production function approach. *Journal of Economic Geography* , 10, 231–255.

Porter, A., Roessner, J., Jin, X. & Newman, N. (2002). Measuring national 'emerging technology' capabilities, *Science and Public Policy*, 29, 189-200.

Rivera-Batiz, A. & Romer, P. (1991). Economic Integration and Endogenous Growth. *Quarterly Journal of Economics*, 106, 531–555.

Schaaper, M. (2009). Measuring China's Innovation System: National Specificities and International Comparisons, *OECD Science, Technology and Industry Working Papers,* 2009/01, OECD Publishing. http://dx.doi.org/10.1787/227277262447

Sigurdson, J. (2004). Regional innovation systems (RIS) in China, *EIJS Working Paper*, No 195.

Vinciguerra, S., Frenken, K., Hoekman, J. & van Oort, F. (2011). European infrastructure networks and regional innovation in science-based technologies, *Economics of Innovation and New Technology*, 20, 517-537.

Verspagen, B. & Schoenmakers, W. (2004). The spatial dimension of patenting by multinational firms in Europe. *Journal of Economic Geography*, 4, 23-42.

Wagner, C. & Leydesdorff, L. (2005). Network structure, self-organization, and the growth international collaboration in science, *Research Policy*, 34, 1608-1618.

Wang, L., Meijers, H. & Szirmai, A. (2013). Technological spillovers and industrial growth in Chinese regions, *UNU-MERIT Working Paper* 2013-044.

Wang, L. & Notten, A. (2010). Benchmark Report: Nano-technology and Nano-science, 1998-2008, the ObservatoryNano project, the 7th Framework Program, *European Commission*.

Zucker, L.G. & Darby, M.R. (1997). The economists' case for biomedical research: Academic scientist-entrepreneurs and commercial success', in biotechnology. The Future of Biomedical Research. In C. E. Barfield & B. Smith (Eds.). Washington, DC: American Enterprise Institute and The Brookings Institution.

Zucker, L.G., M.R Darby & Armstrong, J. (1998). Geographically Localized Knowledge: Spillovers or Markets? *Economic Inquiry*, 36, 65-86.

Gender, Academic Position and Scientific Publishing: a bibliometric analysis of the oeuvres of researchers[1]

Inge van der Weijden* and Clara Calero Medina**

*i.c.m.van.der.weijden@cwts.leidenuniv.nl
**clara@cwts.leidenuniv.nl
CWTS, Leiden University, P.O. Box 905, Leiden, 2300 AX (The Netherlands)

Introduction

Today women are still significantly underrepresented in tenure-track and research university faculty positions. The study of West & Curtis (2006) shows that women represents one-quarter of full professors and earn on average 80% of the salary of men in comparable positions. However, there is increasing gender equity early in the pipeline, on master and PhD degrees. More general gender disparity can be ascribed to a male model of science, including masculinity of organisational, social and cultural norms within academic organization (Van Arensbergen 2014). Literature (qualitative as well as quantitative studies) shows five explanation models that are frequently used in the gender studies in academia: (1) glass ceiling: difficultly identified obstacles that hold women back from accessing the highest position in the hierarchy, (2) leaky pipeline: the pipeline has not advanced women to top-level positions due to leaks and blockages in the pipe; (3) Matthew & Matilde effect: 'the rich get richer' (Merton 1968) and 'the poor get poorer' (Mahbuba & Rousseau 2011); (4) gender myths: persisting myths in favour of men are creating attitudes in relation to the assessment of women's scientific performance' and (5) matching hypothesis: 'tendency of individuals to create ties with similar others' bias. These models are also often mentioned in gender equality debates in higher education in European countries for many years. Since academic publishing is still very important for career opportunities of both males and females in sciences, we focus in this study on comparing with bibliometric methods the oeuvres of female and male scientists.

Data

Our dataset contains in total 1994 researchers, 560 females and 1434 males. These scientists responded in 2011 to an online ACUMEN survey about web-presence (ACUMEN, 2011) conducted by University of Wolverhampton. Respondents are active in 15 different EU countries and four disciplines: (a) astronomy and astrophysics (A&A), (b) public health (PH), (c) philosophy (Phil) including the history and philosophy of science, and (d) environmental engineering (EE). We collected a set of papers from these academic researchers to conduct a bibliometric analysis. The ACUMEN partners already provided part of the set of papers and we completed the set by using the 'Large scale author name disambiguation using rule-based scoring and clustering" algorithm developed at CWTS to detect publications per researcher. The algorithm used the email information for each

[1] This work was supported by the EU FP7 ACUMEN project (Grant agreement: 266632). We would like to thank Mike Thelwall, Kayvan Kousha, Judit Bar-Ilan, Hadas Shema for collecting parts of the data.

researcher to retrieve the publications. The discipline in which gender is most equally represented is PH. The academic position where the female researchers are less represented is full professor (17% of the full professors is female); while the PhD & master student rank has the highest female/males ratio (39% of the students are female).

Results

General Gender Analyses

First we investigated possible gender differences in scientific output. We counted the total number of publications per researcher. A total of 23 researchers have zero publications in the Web of Science (WoS) database, we excluded them from further analyzes. Results show (see table 1) that men produced on average a significant higher number of publications compared to women; 23 publications versus 36 publications (*F=34.2; p=.000*). This suggests that men are more productive. This result confirmed previous findings on productivity and gender (i.e. Fox 2005, Mauleón & Bordons 2006, Leahey 2006, and Larivière et al. 2011).

Table 1. Indicators of output per gender (1980-2011/12)

Gender	Number of researchers	Publications	P per researcher
Female	553	12526	**23.11**(Mdn=10.0)
(Male	1418	48721	**36.58** (Mdn=18.0)

In this study we also pay attention to authorship order, given that the first and sometimes also last author publications are at least as important as raw publication counts for hiring, promotion and tenure (Wren et al. 2007). Table 2 presents the proportion of papers in which the researchers in our sample are mentioned as first author, last author and single author on the publications. Here we show that women are not evenly represented across authorship positions. With regard to first authorship, studies showed that women have been historically underrepresented in the first author position. Recently West et al (2012) showed that these discrepancies have been declined. Interestingly, in our sample women are overrepresented in the first authorship position. On average 37% of the papers in the oeuvres of female researchers consist of first authorships; this is a significantly higher percentage compared to males oeuvres (average 28%; *F=41.9; p=.000*). However, women are significantly underrepresented in the last author position (13% versus 19%; *F=29,8; p=.000*) and single-authored papers (15% versus 25%; *F=25.6; p=.000*). This last finding is in line with earlier studies (West 2012) that also showed that women remain underrepresented as last authors.

Table 2. Indicators of output per gender (1980-2011/12)

Gender	First authorship	Last authorship	Single authorship
Female	**37%** (Mdn=33%)	**13%** (Mdn=10%)	**15%** (Mdn=0%)
Male	**28%** (Mdn=24%)	**19%** (Mdn=15%)	**25%** (Mdn=0%)

Second, we explored gender differences with regard to the impact of publications. Table 3 shows impact indicators and some citation information per gender. The average number of citations per publication of male scientists (Mcs) is 10.13; this is somewhat higher than females (Mcs=9.74). However this difference is very small and non-significant (*F=0.48, p=0.49*). The mean normalized citation score (Mncs) show that both females and males are having an impact around world average. The average Mnjs is 1.09 for males and 1.08 for females, suggesting that both males and females tend to publish in journals with the same impact. Both female and male scientists have the same proportion of papers (11% for both genders) that belong to the top 10%. In sum, there are no gender differences with regard to impact.

Table 3. Indicators of impact per gender (1980-2011/12)

Gender	**Mcs**	**Mncs**	**Mnjs**	**pp top 10%**
Female	9.74 (Mdn=6.73)	1.08 (Mdn=0.88)	1.05 (Mdn=1.04)	11% (Mdn=6.0%)
Male	10.13 (Mdn=7.22)	1.09 (Mdn=0.88)	1.09 (Mdn=1.05)	11% (Mdn=6.5%)

Third, we calculated collaboration indicators and analyzed differences between women and men. Currently, lots of publications are written in teams in which researchers from different national and international institutes are collaborating together. Among female researchers in our dataset, on average 59% of their papers are written and published in collaboration with researchers (co-authors) from other institutions, and 32% is the result of international collaboration. As shown in Table 4, the male researchers in our dataset have a lower percentage of publications in collaboration (55%) compared to females (59%; *F=7.3; p=.007*). Many bibliometric studies show that female researchers are less involved in international collaboration than male researchers (i.e. Lewison 2001, Webster 2001, Larivière et al. 2011 & 2013). Our sample shows small and non-significant gender differences in this respect: on average 35% of the WoS publications in the oeuvres of males are the result of international collaboration compared to 32% females (*F=3.3; p=.07*).

Table 4. Indicators of collaboration per gender (1980-2011/12)

Gender	**pp collab**	**pp int collab**
Female	**59%** (Mdn=67%)	32% (Mdn=24%)
Male	**55%** (Mdn=33%)	35% (Mdn=30%)

Gender Analyses based on Research Disciplines

Table 5 shows the indicators of output per discipline and gender of the researchers. In the discipline labeled as philosophy the number of papers per researcher is low, as we expected in such a field. In all the four disciplines the number of publications per researcher is significant higher per male than per female.

Table 5 Indicators of output per discipline and gender (1980-2011/12)

Discipline	**Gender**	**Number of Researchers**	**P**	**P per Researcher**
A&A	Female	101	3465	**36.0** (Mdn=29.0)
A&A	Male	393	18686	**53.8** (Mdn=35.0)
EE	Female	138	3006	**21.8** (Mdn=10.0)
EE	Male	393	13724	**35.5** (Mdn=21.0)
Phil	Female	86	521	**6.0** (Mdn=3.0)
Phil	Male	354	4879	**13.9**(Mdn=5.0)
PH	Female	228	5560	**24.6** (Mdn=13.0)
PH	Male	278	11721	**42.7** (Mdn=26.0)

In terms of the proportion of papers signed as the first author, for the field of A&A and PH the average percentage is significantly higher for female researchers compared to male researchers. Literature showed that author order varies across disciplines (Waltman 2012). In life sciences and biosciences the last authorship position is a prestige one. Males in each of the four selected disciplines of this study have a higher proportion of last authorship compared to their female colleagues. In the disciplines environmental engineering (EE; *F=20.2; p=.000*) and public health (PH, *F=16.6; p=.000*) these gender differences are also significant.

Table 6. Indicators of output per discipline and gender (1980-2011/12)

Discipline	Gender	First authorship	Last authorship	Single authorship
A&A	Female	**36.4%** (Mdn=32.7%)	18.3% (Mdn=15.4%)	**1.1%** (Mdn=0.0%)
A&A	Male	**31.2%** (Mdn=25.7%)	21.3% (Mdn=16.7%)	**7.9%** (Mdn=0.0%)
EE	Female	41.4% (Mdn=35.0%)	**14.5%** (Mdn=10.0%)	6.6% (Mdn=0.0%)
EE	Male	37.3% (Mdn=33.3%)	**23.9%** (Mdn=19.6%)	6.4% (Mdn=0.0%)
Phil	Female	14.4% (Mdn=0.0%)	9.3% (Mdn=0.0%)	72.8% (Mdn=100%)
Phil	Male	13.1% (Mdn=0.0%)	12.2% (Mdn=0.0%)	71.8% (Mdn=95.5%)
PH	Female	**44.1%** (Mdn=39.1%)	**12.8%** (Mdn=11.1%)	2.4% (Mdn=0.0%)
PH	Male	**32.6%** (Mdn=28.2%)	**19.1%** (Mdn=17.6)	3.8% (Mdn=0.0%)

In philosophy (Phil) the oeuvres of both female and male scientists consist mainly of single author papers. There are no gender differences in this discipline. In each of the disciplines we report no significant differences between both genders with regard to impact indicators.

In terms of collaboration, Table 7 shows the indicators of collaboration per discipline and gender of the researchers. Our results show that depending on the discipline the degree of collaboration in general and internationally specifically varies. Interestingly, females in the discipline A&A show higher percentages of both inter-institutions and international collaborations compared to males. Inter-institutional collaboration is significantly higher for females (81%) than for males (74%; *F=7.3; p=.007*). In contrast, the percentage inter-institutional collaborative publications and the percentage international collaborative publications for the discipline EE is higher for men than for women. International collaboration is significantly higher for males (31%) than for females (26%; *F=4.0; p=.05*). PH shows quite similar percentage of inter-institutional and international collaborations per gender. As the oeuvres of scientists in Phil mainly consist of single author papers, collaboration for both genders is low.

Table 7. Indicators of collaboration per discipline and gender (1980-2011/12)

Discipline	Gender	pp collab	pp int collab
A&A	Female	**81.2%** (Mdn=86.8%)	66.4% (Mdn=71.4%)
A&A	Male	**74.3%** (Mdn=80.0%)	62.9% (Mdn=67.4%)
EE	Female	51.4% (Mdn=51.6%)	**25.9%** (Mdn=18.8%)
EE	Male	55.2% (Mdn=57.1%)	**31.0%** (Mdn=27.5%)
Phil	Female	17.1% (Mdn=0.0%)	10.1% (Mdn=0.0%)
Phil	Male	21.1% (Mdn=0.0%	12.0% (Mdn=0.0%)
PH	Female	70.0% (Mdn=73.3%)	29.8% (Mdn=22.0%)
PH	Male	69.1% (Mdn=71.4%)	32.2% (Mdn=26.7%)

Gender Analyses based on Academic Positions

In terms of the output per researcher and academic position results show that male researchers produce more papers than females regardless their academic position. For the postdoctoral research fellows is where the differences are lower though. Table 8 presents the proportion of papers in which the researchers in our sample are mentioned as first author, last author and single author on the publications. At each level on the career ladder, the papers in the oeuvres of female researchers consist of a higher percentage of first authorships compared to men. For associate professors, assistant professors, postdocs and other research positions these gender differences are also significant. With regard to last authorships, female associate professors are significantly underrepresented in this prestigious authorship position (16% versus 20%; *F=5.8; p=.016*). At the full professor level we find relatively small and non-significant gender differences (females 24% versus males 26%; *F=0.55; p=.46*). As last authorship positions are mainly dedicated to full and associate professors, as indicated by medians of 0% for both genders at lower ranks we can't elaborate on possible gender differences at these lower positions.

Table 8. Indicators of output per academic position and gender (1980-2011/12)

Academic Position	Gender	First authorship	Last authorship	Single authorship
Full Prof	Female	25.0% (Mdn=22.0%)	23.9% (Mdn=19.4%)	**22.4**% (Mdn=0.0%)
Full Prof	Male	21.3% (Mdn=17.4%)	26.01 (Mdn=23.08%)	**36.8%** (Mdn=0.0%)
Associate Prof/ Reader/Sr Lecturer	Female	**32.9%** (Mdn=27.4%)	**15.9%** (Mdn=14.3%)	16.21% (Mdn=0.0%)
Associate Prof/ Reader/Sr Lecturer	Male	**27.8%** (Mdn=25.0%)	**19.7%** (Mdn=17.7%)	22.07% (Mdn=0.0%)
Assistant Prof/Lecturer	Female	**37.9%** (Mdn=33.3%)	10.6% (Mdn=0.0%)	18.22% (Mdn=0.0%)
Assistant Prof/Lecturer	Male	**31.0%** (Mdn=29.0%)	14.7% (Mdn=10.0%)	27.64% (Mdn=0.0%)
Postdoc	Female	**45.7%** (Mdn=45.3%)	8.26% (Mdn=0.0%)	**7.9%** (Mdn=0.0%)
Postdoc	Male	**36.9%** (Mdn=32.6%)	10.8% (Mdn=0.0%)	**17.7%** (Mdn=0.0%)
Student	Female	49.2% (Mdn=50.0%)	6.6% (Mdn=0.0%)	11.0% (Mdn=0.0%)
Student	Male	40.0% (Mdn=40.0%)	11.9% (Mdn=0.0%)	13.6% (Mdn=0.0%)
Other	Female	**36.4%** (Mdn=30.0%)	**13.4%** (Mdn=10.0%)	17.4%(Mdn=0.0%)
Other	Male	**24.5%** (Mdn=22.0%)	**21.5%** (Mdn=18.1%)	22.7%(Mdn=0.0%)

The impact indicators, based on the academic position and gender, show very small differences between female and male researchers. Finally, Table 9 presents the results in terms of collaboration per academic position and gender. At the level of full professors, the percentage of collaboration is higher compared to males who have the same position in academia (57% versus 46%, *F=7.3; p=.007*). At lower rank, the percentage of international collaboration is always lower (although not significantly) for female researchers than for male researchers. In terms of inter-institutional collaboration there are no gender differences among those lower ranked academic positions.

Table 9. Indicators of collaboration per academic position and gender (1980-2011/12)

Academic Position	Gender	pp collab	pp int collab
Full Prof	Female	**56.8%** (Mdn=64.9%)	30.7% (Mdn=25.0%)
Full Prof	Male	**46.0%** (Mdn=50.4%)	28.9% (Mdn=25.4%)
Associate Prof/ Reader/Sr Lecturer	Female	57.2% (Mdn=66.3%)	30.5% (Mdn=23.5%)
Associate Prof/ Reader/Sr Lecturer	Male	56.6% (Mdn=62.5%)	34.6% (Mdn=27.3%)
Assistant Prof/Lecturer	Female	53.3% (Mdn=50.0%)	28.6% (Mdn=18.2%)
Assistant Prof/Lecturer	Male	54.0% (Mdn=58.0%)	31.6% (Mdn=21.9%)
Postdoc	Female	67.1% (M)n=75.0%)	43.2% (Mdn=42.9%)
Postdoc	Male	64.6% (Mdn=69.4%_	46.3% (Mdn=50.0%)
Student	Female	65.8% (Mdn=82.6%)	33.6% (Mdn=19.1%)
Student	Male	58.0% (Mdn=82.6%)	45.0% (Mdn=40.0%)
Other	Female	58.6% (Mdn=66.7%)	**26.3%** (Mdn=15.8%)
Other	Male	57.3% (Mdn=66.7%)	**38.5%** (Mdn=35.1%)

Conclusion and discussion

Our bibliometric analysis confirms the traditional gender pattern; men produce on average a higher number of *publications* compared to women, regardless their academic position and research field. In this report we also pay attention to authorship order, given that the first and sometimes also last author publications are at least as important as raw publication counts for hiring, promotion and tenure (Wren et al. 2007). Our results suggest that women are not evenly represented across authorship positions. In our sample women are overrepresented in the first authorship position, especially in the disciplines A&A and PH. At each level on the career ladder, the papers in the oeuvres of female researchers consist of a higher percentage of first authorships compared to men. With regard to last authorship position, women in all four selected disciplines are significantly underrepresented this prestigious position. Female associate professors are significantly underrepresented. As last authorship positions are mainly dedicated to full and associate professors we can’t elaborate on possible gender differences at these lower positions. Interestingly, we show no gender differences regarding research *impact* in each studied disciplines and positions in academia, as measured by three indicators (MCS, MNCS, and PPtop10%). Our results show that depending on the discipline the degree of *collaboration* in general (inter-institutional) and internationally specifically varies. Interestingly, at the level of full professors, the percentage of collaboration is higher compared to males who have the same position in academia. At lower rank, the percentage of international collaboration is always lower for female researchers than for male researchers. As collaboration is one of the main drivers of research output and scientific impact (Larivière et al 2013), we recommend to develop and promote programs for female early career researchers. To increase internationally collaboration opportunities, female scientists should search for support of an international mentor. In the mentor-mentee conversations, female mentees will also be trained to improve their personal & managerial skills such as negotiation, self-promoting and networking, as research (West et al 2011) suggested that these qualities are necessary in discussions about authorship order. In this way mentorship could contribute to speed up the process of closing the gender gap in science. Female full professors could act as a role model mentor for female early career scientists as there are some expectations in the literature that underrepresented groups are better served with mentors or role models who had similar characteristics of life experiences (Kopia, Melkers & Tanyildiz 2009).

References

Fox, M. F. (2005). Gender, family characteristics, and publication productivity among scientists. *Social Studies of Science*, 35(1), p. 131–150.

Kopia, A., Melkers, J. and Tanyldiz, Z.E. (2009). Women in academic science: mentors and career development. Book chapter in: *Women in science & technology*. Edited by Prpic, K., Liveira, L and S. Hemlin. Institute for Social Research, Zagreb. Sociology of Science and Technology Network of the European Sociological Association.

Larivière, V., Vignola-Gagné, E., Villeneuve, E., Gélinas, P. & Gingras, Y. (2011). Sex differences in research funding, productivity and impact: an analysis of Québec university professors. *Scientometrics*, 87(3), p. 483–498.

Larivière, V., Ni, C., Gingras, Y., Cronin, B. & Sugimoto, C.R. (2013). Global gender disparities in science. *Nature*, 504, p. 211-213.

Leahey, E. (2006). Gender differences in productivity. Research specialization as a missing link. Gender & Society, 20(6), p. 754–780.

Lewison, G. (2001). The quantity and quality of female researchers: a blibliometric study of Iceland. *Scientometrics*, 52(1), p. 29-43.

Mahbuba, D. & Rousseau, R. (2011). Matthew, Matilda and the others. In proceedings of the 7th International Conference of Webometrics, Infometrics and Scientometrics & 12 COLLNET Meeting, 20-23 September 2011, Istanbul Bilgi University, Instanbul.

Mauleón, E. & Bordons, M. (2006). Productivity, impact and publication habits by gender in the area of materials science. *Scientometrics*, 66(1), p. 199–218.

Merton, R.K. (1968). The Matthew effect in science. *Science*, 159, p. 56-63.

Van Arensbergen, P. (2014). The selection of talent in academia. PhD thesis. The Hague: Rathenau Instituut. *Forthcoming*

Waltman, L. (2012). An empirical analysis of the use of alphabetical authorship in scientific publishing. *Journal of Infometrics*, 6 (4), p. 700-711

Webster, B. M. (2001). Polish women in science: a bibliometric analysis of Polish science and its publications, 1980–1999. *Research Evaluation*, 10(3), p. 185–194.

West, M.S. & Curtis, J.W. (2006). AAUP Faculty Gender Equity Indicators 2006, Technical Report. American Association of University Professors. http://www.aaup.org/NR/rdonlyres/63396944-44BE-4ABA-9815-5792D93856F1/0/AAUPGenderEquityIndicators2006.pdf

West, J.D., Jacquet, J., King, M.M., Correll, S.J., and Bergstrom, C.T. (2012). The role of gender in scholarly authorship. *PloS ONE*, DOI: 10.1371/journal.pone.0066212.

Wren, J.D., Kozak, K.Z., Johnson, S.J., Deakyne, L.M., Schilling, L.M., and Dellavalle, R.P. (2007). *The write position. A survey of* perceived *contributions to papers based on byline position and number of authors*. EMBO report. November (8)11: 988-991. Doi: 10.1038/sj.embor.7401095.

Gender Differences in Societal Orientation and Output of Individual Scientists[1]

Inge van der Weijden*, Zohreh Zahedi**, Ülle Must*** and Ingeborg Meijer****

*i.c.m.van.der.weijden@cwts.leidenuniv.nl; **z.zahedi.2@cwts.leidenuniv.nl, ****i.meijer@cwts.leidenuniv.nl
CWTS, Leiden University, P.O. Box 905, Leiden, 2300 AX (The Netherlands)

***Ulle.Must@etag.ee
Estonian Research Council (Estonia)

Introduction

Academic science is currently shaped by pressure towards academic excellence and by (policy) aspirations towards knowledge transfer and research activities beyond academia (van der Weijden et al 2012). Over the years, discussions about the societal value of academic science have become more extensive—and research-funding agencies increasingly ask about the explicit societal relevance or anticipated societal impact of proposed research, e.g. the Dutch Research Council, and the National Science Foundation. Potentially, society can benefit from academic research in various ways, ranging from contributions to culture and education to specific insights or products with economic or socio-political value (Van der Weijden et al 2012). Several policy frameworks have been developed to value societal outputs and impacts (Meijer 2012, Finne et al 2011, ref http://www.ref.ac.uk/), but the actual measuring is still scarce (Mostert et al 2010). Moreover, scientists do not necessarily share the policy focus on societal outputs or impact since they have to find a balance between activities towards academic excellence and other society oriented activities (Niederkrothenthaler et al 2011). From research leaders in the academic system, who have a tenure position, one would expect responsiveness to the societal policy focus, both in terms of opinion and in terms of activities. From postdocs, however, their individual career perspective, which is predominantly guided by academic excellence, might prevail over societal orientation. Here we investigate the opinion and socio-economic oriented activities of researcher leaders and postdocs. And given the stagnation of women in science, we also investigate whether there are gender differences both in opinions and activities. We hypothesize that male scientists are more geared towards the scientific reward and recognition system, and consequently may be less oriented towards societal relevance, In this paper, our aim is to contribute empirical evidence on how the quest for societal relevance is differently taken up by male and female principal investigators.

Women in Science

The stagnation of women in the academic system is considered to be the result of different developments. First, as norms in science are still masculine, this influences rewards and career support and -opportunities for women (Van Arensbergen 2014). Changes in organizational culture alone cannot solve all women career problems. Also policy changes on

[1] The Peer Review Practices questionnaire was supported by the EU FP7 ACUMEN project (Grant agreement: 266632). We are greatly indebted to Rosalie Belder for collecting the Academic Leadership data.

both national and international level are needed to give suitable stereotypes and frameworks. Second, research showed gender differences in salary (Levecque et al 2014), tenure, rank, promotion, mobility and employment outside one's field of training. After controlling for professional characteristics and productivity, the 'pay-gap' between men and women in academia is still prominent (DesRoches et al 2010). Furthermore, females tend to believe that being a woman was a negative factor with respect to academic advancement, leadership, opportunities, salary and resources (Wagner et al 2007). Third, studies show that women have less self-confidence compared to men. Most women believe that they would never reach the professoriate while most men assumed they would (Baker 2010; Chesterman, Ross-Smith & Peters 2005). Fourth, literature shows the impact of partnership and marriage on female academic careers. Family patterns are still hindering career success and reproduce the gender gap. However, having a partner who is also working in academia has a positive impact on performance. The fact that women bear children and take on the majority of childcare responsibilities often leads to career breaks and fewer weekly working hours for women. To come to a suitable work-life balance with children is harder for women than for men, but also depend on personal determination, networks, and institutional conditions (Baker 2010). Compared with marriage, parenthood seems to be more careers hindering than marriage (Wolfinger et al 2008).

Research Questions
In this study we investigate three research questions. What are gender differences with regard to:

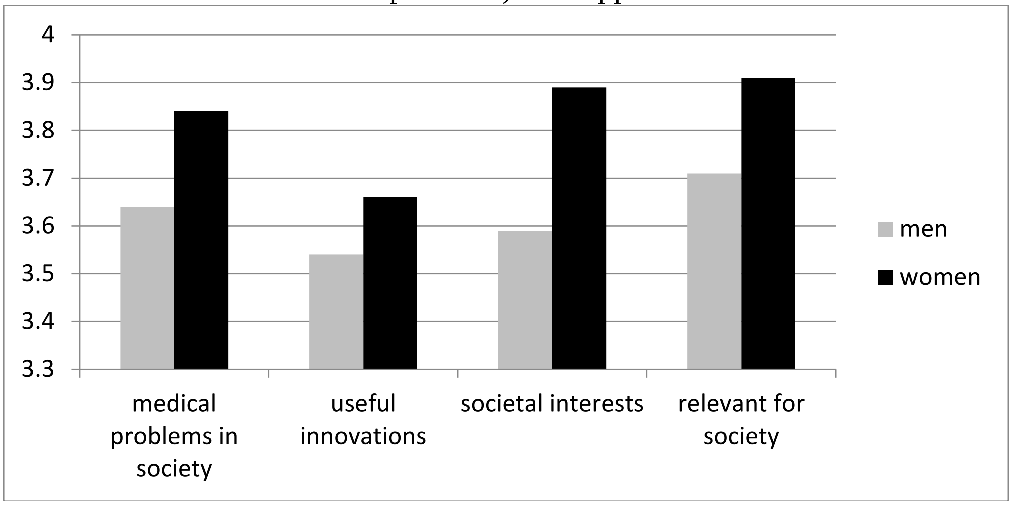
1. Researchers' views about the increasing emphasis on the societal impact;
2. Researchers' attitudes towards the way in which impact of scientific production is measured and evaluated;
3. Different types of direct societal output products produced by research groups.

Data and Methods
We used two datasets to analyze the gender differences in societal orientation and output of individual scholars. First, we used the academic leadership dataset (Belder et al 2012) to analyze societal orientation and output of principal investigators. Data were collected in 2010-2011 in a survey among 458 biomedical and health research leaders (351 males; 107 females) in the Netherlands. Second, we used the Peer Review Practices (PRP) dataset (Must, Otsus & Mustajoki 2012) to analyze gender differences with regard to researchers' attitudes towards the way in which impact of scientific production is measured and evaluated. The PRP is the result of a web-survey conducted in 2011 as part of the FP7 project 'ACUMEN'. In total, 2114 respondents in different phases of their careers and affiliated in 66 countries answered the PRP questionnaire. In both studies we conducted gender analyses by using SPSS.

Results
Societal orientations
First we investigate possible gender differences in societal research goals of research groups. In the academic leadership survey, we ask whether the increasing emphasis on societal impact has implications for research group goals. Do research group leaders take societal relevance into account when formulating the group's research agenda? As shown in Figure 1 the attitude of female research leaders towards societal research goals is positive. Male leaders have more neutral views compared to female leaders. Overall, female respondents have more positive views about the orientation of their research towards: (1) medical problems in society

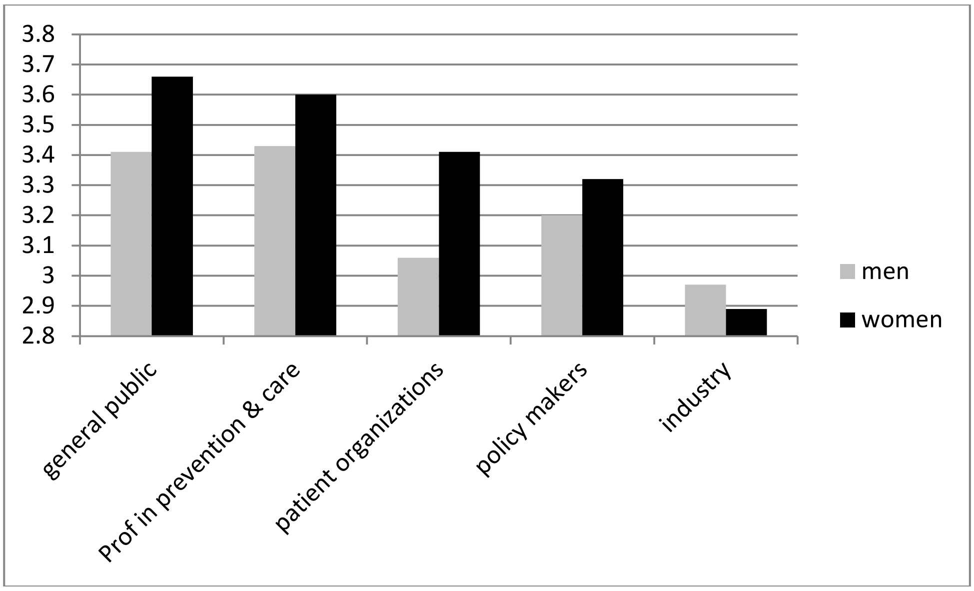
(U=15861, p=.028), (2) production of useful innovations for healthcare practice (U=14619; p=.001) and (3) relevance for society (U=16025, p=.045) compared to males.

Figure 1: Opinions of research leaders on societal research goals measured on a five point scale ranging from 1=strongly disagree till 5=strongly agree (data from the academic leadership dataset). See appendix 1 for more details.



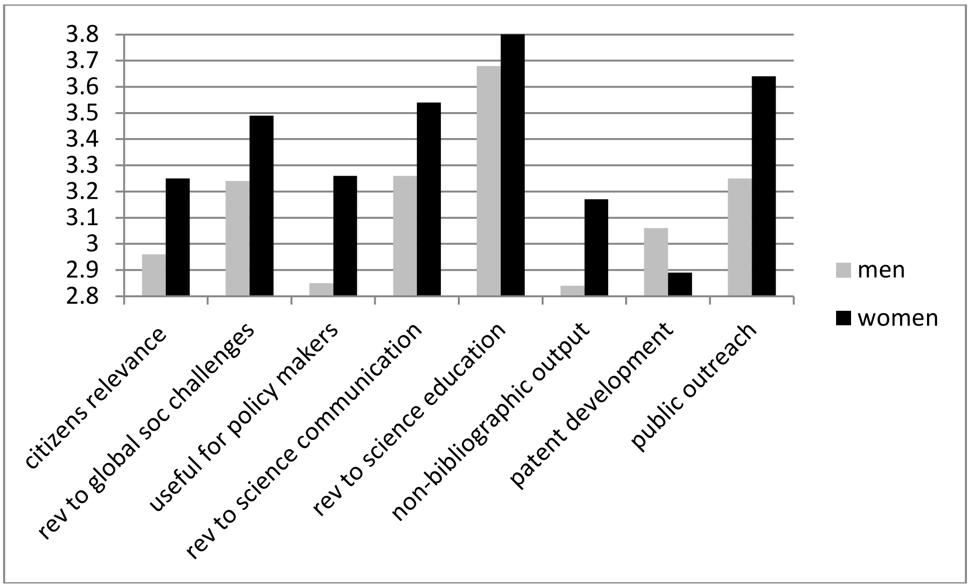
Second, we explore whether this positive attitude towards societal orientation in research also translates into (more? better?) interactions with stakeholders? When asked about the effects of the recent increase in emphasis on societal impact, research leaders reported a slight increase of interactions with various stakeholders (figure 2). Overall, female respondents have more positive views about changes in their interactions with the general public (U=15178, p=.010), professionals in prevention and care (U=16051; p=.050) and patient organizations (U=14145, p=.000) compared to males. Interactions with industry show a slightly different pattern: both female and male group leaders do not report intensified interaction with companies

Figure 2: Opinions of research leaders on research communication with stakeholders measured on a five point scale ranging from 1=strongly disagree till 5=strongly agree. (prof=professionals; (data from the academic leadership dataset). See appendix 1 for more details.



Third, by using the PRP dataset we find that opinions about societal orientation of scientists vary among different career stages. Particularly different are the preferences of postdocs and professors. As shown in Figure 3 gender differences in rating societal indicators are most prominent in the postdoc career phase. Female postdocs give higher values to the: (a) relevance to global societal challenges and (b) science communication; (c) contributing to science education; (d) usefulness to policy decision makers; and (e) relevance to citizens' concerns to these indicators compared to males in the same career phase. Also output indicators as non-bibliographic outputs and public outreach are more important for female postdocs compared to male postdocs, who give on average lower ranks. One exception: patents, where men are more positive then women. This suggests a gender difference in entrepreneurial behaviour and activities of research group leaders. Our data show that groups which are managed by a female leader indeed collaborate less with partners from industry compared to male group leaders (f=37%; m=55%; Chi=10.52; p=.001).

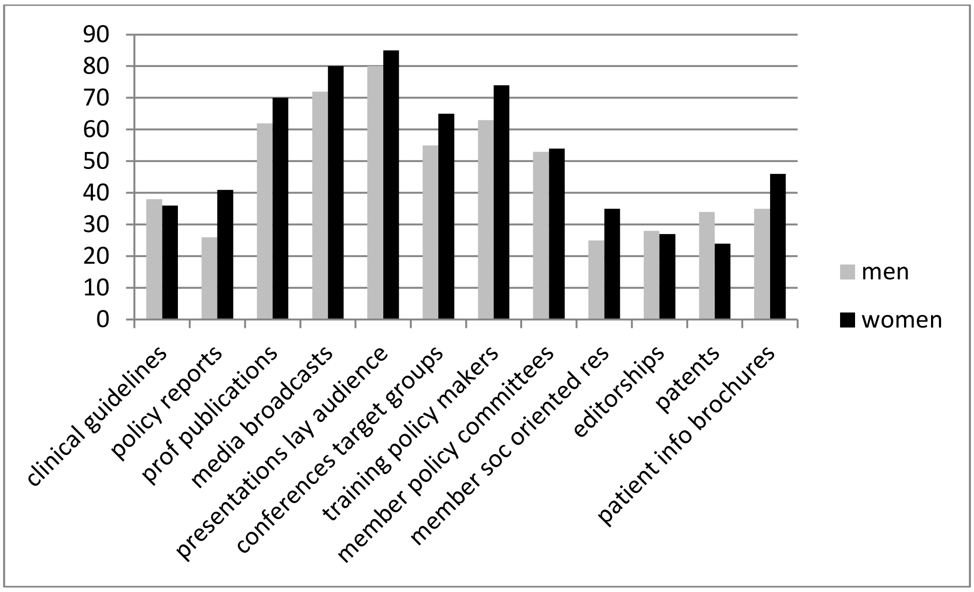
Figure 3: Preferences of postdocs on the use of societal indicators measured on a five point scale ranging from 1=strongly disagree till 5=strongly agree. (rev=relevance; data from the PRP dataset)



Non-scholarly output

Whereas the previous section was about the societal orientation, we turn in this section to the non-scholarly output of research. Biomedical and health groups produced different types of societal research output. Products can be 'tangible', such as reports, but we also included productive interactions with stakeholders as output (Evaluating Research in Context 2010, van der Weijden et al 2012). Figure 4 presents the proportion of research leaders in our academic leadership sample who generate different societal output within a three-year period. We measure 11 different indicators. Here we show that research groups that are managed and lead by women are generally more concentrated towards societal output. On 8 of the 11 indicators, females score higher than males. The three most important gender differences in societal output are: (1) policy reports (f=41% vs m=26%; Chi=8.06; p=.005); (2) patient information brochures (f=46% vs m=35%; Chi 4.17, p=.041); and (3) training for policy makers (f=74% vs m=63%; Chi 4.13, p=.042).

Figure 4: Percentage of male and female PI's who produce societal output (data from the academic leadership dataset)



Conclusion & Discussion

In summary, men and women differ in their societal orientation and societal output. Female research leaders are quite positive about the increased societal orientation of their research. Male research leaders have more neutral views. Females are also overall more active and productive in generating societal output, compared to their male colleagues. From the survey on peer review practices it can be concluded that the new generation of researchers give higher rates to social (relevance) indicators compared to the older generation; women even more than men. Gender differences in rating social indicators are most prominent in the postdoc career phase; female postdocs give higher values to these indicators compared to males in the same career phase. This may indicate a stronger male focus on the scientific reward and recognition system.

To an increasing degree attention is paid to the contribution of academic research to the wider society and societal impact is currently a significant subject of science policy. In the Netherlands for example it recently became an obligatory paragraph in several types of grant applicants (van Arensbergen 2014). This development could contribute to the improvement of incentives for scientists who focus on societal impact. It is also interesting to note that women are more oriented towards various types of societal outputs, whereas men have a stronger focus on patents, and entrepreneurial activities. In the Netherlands, the 'Topsector' policy is steering this economic type of societal activities, which may not be beneficial for closing the gender gap in the Netherlands.

Overall in Europe, reducing the mismatch between academic activities that women prefer (e.g. teaching, supervision, distributing knowledge to society) and activities that are rewarded (publications, citations, grants), could speed up the process of closing the gender gap in science.

References

Baker, M. (2010). Career confidence and gendered expectations of academic promotion. *Journal of Sociology*, 46, p. 317-334.

Belder, R., Verbree, M., Van der Weijden, I. & Van der Meulen, M. (2012). Een andere stijl van academisch leiderschap? Effect van loopbaanfase en gender op het leiderschap van onderzoeksleiders. The Hague: Rathenau Instituut.

Chesterman, C., Ross-Smith, A. & Peters, M. (2005). Not doable jobs! Exploring senior women's attitudes to academic leadership roles. *Women's Studies International Forum*, 28, p. 163-180.

DesRoches, C.M., Zinner, D.E., Rao, S.R., Iezzoni, L. & Campbell, E.G. (2010). Activities, productivity and compensation of men and women in the life sciences. *Academic Medicine*, 85, p. 631-639.

Evaluating Research in Context (ERiC). (2010) Handreiking. Evaluatie van maatschappelijke relevantie van wetenschappelijk onderzoek. Den Haag: ERiC.

Finne, H., Day, A., Piccaluga, A., Spithoven, A., Walter, P. & Wellen, D. (2011). A Composite Indicator for Knowledge Transfer. Report from the European Commission's Expert Group on Knowledge Transfer Indicators

Levecue, K., Baute, S., Van Rossem, R. & Anseel, F. (2014). *Money, money, money…On PhDs, wages and gender*. ECOOM Briefs, 9, January 2014. Ghent University, Belgium.

Meijer, I. (2012). Societal returns of scientific research; how can we measure it. CWTS Working Paper Series CWTS-WP-2012-014.

Mostert, S.P., Ellenbroek, S.P.H., Meijer, I., van Ark, G. & Klasen, E.C. (2010). Societal output and use of research performed by health research groups. *Health Research Policy and Sytems 8:30.*

Must, Ü., Otsus, O. & Mustajok, A. (2012). *ACUMEN survey on Peer Review Practices*. ACUMEN Task 1.2. April 2012.

Niederkrothenthaler, T., Dorner, T.E. & Maier, M. (2011). Development of a practical tool to measure the impact the impact of publications on the society based on focus group discussions with scientists. *BMC Public health 11:588*

Van Arensbergen, P. (2014). The selection of talent in academia. PhD thesis. The Hague: Rathenau Instituut. *Forthcoming*

Van der Weijden, Verbree, M. & Van den Besselaar, P (2012). From bench to bedside: the societal orientation of research leaders: The case of biomedical and health research in the Netherlands. *Science and Public Policy* 39, pp.285-303

Wagner, A.K., Mcelligott, J., Chan, L., Wagner, E., Segal, N.A. & Gerber, L.H. (2007). How gender impacts career development and leadership in rehabilitation medicine: a report from the AAPM&R Research Committee. *Archives of Physical Medicine and Rehabilitation*, 88, 560-568.

Wolfinger, N.H., Mason, M.A. & Goulden, M. (2008). Problems in the pipeline: gender, marriage, and fertility in the ivory tower. *Journal of higher Education*, 79, p. 388-405.

Appendix 1: Survey questions societal orientation

Question: The growing attention focused on the societal impact of health research has meant that research within my group:

	Strongly disagree	**Disagree**	**Neither agree nor disagree**	**Agree**	**Strongly agree**
a) is more oriented towards medical problems in society	1	2	3	4	5
b) produces more useful innovations for healthcare practice	1	2	3	4	5
c) takes better account of societal interests	1	2	3	4	5
d) is more relevant to society	1	2	3	4	5
e) is better communicated to the general public (via the media)	1	2	3	4	5
f) is better communicated to professionals working in healthcare and disease prevention	1	2	3	4	5
g) is better communicated to patients (and their organisations)	1	2	3	4	5
h) is better communicated to policymaking professionals	1	2	3	4	5
i) is better communicated to professionals in industry	1	2	3	4	5

The effects of ‘ready to use’ bibliometric indicators[1]

Lorna Wildgaard*

*pnm664@iva.ku.dk
Royal School of Library and Information Science, Faculty of the Humanities, Copenhagen University, Birketinget 6, 2300 S (Denmark)

Introduction

‘Ready to use’ bibliometric indicators are being used by end users as never before. Administrators and evaluators for assessment purposes, whereas researchers add indicators to their CV in a competitive move to show visibility in the academic community. ‘Ready to use’ indicators are important as they provide researchers with the opportunity to evaluate the effect of their own published works before being evaluated by others. The numerical values of these indicators have personal significance to the individual as they can be interpreted as criteria of success or failure, therefore the bibliometric community must investigate the psychological effects of ‘ready to use’ indicators.

Therefore, the key questions this poster addresses are:

1. What are the psychological effects of ‘ready to use’ indicators on the researcher?
2. Accordingly, which issues, need to be addressed?

Background

Accepting the application of ‘ready to use’ bibliometrics by laymen means it is important to understand their advantages, short comings and effects. Citation databases support ‘ready to use’ indicators with studies validating their robustness and detailed guides on how to collect data. However, the do’s and don’ts in interpreting the values are not as clear cut. Indicators are used in a culture of evaluation that has become a reward system. The indicator values are no longer limited to measuring the effects of a publication strategy, but are used as surrogate measures of “good” science and “bad” science. Table 1 presents the three main sources of “ready to use” indicators for the layman. The first column names the citation index, the second the indicators, the third and fourth columns present the advantages and disadvantages of the index. From this simple overview, it is clear that the same indicator computed in different citation indices can produce different results, as publication and citation counts are affected by indexing policies and technical limitations as well as by variation rates between specialities. The researcher’s success or failure can be interpreted differently in each index which is further extemporised by the individuals cognitive, institutional and social locations, as discussed in Martin and Irvine (1983), Bach (2011), KAOW/ARSOM 2012 and Bornmann (2013). As such, the layman is using incomplete bibliometrics to evaluate scientific contributions, normalized (interpreted) subjectively for their own social and cognitive locations.

[1] This work was supported by the ACUMEN FP7 project. The work presented here is used on the development of Guidelines for Good Evaluation Practice. The ACUMEN collaboration aims at understanding how researchers are evaluated and the science system can be improved and enhanced, www.research-acumen.eu.

Table 1. Three major sources of author-level ready-to-use indicators

Database	Author level Indicators	Advantages	Disadvantages
Google Scholar, searched using Harzing's Publish or Perish (POP)[2]	P, PPA, C, CPP, CPA, CPY, h, g, contemporary h, hi, POP h variation, multi-authored - index, average annual increase in the individual h, AWCR, AW, AWCRpa, APP.	• Free. • Covers non-English language titles. • Covers all types of published and unpublished work, written, audio and video.	• No quality control, errors in data, no subject filtering, unclear coverage. • Limited retrieval (page requests and top 1000 results). • 15 week publication time lag.
Web of Science, Thomsen Reuters [3]	P, PT, PY, C, CPP, C-sc, CPY, h, (citing authors/articles/articles-sc/, countries/institutions), WOS/research categories, language, collaborating authors, collaborating institutions.	• Citation data from 1900 to present. • 12.000 ISI indexed journals. Updated weekly. Excellent coverage of science. • Good author identification tools.	• Access to publication & citation data depends on version. • Strong English language bias. • Limited coverage of conference proceedings and books. • Moderate coverage of arts/humanities.
Scopus, Elsevier[4]	P, PY, PT, C, CPY, h, minus SC/citations from books, source titles/type, affiliation names, authors, language, countries, subject areas, citing areas, citing authors, citing affiliations, citing countries.	• More European literature than WOS. • 18.000 indexed journals. Updated daily. Good coverage of social science titles. • Includes conference proceedings.	• Limited coverage of books, book chapters and dissertations. • Citation data for papers published from 1996 onwards.

Method

This study aims to learn more about possible effects of 'ready to use' bibliometrics by reviewing the psychology and evaluation studies literature. Appendix 1 presents the studies we found in the literature. The first column is the reference for the study, the perspective is in the second column, while in the third column the main conclusions drawn in each paper.

Discussion

Our main findings and conclusions are discussed below:

- Career success[2] contains both subjective aspects, e.g. attitudes to work and career, and objective aspects, e.g. awards, publications and invited talks. The objective aspects are countable using 'ready to use' indicators, and in turn become explicit indicators of success.
- When documenting performance using 'ready to use' indicators, researchers are at the same time presenting for appraisal a snap shot of their self image and core personality traits (Judge & Hurst, 2007).
- 'Self esteem' effects overall evaluation (Vallacher et al, 2012). Academics with high 'self esteem' take more actions to achieve their goals, maximizing potential by consciously choosing to publish in high impact factor journals indexed in citation indices and hence will perform well on 'ready to use' indicators as the sources are more likely to be included in the indexing policies of citation indices.

[2] An extensive overview of CSE literature can be found in (Stump et al, 2010).

- Researchers regulate publishing success or failure to maintain positive ‘self views’ (Nicholls & Stukas, 2011) or fulfill performance quotas. Adapting behaviour to score well on indicators can undermine the purpose of evaluation and validity of the indicator (Dahler-Larsen, 2011) and reward competitive and aggressive researchers (Cheung 2008).
- Using ‘ready to use’ indicators to compare researchers can expose the individual (Crocker et al, 2003; Nicholls & Stukas, 2011). Documenting that a researcher is being out performed is detrimental to his or her ‘self definition’ and is more extreme when the comparisons are with colleagues rather than strangers.
- In the event of discrepancy between the researcher’s expected and actual indicator values, the individual will be more susceptible to influence and interpreting the values will become unstable due to lack of ‘self confidence’ (Misra, 1973), especially when the role of the indicators for the evaluation is unclear.
- Researchers are not the best able to document their performance as they interpret indicator values as what is significant to them but perhaps not significant to the evaluator (Misra, 1973; KAOW/ARSOM, 2012). Individuals will seek and utilize whatever bibliographic information is available that will increase the values of their indicators and thus increase their subjective validity and ‘self worth’. Adding indicator values to CVs can contribute with too much information that can become meaningless, especially considering the many caveats of ready to use indicators. Too much information can be interpreted as lack of ability to ‘self edit’ or lack of confidence.

Conclusions

Ready to use bibliometrics present partial measures of a researcher’s contribution to scientific knowledge however the values of these bibliometric indicators can be used to inform tenure decisions, distribute funds, discriminate between researchers and document activities. By linking these uses to empirical and conceptual personality traits commonly appraised in evaluations of work satisfaction and career success, we began to understand the role indicators can play on 1) ‘self esteem’ (seeing oneself as successful and worthy), 2) self efficacy’ (trust in ones capability to perform in many contexts and believing in one’s ability to control one’s environment), and 3) career success.

References

Bach, J. (2011). *On The Proper Use of Bibliometrics to Evaluate Individual Researchers*. Rapport de l’Académie des sciences - 17 janvier 2011: Report presented on 17 January 2011 to the Minister of Higher Education and Research.

Bornmann, L., & Marx, W. (15th. February 2013). Standards for the application of bibliometrics in the evaluation of individual researchers working in the natural sciences. Retrieved 25th. February 2013 from ArXiv.org: http://arxiv.org/abs/1302.3697v1

Cai, H., Brown, J.D., Deng, C., & Oakes, M.A. (2007) Self-esteem and culture: Differences in cognitive self-evaluations or affective self-regard? *Asian Journal of Social Psychology*. 10, pp. 162-170.

Cheng, C.H.K., & Watkins, D. (2000). Age and gender invariance of self-concept factor structure: An investigation of a newly developed Chinese self-concept instrument. *International Journal of Psychology*, 35 (5), pp.186-193.

Cheung, W. (2008). The Economics of Post-Doc Publishing. Ethics in Science and Environmental Politics, 8, pp. 41-44.

Crocker, J., Luhtanen, R.K., Cooper, M.L., & Bouvrette, S. (2003) Contingencies of self-worth in college students: Theory and measurement. *Journal of Personality and Social Psychology*, 85, pp.894-908

Dahler-Larsen, P. (2011) *The Evaluation* Society. California: Stanford University Press

Judge, T.A., & Hurst, C. (2007) Capitalizing on one's advantages: Role of core self-evaluations. *Journal of Applied Psychology*, 92, pp.1212-1227

KAOW/ARSOM (2012) Conditions for a Good Evaluation. *The Guidelines Project: Providing Instruments for* Evaulating Development *Research* Retreived 15th Febraury 2014 from: http://www. guidelines.kaowarsom.be/en

Koenig, A.M., Eagly, A.H., Mitchell, A.A. & Ristikari, T. (2011) Are Leader Stereotypes Masculine? A Meta-analysis of Three Research Paradigms. *Psychological Bulletin*, 137(4), pp.616-642

Kurman, J. (2002). Measured cross-cultural differences in self-enhancement and the sensitivity of the self-enhancement measure to the modesty response. *Cross Cultural Research*, 36 (1), pp.73-95.

Martin, B.R., & Irvine, J (1983) Assessing Basic Research: Some partial indicators of scientific progress in radio astronomy Research *Policy* 12, pp.61-90.

Misra, S. (1973) Instability in self-evaluation, conformity and affiliation. *Journal of Psychology*, 41(3), pp.361-375

Nicholls, E. & Stukas, A.A. (2011) Narcissism and the Self-Evaluation Maintenance Model: Effects of Social Comparison Threats on Relationship Closeness. *The Journal of Social Psychology*, 151 (2), pp.201-212

RAISE. (2013). Recognizing the acheivements of Women in Science, technology, eningeering, maths and medicine, at http://www.raiseproject.org/index.php [accessed 10 June 2013]

Stumpp, T., Muck, P.M., Hülsheger, U.R., Judge, T.A., & Maier, G.W. (2010) Core self-evaluations in Germany: Validation of a German Measure and its relationships with career success. *Applied Psychology: An International Review*, 59 (4), pp.674-700.

Takata, T. (2003) Self-enhancement and self-criticism in Japanese culture. An experimental analysis. *Journal of Cross-Cultural* Psychology, 34 (5), pp.542-551.

Tyser, M.P., McCrea, S.M., & Knüpfer, K. (2012). Pursuing perfection or pursuing protection? Self-evaluation motives moderate the behavioral consequences of counterfactual thoughts. *European Journal of Social Psychology*, 42, pp.372-382.

Vallacher, R.R., Nowak, A., Froehlich, M., & Rockloff, M. (2002) The dynamics of selfevaluation. *Personality and* Psychology *Review*, 6 (4), pp.370-379.

Yin, P., & Fan, X. (2003) Assessing the factor structure invariance of self-concept measurement across ethnic and gender groups: Findings from a national sample. *Educational and Psychological Measurement*, 63 (2), pp.296-318.

Appendix 1: Studies on the common psychological effects of individual evaluation

References	Perspective	Main Conclusions
(RAISE, 2013; Koenig 2011)	Gender differences	Women believe they are discriminated against in promotion as there are relatively few women employed in high level faculty positions and masculinity lessens for lower level positions.
BURG (2009)[1]	Gender differences	Evaluators rate CVs and journal articles lower on average for women than men
(Koenig 2011; Cai, 2007)	Gender stereotypes	'Communal' qualities, such as being nice or compassionate, are associated with women. Agentic qualities, being assertive or competitive, are associated with men. The results of being competitive or assertive are measurable, e.g. winning awards, initiating projects, where in contrast researchers are not awarded a grant or published because they are 'nice' or 'compassionate'.
(Cai et al, 2007)	Cultural differences	People from some countries evaluate themselves in an excessively less positive manner than others due to cultural differences in modesty, not self esteem
(Kurman, 2002; Takata, 2003)	Cultural differences	Cross cultural differences affect self enhancement
(Cheng, C.H.K., & Watkins, D, 2000; Yin & Fan, 2003)	Cultural differences	Which criteria should be used to account for variance in measures of 'self esteem' across academic lifespan and the effect of age, gender, ethnic groupings and variances in 'self esteem' on academic profiles
(Nicholls & Stukas, 2011)	Self worth	The pressure to publish means that researchers see their 'self worth' as contingent on publication success, which is easy to measure bibliometrically. Researchers can be tempted to 'self regulate' their publishing success or failures to maintain positive 'self views' of themselves. 'Ready to use' indicators do not set publication lists in context of the researcher's gender, seniority, specialty, affiliation and discipline.
(Crocker et al, 2003)	Social comparison	The performance of relevant others is used to inform social comparison and indicators can provide positive self enhancement possibilities. Documenting influence reduces uncertainty in the individual's abilities. The individual will choose not to report the results of the indicators if they are exposed as low achievers compared to their peers resulting in incomplete or missing indicators.
(Misra 1973)	Uncertainty	By providing relevant information uncertainty is reduced. Too much information can quickly become meaningless and hide relevant information
(Crocker et al 2003)	Motive	Motive is circumstantial and depends on how malleable the evaluation is. 'Self improvement' and 'self protection' arise in many situations and can come into conflict
(Tyser et al, 2012)	Motive	In 'self protection' the individual may ignore indicators that result in useful but negative feedback, whereas 'self improvement' would require attention to this information, even though it could be damaging to the researcher's 'self esteem'

[1] An overview of sources is too extensive to list. Please refer to, amongst others, the Boston University Recruitment Guidelines (BURG) and corresponding reference list, available at: http://www.bu.edu/apfd/recruitment/fsm/assumption_awareness/

Scaling Analysis of Author Level Bibliometric Indicators[1]

Lorna Wildgaard* and Birger Larsen**

*pnm664@iva.ku.dk
Royal School of Library and Information Science, Birketinget 6, 2300 Copenhagen (Denmark)

** birger@hum.aau.dk
Department of Communication and Psychology, Aalborg University Copenhagen, A. C. Meyers Vænge 15, 2450 Copenhagen SV (Denmark)

Introduction

Despite of the concerns from the bibliometric community, evaluation of the individual through bibliometric indices is already performed as a form of 'pseudo peer review' in selection of candidates for tenure, in background checks of potential employees' publication- and citation impact, and in appraisal of funding applications. As part of developing the ACUMEN portfolio we therefore undertook an extensive review of 114 bibliometric indicators in Wildgaard, Schneider and Larsen (2014) to identify 1) which author level indices are useful to document the effect of publication performance, 2) identify which scientific activities it is possible to measure and with which indices, 3) analyse the applicability of these indices by discussing the strengths and weakness of each one, and 4) identify if there is a need for any additional novel indicators to measures the performance of individuals. The review confirmed that there is no immediate need to develop new bibliometric indicators. There is a wealth of indicators to choose from, some used in practice and some theoretical only. There is however a need to understand the usefulness of existing indicators and which ones represent independent research activities of authors.

We have begun our investigation into how indicators complement each other, specifically if there is a redundancy among indicators, i.e. two or more indicators measure the same thing, and which indicators are the "best" choice in regards to four predefined disciplines. The main parameter we judge the usefulness of indicators is on their simplicity, understood as the simplicity of data collection and the simplicity of mathematical computation for each indicator (Wildgaard, Schneider & Larsen 2014). The present study is a further investigation into which effects of publishing and citing these simple indicators attempt to capture.

Data

The data is drawn from a set of 2,554 European researchers in four scientific disciplines, *Astronomy*, *Environmental Science*, *Philosophy* and *Public Health*, identified in an online survey of web-presence conducted by Wolverhampton University in 2011. In the survey, the respondents reported their academic discipline and seniority, and these are used to group the researchers in our study. We found 741/2,554 researchers had a curriculum vitae and a publication list on the web. We extracted their publications from the CVs/publication lists and searched the Thomsen Reuters Web of Science (WoS) to identify them. We identified 34,660 citable papers. Additional publication and citation information on articles and reviews in this

[1] This work was supported by the ACUMEN FP7 project. The work presented here is used in the development of Guidelines for Good Evaluation Practice. The ACUMEN collaboration aims at understanding how researchers are evaluated and the science system can be improved and enhanced, www.research-acumen.eu

data set was kindly provided by the Centre for Science and Technology Studies (CWTS) at Leiden University, the Netherlands from their custom version of the WoS. As the CWTS data does not contain data from the Conference Proceedings Citation Indexes we do not have additional data on 3,693 citable papers and these are excluded from the present analysis. Our final data set thus consists of 30,967 publications with additional citation information, Table 1. The table shows the mean and median number of publications and citations, mean number of citations per year and also the meanPage which is an indicator of the mean academic age of the researchers, measured as the number of years since the researcher's first publication registered in WoS. Confidence intervals (CI) are computed to contextualize these averages.

Methods

Bibliometric indicators were derived from a review of the literature (Wildgaard, Schneider & Larsen 2014).

The simplicity of data-collection and calculation of each indicator was assessed, and only indicators that we deemed practically feasible for individual researchers without special bibliometric expertise or access to special datasets are included in the present analysis. This results in 37 potentially useful indicators at the individual level. All these indicators are simple to calculate but in prioritizing simplicity our method may result in choosing coarse measures of performance. These indicators are supplemented by 17 more fine-grained field level performance indicators supplied by CWTS. For an overview, see the Appendix where the indicators are briefly presented.

The set of selected indicators is intended to capture the major output and effects of a researcher's published work, defined as: *publication output*, i.e. counting publications in various ways; *the effect of output* i.e. raw citation or fractionalised counts, as well as average citations of the entire portfolio; *impact over time*, e.g. with citations adjusted for length of academic career and field norms, and finally *citations to core or selected publications*.

Preliminary analyses

IBM SPSS version 19 was used for calculation of statistics.

Table 13. Sample of 741 researchers, distribution of publications and citations across disciplines and seniorities.

Publications						Citations			
Discipline	**Sample**	**Range**	**Median (CI)**	**Mean (CI)**	**MeanPage (CI)**	**Range**	**Median**	**Mean (CI)**	**MeanCPY**
	***Astronomy*, 192 researchers**								
PhD	15	2-36	7(5.0;14.2)	10.8(5.6;15.9)	4.8(3.9;5.7)	8-529	150(27.9;209.7)	149.4 (64;234.7)	36.8(12.8;60.7)
Post Doc	48	3-103	19.5(14;26.5)	26 (19.9;32.1)	8.8(7.9;9.6)	3-3177	201.5(140.4;479.4)	561.1(339,7;782.4)	61.4(36.9;85.8)
Assis Prof	26	10-142	39.5(30;65.9)	51 (37.3;64.8)	12.2(10.6;13.7)	69-4009	702 (432.2;1327.5)	1118,6 (675;1562.1)	84(58.5;109.4)
Assoc Prof	66	7-292	61.5(48.5;75.4)	77.7(63.2;92.2)	19.7(18.1;21.2)	19-9083	1214(783.6;1622.8)	1981.1(1477.8;2484.4)	107(79.9;134.0)
Professor	37	34-327	90(75.2;109.6)	121.3(92.8;149.8)	25.7(23.4;27.9)	177-16481	1889(1292.9;3245.3)	3579.1(2170.9;4988.2)	146(97.5;194.4)
	***Environmental Science*, 195 researchers**								
PhD,	3	3-5	4	4	9.6	16-60	34	36	5.6
Post Doc	17	2-59	9(6;12.9)	12.8(5.6;20)	6.8(4.5;9.0)	10-642	41(25;56)	91.7(11.1;172.2)	10.6(5.8;15.3)
Assis Prof	39	2-46	18(13.9;20)	19(15.6;22.5)	10.7(8.8;12.5)	0-573	148(90.6;167.6)	185.4(133.7;237.1)	16.7(12.5;20.8)
Assoc Prof	85	1-103	29(25;41)	36.8(31.7;42)	16.6(15.2;18)	2-2519	326(232.9;459.4)	520.1(404.4;635.7)	30.2(23.9;36.4)
Professor	51	1-425	51.5(39.3;64.2)	59.7(46.8;72.5)	24.1(21.8;26.3)	6-14141	435(324.5;722.6)	998.1(614.7;1381.5)	48.2(29.8;66.5)
	***Philosophy*, 222 researchers**								
PhD	8	1-5	1(1;4.1)	2(0.6;3.3)	3.5(2.3;4.6)	1-33	0.5(0;13.5)	6.2(-3.2;15.7)	1.7(-0.31;3-71)
Post Doc	22	1-31	4(3;8)	7(3.8;10.1)	6.2(4.8;7.5)	0-235	8(1-10)	21.4(-1.9;44.7)	15.4(2.0;28.7)
Assis Prof	44	1-106	6.5(4;8.9)	10.8(5.7;15.9)	7.6(6.3;8.9)	0-1829	6.5(3;20)	74.3(-11.5;160.2)	6.5(0.6;12.3)
Assoc Prof	73	1-45	7(6;9)	10(7.8;12.1)	11.2(9.6;12.7)	0-565	8(5;13)	50.7(22.7;78.7)	4.2(1.9;6.4)
Professor	75	1-140	18(13.5;23.4)	28.1(21;35.2)	19.6(17.6;21.5)	0-3495	29(20.5;65.6)	157(52.1;262)	7.0(2.6;11.3)
	***Public Health*, 132 researchers**								
PhD	9	4-27	8(7.1;17.8)	12.2(6.6;17.8)	5.6(3.7;7.4)	7-253	60(34.5;146.7)	82.2(23.5;140.8)	17.8(4.5;31.0)
Post Doc	14	1-23	11(8.8;14.4)	12(8.6;15.3)	7.2(4.9;9.4)	0-353	80.5(21.5;203.9)	113.6(49.4;177.6)	14.1(7.9;20.2)
Assis Prof	30	3-288	22(13.1;29.6)	36.2(15.6;56.7)	10.7(8.5;12.8)	10-3796	167(107.8;350.8)	417.4(131.4;703.3)	34.4(17.8;50.9)
Assoc Prof	50	4-221	43(30.6;56.3)	54.6(41.6;67.7)	16(14.2;18.5)	4-3649	518(312.6;701.7)	778.5(539.4;1017.5)	46.7(33.6;59.7)
Professor	29	5-661	76(53.6;107.6)	110.2(62.7;157.7)	17.4(14.7;20.0)	13-13520	954(554,2;2394.7)	2104(1065.3;3142.6)	109.8(62.1;157.4)

Predicting the usefulness of indicators at the seniority level

In order to investigate the usefulness of indicators for different levels of academic seniority we computed a cross-correlation matrix (per discipline) for the indicators using Kendall's tau rank correlation coefficient, and *gamma* as the symmetric measure of association. Across all four disciplines the association between seniority and the h-type indicators was minimal or none existent. This lack of association makes sense, as h-type indicators are dependent on citations and publications also having specific seniority level values, and clearly this is not the case as the range of publications and citations as well as the confidence intervals around the averages document, Table 1.

Identifying central and isolated indicators across disciplines

So far our analysis shows that publication and citation data between scholars within seniority is so varied that recommending any of our 52 sampled indicators as preferred "seniority level indicators" is unwise. We take the analysis up a level, from seniority to discipline, to investigate if the indicators are able to represent disciplinary traits. Inspired by Franceschet (2009) we begin by analysing if indicators display high correlations to other indicators, and identifying indicators that practically measure the same inherent properties. If indicators can be grouped by such an analysis into "clusters" of highly similar indicators, then the simpler alternatives from each cluster can be recommended over more complex ones.

Table 2 uses data from the correlation matrices to highlight central and isolated indicators. Isolated indicators are defined as having any only moderate or weak links, strength of association ≤0.7, to any of the other of the 51 indicators in the correlation. Central indicators are the indicators that have the highest number of links, over 0.7, to the other 51 indicators in the matrix (indicated in Table 2, column 4).

Table 14. Isolated and Central indicators across disciplines.

Discipline	Isolated Indicators	Central Indicators	Number of links to other indicators
Astronomy	App, sum sc, AWCR_pp, fp, %nc, average mjs mcs, min mjs mcs, maxs mjs mcs, average mnjs, h norm, wu	Hg IQP, AR	25 24
Environmental Science	Pyrs, App, %sc, Fp, nnc, %nc, Cage, AWCR_pp, PI, average mnjs, min mjs mcs, maxs mjs mcs, nproductivity adjusted papers, wu, AR	H, h2 popH, Q2, e, IQP	26 25
Philosophy	App, %sc, nnc, &nc, PI, sum pp top prop, average mjs mcs, max mjs mcs, average mnjs, nproductivity adjusted papers, hnorm, Wu	IQP AR, h2, Q2, e, g, h	28 27
Public Health	Pyrs, app, %sc, nnc, %nc, cage, AWCR_pp, minC, PI, min mjs mcs, average mnjs, nproductivity adjusted papers, hnorm, Wu	g Hg, ħ, h2	23 22

To investigate the role of the identified central and isolated indicators, we ranked researchers within disciplines and mapped how their position in the ranks changes when using these indicators as the control. We identified the top 10%, top 25%, middle 50% and bottom 25% in each set. We noticed that the isolated indicators produce a very random rank, placing a researcher sometimes in the top 10% and sometimes in the bottom 25%. The central indicators are all hybrid indicators. In *Astronomy* we used the **hg** index as the ranking factor, in *Environmental Science* the **h** index, in *Philosophy* the **IQP** index and in *Public Health* we used the **g** index. Across all disciplines we observed the same trend. If a researcher is placed in the top 10% of the sample by the central indicator, the researcher is placed in the top 10% using the other indicators that the central indicator has strong links to. Likewise, for researchers in the top 25%, middle 50% and bottom 25%.

To continue the analysis of how the central indicators gather other indicators around them we used the ALSCAL procedure in SPSS. This model allows us to visualize groupings of indicators as well as measure the distance between them. This is a good method of analysis of our skewed bibliometric dataset, as it accommodates interval and ratio scales, missing objects as well as symmetric and non-symmetric data. To get an idea of how well the model fits the data, we use the S-stress as a measure of fit ranging from 1 (worst possible fit) to 0 (perfect fit) and R-square to illustrate how much of the variance in the model is explained by these two dimensional models of Euclidean distance. The results present a low fit and high stress indicating that the maps are not very successful in capturing the complexity of higher dimensions and only coarsely group the indicators, Table 3 and Figures 1-4.

Table 3. MDS model fit

Discipline	Central Indicator	S-stress (R^2)	% variance explained (R^2)
Astronomy	hg	0.375	25
Environmental Science	H, h2	0.378	24
Philosophy	IQP	0.380	47
Public Health	g	0.499	38

Figures 1-4. Multidimensional Scaling maps of the studied bibliometric indicators in each of the four fields.

Fig.1. Astronomy

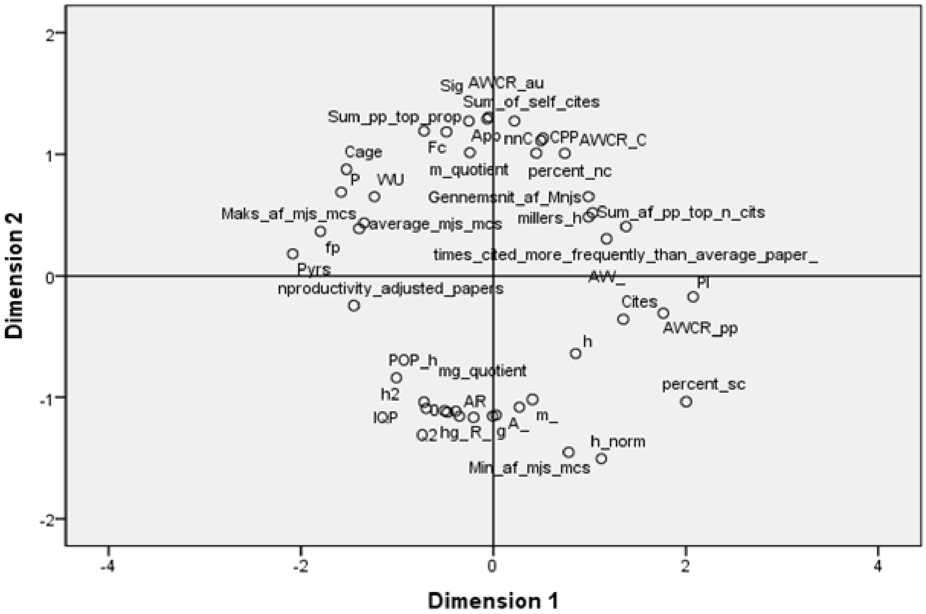


Fig. 2. Environmental Science

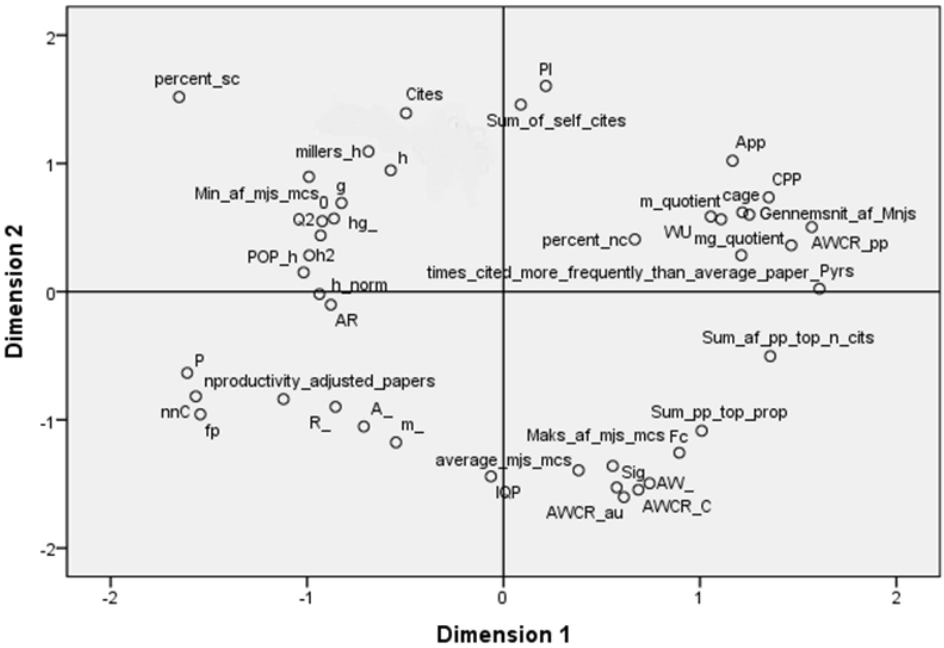


Fig. 3. Philosophy

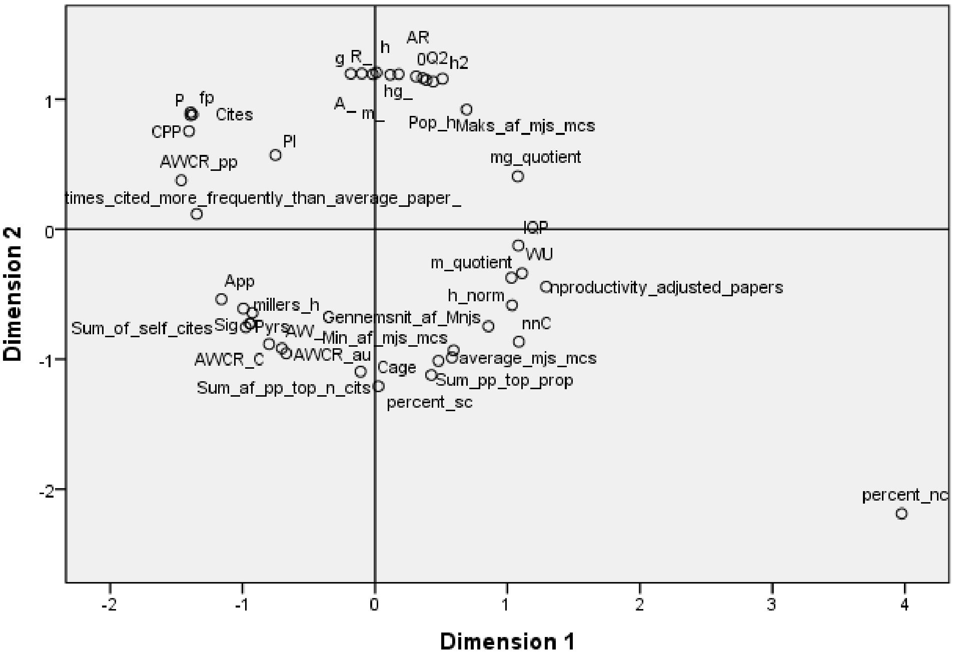
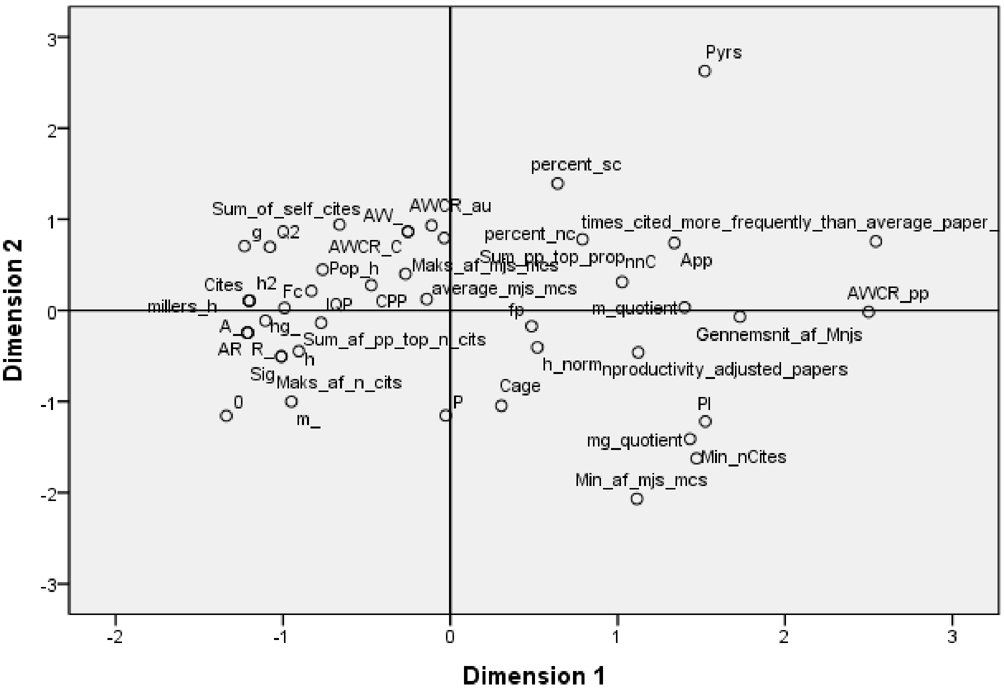


Fig. 4. Public Health



Next steps

The MDS maps show some overall structure, but the goodness of fit in the models is not high and needs improving. Across *Astronomy*, *Environmental Science and Philosophy* the indicators cluster in separate groups of hybrid, publication based or citation based (weighted or not weighted) indicators. In *Public Health* there are no clear groups. Depending on the indicators in each group, research may be appropriately evaluated in a more nuanced way, and it is therefore interesting to continue this study. We plan to supplement the maps with a hierarchical clustering analysis, resulting e.g. in a dendrogram, that will allow us to trace backward or forward to any individual indicator or cluster at any level. In addition, this may give an idea of how great the distance is between indicators or groups that are clustered in a particular step. This will help us understand which aspects of the effect of a researchers' production the central and isolated indicators capture as well as the strength of the role of the indicator. Particularly 1) if the isolated indicators indicate activities not covered by the central indicators, and 2) if the overlap between the central indicators and the indicators they link to means they measure the same thing.

References

Wildgaard, L. Schneider, J.W & Larsen, B (2014) *Bibliometric Self-Evaluation: A review of the characteristics of 114 indicators of individual performance*. Manuscript submitted for publication and under revision

Franceschet, M (2009) A cluster analysis of scholar and journal bibliometric indicators. *Journal of the American* Society *for Information Science and Technology* 60(10). Pp.1950-1964

Appendix: Indicators of individual impact as well as discipline benchmarks analysed in this study.

ID	Type	Abbr.	Indicator	Intention
Productivity metrics				
1	Publication	**P**	Publication count	Total count of production used in formal communication. Limited in our dataset to ISI processed publications
2	Publication	**Fp**	Fractionalized publication count	Each of the authors receive a score equal to 1/n to give less weight to collaborative works
3	Publication	**App**	Average papers per author	Average number of authors per paper over all publications
4	Publication/time	**Pyrs**	Years since first publication	Length of publication career from 1st article in dataset to 2013
Impact metrics				
5	Citation	**C**	Citation count	Use of all publications
6	Citation	**C-sc**	Citation count minus self-citations.	Use of publications, minus self-use.
7	Citation	**Sig**	Highest cited paper	Most significant paper
8	Citation	**minC**	Minimum citations	Minimum number of citations
9	Citation	**%sc**	Percent self-citations	Disambiguate self-citations from external citations
10	Citation/author	**Fc**	Fractional citation count	Remove dependence of co-authorship, all authors receive equal share of citations.
11	Citation/time	**C<5**	Citations less than 5 years old	Age of citations
Hybrid metrics				
12	Citation/publication/field	**IQP**	Index of Quality & Productivity	Number of citations a scholar's work would receive if it is of average quality in the field
13	Citation/publication/field	**Tc>a**	(part of IQP)	Actual times scholar's core papers are cited more than average quality of field
14	Citation/publication/field	**H norm**	Normalized h	Normalizes h-index (to compare scientists across fields).
15	Citation/publication	**Cage**	Age of citation	If citations are due to recent or past articles
16	Citation/publication	**%PNC**	Percent not cited	If citations are due to a few or many articles
17	Citation/publication	**CPP**	Citations per paper	Average citations per paper
18	Citation/publication	**h**	h index	Cumulative achievement
19	Citation/publication	**g**	g index	Distinction between and order of scientists
20	Citation/publication	**m**	m index	Median citations to publications included in h to reduce impact of highly cited papers
21	Citation/publication	**e**	e index	Supplements h, by calculating impact of articles with excess h citations
22	Citation/publication	**w**	wu index	Impact of researcher's most excellent papers
23	Citation/publication	**hg**	Hg index	Balanced view of production by keeping advantages of h and g, and minimizing their disadvantages
24	Citation/publication	$\mathbf{H^2}$	Kosmulski index	Weights most productive papers
25	Citation/publication	**A**	A index	Magnitude of researcher's citations to publications
26	Citation/publication	**R**	R index	Improvement of A-index
27	Citation/publication	**AR**	AR-index	Citation intensity and age of articles in the h core
28	Citation/publication	**ħ**	Miller's h	Overall structure of citations to papers
29	Citation/publication	$\mathbf{Q^2}$	Quantitative & Quality index	Relates the number of papers and their impact
30	Citation/publication/author	**hi**	individual h	Number of papers with at least h citations if researcher had worked alone
31	Citation/publication/author	**POP h**	Harzing's publish or perish h index	Accounts for co-authorship effects
32	Citation/publication/author/time	**AWCR**	age weighted citation rate	Number of citations to all publications adjusted for age of each paper
33	Citation/publication/author/time	**AW**	Age weighted h	Square root of AWCR to avoid punishing researcher's with few very highly cited papers. Approximates h index
34	Citation/publication/author/time	**AWCRpa**	Per-author AWCR	Number of citations to all publications adjusted for age of each paper and number of authors
35	Citation/publication /time	**M quotient**	m-quotient	Age weighted h. H divided by years since first publication

36	Citation/publication/time	**Mg**	Mg-quotient	Age weighted g. G divided by years since first publication
37	Citation/publication/time	**PI**	Price Index	Percentage references to documents not older than 5 years at the time of publication of the citing sources
Journal-field benchmarks, calculated by CWTS				
38		**mcs**	Mean citation score	Average citation score
39		**mncs**	Mean normalized citation score.	Shows relation to world average in regards to document type, publishing year and field.
40		**pp top n cites**	Proportion of top papers	Proportion of papers that have received more than 10 citations
41		**pp top prop**	Proportion in top 10% of world	If the article is cited in the top 10% of its field
42		**pp uncited**	Proportion uncited	Proportion uncited papers
43		**mjs mcs**	Crown-type indicator	Average number of citations of the journal the article is published in
44		**mnjs**	Mean normalized journal score	Performance of the journal the article is published in normalized to mncs
45		**mjs pp top n cits**	Crown-type indicator	Proportion of papers that have received more than 10 citations in the publishing journal
46		**mnjs pp top prop**	Crown-type indicator	Proportion of papers in the journal that are in the world pp top %
47		**mjs pp uncited**	Crown type indicator	Percent uncited on average in the publishing journal
48		**prop self cits**	Proportion self-citations	Self citations
49		**int coverage**	Internal coverage.	% cited references in the paper linking to WOS publications since 1980
50		**pp collaboration**	collaboration	Proportion collaboration outside of authors affiliated institution
51		**pp int collab**	International collaboration	Proportion international collaboration
52		**n self cites**	Number of self-citations	Count of self citations

A Preliminary Examination of the Relationship between Co-Publications with Industry and Technology Commercialization Output of Academics: The Case of the National University of Singapore

Poh Kam Wong *, Annette Singh **

* *pohkam@nus.edu.sg*
NUS Entrepreneurship Centre and NUS Business School, National University of Singapore, Level 5, 21 Heng Mui Keng Terrace, Singapore 119613

** *annette_singh@nus.edu.sg*
NUS Entrepreneurship Centre, National University of Singapore, Level 5, 21 Heng Mui Keng Terrace, Singapore 119613

Introduction

Amidst increasing research interest in university-industry collaboration (UIC), this paper examines the relationship between UIC as proxied by co-publications and technology commercialization at the level of the individual faculty member. Preliminary results of analysis on Engineering and Medical faculty at a university in Singapore suggest a positive relationship between authoring university-industry co-publications and patenting.

Conceptual Framework

Research interest in university-industry collaboration has grown, particularly in the light of the enhanced role that universities play in innovation and economic development as held by the Triple Helix thesis (Etzkowitz and Leydesdorff, 2000; Calvert and Patel, 2003). R&D collaboration between universities and industry has been identified as an important part of the "third mission" of universities due to the role that such collaboration can play as a source of innovation and in generating technological spillovers (D'Este and Patel, 2007; Ambos et al., 2008).

As noted by Beaudry and Kananian (2012), the entrepreneurial university generally begins with "…academic inventors that generate or contribute to patents" (p. 1). An understanding of academic inventors – who they are, and why they engage in technology commercialization - is therefore needed for a comprehensive understanding of the implications of the entrepreneurial university, and to facilitate the design of effective policies for developing the third mission of universities (Ambos et al., 2008; D'este and Perkmann, 2011). Yet, most studies on university technology commercialization are conducted at the level of the organization (Azoulay, Ding & Stuart, 2007). A smaller body of literature has identified characteristics influencing the commercialization output of individual researchers (eg Landry, Amara & Saïhi, 2007; Stephan et al., 2007; Azoulay, Ding & Stuart, 2007). Further, much of the literature in this area focuses on the US, particularly on the development of academic technology commercialization following the passage of the Bayh-Dole Act (Beaudry and Kananian, 2012). While other studies focus on universities in other regions, those in Asia are less prevalent. This study thus aims to contribute to the literature by adding empirical evidence from a university in Singapore.

The Impact of R&D Collaboration on University Technology Commercialization

R&D collaboration with industry can be expected to positively influence researchers' technology commercialization performance. Through interaction with industry researchers, university researchers gain experience and competencies that are conducive to engaging in technology commercialization (Ambos et al., 2008), as well as building a network of contacts which may facilitate inventing activity and later commercialization efforts (Beaudry and Kananian, 2012). Further, exposure to technological problems faced by firms may open research areas with commercial potential which academic researchers may not otherwise have encountered (D'Este and Patel, 2007). All this may thus induce academic researchers to an orientation towards research which can be commercialized (Audretsch and Aldridge, 2012).

More specifically, there are two pathways through which upstream university-industry research collaboration could result in downstream technology commercialization activities in the form of patenting, licensing and spin-offs (Wong and Singh, 2013):

(i) *direct* commercialization pathways, whereby the industry collaboration creates patented knowledge that is directly commercialized by the industry partners themselves through joint patent ownership or licensing;

(ii) *indirect* commercialization pathways, whereby the collaboration generates university-owned patents that are non-exclusive and could be licensed to other industrial firms not involved in the original collaboration; in addition, there could be wider knowledge spillover effects, whereby university researchers gain knowledge of industry needs and technology development directions through their industry collaboration, and use that knowledge to independently create new patented knowledge that is then commercialized by licensing to other industry firms, or through spin-offs by the university researchers themselves or their students.

Co-publications as a Measure of University-Industry R&D Collaboration

Empirically, a number of studies have examined the influence of industry linkages on the technology commercialization of academic researchers (eg Landry, Amara & Saïhi, 2007; Beaudry and Kananian, 2012; Ambos et al., 2008), and most of these find such linkages to have a positive effect. However, the indicators of linkages used (funding received by industry and number of research agreements) do not differentiate between the different types of university-industry linkages (eg consulting, contract research and joint research). This may mask differences that arise due to the variety of university-industry interaction involved. Linkages in which a firm merely contributes finances or equipment to a university produces relatively little interactivity between the partners. Collaborative research on the other hand, involves a much higher degree of interactivity, and can be expected to have stronger effects on academic entrepreneurship (Ambos et al. 2008; Audretsch and Aldridge 2012).

Empirically, one proxy measure of university-industry R&D collaboration is through the use of university-industry co-publications (UICPs) - publications that have been co-authored by university and industry researchers. Co-publications can be seen as an output of research cooperation where some diffusion of knowledge and skills has taken place, and give access to even informal networks between university and industry researchers (Lundberg et al 2006, Tijssen 2006). Generally, the consensus is that despite their limitations, co-publications provide a reasonable proxy measure of successful university-industry research cooperation (Calvert and Patel, 2003; Tijssen, 2006).

Audretsch and Aldridge (2012) and Azoulay, Ding & Stuart (2007) have both employed UICPs in studies of university life science researchers, finding those with UICPs to have a higher likelihood to engage in patenting. Wong and Singh (2013) conducted an organizational-level study of 82 North American universities, finding UICPs to positively influence universities' technology commercialization output in terms of patenting (simple patent counts and quality-adjusted counts), spin-off formation, and technology licensing. In this present study we now turn to the level of the individual, and to a university in Asia. Including researchers from both the Engineering and Medicine faculties of the National University of Singapore (NUS), we examine the relationship between their technology commercialization activity and R&D collaboration as proxied by university-industry co-publications (UICPs).

As a preliminary, we begin by focusing on the initial stage of the technology commercialization process – academic patenting. Prior studies have used patenting as an operational measure for academic entrepreneurship (eg Audretsch and Aldridge, 2012), and in fact patents are the most frequently-used indicator to reflect entrepreneurial activities of university inventors (Landry and Amara, 2012). Further, whilst it is true that patenting does not necessarily lead to actual technology commercialization activity such as licensing or spin-off formation we would suggest that these latter outcomes depend heavily on the initial step of patenting (Ambos et al., 2008).

Our first hypothesis is thus as follows:

H1: University–industry co-publication (UICP) propensity is positively associated with commercialization output as measured by patenting.

The Moderating Effect of Faculty Rank on the Link between UICP and Technology Commercialization

The literature has identified seniority, or rank of faculty members amongst the various individual-level characteristics which influence technology commercialization. Rank reflects, among other things, the stage faculty are at in their career, and the experience they have accumulated over their years of research. One argument is that early in their careers, researchers are working to establish their academic reputations and obtain tenure, and may thus be less willing to become involved in activities that would take time away from their research. Conversely, at later stages in their career, faculty members have already established their reputations and so may be less motivated by such traditional academic incentives as promotions and tenure. For them, the potential financial gain from engaging in knowledge commercialization - especially with a view to generating revenues after retirement - may be a more compelling motivation (Stephan et al., 2007). Furthermore, experienced researchers may have acquired more intellectual capital with the potential to be commercialized than those who are less experienced (Geuna and Nesta, 2006). If this is true, technology commercialization can be expected to increase with seniority. However, we suggest that this may be moderated by other career effects. Some faculty members, having attained tenure, may not be able to achieve promotion to full professorship. Associate Professors who realize they are in this position may then divert more of their attention towards industrial research collaboration and technology commercialization as an alternative means of career development. In these cases, we may expect the positive effect of R&D collaboration on technology commercialization. Empirically, Landry, Amara & Saïhi (2007) found that Associate Professors in the life sciences do indeed exhibit higher patent activity relative to full professors, although they did not specifically test the additional impact of industry research collaboration on this group.

Our second hypothesis is then:

H2: The positive effect of UICP propensity on commercialization output is stronger for Associate Professors relative to other ranks.

Methodology and Data

Our study comprises faculty members from the two departments at the NUS which were found to have among the highest volumes of UICPs in the Schools of Medicine and Faculty of Engineering. Specifically, all faculty members of the Department of Pharmacology, along with half of the faculty members of the Department of Electrical & Computer Engineering (ECE), with the latter being randomly selected. This resulted in a sample of n=32 faculty members from Pharmacology and n=62 faculty members from ECE.

Dependent Variables

The dependent variable is the number of NUS-assigned USPTO patents issued over 2009-2012 on which the faculty members are listed as inventors.

Predictor Variable

For the UICP explanatory variable we drew upon a database of UICPs published between 2004-2011 and having at least one co-author from NUS. This was obtained from the Centre for Science and Technology Studies (CWTS) of Leiden University. The variable takes the form of a binary variable with the value of 1 if the faculty member produced a UICP during this period, and 0 otherwise. In order to gain some indication of the effect of different time-periods, this period was divided into two sub-periods: 2004-2007 and 2008-2011. 37.2% of the faculty members in our sample co-authored at least one UICP during 2004-2011. 24.5% co-authored at least one UICP during the earlier period and 21.3% co-authored at least one UICP during the later period.

Control Variables: Publication-Related Control Variables

Our analysis further included a number of control variables for relevant inventor characteristics as identified in the literature. Two of these are publications-related variables: publication quantity and quality (Beaudry and Kananian, 2012; Landry, Amara & Saïhi, 2007; Audretsch and Aldridge 2012). The number of publications over 2008-2011 for each faculty member was obtained from Thomson Reuter's *Web of Science*, along with the number of citations (up to the end of 2013) to these publications.

Control Variables: Other Control Variables

Non-publication-related control variables which have been identified in the literature include the department, rank and experience of the faculty members (see eg Landry, Amara & Saïhi, 2007; Ambos et al., 2008; Audretsch and Aldridge, 2012). These characteristics were included into our study:

i) 'Department' is a binary coded variable taking the value of 1 for pharmacology faculty and 0 for ECE
ii) The rank of each faculty member is incorporated through a series of dummy variables: 'Assistant Professor' was coded 1 if the faculty member is an Assistant Professor and coded 0 otherwise; similarly 'Assoc Professor' was coded 1 if the faculty member is an Associate Professor (and 0 otherwise). 'Prof' was coded 1 for Professors and 0 otherwise, and this is taken as the reference category. 'Other' was coded 1 if the faculty members did not fall into any of the aforementioned categories, and 0 otherwise.
iii) 'Experience' captures the level of experience gained by the faculty member. Following Landry, Amara & Saïhi (2007) and Stephan et al. (2007), this variable is operationalized as the number of years since completion of PhD (counted up to 2012).

The following model is estimated using multiple regression analysis for H1:

$$\text{Number of patents} = \beta_0 + \beta_1\,(\text{With UICP in 2004-2007}) + \beta_2\,(\text{With UICP in 2008-2011}) + \sum \beta_j(\text{control variables}) + \varepsilon$$

The following model is estimated using multiple regression analysis for H2:

$$\text{Number of patents} = \beta_0 + \beta_1\,(\text{With UICP in 2004-2007}) + \beta_2\,(\text{With UICP in 2008-2011}) + \sum \beta_3\text{Assoc Professor* (With UICP in 2008-2011)} + \sum \beta_j(\text{control variables}) + \varepsilon$$

Results

Descriptive statistics for the sample can be seen in **Table 1** and the Pearson correlations in **Table 2.** As can be seen from **Table 2**, the correlation between the incidence of a faculty member co-authoring a UICP during 2004-2007 is only weakly correlated with the incidence of a faculty member co-authoring a UICP during 2008-2011 ($r=.188$, $p<0.1$). Further, the number of active invention disclosures is positively correlated with the incidence of co-authoring a UICP during the second period ($r=.451$, $p<.01$) but not during the first period. A similar result can be observed for the correlation of patents and the incidence of co-authoring a UICP ($r=.310$, $p<.01$ for incidence of UICP co-authorship over 2008-2011).

Table 1. Summary Statistics

	Total sample					ECE					Pharmacology				
	n	Min	Max	Mean	Std. Dev.	n	Min	Max	Mean	Std. Dev.	n	Min	Max	Mean	Std. Dev.
Continuous variables															
No. of patents 2009-12	94	0.0	4.0	.16	.57	62	0.0	4.0	.19	.65	32	0.0	2.0	.09	.39
Experience	94	0.0	42.0	16.6	10.0	62	0.0	40.0	15.82	10.04	32	4.0	42.0	18.13	9.98
No. publications 2008-11	94	0.0	155.0	17.0	22.77	62	0.0	155.0	18.47	26.00	32	0.0	62.0	14.19	14.60
Average citations/publication	94	0.0	32.5	7.5	7.49	62	0.0	22.1	4.91	5.17	32	0.0	32.5	12.59	8.68
Dichotomous variables															
Rank	• Professor: 23.4% • Assoc Professor: 43.6% • Assistant Professor: 19.1% • Other: 13.8%					• Professor: 25.8% • Assoc Professor: 40.3% • Assistant Professor: 27.4% • Other: 6.5%					• Professor: 18.8% • Assoc Professor: 50.0% • Assistant Professor: 3.1% • Other: 28.1%				
With UICP in 2004-2007	24.5%					24.2%					25.0%				
With UICP in 2008-2011	21.3%					19.4%					25.0%				

Table 2. Pearson Correlations

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
No. of Patents 2009-12 (1)	1									
Department (2)	-.083	1								
Assoc Professor (3)	-.168	-.092	1							
Assistant Professor (4)	-.089	-.293**	-.428**	1						
Other (5)	-.112	.297**	-.352**	-0.195†	1					
Experience (6)	.213*	.109	.147	-.539**	-0.182†	1				
No. publications 2008-11 (7)	.054	-.090	.038	-.285**	-.140	.112	1			
Average citations/publication (8)	.054	.488**	.088	-.326**	0.190†	0.173†	0.186†	1		
With UICP in 2004-2007 (9)	.144	.009	.048	-.214*	-.085	.285**	.315**	.204*	1	
With UICP in 2008-2011 (10)	.310**	.065	.014	-.121	.018	.130	.232*	.350**	0.188†	1

** Significant at the 1% level * Significant at the 5% level † Significant at the 10% level

Hypothesis 1
A Mann-Whitney U test[1] revealed the number of patents for UICP co-authors (i.e. those who have published at least one UICP between 2004-2011) to be significantly different from the number patents by faculty members who are not UICP co-authors (U=812.00, p=0.001).

The patterns observed in the correlations are corroborated by the results of the regression analysis, as reported in **Table 3**. Incidence of UICP co-authorship during 2008-2011 is positively associated with patenting (b=.42, p < 0.01). However, being a UICP co-author in the earlier period yields no significant effect. In terms of the control variables, the rank of the faculty member has a significant effect, with Associate Professors being more likely to have patents than Professors (b=.27, p<0.1). Faculty member experience has a significantly positive effect on patenting (b=0.01, p<0.1)

Table 3. Regression of no. of Patents 2009-12

	Model 1	Model 2
	B	B
Constant	-0.292 (0.266)	-0.146 (0.249)
Department	-0.14 (0.146)	-0.244 † (0.138)
Assoc Professor	0.274 † (0.162)	-0.009 (0.167)
Assistant Professor	0.234 (0.258)	0.13 (0.241)
Other	0.141 (0.24)	0.077 (0.222)
Experience	0.014 † (0.008)	0.014 † (0.007)
No. publications 2008-11	0 (0.003)	-0.003 (0.003)
Average citations/publication	-0.002 (0.01)	0.01 (0.01)
With UICP in 2004-2007	0.081 (0.145)	0.046 (0.135)
With UICP in 2008-2011	0.424** (0.151)	-0.103 (0.195)
Assoc Professor * With UICP in 2008-2011		1.076** (0.277)
Adj R^2	.090	.220
F	2.023*	3.631**
N	94	94

Standard errors in brackets
** Significant at the 1% level * Significant at the 5% level † Significant at the 10% level

Hypothesis 2
The introduction of the interaction term in Model 2 reveals that Associate Professors with UICPs do have a significantly higher number of patents (b=1.08, p<.001) (**Table 3**). Amongst the control variables the positive effect of experience on patenting remains significant (0.01, p<.1). In addition, the department is also found to influence patenting, with pharmacology academics having fewer patents relative to those in ECE (b=.24, p<.1).

Discussion

Our preliminary results support our hypothesis: R&D collaboration between university and industry, as measured by university-industry co-publications, is positively associated with active invention disclosure and patenting output of faculty members. This may provide some support for the argument that collaboration with industry facilitates commercialization activity of university researchers by granting them access to resources, research ideas and networks.

[1] ANOVA could not be used due to the unequal sample sizes and lack of homogeneity of variances in the data

However the regressions, together with the correlations, fail to point to UICPs published over 2004-2007 having a significant effect on the initial stages of technology commercialization for academics. This lack may be due to our relatively small sample size. Indeed, some evidence of the direct effect of UICPs published in the earlier period on patenting can be found amongst our database. US patent no. 7504329 was issued in 2009 to NUS as a co-assignee along with Interuniversitair Microelektronica Centrum (IMEC) and Texas Instruments. The inventors of the patented technology include Prof Kwong Dim-Lee in the ECE department. Amongst the non-patent references in the patent is a citation to a UICP co-authored by Prof Kwong and researchers from Texas Instruments and the University of Texas (Liu et al. 2005). This may be an example of direct commercialization, in which collaborative research between NUS and Texas Instruments resulted in the creation of patented knowledge as a first step to commercialization.

The results also provide some evidence in support of our second hypothesis – Associate Professors who are also UICP co-authors are found to have higher levels of patenting activity. Some corroboration of this can be found in the fact that 37.5% of Associate Professors who are also UICP co-authors have patents, whereas this is only true of 15.8% of UICP co-authors from other ranks. Further, the average PhD age of Associate Professors who are also UICP co-authors with patents is 22.5 years, vs. 17.0 for Associate Professors without UICPs. Although this evidence remains preliminary, it is suggestive that these are older Associate Professors having less expectation of attaining full professorship and so are engaging more heavily in industrial research collaboration and technology commercialization as an alternative means to career development.

In view of the promising findings from this preliminary study's relatively small sample covering only two departments, future research will expand the coverage of patents over a longer period, and cover faculties in the university. This may be especially needed due to the relative recentness of university technology commercialization in Singapore in general and NUS in particular. As such, there is a smaller volume of technology commercialization output in the form of patents, licenses and spin-off companies; none of the faculty members in our sample had licensed technologies over 2009-12 and only one had founded a spin-off company during this period. This further highlights the need to supplement analysis with case studies tracing individual UICPs that were subsequently cited in patents or led to licenses or spin-offs.

References

Ambos, T. C., Mäkelä, K., Birkinshaw, J. & D'Este, P. (2008). When does university research get commercialized? Creating ambidexterity in research institutions. *Journal of Management Studies*, 45(8), 1424-1447.

Audretsch, D. & Aldridge, T. (2012). Transnational social capital and scientist entrepreneurship. *Journal of Management &* Governance, 16(3), 369-376.

Azoulay, P., Ding, W. & Stuart, T. (2007). The determinants of faculty patenting behavior: Demographics or opportunities?. *Journal of* Economic *Behavior & Organization*, 63(4), 599-623.

Beaudry, C. & Kananian, R. (2012). Impact of university-industry contracts resulting in patents on the quality of patenting in biotechnology. Presented at *DRUID 2012 Conference*, Copenhagen, Denmark, June 2012.

Calvert, J., & Patel, P. (2003). University-industry research collaborations in the UK: Bibliometric trends. *Science and Public Policy*, 30(2), 85-96.

D'Este, P., & Patel, P. (2007). University-industry linkages in the UK: What are the factors underlying the variety of interactions with industry? *Research Policy*, 36, 1295-1313.

D'Este, P. & Perkmann, M. (2011). Why do academics engage with industry? The entrepreneurial university and individual motivations. *The Journal of Technology Transfer*, 36(3), 316-339.

Etzkowitz, H., & Leydesdorff, L. (2000). The dynamics of innovation: from National Systems and "Mode 2" to a Triple Helix of university–industry–government relations. *Research policy*, 29(2), 109-123.

Geuna, A. and Nesta, L. (2006). University patenting and its effects on academic research: The emerging European evidence. Research *Policy*, 35(6), 790–807.

Landry, R., Amara, N., & Saïhi, M. (2007). Patenting and spin-off creation by Canadian researchers in engineering and life sciences. *The Journal of Technology Transfer*, *32*(3), 217-249.

Landry, R., & Amara, N. (2012). Why and how do academics bridge the gap between invention and innovation? *International* Journal *of Technology Management*, *58*(3), 174-212.

Liu, J., Wen, H.C., Lu, J.P., & Kwong, DL. (2005). Dual-work-function metal gatesby full silicidation of poly-Si with Co-NiBi-layers. IEEE Electron Device Letters, 26(4), 228-230

Lundberg, J., Tomson, G., Lundkvist, I., Skår, J., & Brommerls, M. (2006). Collaboration uncovered: Exploring the adequacy of measuring university-industry collaboration through co-authorship and funding. *Scientometrics*, 69(3), 575-589.

Stephan, P. E., Gurmu, S., Sumell, A. J., & Black, G. (2007). Who's patenting in the university? Evidence from the survey of doctorate recipients. *Economics of Innovation and New Technology.*, *16*(2), 71-99.

Tijssen, R.J.W. (2006). Universities and industrially relevant science: Towards measurement models and indicators of entrepreneurial orientation. *Research Policy*, 35, 1569-1585.

Wong, P. K., & Singh, A. (2013). Do co-publications with industry lead to higher levels of university technology commercialization activity? *Scientometrics*, *97*(2), 245-265.

University-Industry dual appointments: global trends and their role in the interaction with industry

Alfredo Yegros-Yegros, Robert Tijssen

{a.yegros;tijssen}@cwts.leidenuniv.nl
Centre for Science and Technology Studies (CWTS), Leiden University,
PO Box 905, 2300 AX Leiden (the Netherlands)

Introduction

Mobility of R&D staff is considered one of the multiple mechanisms of interaction and transfer of knowledge between universities and companies. A wide definition of mobility should consider a University-Industry combined position or dual appointments of researchers.
A recent report by the European Science Foundation (2003, p.5) indicates that combined positions, in general, "*will facilitate knowledge transfer between the institutions 'in person'. Combined, part-time researcher positions will allow mobility and direct knowledge transfer and cooperation and may link institutions, disciplines, countries and sectors (industry/academia/public)*".
Indeed, one could argue that compared to other channels of knowledge transfer, University-Industry dual appointments' (U-I DAs) might facilitate the transmission of knowledge as it is embedded in research staff, thus contributing to both a more effective communication process and also the transformation of such knowledge into new products, processes, etc. For these reasons, from a policy perspective combined positions, including U-I DAs, are seen as a key element to foster innovation.
On the other hand, according to the OECD, these 'personnel exchange' or 'inter-sectoral mobility' channels, and their knowledge flows, might be not so effective as are characterized by a medium degree of 'relational intensity', i.e. the degree of interaction between knowledge senders and receivers, a low significance for industry, a high degree of formalization (i.e. contract-based), and low levels of finalization of end products (OECD, 2012; 2013).
There seems to be then a conflicting view on U-I DAs when comparing the policy discourse and how these combined positions are perceived by the private business companies. Also, when reviewing the literature, we observed that U-I DAs have received negligible attention in the international research literature on knowledge flows, research mobility and University-Industry relationship (Yegros-Yegros and Tijssen, 2014), therefore with this study we aim at contributing to the study of U-I DAs and the specific role they play in University-Industry relationships.
We will base our approach on those on scientific publications in which a given author indicate simultaneously at least two affiliations (one referring to a University and another to a private company). This analytical approach was first mentioned by Tijssen (2012) as a possible add-on to the analysis of *university-industry co-authored research publications* (UICs). UICs are one of the very few sources for gathering aggregate-level proxy measures of university-industry research cooperation and interaction patterns and trends (Tijssen et al., 2009; Tijssen, 2011).
Different situations might lead to the existence of U-I DAs: (a) interchange of students;
(b) a PhD graduate or university finds employment at a private company; (c) researchers from companies who move to a university; (d) the person holds a (semi-)permanent combined position at the university and a company. A given researcher in any of these cases would in principle include at least two different affiliations in the scientific publications, which could

be interpreted as an existing interaction between the two institutional sectors involved (i.e. universities and companies).
The different situations producing U-I DAs can be seen as modalities of 'knowledge flows pathways', which represent different degrees of relational intensity across the two sectors involved. The strongest and with most potential to be mutually beneficial linkage would be represented by researchers with (semi-)permanent combined University-Industry positions as would represent a bi-directional flow of knowledge. However, at the same time researchers holding these more relatively stable dual appointments could also be more likely to be exposed to situations of conflict due to the different goals of universities and companies. Other types of moves by researchers (i.e. from academia to industry and *vice versa)* would represent a directed transfer of knowledge embodied in the mobile researcher.
The objective of this empirical study is twofold. First, contextualize these dual appointments within a broader classification system based on institutional affiliations. Second, determine whether or not those who are University-Industry mobile researchers (i.e. (1) semi(permanent) combined University-Industry positions; (2) moves from academia to industry; and (3) moves from industry to academia) produce more University-Industry co-publications compared to other researchers in the same country and research area(s). This would confirm - or refute - their alleged role as transfer agents, and in some cases perhaps even 'gatekeepers', that populate one of the major knowledge transfer channels between industry and academia.
Our analysis will focus on the four European countries with a higher number of university-industry co-publications during the period 2008-2012 (United Kingdom, Germany, France and Netherlands).

Data and methodology

From 2008 onwards the Thomson Reuters *Web of Science* database (WoS) contains all institutional affiliations in the bibliographical records linked to the authors. Thanks to this bibliographic feature of the database it is possible to know, for each and every author, to which specific institution or institutions they are affiliated.
The classification of organizations into main institutional sectors (e.g. university, public research institutes, business enterprise, among others) enables us to detect publications in which at least one of the authors is affiliated simultaneously to a university and to one private sector organization. These cases are classified as *university-industry dual appointments* (U-I DAs).
Also, an in-house algorithm developed at CWTS (Caron and van Eck, 2014) allows us to collect the entire WoS-indexed research publication oeuvre of a given researcher, so for each given author with U-I DA we can trace other affiliations – both former and subsequent affiliations, in the case of extensive time-periods - and thus classify that particular author according to the following person-embodied time-bounded 'step-stage' categories: e.g. (1) semi(permanent) combined University-Industry positions; (2) moves from academia to industry; and (3) moves from industry to academia. Over time, a person can shift from one stage into another. The first three are 'transactional' categories representing 'knowledge flow pathways'.
Table 1 illustrates different situations when classifying authors with U-I DAs. Author1 would represent a clear example of a researcher that potentially holds a semi(permanent) combined U-I position while Author2 would represent a move from university to industry or *vice versa*. Authors 2 and 3 would be more problematic as they might belong to the same category as Author2 but it might also be the case that they held semi(permanent) combined positions in the past (Author3) or that they recently started this type of more stable dual appointment (Author4).

Table 1 Publications indicating a U-I DA

	Year_1	Year_2	Year_3	Year_4	Year_5
Author1					
Author2					
Author3					
Author4					

Source: Thomson Reuters/CWTS Web of Science database.

Findings

The findings we present at the conference are exploratory. They will cover the three classes of 'transactional' U-I DA personnel, focusing on their frequency of occurrence in scientific publications.

Figure 1 shows country-level statistics on UICs and U-I DA for the top European countries in terms of the UIC volume in the period 2008-2012.

Our results indicate that in the most outstanding countries, R&D staff with University-Industry dual appointments participates in around 20% of the UICs.

Table 2. UIC and U-I DA per country (2008-2012)

	# UICs	# U-I DA	% U-I DA
France	9,863	1,318	13.3
Germany	20,657	2,710	13.1
Netherlands	20,855	2,617	12.5
United Kingdom	9,826	2,345	23.8

Source: Thomson Reuters/CWTS Web of Science database.

Conclusions

We conclude that these double appointments, as identified within the author affiliate addresses of scientific publications, offer a new and interesting perspective on the propensity for University-Industry interactions within countries, fields, sectors, or organizations. This study represents an effort on validation of U-I DAs in order to disentangle microlevel types of relationships involved, and assess their contribution to comparative empirical studies of knowledge transfer channels, modes and processes.

References

Boyle, P. (2013). Policy: A single market for European research. *Nature*, 501(7466): 157-158.

Brennenraedts, R., Bekkers, R. and Verspagen, B. (2006). The different channels of university industry knowledge transfer: empirical evidence from biomedical engineering. Ecis.

Caron, E., van Eck, N.J. (2014). Large scale author name disambiguation using rule-based scoring and clustering. STI Conference, 3-5 September, Leiden, The Netherlands

European Science Foundation - ESF (2013). New concepts of research mobility: a comprehensive approach including combined/part-time positions.

Herrera, L, Munoz-Doyague, M.F., Nieto, M. (2010). Mobility of public researchers, scientific knowledge transfer and the firm's innovation process. *Journal of Business Research*, 63, 510-518.

OECD (2012). Knowledge Networks and Markets in the Life Sciences, OECD Publishing.

OECD (2013). Knowledge Networks and Markets, OECD Science, Technology and Industry Policy Papers, No. 7, OECD Publishing.

Tijssen, R.J.W., Van Leeuwen, T.N., van Wijk, E (2009). Benchmarking university-industry research cooperation worldwide: performance measurements and indicators based on co-authorship data for the world's largest universities. *Research Evaluation*, 18(1), 13-24.

Tijssen, R.J.W. (2012). Co-authored research publications and strategic analysis of public-private collaboration. *Research Evaluation*, 21(3), 204-215.

Yegros-Yegros, A., Tijssen, R.J.W. (2014). A global view of University-Industry interactions and dual appointments: quantitative indicators, empirical evidence and new insights. University-Industry Interaction Conference, 23-25 April, Barcelona.

Do University-Industry co-publication volumes correspond with university funding from business firms?

Alfredo Yegros-Yegros*, Joaquín M. Azagra-Caro**, Mayte López-Ferrer** and Robert J.W. Tijssen*

*{a.yegros;tijssen}@cwts.leidenuniv.nl
Centre for Science and Technology Studies (CWTS), Leiden University,
PO Box 905, 2300 AX Leiden (the Netherlands)

**{jazagra;mayte.lopez}@ingenio.upv.es
INGENIO (CSIC-UPV), Universitat Politècnica de València, Camino de Vera s/n, E-46022 Valencia, Spain

Introduction

It is nowadays almost taken for granted that interaction between universities and companies conduct to a variety of potential benefits, including the contribution to the economic growth. For this reason during the last decades this topic has attracted the attention of academics and increasingly also of policy makers, trying to understand the conditions in the University-Industry relationship leading to beneficial results in order to promote and boost this type of interactions.

An essential tool in the study of University-Industry interactions is the design of indicators able to capture accurately different aspects of this complex process. Scott et al. (2001) classified into four main categories the different channels through which university and industry interact: 1) codification/artefacts (e.g. publication or patents); 2) Cooperation (e.g. joint ventures or exchange of personnel); 3) Contacts (e.g. meetings or informal contacts) and 4) Contracts (e.g. licenses or contract research). This variety of channels suggest that to capture in a comprehensive way the interactions between universities and industries it would be necessary the use of an ample battery of indicators. However, the most problematic aspect in the measurement of University-Industry interactions is probably that the information required is not publicly available (for instance value of R&D contracts with industry, value of patents licensed to companies, etc.). Moreover, it is difficult to make comparisons across countries due to the lack of standardization in concepts and measures.

One indicator based on scientific publications in scholarly journals has been proposed to partially overcome these problems associated to the measurement of University-Industry interactions. More specifically, the indicator refers to those publications in which the author addresses include at least one university and one private sector organization: University-Industry co-publications (hereinafter UICs). These joint publications, compared to other indicators, are easily accessible and ensure to some extent comparability. They are one of the very few sources for gathering aggregate-level proxy measures of University-Industry research cooperation and interaction patterns and trends (Tijssen et al., 2009; Tijssen, 2011).

Some previous studies have already used UICs as proxy of University-Industry collaborations (e.g. Calvert and Patel, 2003). However, according to the literature it seems this particular indicator has not been extensively used for these purposes. As a matter of fact UICs sometimes are not even considered to be an important indicator of knowledge transfer in universities (e.g. Palomares-Montero and Garcia-Aracil, 2011).

The not so extensive use of UICs might be partially due to the fact that it is not clear the extent to which they represent actual interactions between universities and companies. Only a few studies have tried to seed light about this question. Lundberg et al. (2006) compared companies funding research conducted in Karolinska Institute with those co-publishing with the university and found that one third of the companies that had provided funding to the university did not co-publish any scientific paper with the university, concluding thus that UICs provided incomplete results on the actual collaborations between university and industry.

More encouraging are the results of a recent study conducted by Wong and Singh (2013), who found a significant positive influence on universities' technology commercialization outputs, including patenting, spin-off formation, and technology licensing, providing empirical evidence supporting the usefulness of UICs as a tool to measure the interactions between universities and companies.

Following a similar line of inquiry, the objective of this study is to provide new insights on the validity of UICs as indicator of the interactions between Universities and Industry by analyzing the relationship between UICs and a widely accepted indicator of these interactions: direct investment of private companies in university research.

Data and Methods

In our study we will focus on the Technical University of Valencia (UPV) to analyse the extent to which the amount of UICs correspond to the university funding coming from business firm.

The UPV is a Spanish public university founded in 1971. It is among the top three national universities in terms of Spanish issued patents and often the first in the EPO and PCT rankings. It is also representative of young European universities, characterized by their small size, technological research and less consolidated public funding, which made them prone to heavy dependence on industry. The UPV has engaged in increasing interaction activities through a relatively well-endowed industrial liaison office and a pioneering program to support the creation of spin-off companies. However, public funding has grown at a faster rate than private funding, as an internal policy response to keep up a certain standard of quality in research.

We collected two sets of data to develop this study. First, UICs published by the UPV in the period 2008-2011 were extracted from the Web of Science, including articles, reviews and letters. The second data set refers to external funding sources of the UPV and comes from the Centre for Innovation, Research and Technology Transfer (CTT), the technology transfer office of the UPV. It covers the period 2000-2013 and 7,110 funding agreements. They include project funding (collaborative and non-collaborative) and contract funding (research, development, technical support, professional works, etc.). The database contains fields on the geographic origin of funding (domestic, foreign) and institutional source of funding (public administration, company, etc.). The agreements involve over 1,700 principal investigators from UPV.

Thus in the first data set we have a=1,…,A number of UICs and each UIC with b=1,…,B number of UPV authors, i.e. a x b UIC-authors. In the second dataset (i.e. income from private companies) we have c=1,…,C number of funding agreements and d=1,…,D, principal investigators (PIs), i.e. c x d agreements-PIs. If we cross both data sets, we can create a matrix of a x b x c x d observations. Now we can define:

$$\begin{aligned} Prob(matching_{abcd}) &= Prob(author\ a\ of\ UIC\ b \\ &= principal\ investigator\ c\ of\ funding\ agreement\ d) \end{aligned}$$

Where matching=1 if author b of UIC a is principal investigator d of funding agreement c, 0 otherwise.

At aggregate level, e.g. at scientific area or at department level, the higher the value of this variable, the better UIC data reflect or represent R&D-based links between UPV and industry. Descriptive results will answer the question, how good are UICs as a proxy of interactions between a university and its R&D partners in the business sector?
Econometric results will tell us whether or not UIC volume corresponds with university income from private companies. The econometric model will be:

$$Prob(matching_{abcd}) = f(\alpha X_{ab}, \beta X_{cd}, \gamma X_{abcd})$$

Findings
Among the many types of funding agreements used at the UPV, the most frequent are competitive research projects (30%), technical support contracts (21%), contract R&D (21%) and competitive collaborative R&D (6%).
Regarding the matching process of researchers at UPV and (co-)authors in UICs, contract R&D increases the matching probability. The matching probability is increasing in time, in longer funding agreements and in those that involve larger amounts of money.
The number of institutions involved in the agreement is not influential, but if there is at least one foreign institution, the matching probability increases. If there is at least one firm, the probability decreases.
When we include in our regression analysis funding agreements with firms only, the effect of budget is not significant.

Table 1 Probit regression of the probability of a matching between UPV project members and UIC authors

Variable	Firms only
Competitive research projects	0.09
	(0.09)
Technical support contracts	-0.10
	(0.07)
Contract R&D	0.32***
	(0.07)
Competitive collaborative R&D	0.08
	(0.25)
Start year	0.02***
	(0.01)
Duration	0.02***
	(0.01)
Number of institutions	0.00
	(0.06)
At least one foreign institution	0.40***
	(0.14)

Member number of projects	0.00***
	(0.00)
PI	0.28***
	(0.05)
Male	-0.49***
	(0.06)
Age	-0.01**
	(0.00)
Budget	-0.03
	(0.21)
Constant	-48.12***
	(14.76)
Observations	8,377
Log likelihood	-1,135
Chi2	256
Prob_chi2	0.00

* p<0.1; ** p<0.05; *** p<0.01. Robust standard errors in parenthesis. No multicollinearity according to VIF. Weighting variable: share of member number.

References

Calvert J., Patel. P. (2003) University-industry research collaborations in the UK: bibliometric trends. *Science and Public Policy*, 30, 85-96.

Lundberg, J., Tomson, G., Lundkvist, I., Skar, J., & Brommerls, M. (2006). Collaboration uncovered: Exploring the adequacy of measuring university–industry collaboration through co-authorship and funding. *Scientometrics*, 69(3), 575–589.

Palomares-Montero, D; Garcia-Aracil, A (2011) What are the key indicators for evaluating activities of universities? *Research Evaluation*, 20(5):353-363.

Scott, A., Steyn, G., Geuna, A, Brusoni, S., Steinmueller, E. (2001). The economic returns to basic research and the benefits of university-industry relationships: a literature review and update of findings. Science Policy Research Unit.

Tijssen, R.J.W., Van Leeuwen, T.N., van Wijk, E. (2009) Benchmarking university-industry research cooperation worldwide: performance measurements and indicators based on co-authorship data for the world's largest universities.' *Research Evaluation*, 18(1), 13-24.

Tijssen, R.J.W. (2011) Joint research publications: a performance indicator of university-industry collaboration. *Evaluation in Higher Education*, 5(2), 19-40.

Wong, PK, Singh, A. (2013). Do co-publications with industry lead to higher levels of university technology commercialization activity? *Scientometrics*, 97, 245-265.

Broad altmetric analysis of Mendeley readerships through the 'academic status' of the readers of scientific publications

Zohreh Zahedi*, Rodrigo Costas** and Paul Wouters***

*z.zahedi.2@cwts.leidenuniv.nl; **rcostas@cwts.leidenuniv.nl; ***p.f.wouters@cwts.leidenuniv.nl
CWTS, Leiden University, P.O. Box 905, Leiden, 2300 AX (The Netherlands)

Introduction

Mendeley is a major multidisciplinary source of readership counts for scholarly publications (Zahedi, Costas & Wouters, 2014) and also it is one of the most promising tools for altmetrics research (Li, Thelwall and Giustini, 2012; Wouters & Costas, 2012). As a popular reference management tool, Mendeley has become an interesting and rich source of readership altmetrics. Mendeley collects wide variety of different metadata[1] for each publication saved by the different types of users; among these metadata we have the 'academic status'[2], 'discipline' and the 'country' of the users. But these statistics are available only for the top 3% typologies per publication (Gunn, 2013); this means that for some typologies they are not fully reported. We call this non-identified users 'unknown users' – cf. Zahedi, Costas & Wouters (2013).

Previous altmetrics studies on Mendeley reported moderate correlations between Mendeley readership and citation counts (Priem, Piwowar & Hemminger, 2012; Bar-Ilan et al., 2012; Haustein et al., 2012; Thelwall et al., 2013; Haustein et al., 2013; Zahedi, Costas & Wouters, 2014; Mohammadi & Thelwall, 2014). It has also been observed that Mendeley has different coverage and proportion across different fields (Zahedi, Costas & Wouters, 2014). Other Mendeley studies have analyzed the countries of the users and reported weak correlations among number of authors, departments, institutions and countries and readership and citation counts for WOS publications from biochemistry research published in 2011 (Sud & Thelwall, in press; Thelwall & Maflahi, in press). In a previous study (Zahedi, Costas & Wouters, 2013), we analysed the Mendeley users for a set of 200,000 WOS publications published between 2011 and 2012 also from all disciplines. The result showed that PhD students are the most common types of users per publication in Mendeley (Similar results have been confirmed recently by Mohammadi et al., in press). In terms of highly cited publications, apart from the unknown (i.e. non unidentified) users, Post Docs and PhD students tend to read papers with higher impact than other users in Mendeley.

The current study is built upon the previous study of analyzing Mendeley users with focus on the types of the different Mendeley users (known users) in order to explore their patterns of saving publications in terms of subject fields, citation and readership impact. Particular attention will be paid to the extent to which the readerships of the publications saved by the different types of users in Mendeley correlate with their citation indicators and across 5 major fields of science in the Leiden Ranking (LR). For this reason, we present an exploratory analysis of the patterns of reading of the different types of users in Mendeley and we study their relationship with citations and across LR fields.

---

1 See: http://apidocs.mendeley.com/home/user-specific-methods/user-library-document-details

2 With 'academic status' we refer to the category of different types of users in Mendeley (i.e. PhD students, Professors, Post doc researchers, Students (under graduates and post graduates), Librarians, Lecturers, Other Professionals and Academic and non-Academic researchers) who have saved publications in their individual libraries. This information allows identifying users of scientific publications but this information is not free of limitations. For example, it is not clear whether the academic status of the users are updated regularly or how to distinguish users who could belong to more than one category (e.g. a librarian who is also a PhD student)

Research questions

- What is the distribution of Mendeley readerships across fields and by different academic status? Are there any differences among the different types of users and across fields?
- To what extent do the readerships of the publications saved by the different types of users in Mendeley correlate with their citation indicators? What are the differences in correlation by academic status and across 5 LR fields?

Data and Methodology

For this study, we used a dataset of 1,107,917 publications (review and articles) in WOS published in 2011 with DOIs. The DOIs were used as the basis to extract readership altmetrics from Mendeley by using the Mendeley API in November 2013. The data from Mendeley has been matched back with CWTS in house Web of Science in order to add citation data. Citations have been calculated up to 2012. A total of 748,541 (67.6%) publications have at least one reader in Mendeley (with 7,661,514 total readership and 1,853,045 total citations). Publications with readerships have higher citation scores than publications without readerships (table 1) and also these publications have been published in journals with higher journal citation scores (2.4 vs. 1.6 JCS). Also comparing citations per publication (CPP) and readerships per publication (RPP), on average, these publications have higher RPP (10.2) than CPP (2.4). Out of the 748,541 publications in the dataset with at least one reader in Mendeley, for 339,789 (45.4%), all Mendeley users are known as they lay within the limits of the top three users. These are the publications finally considered for our analysis. In this set we also find higher RPP scores (2.6) than CPP scores (1.1), however it is noticeable both the lower impact and readership density of these publications as compared to the total set of publications and to the publications with unknown users. The explanation is simple, the more readerships of a publication the higher the chances of having readers from more than 3 academic status and as a result the higher the RPP values for publications with unknown users exist(the correlation between Mendeley and citations also explains the higher CPP values for these publications).

Table 1. General distribution of publications

	pubs	%	Total Citation Score (TCS)	CPP	Total Readership Score (TRS)	RPP	Journal Citation Score (JCS)
Pubs with some unknown users	408,752	54.6	1,447,191	3.5	6,755,772	16.5	3
Pubs with only known users	339,789	45.4	405,854	1.1	905,742	2.6	1.6
Total Pubs with readerships	748,541	67.6	1,853,045	2.4	7,661,514	10.2	2.4
Total Pubs without readerships	359,376	32.4	545,411	1.5			1.6
Total	1,107,917	100	2398456	2.16			2.1

For every publication we have calculated the actual number of the different types of Mendeley users (basically dividing the proportion of each type of user provided by Mendeley by the total number of readerships). At the same time, bibliometrics indicators have also been calculated for each publication. Our results show that the most common types of Mendeley users per publication are PhD (42%) and Students (21.3%) and the least common user types are Lecturers (1.6%) and Librarians (1.3%) of all known readerships.

Analysis and Results

Distribution of Mendeley readerships across 5 LR major fields

For each publication the total citation and total readership scores calculated for the 5 major fields of science in the LR[3]. In general for all the fields, publications with Mendeley readerships have higher citation scores than publications without readerships. Table 2 shows that Biomedical & health sciences (37.5%) have the highest share of publications with readerships while Mathematics and computer science (6.4%) have the lowest share. In terms of readership density (i.e. RPP scores) the Life & earth sciences have the highest values (14.1) followed by the Social science & humanities (13.3). Mathematics and computer science (8.1) and Natural sciences & engineering (8.1) exhibit the lowest readerships density. This is in line with both the Mendeley global report[4] and our previous study which showed that the second best covered publications in Mendeley are publications from Medical and life sciences (Zahedi, Costas & Wouters, 2014). Also, on average, all fields show higher RPP scores than CPP scores. This is reasonable since these recent publications (from 2011) still need some time to get their optimum levels of citations, while in terms of social media, the uptake is much faster (Haustein et al, 2013), although we still lack information on the readership pace for publications.

Table 2. Mendeley readerships distribution across 5 major fields of science in LR

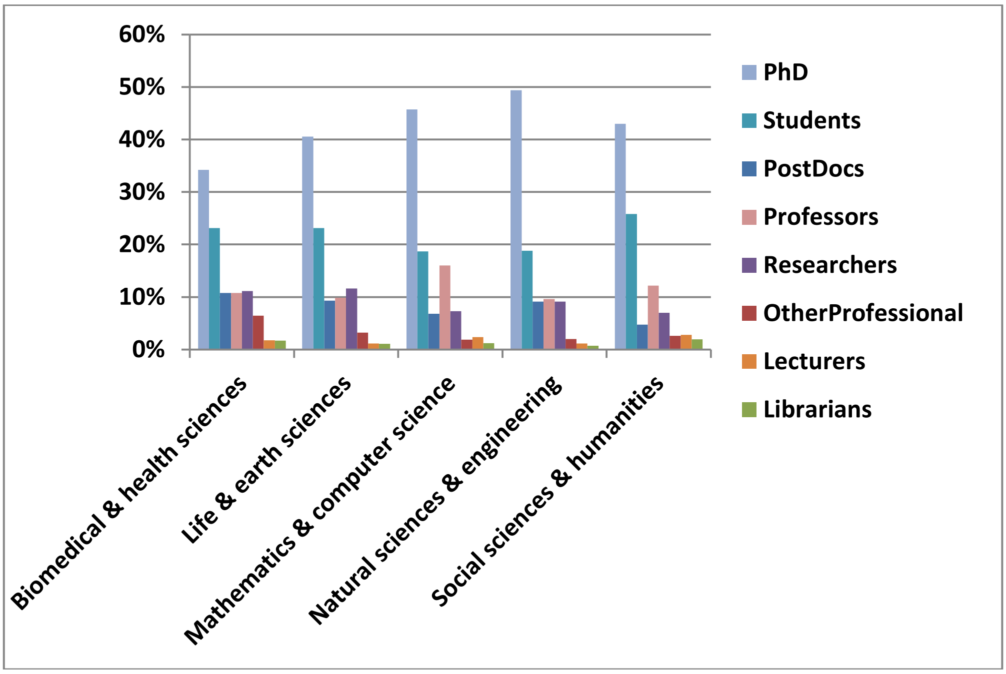
LR Main fields of pubs	P	%	TCS	%	CPP	Total readerships	%	RPP
Biomedical & health sciences	334,268	37.5	1,016,591	45.2	3.0	3,787,811	38.9	11.3
Life & earth sciences	165,969	18.6	411,624	18.3	2.5	2,335,956	24.0	14.1
Mathematics & computer science	57,344	6.4	61,991	2.8	1.1	466,038	4.8	8.1
Natural sciences & engineering	245,857	27.6	652,492	29.0	2.7	1,989,821	20.4	8.1
Social sciences & humanities	87,168	9.8	107,599	4.8	1.2	1,156,794	11.9	13.3
Total		100		100			100	
LR Main fields of Pubs With only Known users	**P**	**%**	**TCS**	**%**	**CPP**	**Total readerships**	**%**	**RPP**
Biomedical & health sciences	133,957	34.3	192,181	41.0	1.4	353,930	33.7	2.6
Life & earth sciences	56,325	14.4	62,617	13.4	1.1	159,911	15.2	2.8

3 http://www.leidenranking.com/ranking/2013
4 http://www.mendeley.com/global-research-report/#.Ux3Tw_k2xv8

Mathematics & computer science	34,482	8.8	21,593	4.6	0.6	82,034	7.8	2.4
Natural sciences & engineering	136,594	35	177,989	38.0	1.3	372,183	35.4	2.7
Social sciences & humanities	28,995	7.4	13,878	3.0	0.5	82,148	7.8	2.8
Total		100		100			100	

Figure 1 shows the proportion of readerships by the different types of Mendeley users across LR fields. Although there are some differences across the fields, in general we find that PhD and students are the most common types of users while Lecturers and Librarian are the least common types of users across all LR fields.

Figure 1. Distribution of Mendeley readerships by the different types of known users across LR fields



Relationship of Mendeley readerships with bibliometric indicators

The focus here is to explore to what extent do the readerships for the publications saved by the different users in Mendeley related with their citation and journal indicators. We have calculated a factor and correlation analysis that we present here. Due to space limits for this short paper, other more detailed results (including correlations among readerships and bibliometric indicators by the different types of known users across LR Fields and precision-recall analysis – as introduced by Waltman & Costas (2014) – to analyze the ability of Mendeley readerships to filter highly cited publications) will be presented in the full version of this paper. A factor analysis based on the Principal Component Analysis (PCA) used to learn more about the underlying structure, dimensions and any relationship among the variables, revealed the presence of 7 main components (dimensions) with eigenvalues exceeding 1 and explaining 74.8% of the total variance. The first dimension is dominated by the bibliometric indicators in which both direct and normalized citation indicators and journal indicators are grouped together (Table 4). In a second dimension we have ‘total readerships’ together with PhDs to a some lesser degree with students; other dimensions related to different types of users in Mendeley: the third, fourth and the fifth ones refers to the scientific users (PostDocs, Professors and Researchers with the highest loadings). The sixth dimension related to Professional users with higher loadings of Other Professionals vs Librarians and the seventh one include Lecturers. These results suggest that considering different types of users in each dimension may represent similar types of impact.

Table 4. Factor analysis of Mendeley main types of users and bibliometric indicators

	Components						
	1	2	3	4	5	6	7
TCS	**.850**	.046	-.151	.049	.029	.007	.080
JCS	**.779**	.026	.282	-.075	.006	.012	-.121
MNJS	**.776**	.042	.219	-.051	-.025	-.011	-.097
NCS	**.774**	.059	-.245	.081	.015	-.007	.124
Total Readers	.075	**.977**	.071	.005	.085	.082	-.019
PhD	.057	**.826**	.119	.047	.019	-.074	.051
Post Docs	.043	.195	**.770**	-.061	-.032	-.043	-.016
Professors	.009	.094	-.091	**.893**	-.149	-.053	-.104
Students	.014	**.406**	-.460	-.498	-.333	-.105	-.227
Researchers	.013	.082	.001	-.110	**.873**	-.006	-.074
Other Professionals	.018	.016	-.175	.047	.262	**.726**	.033
Librarians	-.019	-.029	.149	-.091	-.314	**.685**	-.051
Lecturers	-.003	.026	-.001	-.065	-.065	-.014	**.947**

Extraction Method: Principal Component Analysis.
Rotation Method: Varimax with Kaiser Normalization.
Rotated Component Matrix: Rotation converged in 9 iterations.

The spearman correlation analysis (Table 5) (with 95% Confidence Interval) shows that the overall correlation scores among total readerships and bibliometrics indicators are low (varies between .1 and .2) although the correlations are better with journal indicators (JCS and NJCS) than with article level impact indicators (CS and NCS).

Regarding the different types of users, citations have a higher correlation with PhDs followed by Postdocs. Different patterns also observed in terms of correlations among the readerships and the other different types of users. For example, Professors and Students have a slight negative correlation among them; this may suggest that they have different readership patterns and potentially different readership interest.

Table 5. Correlation analysis of the rank values of citation and altmetrics variables (values >.150 highlighted in bold)

	NCS	**JCS**	**NJCS**	**Total readers**	**Professors**	**PhD**	**Lecturer**	**Students**	**Researchers**	**Librarians**	**Other Prof.**	**Post Docs**
CS	**.963 (.963-.963)**	**.4 (.402-.407)**	**.274 (.271-.277)**	**.165 (.162-.168)**	.005 (.002-.008)	.107 (.104-.111)	-.007 (-.01--.004)	.039 (.035-.042)	.036 (.033-.039)	-.012 (-.015--.008)	.026 (.023-.03)	.073 (.07-.077)
NCS		**.307 (.304-.031)**	**.303 (.3-.306)**	**.155 (.152-.158)**	.009 (.006-.012)	.109 (.106-.122)	-0.001 (-.005-.002)	.038 (.035-.041)	.028 (.025-.031)	-.009 (-.013--.006)	.015 (.012-.018)	.052 (.049-.055)
JCS			**.698 (.696-.7)**	**.209 (.206-.212)**	-.007 (-.011--.004)	.122 (.119-.125)	-.020 (-.023--.016)	.041 (.038-.044)	.066 (.063-.069)	-.013 (-.016--.01)	.039 (.036-.042)	.130 (.127-.133)
NJCS				**.194 (.191-.198)**	.022 (.019-.025)	.149 (.146-.152)	.007 (.004-.01)	.036 (.033-.039)	.035 (.032-.038)	-.006 (-.009-.003)	-.006 (.009-.003)	.056 (.053-.059)
Total readers					**.138 (.135-.141)**	**.59 (.588-.592)**	.049 (.047-.052)	**.338 (.335-.34)**	**.190 (.187-.193)**	.033 (.03-.036)	.081 (.078-.083)	**.211 (.208-.214)**
Professors						-.116 (-.12--.013)	-.037 (-.04-.035)	**-.150 (-.153--.147)**	-.102 (-.104--.099)	-.039 (-.042--.037)	-.060 (-.062-.057)	-.074 (-.077--.072)
PhD							-.052 (-.055--.048)	-.058 (-.061--.055)	-.082 (-.085--.079)	-.070 (-.073--.067)	-.118 (-.121--.115)	-.029 (-.032--.026)
Lecturers								-.05 (-.053--.047)	-.038 (-.041--.035)	-.011 (-.013--.008)	-.016 (-.018--.013)	-.038 (-.041--.036)
Students									-.119 (-.122--.116)	-.039 (-.042--.036)	-.063 (-.066--.061)	-.116 (-.119--.113)
Researchers										-.025 (-.028--.022)	-.016 (-.019--.013)	-.051 (-.054--.048)
Librarians											-.005 (-.008--.002)	-.034 (-.037--.032)
Other Prof.												-.050 (-.052--.047)

Correlation is significant at the 0.01 level (2-tailed). Correlation is significant at the 0.05 level (2-tailed).

Discussion

Mendeley is global data source of readerships data for scholarly outputs; it collects vide variety of metadata per publication saved by different users. The statistics for 'Career Stage' of users is a valuable source of information provided by Mendeley. Although at the moment the 100% of users of all publications saved in Mendeley are not available, their availability would make possible to learn more about the academic and non-academic positions of readers of scientific outputs, thus opening the possibility of more thorough study of the possible different types of impact that these different types of users main entail.

The current study has focused on the idea of analyzing and comparing the readership and citation impact of the scholarly publications used by the different Mendeley users in terms of their typology and across different LR fields. The findings showed that in general, the publications with Mendeley readerships received higher citation impact per publication than those without readerships. Also, in terms of readership density across the 5 major LR fields, on average, all fields show higher RPP scores than CPP scores. Regarding the types of users, the most common types of users in Mendeley are PhDs and Students, and the same proportions are observed for all the LR fields. The correlation analysis shows relatively low relationships among the users with different types of users. This suggests that the different types of Mendeley users could be reading different publications and this could justify the use of these "career stages" to detect different typologies of impact. In follow up research we will deal particularly with this issue.

References

Bar-Ilan, J., Haustein, S., Peters, I., Priem, J., Shema, H., & Terliesner, J. (2012). Beyond citations: Scholars' visibility on the social Web. Retrieved from:

http://arxiv.org/ftp/arxiv/papers/1205/1205.5611.pdf

Costas, R., Zahedi, Z. & Wouters, P. (2014). Do altmetrics correlate with citations? Extensive comparison of altmetric indicators with citations from a multidisciplinary perspective. Retrieved from: http://arxiv.org/abs/1401.4321v1

Gunn, W. (2013). Social Signals Reflect Academic Impact: What it Means When a Scholar Adds a Paper to Mendeley. Information Standards Quarterly. 25(2): 33-39.

Haustein, S., Peters, I., Bar-Ilan, J., Priem, J., Shema, H., & Terliesner, J. (2013). Coverage and adoption of altmetrics sources in the bibliometric community. 14th International Society of Scientometrics and Informetrics Conference (pp. 1–12). Digital Libraries. Retrieved from: http://arxiv.org/abs/1304.7300

Li, X., Thelwall. M., & Giustini, D. (2012). Validating online reference managers for scholarly impact measurement. Scientometrics. 91(2). 461-471.

Mohammadi, E. & Thelwall, M. (in press). Mendeley readership altmetrics for the social sciences and humanities: Research evaluation and knowledge flows. Journal of the American Society for Information Science and Technology.

Mohammadi, E., Thelwall, M., Haustein, S., & Larivière, V. (in press). Who Reads Research Articles? An Altmetrics Analysis of Mendeley User. Journal of the Association for Information Sciences and Technology.

Priem, J., Piwowar. H., & Hemminger, B. H. (2012). Altmetrics in the wild: Using social media to explore scholarly impact. Retrieved from: ArXiv: 1203.4745v1

Sud, P. & Thelwall, M. (in press). Not all international collaboration is beneficial: The Mendeley readership and citation impact of biochemical research collaboration. Journal of the Association for Information Science and Technology.

Thelwall, M., Haustein, S. Larivière, V. & Sugimoto, C. (2013). Do altmetrics work? Twitter and ten other candidates. PLOS ONE. 8(5): e64841. doi:10.1371/journal.pone.0064841

Thelwall, M. & Maflahi, N. (in press). Are scholarly articles disproportionately read in their own country? An analysis of Mendeley readers. Journal of the American Society for Information Science and Technology.

Thelwall, M. & Maflahi, N. (in press). Are scholarly articles disproportionately read in their own country? An analysis of Mendeley readers. Journal of the Association for Information Science and Technology.

Waltman, L., Van Eck, N.J., Van Leeuwen, T.N., Visser, M.S. & Van Raan, A.F.J. (2011). Towards a new crown indicator: Some theoretical considerations. Journal of Informetrics. 5(1). 37–47.

Waltman, L. & Costas, R. (2014). F1000 Recommendations as a Potential New Data Source for Research Evaluation: A Comparison with Citations. Journal of the Association for Information Science and Technology. 65: 433–445.DOI: 10.1002/asi.23040

Wouters, P. & Costas, R. (2012). Users, narcissism and control: Tracking the impact of scholarly publications in the 21st century. Utrecht: SURF foundation. Retrieved from:

http://www.surffoundation.nl/nl/publicaties/Documents/Users%20narcissism%20and%20control.pdf

Zahedi, Z., Costas, R. & Wouters, P. (2013). What is the impact of the publications read by the different Mendeley users? Could they help to identify alternative types of impact? In PLoS ALM Workshop. 7-9 October. 2013. San Francisco. US.

Retrieved from: http://article-level-metrics.plos.org/alm-workshop-2013/ Available: https://openaccess.leidenuniv.nl/handle/1887/23579 (abstract) and https://www.academia.edu/6670997/What_is_the_impact_of_the_publications_read_by_the_different_Mendeley_users_Could_they_help_to_identify_alternative_types_of_impact (slides)

Zahedi, Z., Costas, R. & Wouters, P. (2014). How well developed are Altmetrics? Cross-disciplinary analysis of the presence of 'alternative metrics' in scientific publications. Scientometrics. DOI:10.1007/s11192-014-1264-0

Author index

STI 2014 Leiden

Context counts: pathways to master big and little data
Proceedings of the science and technology indicators conference 2014 Leiden

ISBN 978-90-817527-1-8
Cover design by: Ferdy van Gool (KijkMijnHuis B.V.)
Printed by: PP-Offset
*Website: **sti2014.cwts.nl***

STW Enabling new technology
BRILL
ELSEVIER
THOMSON REUTERS