Jointly published by Akadémiai Kiadó, Budapest and Springer, Dordrecht

Scientometrics, Vol. 79, No. 2 (2009) 377–388
DOI: 10.1007/s11192-009-0425-z

A structural analysis of benchmarks on different bibliometrical indicators for European research institutes based on their research profile

BART THIJS,[a] WOLFGANG GLÄNZEL[a,b]

[a] *K.U. Leuven, Steunpunt O&O Indicatoren and Faculty ETEW, Dept. MSI, Dekenstraat 2, B-3000 Leuven, Belgium*
[b] *Hungarian Academy of Sciences, Institute for Science Policy Research, Budapest, Hungary*

The present study is part of an ongoing project on clustering European research institutions according to their publication profiles. Using hierarchical clustering eight clusters have been found the optimum solution for the classification. Aim of the present study is a structural analysis for the evaluation of research performance of specialised and multidisciplinary institutions. A breakdown by subject fields is used to characterise field-specific peculiarities of individual clusters by bibliometric indicators and to allow comparison within the same and among different clusters. Finally, benchmarks can then be used to study national research performance on basis of the institutional classification.

Introduction

Comparative studies of institutional research performance often face the problem of incomparability of universities and research institutes on bibliometric indicators. Strict normalisation of bibliometric measures and the applications of relative indicators can compensate for this effect to a certain extent but not completely overcome incommensurability. Although these differences can partly be explained by the different mission, funding or tasks of institute, scientometric literature attributes a great deal of these differences in publication and citation indicators to the variation in field activity (see, for instance, [VAN RAAN, 1999; GLÄNZEL, 2001]). Publication and citation behaviour considerably varies among different fields of science, social sciences and humanities. To overcome these comparability problems in evaluative studies we propose the usage of specific benchmarks for a set of likewise institutes. In an earlier study [THIJS & GLÄNZEL, 2008] we constructed 8 sets of institutes that are active in the same fields, they share the same research profile: 1. Biology, 2. Agricultural Sciences, 3. Multidisciplinary, 4. Geo & Space Sciences, 5. Technical and natural Sciences, 6. Chemistry, 7. General and Research Medicine, 8. Specialised Medicine.

Received December 5, 2007

Address for correspondence:
BART THIJS
E-mail: Bart.Thijs@econ.kuleuven.be

0138–9130/US $ 20.00

In the present paper we want to investigate whether this clustering of research institutes into groups of likewise institutes active in the same areas really can create a better set of specific benchmarks for comparable institutes. Therefore we will look at several commonly used publication and citation indicators and apply statistical tests. In particular, we first focus on the differences in bibliometric indicators among clusters. Secondly, we would like to answer the question wether there is still a difference between groups within countries and do these differences change over countries? In a third part, finally, we will analyse if these differences between groups in performance still can be found within a specific field. For this purpose, we have chosen the field of chemistry.

Data sources and processing

Data were extracted from the yearly updates of the Web of Science database of Thomson-ISI (Philadelphia, PA, USA). Only papers of the document type article, letter, and review indexed in the 2001 to 2005 volumes have been selected. This data has undergone a detailed cleaning and was processed to bibliometric indicators.

Publications were assigned to countries and institutions according to the address in the by-line of the paper. A 3-step assignment procedure was developed. For each country under study, a list of distinct names of institutes as occurring in the extracted addresses is compiled. This list contains thus all possible synonyms and spelling variance/errors of research institutes. Secondly, each entry in this list with a number of publications above a certain threshold was assigned, if possible, to a unique, known institute. Finally, this thesaurus is matched with all addresses in a paper's by-line. For this study we selected institutes from 15 European countries (EU15 without Greece but including Switzerland). All institutes have at least 10 publications in the considered time frame (2001–2003). This resulted in a set of 1988 institutes.

Subject classification of the publications was based on the field assignment of journals according to sixteen major fields of science developed in Leuven and Budapest [GLÄNZEL & SCHUBERT, 2003]. These fields are Agriculture & Environment, Biology (Organismic & Supraorganismic Level), Biosciences (General, Cellular & Subcellular Biology, Genetics), Biomedical Research, Clinical and Experimental Medicine I (General & Internal Medicine), Clinical and Experimental Medicine II (Non-Internal Medicine Specialties), Neuroscience & Behaviour, Chemistry, Physics, Geosciences & Space Sciences, Engineering, Mathematics, Social Sciences I (General, Regional & Community Issues), Social Sciences II (Economical & Political Issues) and Arts & Humanities. Journals can also be classified as Multidisciplinary.

Citations received by these papers have been determined using an item-by-item procedure using special identification-keys (so-called cluster-keys) made up of bibliographic data elements. A three year citation window is used beginning with the year of publication. Citation data up to 2005 has been extracted from the WoS.

All 1988 institutes were assigned to one out of eight different groups based on their research profile using a classification model developed THIJS & GLÄNZEL [2007]. Each group can be characterised by an average research profile of each group. Table 1 lists the possible groups together with their share in the total set of institutes.

Table 1. Eight institutional clusters and their share in the total set of institutes

Group	Code	Share
Group 1 (Biology)	BIO	6.7%
Group 2 (Agriculture)	AGR	4.0%
Group 3 (Multidisciplinary)	MDS	30.0%
Group 4 (Geo- & Space Science)	GSS	3.3%
Group 5 (Technical & Natural)	TNS	14.1%
Group 6 (Chemistry)	CHE	4.1%
Group 7 (General & Internal Med.)	GRM	11.1%
Group 8 (Non-internal Med. Spec.)	SPM	26.7%

Bibliometric indicators used in the study

A set a 13 commonly used bibliometric indicators are used to get more insight into the differences between the 8 groups.

- *Publications* assigned to the institute
- *Citations* received by those publications in a 3-year citation window starting at the year of publication.
- *Share of cited papers* (Cited%) is the share of papers receiving at east one citation in a 3-year citation window beginning with the year of publication.
- *Share of international co-authorship* (Coop%) is the percentage of internationally co-authored papers in all publications. Additional to the country of the institute at least one other country is mentioned in the by-line of the paper.
- *Share of author self-citations* (SCit%) is the percentage of author self-citations in all citations received in the same period.
- *Mean Observed Citation Rate* (MOCR) is defined as the ratio of citation count to publication count. MOCR is here based on 3-year citation windows.
- *Mean Observed Citation Rate eXcluding self-citations (MOCRX)* is the ratio of citations excluding self-citations to the total number of publications. This equals (1-SCit%)*MOCR.

- *Mean Expected Citation Rate (MECR)* is a journal-based indicator expressing an expected citation rate of a given paper set. The (journal-based) expected citation rate of a single paper is defined as the average citation rate of all papers published in the same journal in the same year. Instead of the one-year citation window to publications of the two preceding years as used in the Journal Citation Report (JCR), a three-year citation window to one source year is used, as explained above.
- *Mean Expected Citation Rate eXcluding self-citations (MECRX)* is a special journal impact measure expressing the expected rate of citations excluding the self-citations. It's calculated similar to the MECR and has the same implications as the previous one.
- *Relative Citation Rate (RCR)* is the ratio of MOCR and MECR. RCR thus compares the observed citation rate with the expected one. RCR = 0 corresponds to uncitedness, RCR < 1 means lower-than-average, RCR > 1 higher-than-average citation rate, RCR = 1 if the set of papers in question attracts just the number of citations expected on the basis of the average citation rate of the publishing journals.
- *Relative Citation Rate eXcluding self-citations (RCRX)* is build analogously to the previous indicator, particularly, RCRX = MOCRX/MECRX.
- *Normalised Mean Citation Rate (MNCR)* compares analogously to the RCR the observed citation rate (MOCR) with an expected citation rate. Here, however, a field based expected citation rate is used expressing the average number of citations received over all papers published within a specific field within one year. NMCR = 0 corresponds to uncitedness, NMCR < 1 means lower-than-average, NMCR > 1 higher-than-average citation rate, NMCR = 1 if the set of papers in question attracts just the number of citations expected on the basis of the average citation rate of the field to which the papers belong.
- *NMCR/RCR* gives the ratio between the two expected citation rates for a given set of publications. A value above 1 means that the journal based expected rate is higher than the rate based on the field. As journals belong –at least partially- to the given field this value higher than 1 corresponds to the fact that the set of papers is published in journals more visible than the field in average. A value lower than 1 means thus that the papers are published in journals belonging to the lower segment within the field.

The above set of indicators has been calculated for each of these 1988 institutes. The mean values of indicators for each of the eight groups are listed in Table 2.

Table 2. Mean values of indicators by profile clusters

Indicator	AGR	BIO	MDS	GSS	TNS	CHE	GRM	SPM
Publications	248.5	113.8	1226.4	150.8	528.1	87.3	120.3	209.4
Citations	1159.7	387.0	7065.2	1026.8	2140.9	305.0	927.8	1200.2
Cited%	78%	75%	77%	78%	66%	65%	78%	71%
SCit%	32%	33%	28%	37%	37%	32%	19%	19%
Coop%	44%	34%	39%	65%	44%	36%	29%	23%
MOCR	4.41	3.21	5.41	5.41	3.27	3.07	6.55	4.68
MOCRX	3.12	2.18	4.03	3.41	2.15	2.16	5.30	3.78
MECR	4.15	2.94	4.82	5.06	2.94	3.08	5.62	4.05
MECRX	3.03	2.03	3.63	3.53	1.94	2.02	4.60	3.28
RCR	1.04	1.09	1.07	1.04	1.07	0.97	1.14	1.13
RCRX	1.00	1.07	1.05	0.95	1.05	1.04	1.13	1.13
NMCR	1.08	1.04	1.12	1.16	1.11	0.90	1.10	1.01
NMCR / RCR	1.03	0.96	1.02	1.10	1.02	0.93	0.95	0.88

Data analysis

Bibliometric characteristics of profile clusters

The results in Table 2 show that institutes from the multidisciplinary group are the main producers of publications. In the same period they have, on average, more than twice as much papers as the Technical and Natural Science group and even more than ten times the average of the Biology, Chemistry or General & Research Medicine group. For international collaboration, the Geo & Space Science group takes the lead as expected while institutes for both medical groups have on average less than 30% of internationally collaborated papers. For citation-indicators we find also differences in line with what we might expect from field specific indicators. Institutes from the General & Research Medicine group have not only the highest MOCR and MECR but also the highest RCR and RCRX. Also the Specialised Medicine group ranks highest on the latter two indicators. The low NMCR/RCR value for both groups however point out that institutes from these groups tend to publish in the lower segment of journals within their research field. The only group with a RCR below 1 is the Chemistry group. They also have the lowest average on share of cited papers, MOCR and NMCR.

The question arises of whether these differences are indeed statistically significant. Several steps were taken to come up with a solid answer. First, Levene's test for the homogeneity of variance was used. This revealed unequal variance for all indicators but NMCR. Therefore it was not safe to use ANOVA as statistical test and two non-parametric tests were chosen (Kruskal-Wallis and Median test). The Chi-Square statistics from both tests are shown in table 3, degrees of freedom is 7. To be consistent we used these tests also for the NMCR although it had homogeneity of variance.

All of the 13 indicators show a significant difference over the eight groups with significance level lower than 0.01, that is, the a positive answer can be given on the first question: bibliometric characteristics of clusters do differ. After this, ANOVA Post-Hoc tests according to Games-Howell were performed (Games-Howell test does not assume equal variances). The results, which illustrate these between-group differences, can be found in Appendix I.

Table 3. Non-parametric tests for all indicators

Indicator	Kruskal-Wallis Chi-Square	Median Test Chi-Square
Pub	274.97	201.46
Cit	264.10	172.61
Cited%	215.19	198.34
SCit%	870.01	770.79
Coop%	549.22	421.14
MOCR	300.15	226.94
MOCRX	414.88	332.51
MECR	436.61	327.32
MECRX	579.93	400.40
RCR	34.38	28.19
RCRX	43.09	37.94
NMCR	42.93	29.40
NMCR / RCR	172.67	141.24

National characteristics of profile clusters

As groups are not evenly distributed over countries – a Likelihood Ratio Chi-Square test reveals a clear significant relation – one could argue that country-specific particularities are reflected in the above mentioned group differences. Two different tests were run to tackle this question. First a General Linear Model was constructed with groups and countries as fixed factors. The interaction between both factors was also included in the model. Table 4 shows the result of this analysis. Except for the number of publications and citations, an interaction was found for all indicators.

In a second step, the above mentioned Kruskal Wallis test was run on each country separately. The results can be found in Appendix II. This table gives for each indicator in each country the level of significance for the differences. Values below 0.05 (5%) indicate the rejection of the hypotheses that there are no differences on the respective indicator within the relevant country. The table shows that all indicators have significant differences in at least some countries. The shares and the mean citation rates (both observed and expected) remain significant in even at least 12 countries. The RCR, RCRX and NMCR show significant differences in less countries (3 to 5). The two tests allow us to conclude that although there is some relation between the grouping and the countries it is not possible to reduce the group effect to a side-effect from general national characteristics as reflected by the national indicators.

Table 4. F-Statistic for Group*Country Interaction all indicators

Indicator	F	Sig.
Pub	0.721	0.98
Cit	0.786	0.93
Cited%	2.521	0.00
SCit%	1.706	0.00
Coop%	2.113	0.00
MOCR	1.932	0.00
MOCRX	1.698	0.00
MECR	2.024	0.00
MECRX	1.907	0.00
RCR	1.814	0.00
RCRX	1.690	0.00
NMCR	1.949	0.00
NMCR / RCR	2.288	0.00

Group-specific characteristics in subject fields. The case of chemistry

In order to get more insight in the underlying differences we have broken down the profiles of the individual clusters by subject fields. The question that we want to answer is whether the group differences only occur when the complete set of publications is taken into account or also when looked within one specific field or even within subfields. In what follows, we present *chemistry* as an example. 77% of the total publication output of group CHE is devoted to the field of chemistry. The share of this field is still significant in clusters MDS (17%) and TNS (30%), less significant in BIO (10%) and AGR (7%) and almost marginal in the two medical groups and in GSS. Nonetheless we are interested in the performance of chemistry research in the other clusters too. Moreover, we would like to see a finer grained profile of chemistry research in the different institutional groups. Table 5 presents the citation-based indicators in this field organised by the eight clusters. In this table and following one the reported results are not calculated as mean values over the institutes but as indicators over the total set of chemistry papers produced by institutes within the group.

Table 5. Citation-based indicators in chemistry by groups

	AGR	BIO	MDS	GSS	TNS	CHE	GRM	SPM
% Chemistry	7%	10%	17%	3%	30%	77%	1%	1%
% SCit	33.6%	33.4%	36.2%	34.3%	37.3%	31.7%	30.3%	30.1%
MOCR	4.10	2.96	4.32	3.96	3.75	3.40	4.28	4.04
RCR	1.02	0.98	1.01	1.40	1.06	1.01	0.94	0.97
NMCR	1.13	1.00	1.14	1.22	1.17	1.07	0.88	0.89
NMCR/RCR	1.10	1.02	1.14	0.87	1.11	1.06	0.94	0.92

Table 6. Subfield profile of chemistry research in different clusters (C0: MultiDisciplinary Chemistry, C1: Analytical, Inorganic & Nuclear Chemistry, C2: Applied Chemistry & Chemical Engineering, C3: Organic & Medicinal Chemistry, C4: Physical Chemistry, C5: Polymer Science, C6: Materials Science)

Subfield	MDS	TNS	CHE	GRM	SPM	Total
C0	13.1%	8.5%	8.9%	18.9%	17.3%	13.8%
C1	25.0%	18.3%	12.5%	40.2%	40.5%	21.8%
C2	10.4%	12.1%	17.3%	5.4%	8.0%	12.0%
C3	18.2%	6.1%	9.2%	42.6%	24.2%	14.4%
C4	22.5%	23.9%	22.1%	4.1%	7.6%	21.4%
C5	5.9%	8.0%	13.1%	3.0%	4.2%	6.6%
C6	23.0%	39.5%	36.0%	2.4%	10.0%	27.2%

Chemistry papers published by multidisciplinary and by medical institutes have a higher citation impact than those by chemistry institutes themselves. The corresponding expected values (based on journal impact measures) reflect the same tendencies. Multidisciplinary and medical institutes publish on average in higher-impact journals than scientists at chemistry institutes. These data reveal that the differences between groups not only exist in the total set but also prevail within a specific field.

A remarkable difference in the NCMR can be found in Table 5. The multidisciplinary group has a MOCR higher than the Chemistry group but the NMCR is lower. This difference is even more striking when looking at the medical groups. A closer look at subfield profile reveals important reasons for these deviations. The different groups are active in other subfields of chemistry. Each subfield has its own field expected citation rate. The two medical clusters show a clear similarity, particularly, the predominance of subfields C1 and C3, while the cluster CHE (and similarly TNS) are characterised by a more applied profile. Only the multidisciplinary cluster is characterised by a more balanced subfield representation. The deviations of the citation indicators from the CHE standard, on one hand, and from the European total, on the other hand, can largely be explained by these profiles. The multidisciplinary cluster deserve its name at this level, too, since it profile largely coincides with that of the European total in the field. The two medical clusters do research especially in those disciplines which are relevant for their profile, namely, analytical, inorganic & nuclear chemistry and organic & medicinal chemistry.

Comparing the percentages in the TNS and CHE cluster results in a χ^2-value of 4.08 with 5 degrees of freedom. This means that the two distributions are not significantly different. There is also no difference between the distribution of the MDS group and the total set of chemistry publications (χ^2=1.21). However, comparing other pairs results in significant χ^2-values. The pair GRM and SPM has a χ^2-value of 11.43 which is significant at 0.05 level.

Conclusion

This research shows that it is valid to group likewise institutes on basis of their research profile in order to create different benchmarks that can be used in evaluative studies. The clustering proposed earlier by the authors is a valid one as all 13 used indicators show significant differences across groups. This study tried subsequently to get a better understanding of these differences. Although some interaction was found between the grouping and country assignment of institutes it was clear that the grouping stands as an independent main effect.

When digging deeper within a field or even within subfields it becomes clear that the differences in performance between the groups remain and the particular activity profiles of different groups within the same field differ as well. In other words, the profile of chemistry research in the chemistry group differs from that of the multidisciplinary cluster. This again illustrates the advantage of clustering institutes by their research profile in order to improve the validity of comparative studies even within one subject area.

References

AKSNES, D.W. (2003), A macro-study of self-citations. *Scientometrics*, 56 (2) : 235–246.

GLÄNZEL, W. (2001), National characteristics in international scientific co-authorship, *Scientometrics*, 51 (1) : 69–115.

GLÄNZEL, W., SCHUBERT, A.(2003), A new classification scheme of science fields and subfields designed for scientometric evaluation purposes, *Scientometrics*, 56 (3) : 357–367.

SNYDER, H., BONZI, S. (1998), Patterns of self-citation across disciplines. *Journal of Information Science*, 24 : 431–435.

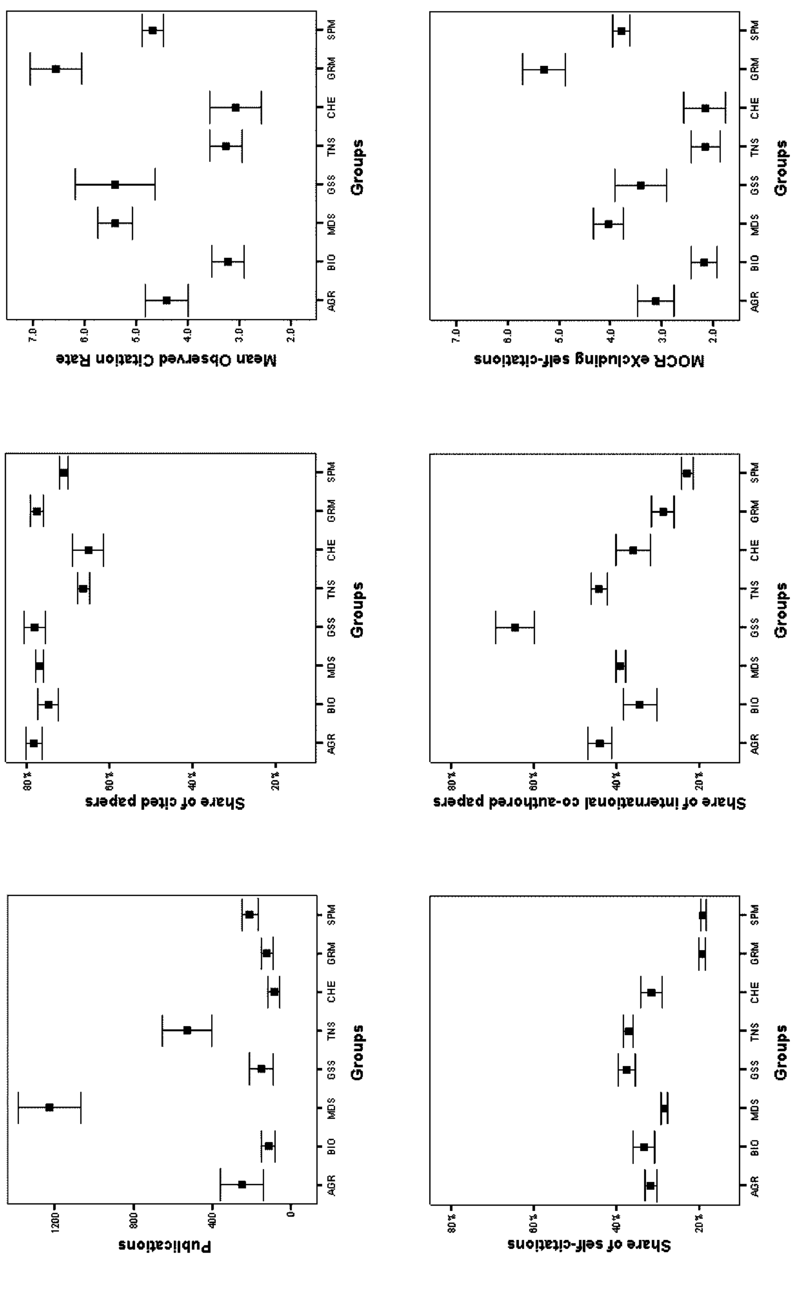
THIJS, B., GLÄNZEL, W. (2006), The influence of author self-citations on bibliometric meso-indicators. The case of European universities. *Scientometrics*, 66 (1) : 71–80.

THIJS, B., GLÄNZEL, W. (2008), A structural analysis of publication profiles for the classification of European research institutes. *Scientometrics*, 74 (2) : 223–236.

VAN RAAN, A. (1999), Advanced bibliometric methods for the evaluation of universities. *Scientometrics*, 45 (3) : 417–423.

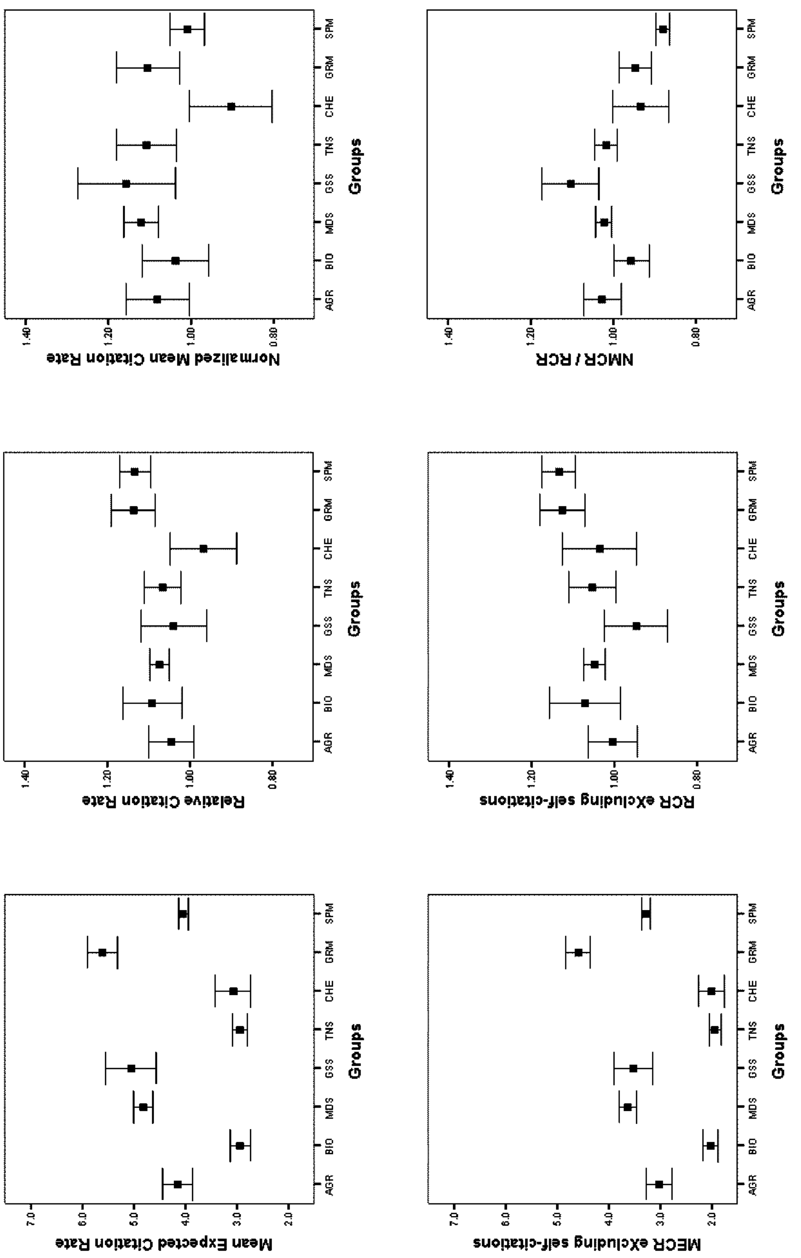
Appendix I

Error bar graphs representing the 95% CI for the mean for 12 of the 23 indicators used for the study



Appendix I

(continued)



Appendix II
Level of significance for chi-square for each indicator within the given country

Country	Pub	Cit	Cited%	SCit%	Coop%	MOCR	MOCRX	MECR	MECRX	RCR	RCRX	NMCR	NMCR / RCR
Austria	0.196	0.076	**0.006	**0.003	**0.002	*0.016	**0.007	**0.002	**0.001	0.579	0.555	0.092	0.427
Belgium	*0.031	*0.010	*0.013	**0.000	0.632	**0.000	**0.000	**0.000	**0.000	0.052	0.057	0.054	0.242
Denmark	0.280	0.313	*0.018	*0.040	0.082	0.098	*0.041	0.050	**0.008	0.339	0.371	0.810	0.065
Finland	0.860	0.732	**0.000	**0.000	**0.004	**0.000	**0.000	**0.000	**0.000	0.128	0.065	*0.029	0.113
France	*0.019	*0.012	**0.000	**0.000	**0.000	**0.000	**0.000	**0.000	**0.000	**0.000	**0.000	0.410	**0.000
Germany	**0.000	**0.000	**0.000	**0.000	**0.000	**0.000	**0.000	**0.000	**0.000	0.126	*0.024	0.148	**0.000
Ireland	0.370	0.101	0.325	*0.018	*0.035	*0.020	*0.020	*0.026	*0.011	*0.013	0.071	**0.008	*0.026
Italy	**0.000	**0.000	**0.000	**0.000	**0.000	**0.000	**0.000	**0.000	**0.000	0.157	0.076	0.385	**0.000
Luxemburg	0.301	0.301	0.223	0.741	0.223	0.368	0.368	0.741	0.741	0.819	0.819	0.301	0.301
Netherlands	0.220	0.098	**0.000	**0.000	**0.000	**0.000	**0.000	**0.000	**0.000	*0.045	0.171	0.456	0.906
Portugal	0.229	0.203	*0.049	**0.000	*0.026	**0.003	**0.001	**0.004	**0.001	0.053	0.102	0.190	0.616
Spain	**0.000	**0.000	**0.000	**0.000	**0.000	**0.000	**0.000	**0.000	**0.000	*0.036	*0.040	**0.000	**0.000
Sweden	**0.000	**0.000	**0.000	**0.000	**0.002	**0.000	**0.000	**0.000	**0.000	0.164	0.402	0.069	*0.019
Switzerland	*0.012	**0.007	*0.012	**0.005	*0.017	**0.008	**0.004	**0.005	**0.003	0.127	0.079	**0.008	*0.013
UK	**0.000	**0.000	**0.000	**0.000	**0.000	**0.000	**0.000	**0.000	**0.000	0.168	0.505	**0.000	**0.000

* indicates a level below 5%, ** below 1%