Jointly published by Akadémiai Kiadó, Budapest and Springer, Dordrecht

Scientometrics, Vol. 71, No. 3 (2007) 423–441
DOI: 10.1007/s11192-007-1682-3

Import-export of knowledge between scientific subject categories: The iceberg hypothesis

VICENTE P. GUERRERO-BOTE,[a] FELIPE ZAPICO-ALONSO,[a] MARÍA EUGENIA ESPINOSA-CALVO,[a] ROCÍO GÓMEZ-CRISÓSTOMO,[a] FÉLIX DE MOYA-ANEGÓN[b]

[a] *Facultad de Biblioteconomía y Documentación, Universidad de Extremadura, Badajoz (Spain)*
[b] *Facultad de Biblioteconomía y Documentación. Universidad de Granada, Campus Cartuja, Granada (Spain)*

The capacity to attract citations from other disciplines – or knowledge export – has always been taken into account in evaluating the quality of scientific papers or journals. Some of the JCR's (ISI's *Journal Citation Report*) Subject Categories have a greater exporting character than others because they are less isolated. This influences the rank/JIF (ISI's Journal Impact Factor) distribution of the category. While all the categories fit a negative power law fairly well, those with a greater External JIF give distributions with a more sharply defined peak and a longer tail – something like an iceberg. One also observes a major relationship between the rates of export and import of knowledge.

1. Introduction

Since publication, and in particular publication in scientific journals, is one of the principal forms of scientific output, it is not surprising that this aspect is often the focus of the evaluation of research activity. But not all papers have the same value or quality, and in this sense, citations, notwithstanding the different motivations that can be attributed to them (BROOKS, 1995), are accepted as an indicator of a paper's quality.

Received July 17, 2006

Address for correspondence:
VICENTE P. GUERRERO-BOTE
Facultad de Biblioteconomía y Documentación, Universidad de Extremadura
06071 Badajoz, Spain
E-mail: guerrero@unex.es

0138–9130/US $ 20.00

Citations are not solely used as an indicator of the quality of the work, but also of that of the journals, hence such citation indices as ISI's Journal Impact Factor (JIF) in the *Journal Citation Report* (JCR). Simultaneously, sometimes due to the time that one must wait in order to have the profile of citation of a given paper and to the randomness factor present in the process of citing papers (VAN LEEUWEN & MOED, 2002), these journal citation indices are also used as indicators of the quality and visibility of the papers they publish. Since all scientific journals apply a selection process, it is to be supposed that in those of greater quality this process is stricter and more competitive, and that a paper's having successfully passed this process is a guarantee of a certain quality. The JIF is an average of the citations received by articles in the journal, and therefore of the international visibility of those articles. One could therefore say that it could be regarded as the expected impact of a paper published in that journal (as against what would eventually be obtained from the citations that it receives).

The citations obtained by a paper or a journal also depend on the corresponding discipline. One factor is the size of the discipline – if the audience is large, there exists a greater possibility of citation, although the field citation average does not depend on it because there are also many more different papers that might be cited. A second consists of the differing habits of citation – in some disciplines (in the humanities mainly), citations of recent documents are combined with those of original sources. And a third is the size of the list of references in a typical article, which is known to have a major influence on the impact (PINSKY & NARIN, 1976).

Together with the impact and other journal indices, the JCR also publishes a classification of journals by "Subject Categories", with each journal being assigned to a minimum of one category. According to PUDOVKIN & GARFIELD (2002): "The categories have been developed by manual methods started over 40 years ago" – we understand by principally taking into account the scope of the journals. When these categories grew, subdivisions were established. "Once the categories were established, new journals were assigned one at a time. Each decision was based upon a visual examination of all relevant citation data", and with the aid of the Hayne-Coulson algorithm (PUDOVKIN & GARFIELD, 2002).

These Subject Categories are currently used, for example, in the evaluation of science (PODLUBNY, 2005; VAN LEEUWEN & MOED, 2002; VAN LEEUWEN et al., 2003; SOMBATSOMPOP & MARKPIN, 2005; MOYA-ANEGÓN et al., 2005), or in constructing maps of science (MOYA-ANEGÓN et al., 2004) by applying citation analysis and such pruning techniques as the Pathfinder algorithm (SCHVANEVELDT et al., 1988; GUERRERO-BOTE et al., 2006). But it is broadly accepted today that, for the reasons expounded above, the JIFs are comparable only within the same category, since the mean JIF of the journals in each category and the range of values that the JIFs take can vary greatly from one category to another. Hence, processes of evaluation sometimes normalize the JIFs for each category, so that the values corresponding to different categories can be compared (PODLUBNY, 2005; VAN LEEUWEN & MOED, 2002; VAN LEEUWEN et al., 2003; SOMBATSOMPOP & MARKPIN, 2005; MOYA-ANEGÓN et al., 2005).

Other indices used to evaluate the quality of journals include the self-citation rate which, as PICHAPPAN (1995) indicates, "show the level of its contribution to the speciality it belongs to", and knowledge export which indicates whether a publication has sufficient scientific impact to break through the borders of its Subject Category and attract citations from neighbouring categories (WORMELL, 1998; STEGMANN & GROHMANN, 2001).

The way that citations act means that citing a document or a journal increases its likelihood of being cited again. This has been termed the "Matthew effect" (MERTON, 1968), and is, as LEYDESDORFF & BENSMAN (2006) note, the mechanism of preferential attachment which is well known for generating negative power laws (BARABÁSI, 2002; BARABÁSI et al., 2002; KATZ, 1999, 2000; WAGNER & LEYDESDORFF, 2006). One can therefore expect the rank/JIF distributions to be exponential, both as totals and within each Subject Category.

Our hypothesis is that, besides the aforementioned factors of the habits of citation and the number of references, the export of knowledge will have a definitive influence on a category's distribution of JIFs. We shall show that this export of knowledge increases the differences between the highest and the lowest impact journals, and thereby leads to sharper rank/impact distributions. We think that this is because there is a larger potential external audience for the papers of that category, and hence that the most cited papers can receive yet more citations, thus widening the difference.

In these distributions, the impact will fall rapidly as a function of rank, giving rise to a peak with a long tail. Thus, in the more exporting categories, the peak will be more pronounced, something like an iceberg, whose part above the surface is visible from a distance whereas the submerged part is only seen from the iceberg itself.

Thus, in this work we shall focus on the categories' rank/impact distributions, trying to discover whether these distributions follow some law, whether there are different types of distributions, and the factors that influence them.

Method and data

The data used correspond to those of the JCR of 1997, accessible through the ISI Web of Knowledge, including both the published JIFs and the Classification by Subject Categories. Likewise, we used the data of the Web of Science, also accessible through the ISI Web of Knowledge, for the calculation of the rest of the indices, considering only citations made in papers published in 1997 to papers published in 1995 or 1996.

For each Subject Category, we calculated the number of citable papers (Articles, Notes, and Reviews published in 1995 or 1996), the total number of citations received (made in 1997 to papers published in 1995 or 1996) (*TotCites*), the number of self-citations received (made in 1997 to papers published in 1995 or 1996, and from journals of the same category) (*SubCites*), the number of reviews published in the category, the number of citing papers of the category, and the number of references of the category (of papers of 1997 to papers published in 1995 or 1996).

For each journal, we calculated the number of citable papers (Articles, Notes, and Reviews published in 1995 or 1996), the total number of citations received (made in 1997 to papers published in 1995 or 1996), and the number of self-citations received (made in 1997 to papers published in 1995 or 1996, and from journals of the same category).

From those data, we calculated the following indices:

– Rate of Export of Knowledge (WORMELL, 1998) (*%export*):

$$\%export = \frac{100 \cdot (TotCites - SubCites)}{TotCites}$$

– Rate of Import of Knowledge (*%import*):

$$\%import = \frac{100 \cdot (References - SubCites)}{References}$$

– Internal Journal Impact Factor: similar to the JIF, but only taking into account the citations from that same Subject Category.

– External Journal Impact Factor: similar to the JIF, but only taking into account the citations not from that same Subject Category.

For all these measures, keep in mind that many journals are assigned to more than one category. This mean that the references from one of these journals are Subcites (internal cites) for all these categories.

For each category, we plotted the impacts against the descending-order ranks, and calculated the best-fit exponential trend line. We used as indices of the distribution the slope *a* and offset *b* of these trend lines, and, from the fits, the R^2 coefficient and the *deviation squared* divided by the mean JIF. This last index was particularly useful in detecting the categories whose distribution has a sharper peak than the exponential.

Finally, we correlated the different indices in order to study their influence on the fit of each distribution to the logarithmic trend.

Results and discussion

As was expected, the distributions of the JIFs were fairly close to an exponential in all the categories, so that in the first JIFs there was a major jump from one to the next, with these jumps getting smaller as one advanced in the ranking, until reaching the tail where the variations were very small.

Figure 1 shows the distribution of the Subject Category Agriculture, Dairy & Animal Science. This is one which most closely fits the logarithmic distribution. For illustration, the figure includes the External JIF, the Export rate, and an exponential trend line with respect to the JIF. This trend line corresponds to the expression:

$$y = -0.51 \cdot \ln(x) + 2.09$$

The coefficient $a = -0.51$ characterizes the sharpness of the trend, and the coefficient $b = 2.09$ the intersect, and hence the height of the peak. The goodness of the fit to this logarithmic approximation can be characterized by the values of R^2 and the squared deviation:

$$R^2 = 0.9948$$

$$\frac{\overline{Diff^2}}{\overline{JIF}} = 0.0019$$

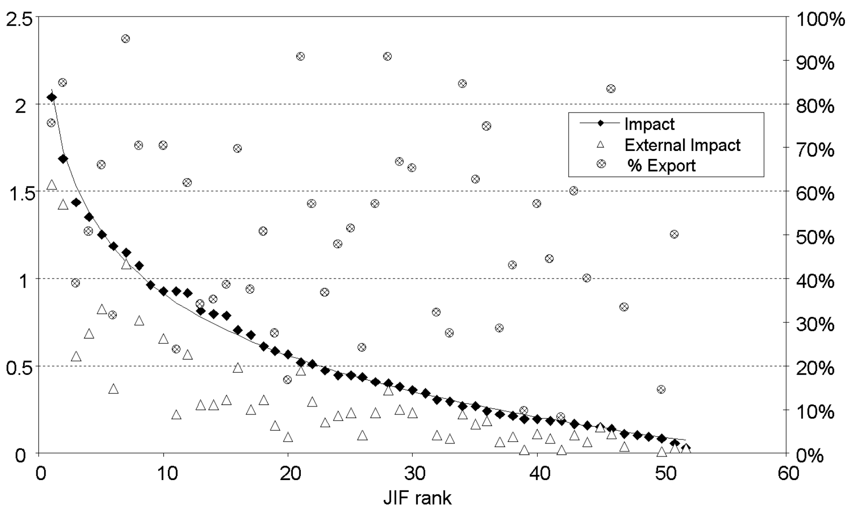

Figure 1. Distribution of the JIFs of the Subject Category Agriculture, Dairy & Animal Science. Added are the curve corresponding to the logarithmic trend of the distribution, and the External JIFs and export rates of each of the journals

With respect to the External JIF, one observes its close relationship with the JIF – larger JIFs usually correspond to large External JIFs.

Figure 2 shows the same data but with the ranks on a logarithmic scale to better appreciate the goodness of the fit of the distribution to a logarithmic function. The slope and intersect are the values of the coefficients given above in the description of the logarithmic trend: $a = -0.51$ and $b = 2.09$.

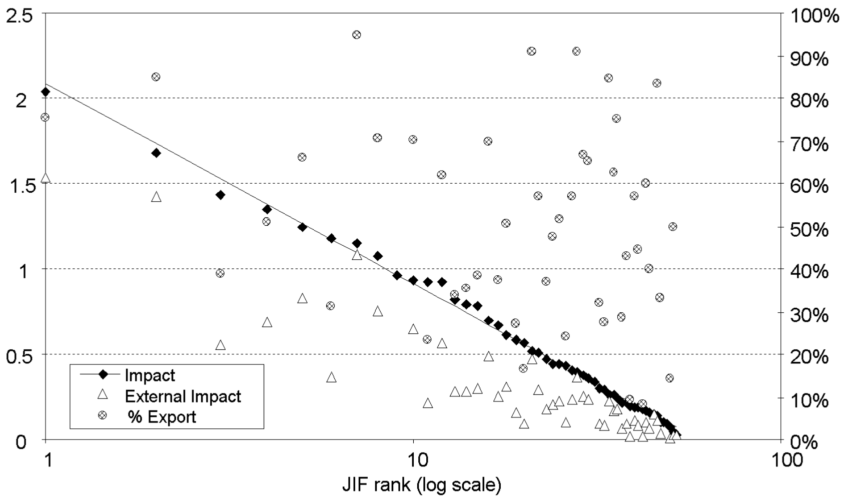

Figure 2. As Figure 1, but with a logarithmic scale for the ranks

No relationship was observed between the export rate and the JIF. This contradicts the hypothesis that journals with a greater JIF also have a higher export rate and *vice versa*. In the External JIFs, the Matthew effect is even more marked, as can be seen in Figure 3 which is a plot of the same data, but ranked by External JIF. The trend of the distribution is sharper than a negative power law, and one can see that the external component accounts for most of the slope of the JIFs (65% average in the impacts superior to 1).

Figure 4 shows the distribution of the Subject Category Biochemistry & Molecular Biology. This is one which least well fits the logarithmic approximation. This trend line corresponds to the expression:

$$y = -4.32 \cdot \ln(x) + 22.68$$

The goodness of the fit to this logarithmic approximation can be characterized by the values of R^2 and the squared deviation:

$$R^2 = 0.7869$$

$$\frac{\overline{Diff^2}}{\overline{JIF}} = 1.5001$$

One observes that this distribution has a sharper initial descent than the logarithmic line, and a longer tail.

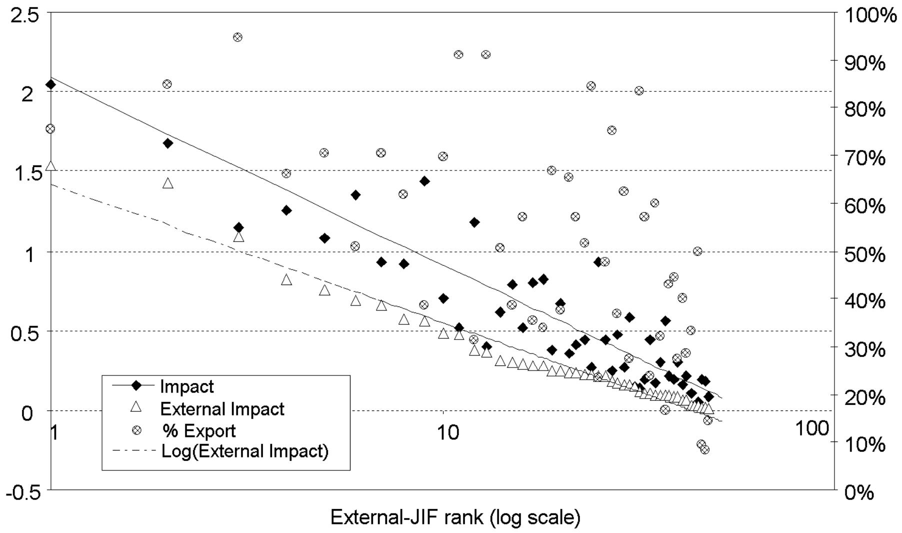

Figure 3. Distribution of the External JIFs of the Subject Category Agriculture, Dairy & Animal Science on a logarithmic scale. Added are the export rates and the straight line trends corresponding to the External JIFs and the JIFs

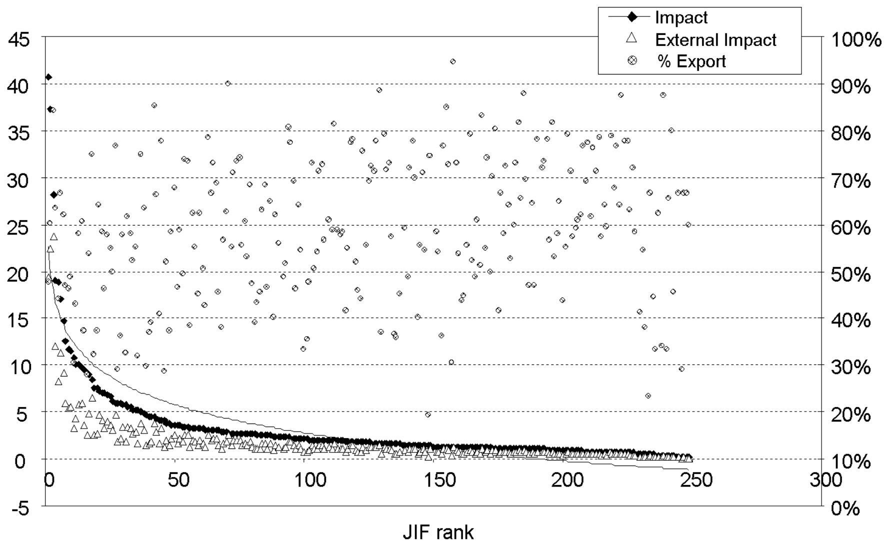

Figure 4. Distribution of the JIFs of the Subject Category Biochemistry & Molecular Biology. Added are the curve corresponding to the logarithmic trend of the distribution, and the External JIFs and export rates of each of the journals

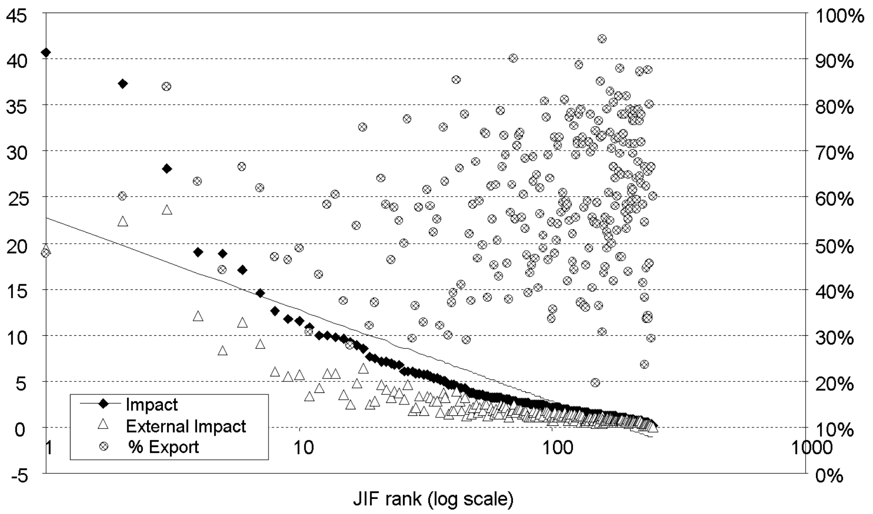

Figure 5. As Figure 4, but with a logarithmic scale for the ranks

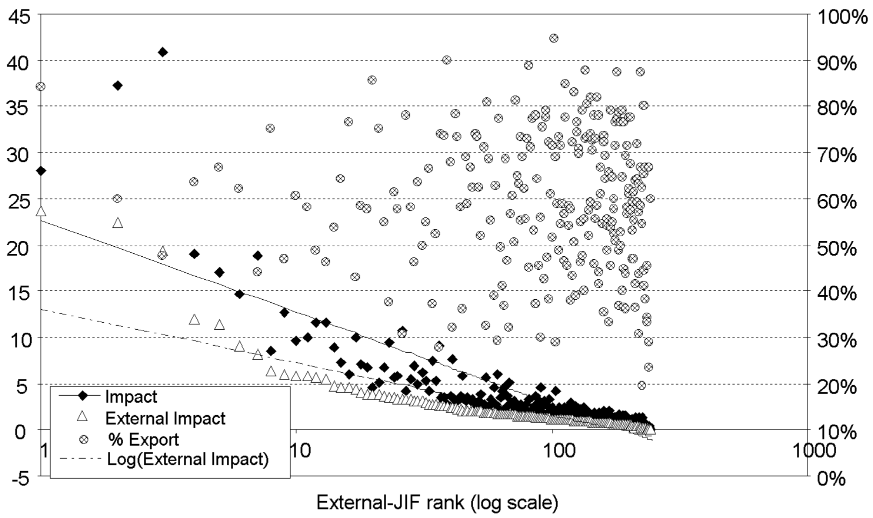

Figure 6. Distribution of the External JIFs of the Subject Category Biochemistry & Molecular Biology on a logarithmic scale. Added are the export rates and the straight line trends corresponding to the External JIFs and the JIFs

This can also be seen in Figure 5, which is a plot of the same data but with a logarithmic scale for the ranks.

The features shown for this Subject Category are typical of all the categories that do not give a good fit to the logarithmic approximation, as is reflected more in the squared deviation than in the R^2 coefficient.

Figure 6 shows the distribution of the external component of the JIFs. One observes in this case that it accounts for more than half the slope of the logarithmic approximation (53% average in the impacts superior to 10), and most of the deviation with respect to that approximate.

Table 1 lists the coefficients of the fits to the logarithmic approximation for all the Subject Categories with an impact distribution. One observes that the category which least well fits the logarithmic approximation by far is Multidisciplinary Sciences. This Subject Category is composed of journals that publish papers in a wide range of fields (not by journals that publish only (or primarily) papers that are themselves multi-disciplinary). The same occurs with other subject categories like "Chemistry Multidisciplinary", "Physics Multidisciplinary", or "Material Science Multi-disciplinary" with regard to the composition, although not with respect to the fit. This is hardly surprising, since this special category of Multidisciplinary Sciences is formed with a few journals that have a very high JIF (i.e. *Science* or *Nature*), while the rest have a low JIF (i.e.: *Revista Chilena de Historia Natural*, or *Interciencia*).

Table 1. R^2 coefficient and squared deviation for the Subject Categories with an impact distribution in 1997. Included are the number of journals in 1997, the mean of papers and articles for 1995–1997, and the number of references from 1997 to 1995–1996

Subject Category	R^2	$Diff^2$	Journals	Papers	Articles	Refs. 1997→95–96
ACOUSTICS	0.9661	0.0107	20	3476	2818	2741
AGRICULTURAL ECONOMICS & POLICY	0.9596	0.0025	13	1054	306	149
AGRICULTURE	0.9849	0.0044	112	10595	9967	8031
AGRICULTURE, DAIRY & ANIMAL SCIENCE	0.9948	0.0019	52	4569	4091	3902
AGRICULTURE, SOIL SCIENCE	0.9499	0.0141	31	3408	2794	2289
ALLERGY	0.9859	0.0102	18	3602	1879	4640
ANATOMY & MORPHOLOGY	0.9308	0.0600	16	1211	1103	2533
ANDROLOGY	0.9737	0.0091	4	275	257	389
ANESTHESIOLOGY	0.9585	0.0337	21	6653	2905	7257
ANTHROPOLOGY	0.9904	0.0034	48	3382	1179	992
AREA STUDIES	0.9383	0.0079	35	3101	794	243
ASTRONOMY & ASTROPHYSICS	0.8912	0.2059	38	10225	9794	25664
AUTOMATION & CONTROL SYSTEMS	0.9853	0.0037	41	3385	2366	1553
BEHAVIORAL SCIENCES	0.7297	0.2343	38	3676	2670	6400
BIOCHEMICAL RESEARCH METHODS	0.8874	0.0412	33	5718	6064	15407
BIOCHEMISTRY & MOLECULAR BIOLOGY	0.7869	1.5001	253	63009	42450	215615
BIOLOGY	0.8233	0.5697	63	13457	5308	16448
BIOLOGY, MISCELLANEOUS	0.9641	0.0381	78	3240	3668	10745
BIOPHYSICS	0.9114	0.2318	46	11756	7524	29555
BIOTECHNOLOGY & APPLIED MICROBIOLOGY	0.9597	0.0702	94	13046	10900	31101
BUSINESS	0.9805	0.0159	51	3874	2000	1547
BUSINESS, FINANCE	0.9527	0.0206	33	2914	1251	874

(continued on next page)

Table 1. (cont.)

Subject category	R^2	$Diff^2$	Journals	Papers	Articles	Refs. 1997→95–96
CARDIAC & CARDIOVASCULAR SYSTEMS	0.8923	0.2443	62	18688	11471	25059
CELL BIOLOGY	0.8155	1.4078	130	23828	14941	93054
CHEMISTRY, ANALYTICAL	0.9794	0.0123	61	12560	12093	26276
CHEMISTRY, APPLIED	0.9754	0.0117	44	5112	4620	5626
CHEMISTRY, INORGANIC & NUCLEAR	0.7301	1.3479	36	9455	8712	18878
CHEMISTRY, MEDICINAL	0.9810	0.0159	28	3775	3515	11649
CHEMISTRY, MULTIDISCIPLINARY	0.7451	1.0437	111	32672	18825	41604
CHEMISTRY, ORGANIC	0.9801	0.0158	43	14656	13898	35658
CHEMISTRY, PHYSICAL	0.8860	0.1841	86	20519	19290	42530
CLINICAL NEUROLOGY	0.9550	0.0583	107	19247	10934	27507
COMMUNICATION	0.8806	0.0169	36	1925	812	597
COMPUTER SCIENCE, ARTIFICIAL INTELLIGENCE	0.9382	0.0323	54	3463	2363	1917
COMPUTER SCIENCE, CYBERNETICS	0.9832	0.0056	18	817	606	455
COMPUTER SCIENCE, HARDWARE & ARCHITECTURE	0.9527	0.0196	43	4841	2717	950
COMPUTER SCIENCE, INFORMATION SYSTEMS	0.9882	0.0051	52	4297	2666	1870
COMPUTER SCIENCE, INTERDISCIPLINARY APPLICATIONS	0.9558	0.0240	63	5348	4461	4122
COMPUTER SCIENCE, SOFTWARE, GRAPHICS, PROGRAMMING	0.9748	0.0063	61	5666	3207	1513
COMPUTER SCIENCE, THEORY & METHODS	0.9641	0.0080	59	5079	2869	1615
CONSTRUCTION & BUILDING TECHNOLOGY	0.9839	0.0024	20	1985	1088	446
CRIMINOLOGY & PENOLOGY	0.9914	0.0026	19	1009	598	318
CRYSTALLOGRAPHY	0.8877	0.0465	18	5213	4723	5776
DEMOGRAPHY	0.9362	0.0430	18	884	461	429
DENTISTRY, ORAL SURGERY & MEDICINE	0.9547	0.0172	42	10431	3608	5113
DERMATOLOGY & VENEREAL DISEASES	0.9054	0.0625	32	7088	3973	7831
DEVELOPMENTAL BIOLOGY	0.7834	1.3519	30	3411	2804	18300
ECOLOGY	0.9792	0.0225	86	7709	7022	11896
ECONOMICS	0.7393	0.3556	161	9959	6064	3793
EDUCATION & EDUCATIONAL RESEARCH	0.9428	0.0194	102	6035	2943	1691
EDUCATION, SCIENTIFIC DISCIPLINES	0.9198	0.0206	10	1679	1181	753
EDUCATION, SPECIAL	0.9784	0.0063	21	954	567	627
ELECTROCHEMISTRY	0.8603	0.0468	9	3132	2722	3498
EMERGENCY MEDICINE & CRITICAL CARE	0.9857	0.0144	20	4085	2980	7083
ENDOCRINOLOGY & METABOLISM	0.7665	0.8530	81	14746	9159	36134
ENERGY & FUELS	0.9776	0.0062	58	7453	4776	1984
ENGINEERING	0.9722	0.0059	56	6093	3882	1656
ENGINEERING, AEROSPACE	0.9413	0.0164	25	5508	2332	1200
ENGINEERING, BIOMEDICAL	0.6047	0.6676	41	3421	3277	3858
ENGINEERING, CHEMICAL	0.9456	0.0216	105	19300	10323	7786
ENGINEERING, CIVIL	0.9802	0.0056	58	6891	4963	2118
ENGINEERING, ELECTRICAL & ELECTRONIC	0.9665	0.0183	193	28154	20557	17432
ENGINEERING, ENVIRONMENTAL	0.9152	0.0785	27	3028	2211	7853
ENGINEERING, GEOLOGICAL	0.8312	0.0151	14	465	522	348
ENGINEERING, INDUSTRIAL	0.9751	0.0052	24	3023	1757	657
ENGINEERING, MANUFACTURING	0.9637	0.0040	26	1781	1441	619
ENGINEERING, MARINE	0.9464	0.0113	16	1744	774	360
ENGINEERING, MECHANICAL	0.8964	0.0332	87	9566	5519	3467
ENGINEERING, PETROLEUM	0.9310	0.0241	21	3736	1358	242

(continued on next page)

Table 1. (cont.)

Subject Category	R^2	$Diff^2$	Journals	Papers	Articles	Refs. 1997→95–96
ENTOMOLOGY	0.7878	0.2064	64	4104	3880	3562
ENVIRONMENTAL SCIENCES	0.9912	0.0047	117	13593	11550	19179
ENVIRONMENTAL STUDIES	0.9774	0.0065	42	2255	1222	1061
ERGONOMICS	0.8668	0.0152	14	851	656	292
ETHNIC STUDIES	0.9837	0.0065	6	463	133	34
FAMILY STUDIES	0.9512	0.0148	33	1579	869	969
FISHERIES	0.9787	0.0060	28	2691	2004	1947
FOOD SCIENCE & TECHNOLOGY	0.9653	0.0128	87	9247	8520	8176
FORESTRY	0.9481	0.0130	30	2512	2251	1895
GASTROENTEROLOGY & HEPATOLOGY	0.9088	0.2045	42	16546	6160	19075
GENETICS & HEREDITY	0.7347	1.9762	90	15240	10706	52628
GEOCHEMISTRY & GEOPHYSICS	0.9786	0.0133	44	5698	4401	6385
GEOGRAPHY	0.9759	0.0077	50	3063	1424	1797
GEOLOGY	0.9371	0.0198	32	1742	1325	1965
GEOSCIENCES, INTERDISCIPLINARY	0.9758	0.0168	104	9814	8071	8400
GERIATRICS & GERONTOLOGY	0.9787	0.0133	43	3571	2030	4808
HEALTH CARE SCIENCES & SERVICES	0.8740	0.0515	16	735	372	1207
HEALTH POLICY & SERVICES	0.9852	0.0076	39	2527	1163	2088
HEMATOLOGY	0.9375	0.1406	59	21356	7642	35080
HISTORY	0.9852	0.0044	17	23136	551	77
HISTORY & PHILOSOPHY OF SCIENCE	0.9378	0.0053	29	1706	497	325
HISTORY OF SOCIAL SCIENCES	0.9837	0.0049	19	1640	321	361
HORTICULTURE	0.9360	0.0074	10	1128	1255	579
IMAGING SCIENCE & PHOTOGRAPHIC TECHNOLOGY	0.8596	0.0807	11	1119	712	750
IMMUNOLOGY	0.6808	2.4342	117	20021	16443	70933
INDUSTRIAL RELATIONS & LABOR	0.9367	0.0164	17	1116	577	210
INFECTIOUS DISEASES	0.9757	0.0216	34	6996	5491	18422
INFORMATION SCIENCE & LIBRARY SCIENCE	0.9731	0.0109	56	12441	2791	1169
INSTRUMENTS & INSTRUMENTATION	0.9946	0.0028	46	9375	7836	6646
INTERNATIONAL RELATIONS	0.8834	0.1566	50	3626	1420	769
LANGUAGE & LINGUISTICS	0.8927	0.0614	40	5165	936	818
LAW	0.9857	0.0166	106	4915	2686	3030
LIMNOLOGY	0.9496	0.0299	12	1203	921	6318
MANAGEMENT	0.9759	0.0160	60	4464	2511	1951
MARINE & FRESHWATER BIOLOGY	0.9167	0.0253	68	6262	5233	6997
MATERIALS SCIENCE, BIOMATERIALS	0.8272	0.0253	9	949	1352	1161
MATERIALS SCIENCE, CERAMICS	0.9579	0.0125	14	3213	2065	1579
MATERIALS SCIENCE, CHARACTERIZATION & TESTING	0.9382	0.0049	14	1556	850	196
MATERIALS SCIENCE, COATINGS & FILMS	0.9788	0.0088	13	2919	2603	4507
MATERIALS SCIENCE, COMPOSITES	0.7897	0.0265	13	1172	1167	576
MATERIALS SCIENCE, MULTIDISCIPLINARY	0.9288	0.0851	111	23771	18011	24911
MATERIALS SCIENCE, PAPER & WOOD	0.9414	0.0087	16	2413	1450	486
MATERIALS SCIENCE, TEXTILES	0.9078	0.0044	8	956	411	201
MATHEMATICS	0.9615	0.0082	136	12056	10520	3972
MATHEMATICS, APPLIED	0.9737	0.0051	117	10624	8958	5431
MATHEMATICS, MISCELLANEOUS	0.9320	0.0269	18	763	675	590
MECHANICS	0.8394	0.1109	79	7494	6364	5802
MEDICAL INFORMATICS	0.9689	0.0080	19	1857	1260	1344
MEDICAL LABORATORY TECHNOLOGY	0.9720	0.0235	21	3349	1851	6825
MEDICINE, GENERAL & INTERNAL	0.6644	2.6064	100	32051	13954	38600
MEDICINE, LEGAL	0.9817	0.0067	25	1718	1096	1338
MEDICINE, RESEARCH & EXPERIMENTAL	0.7189	2.0991	61	10467	6300	30415

(continued on next page)

Table 1. (cont.)

Subject Category	R^2	$Diff^2$	Journals	Papers	Articles	Refs. 1997→95–96
METALLURGY & METALLURGICAL ENGINEERING	0.8664	0.1288	53	8875	6021	5098
METEOROLOGY & ATMOSPHERIC SCIENCES	0.9755	0.0128	37	5565	5110	6496
MICROBIOLOGY	0.8077	0.6502	73	11419	10038	32091
MICROSCOPY	0.9599	0.0099	10	977	905	1135
MINERALOGY	0.9649	0.0164	22	1594	1350	1777
MINING & MINERAL PROCESSING	0.9214	0.0151	20	1608	1245	799
MULTIDISCIPLINARY SCIENCES	0.5193	7.0519	56	21206	10778	34699
MYCOLOGY	0.9392	0.0237	15	1126	978	1584
NEUROSCIENCES	0.8844	0.3785	150	27249	19035	78103
NUCLEAR SCIENCE & TECHNOLOGY	0.7464	0.2774	34	8069	7026	6689
NURSING	0.7958	0.0755	41	3482	1557	1825
NUTRITION & DIETETICS	0.9477	0.0550	50	5212	4295	10015
OBSTETRICS & GYNECOLOGY	0.9535	0.0266	54	7858	5957	11941
OCEANOGRAPHY	0.9778	0.0153	42	3897	2298	2634
ONCOLOGY	0.9420	0.1276	102	18382	14636	59866
OPERATIONS RESEARCH & MANAGEMENT SCIENCE	0.9635	0.0048	40	3528	2790	1779
OPHTHALMOLOGY	0.9237	0.0730	39	11499	4304	8331
OPTICS	0.8999	0.0975	45	10198	9105	13687
ORNITHOLOGY	0.9848	0.0036	14	914	798	830
ORTHOPEDICS	0.9760	0.0094	37	5245	4108	3616
OTORHINOLARYNGOLOGY	0.9806	0.0042	26	3606	3188	3056
PALEONTOLOGY	0.9358	0.0294	24	1338	965	1242
PARASITOLOGY	0.8743	0.1538	21	2205	1976	4346
PATHOLOGY	0.9825	0.0204	66	9183	6425	19203
PEDIATRICS	0.9848	0.0098	66	12062	7655	13008
PERIPHERAL VASCULAR DISEASE	0.9659	0.0821	38	12608	5336	21343
PHARMACOLOGY & PHARMACY	0.7833	0.6630	157	25189	20653	58171
PHILOSOPHY	0.9595	0.0122	21	6334	477	207
PHYSICS, APPLIED	0.9427	0.0293	62	22464	18452	32479
PHYSICS, ATOMIC, MOLECULAR & CHEMICAL	0.9677	0.0378	31	9040	8914	22909
PHYSICS, CONDENSED MATTER	0.5934	1.6061	45	19144	16660	35657
PHYSICS, FLUIDS & PLASMAS	0.9169	0.0764	19	4536	4633	9521
PHYSICS, MATHEMATICAL	0.8863	0.0284	25	5467	5278	11113
PHYSICS, MULTIDISCIPLINARY	0.8384	0.5570	63	19153	16867	35538
PHYSICS, NUCLEAR	0.9661	0.0367	21	5261	5380	17017
PHYSICS, PARTICLES & FIELDS	0.9477	0.0467	16	5590	5631	20040
PHYSIOLOGY	0.7480	1.2927	65	11423	6578	17252
PLANNING & DEVELOPMENT	0.9727	0.0058	37	2550	1394	749
PLANT SCIENCES	0.7500	0.7197	139	15030	12920	27190
POLITICAL SCIENCE	0.9727	0.0086	73	9178	3210	1098
POLYMER SCIENCE	0.9062	0.0677	52	11258	9563	14091
PSYCHIATRY	0.8990	0.1575	114	13722	7332	18569
PSYCHOLOGY	0.9307	0.0918	132	11765	5577	6326
PSYCHOLOGY, APPLIED	0.9750	0.0090	50	2377	1378	1416
PSYCHOLOGY, BIOLOGICAL	0.7217	0.5545	16	1966	924	1618
PSYCHOLOGY, CLINICAL	0.9882	0.0072	85	4722	2953	4716
PSYCHOLOGY, DEVELOPMENTAL	0.9797	0.0122	49	2866	1844	2703
PSYCHOLOGY, EDUCATIONAL	0.9793	0.0061	39	1512	971	999
PSYCHOLOGY, EXPERIMENTAL	0.9924	0.0059	65	3863	2722	4342
PSYCHOLOGY, MATHEMATICAL	0.6820	0.1121	11	517	387	306
PSYCHOLOGY, PSYCHOANALYSIS	0.9737	0.0080	13	957	446	482
PSYCHOLOGY, SOCIAL	0.9510	0.0269	40	2231	1761	2091
PUBLIC ADMINISTRATION	0.9868	0.0027	24	1169	524	324
PUBLIC, ENVIRONMENTAL & OCCUPATIONAL HEALTH	0.9886	0.0073	115	12147	8022	13654

(continued on next page)

Table 1. (cont.)

Subject Category	R^2	$Diff^2$	Journals	Papers	Articles	Refs. 1997→95–96
RADIOLOGY, NUCLEAR MEDICINE & MEDICAL IMAGING	0.8216	0.3477	75	15720	9887	18126
REHABILITATION	0.9730	0.0075	57	3589	1994	2731
REMOTE SENSING	0.9831	0.0060	8	1043	1197	1121
REPRODUCTIVE SYSTEMS	0.9146	0.0494	21	3801	2477	7264
RESPIRATORY SYSTEM	0.9730	0.0263	23	4833	4197	9898
RHEUMATOLOGY	0.8720	0.1368	18	4581	1947	5761
SOCIAL ISSUES	0.9696	0.0069	33	2409	922	370
SOCIAL SCIENCES, BIOMEDICAL	0.9755	0.0053	20	1320	749	1058
SOCIAL SCIENCES, INTERDISCIPLINARY	0.8862	0.0541	56	2846	1502	1033
SOCIAL SCIENCES, MATHEMATICAL METHODS	0.9715	0.0071	25	1173	888	548
SOCIAL WORK	0.9650	0.0081	32	1897	819	775
SOCIOLOGY	0.8220	0.1462	95	5722	2217	1668
SPECTROSCOPY	0.9888	0.0099	35	5587	5255	9538
STATISTICS & PROBABILITY	0.9915	0.0038	53	4205	3282	1931
SUBSTANCE ABUSE	0.9238	0.0243	24	1949	1338	2367
SURGERY	0.9761	0.0195	114	24218	17879	26331
TELECOMMUNICATIONS	0.9681	0.0159	38	5070	3230	1709
THERMODYNAMICS	0.9448	0.0093	32	2954	2497	1912
TOXICOLOGY	0.5871	1.0898	63	6156	5745	12568
TRANSPLANTATION	0.9742	0.0131	9	4140	3575	7297
TRANSPORTATION	0.9072	0.0143	21	671	495	233
TROPICAL MEDICINE	0.9513	0.0137	16	1450	1270	1572
URBAN STUDIES	0.9907	0.0018	26	1410	643	554
UROLOGY & NEPHROLOGY	0.9511	0.0495	37	12078	6011	16701
VETERINARY SCIENCES	0.9631	0.0106	104	12491	8866	8874
VIROLOGY	0.9796	0.0199	24	3965	3695	20228
WATER RESOURCES	0.9443	0.0152	44	4918	4106	8241
WOMEN'S STUDIES	0.9737	0.0055	20	1497	543	404
ZOOLOGY	0.9824	0.0085	119	6671	6753	10730

Another thing that was observed was the strong correlation between the Knowledge Export and the Knowledge Import rates, i.e. between the percentage of citations received from other Subject Categories (with respect to all the citations received) and the percentage of references to other Subject Categories (with respect to the all the references). This could be interpreted in terms of the unequal division of science that was made in establishing the Subject Category classification. While some categories correspond to totally independent disciplines of an appropriate size, other disciplines, because of their size, have had to be divided into different Subject Categories, even though they still maintain a fairly close relationship with each other.

Such is the case, for example, of Medicine, which has had to be divided into different categories, but which are still interrelated, and therefore have a flow of incoming and outgoing knowledge with the rest of the Medicine Subject Categories. In this case too, they are not only related to others of Medicine, but to all the Life Science Categories in general, and to those which contribute methods and tools that are applied in Medicine such as Mathematics, Physics, and Chemistry. Figure 7 shows the Knowledge Export and Import rates ordered by the Export rates.

The Importation rate tends to be slightly superior to the exportation rate; this is due to the effect of the formula of both rates which do not imply equality and of the journals which are assigned doubles or triples categories.

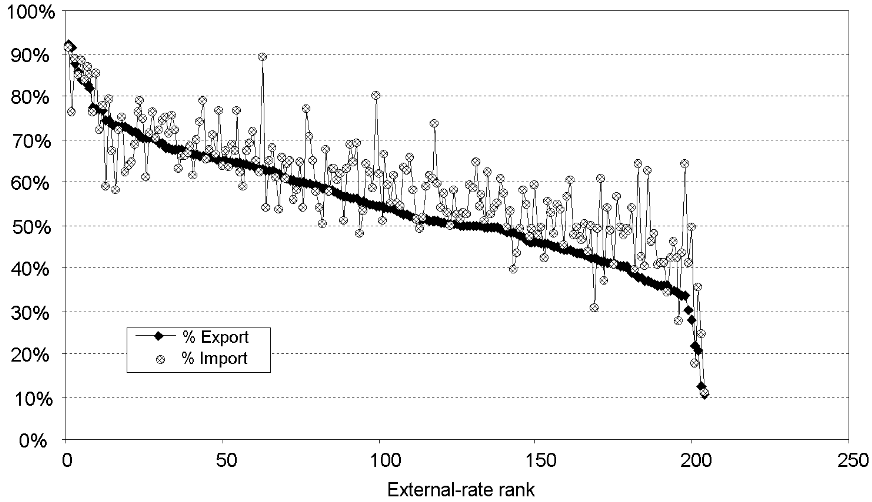

Figure 7. Knowledge Export rate vs Knowledge Import rate, by decreasing order of the Export rate

Returning to the subject of the JIF distributions, in order to attempt to determine the factors involved in a category fitting the logarithmic approximation or, because of a sharper peak, being distant from it, we calculated the correlations between different indices. Those whose absolute value was greater than 0.65 are listed in Table 2 (essentially symmetric). We have cut much higher than the statistical significance for a sample of more than 200 cases to highlight the highest correlations.

Our first comment about this table has to be the strong correlation (a coefficient of 0.85) already remarked on above between Knowledge Import and Knowledge Export. This indicates, as we observed, that Subject Categories with a high export rate also have associated a high import rate, and *vice versa*.

Continuing with the following column, one observes that the External JIF was strongly correlated with the Mean JIF, its Standard Deviation (SD), the two coefficients of the logarithmic approximation, and the Squared Deviation. This indicates that, in a Subject Category, as the External JIF increases, as is logical, so does the Mean JIF and its SD, and the slope of the logarithmic approximation, and, above all, the fit becomes poorer, with the distributions having a sharper descent and longer tails.

Table 2. Coefficients of correlation between the different indices calculated

	% Export	% Import	External Impact	Internal Impact	Mean JIF	S.D. JIF	S.D. JIF/Mean JIF	a	b	R^2	E^2/ Mean JIF
Total Cites				0.66	0.67	0.77		–0.74	0.83		
Self-Cites				0.71					0.74		
Citables									0.69		
Reviews						0.68		–0.66	0.77		
Self-Citers						0.66			0.74		
References				0.70	0.70	0.74		–0.73	0.83		
Journals											
%Export	1.00	0.85									
%Import	0.85	1.00									
External Impact			1.00		0.77	0.84		–0.84	0.79		0.75
Internal Impact				1.00	0.83	0.71		–0.74	0.75		
Mean JIF			0.77	0.83	1.00	0.89		–0.93	0.90		
S.D. JIF			0.84	0.71	0.89	1.00	0.78	–0.99	0.98		0.75
S.D. JIF/Mean JIF						0.78	1.00	–0.73	0.72	–0.68	0.80
a			–0.84	–0.74	–0.93	–0.99	–0.73	1.00	–0.97		–0.68
b			0.79	0.75	0.90	0.98	0.72	–0.97	1.00		0.65
R^2							–0.68			1.00	–0.70
E^2/Mean JIF			0.75			0.75	0.80	–0.68	0.65	–0.70	1.00

The Internal JIF was strongly correlated with the Number of References of the category. This was to be expected, since it has long been known (PINSKI & NARIN, 1976) that the JIF is directly related to the typical size of the list of references of the discipline and of the journal itself. We here observed, as was logical, that this relationship was with the internal component of the JIF. Compared with the External JIF, the correlation of the Internal JIF is stronger with the Mean JIF but weaker with its SD, and also weaker (but still fairly strong) with the two coefficients of the logarithmic approximation.

Although it is not shown in the table, the correlation between external or internal citation rate and total citation rate is high (higher than 0.9). This data is not shown in the table because we think it is obvious, since the former are components of the latter, and moreover an important correlation exist between external and internal citation rate (higher than 0,8).

As well as with the Total of Citations Received, the Mean JIF was correlated with the Number of References of the category, with the Internal and External JIFs as was seen above, with the SD, and very strongly with the coefficients of the logarithmic approximation (as was to be expected given what is indicated by these two coefficients).

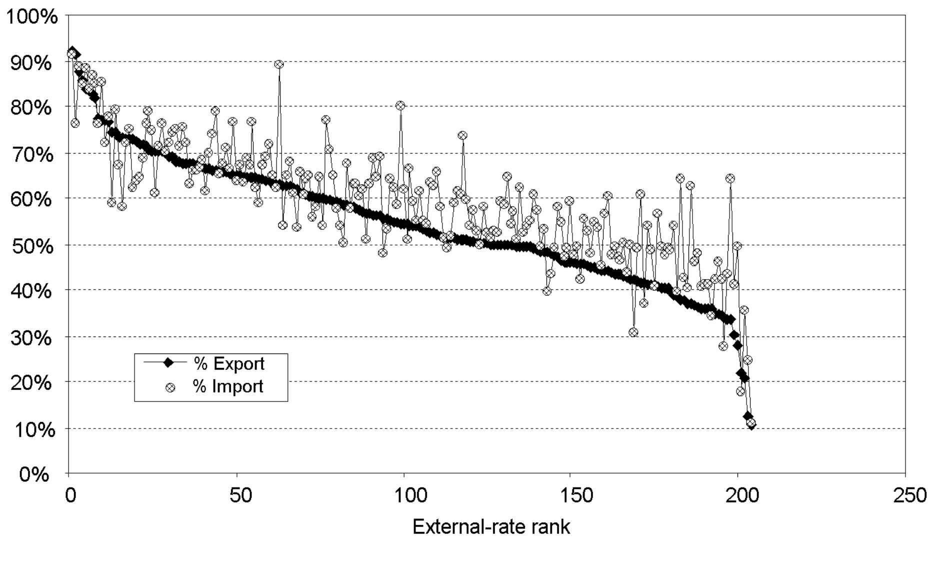

Figure 8. E2/Mean JIF of the different Subject Categories vs other related indices

The absolute SD of the JIF was strongly correlated with many indices. However, its value relative to the Mean JIF was only correlated with the coefficients of the logarithmic approximation, with the Squared Deviation (indicating that the greater the deviation with respect to the mean, the less the distribution approximates the logarithmic), and negatively with the R^2 coefficient.

The coefficients of the logarithmic approximation were strongly correlated with most of the indices, and also with each other (a correlation coefficient of 0.97), indicating that the height of the peak and the slope of the approximation usually increase together.

The R^2 coefficient had a strong correlation curiously with the SD as we remarked above, and with the Squared Deviation – in both cases negative.

The Squared Deviation (normalized to the Mean JIF), which is an indicator of the fit to the logarithmic approximation, was strongly correlated with the SD of the JIF, as is logical since the greater the dispersion of the JIFs around the mean, the greater also will be the deviations from the logarithmic approximation. It also had a strong correlation with the coefficients of the logarithmic approximation, indicating that the fits were poorer in the distributions with a steeper slope. It had a negative correlation with R^2, and, as we have indicated above, a strong correlation with the External JIF. Figure 8 shows plots of the Squared Deviation vs the External JIF, the SD of the JIF normalized to the Mean JIF, and R^2.

Conclusions

The first conclusion is that, as was to be expected, the rank/JIF distributions in each of the Subject Categories fit a negative power law fairly well. This could be, among other things, due to the Matthew effect – i.e., a citation to a paper or a journal increases the probability of its being cited again.

While all the Subject Categories follow a negative power law, some fit less well. This is because they have a sharper descent than the exponential and a longer tail. These poorer fits were locatable mainly by means of the Squared Deviation (normalized with respect to the Mean JIF) resulting from the logarithmic approximation.

The factors with the greatest influence on this deviation were the External JIF and the SD of the JIF (normalized with respect to the Mean JIF). The External JIF in turn had a major influence on the SD of the JIF, from which one can infer that the External JIF is concentrated in very few journals. In particular, the effect of this concentration in Subject Categories with a high External JIF is that the initial descent of the rank/JIF distribution is sharper than the exponential, which leads to a greater SD.

As well as with the Mean JIF, the coefficient *a* (corresponding to the slope of the power law) had a strong correlation with the External JIF and with the SD. As was logical, there was also a strong interdependence with the offset of the curve, *b*.

There also stood out the correlation between a category's Number of References and its Internal JIF, confirming the observations of previous workers (PINSKY & NARIN, 1976)

Finally, it seemed to us that the correlation between the Knowledge Export and Import rates was particularly worthy of note. This indicates that there are Subject Categories which are more independent, importing and exporting little knowledge, and others with greater flows of knowledge across subject boundaries.

In sum, scientific categories can be seen as icebergs, whose tips above the surface can on occasions be sighted in the distance from some other category (iceberg), while at other times the emerged part hardly exists and the iceberg (the category) is only visible being on (or in) it. This metaphor would explain how some scientific disciplines are exporters of ideas, because the knowledge that is generated within them is visible from other disciplines that then import it. Other disciplines are generators of knowledge for self-consumption, although this is not incompatible with the possibility of occasionally importing knowledge from other disciplines. In any case, continuing with our metaphor, the more the iceberg emerges, the more possibilities exist that, being on that iceberg, the tips of other icebergs will be seen. What is more visible from the exterior also allows a better view of the surroundings from the interior.

*

This work was financed by the Junta de Extremadura – Consejería de Educación Ciencia & Tecnología and the Fondo Social Europeo as part of research project 2PR02A041, and by the Plan Nacional de Investigación Científica, Desarrollo e Innovación Tecnológica 2004-2007 and the Fondo Europeo de Desarrollo Regional (FEDER) as part of research projects SEJ-2004-08358-C02-01 and SEJ2004-08358-C02-02.

References

BARABÁSI, A.-L. (2002), *Linked: The New Science of Networks*. Cambridge, MA: Perseus Publishing.

BARABÁSI, A.-L., JEONG, H., NEDA, Z., RAVASZ, E., SCHUBERT, A., VICSEK, T. (2002). Evolution of the social network of scientific collaborations. *Physica A*, 311 (3–4) : 590–614.

BROOKS, T. A. (1985), Private acts and public objects: an investigation of citer motivations. *Journal of the American Society for Information Science*, 36 (4) : 223–229.

GUERRERO-BOTE, V., ZAPICO-ALONSO, F., ESPINOSA-CALVO, M. E., GÓMEZ-CRISÓSTOMO, R., MOYA-ANEGÓN, F. (2006). *Binary Pathfinder: An improvement to the Pathfinder Algorithm.* (in press).

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 & Public Policy*, 27 (1) : 23–36.

LEYDESDORFF, L., BENSMAN, S. (2006), Classification and powerlaws: The logarithmic transformation. *Journal of the American Society for Information Science and Technology* (in press).

MERTON, R. K. (1968), The Matthew Effect in science. *Science*, 159 : 56–63.

MOYA-ANEGÓN, F., VARGAS-QUESADA, B., HERRERO-SOLANA, V., CHINCHILLA RODRÍGUEZ, Z., CORERA ÁLVAREZ, E., MUÑOZ FERNÁNDEZ, F. J. (2004), A new technique for building maps of large scientific domains based on the cocitation of classes and categories. *Scientometrics*, 61 : 129–145.

MOYA-ANEGÓN, F., CHINCHILLA RODRÍGUEZ, Z., CORERA ÁLVAREZ, E., HERRERO-SOLANA, V., MUÑOZ FERNÁNDEZ, F. J., VARGAS-QUESADA, B. (2005), *Indicadores bibliométricos de la actividad científica española-2004*. FECYT: Madrid.

PICHAPPAN, P. (1995), A dual refinement of journal self-citation measures. *Scientometrics*, 33 (2) :13–21.

PINSKI, G., NARIN, F. (1976), Citation influence for journal aggregates of scientific publications: Theory, with application to the literature of physics. *Information Processing and Management*, 12 : 297–312.

PODLUBNY, I. (2005), Comparison of scientific impact expressed by the number of citations in different fields of science. *Scientometrics*, 64 (1) : 95–99.

PUDOVKIN, A., GARFIELD, E. (2002), Algorithmic procedure for finding semantically related journals. *Journal of the American Society for Information Science and Technology*, 53 (13) : 1113–1119.

SCHVANEVELDT, R. W., DEARHOLT, D. W., DURSO, F. T. (1988), Graph theoretic foundations of Pathfinder networks. *Computers & Mathematics with Applications*, 15 (4) : 337–345.

SOMBATSOMPOP, N., MARKPIN, T. (2005), Making an equality of ISI impact factors for different subject fields. *Journal of the American Society for Information Science and Technology*, 56 (7) : 676–683.

STEGMANN, J., GROHMANN G. (2001), Citation rates, knowledge export and international visibility of dermatology journals listed and not listed in the Journal Citation Reports. *Scientometrics*, 50 (3) : 483–502.

VAN LEEUWEN, T. N., MOED, H. F. (2002), Development and application of journal impact measures in the Dutch science system. *Scientometrics*, 53 (2) : 249–266.

VAN LEEUWEN, T. N., VISSER, M. S., MOED, H. F., NEDERHOF, T. J., VAN RAAN, A. F. J. (2003), The Holy Grail of science policy: Exploring and combining bibliometric tools in search of scientific excellence. *Scientometrics*, 57 (2) : 257–280.

WAGNER, C. S., LEYDESDORFF, L. (2006), Network Structure, Self-organization and the Growth of International Collaboration in Science Research Policy (in press).

WORMELL, I. (1998), Informetric analysis of the international impact of scientific journals: How 'international' are the international journals? *Journal of Documentation*, 54 (5) : 584–605.