Redefining the Field of Economics: Improving Field Normalization for the Application of Bibliometric Techniques in the Field of Economics.

Thed N. van Leeuwen and Clara Calero Medina

leeuwen@cwts.leidenuniv.nl and clara@cwts.leidenuniv.nl
Leiden University, Centre for Science & Technology Studies (CWTS)
Wassenaarseweg 62a, Postbus 905, 2300 AX Leiden (The Netherlands)

Abstract

Field normalization, and its effect of bibliometric indicators is a widely discussed topic among bibliometricians. It is not the necessity of field normalization around which the debate evolves, but the way how to field normalize bibliometric indicators. In this paper the authors present the results of a study in which publication data of a large disciplinary database in economics (EconLit) is combined with the multidisciplinary citation indexes produced by Thomson Scientific. Main purpose of the study is to investigate whether it would be possible to combine the classification scheme of the economics database with the advantages of the citation indexes (both multiple addresses and citation data), in order to improve the possible applicability of the citation indexes in research performance studies in the field of economics and its periphery. The authors show the starting points of both database, the outcome of the matching and combining of both sets of publications, the effects of EconLit field classification in terms of differences in impact levels. The study clearly shows that research performance exercises conducted in the field of economics would benefit from the labeling of publications in the citation indexes with a more detailed classification scheme as found in EconLit.

Introduction

Field normalization has proven to be a valid addition in applications of bibliometric techniques in research evaluation exercises (Moed et al., 1995, Bassecoulard and Zitt, 1999). Large-scale applications of bibliometric techniques in research domains, very often covering scientists from different research specialties, and publishing across the boundaries of research fields, have shown that the field normalization applied in these studies made it possible to make comparisons on the level of research groups in biology, physics, or chemistry (van Leeuwen et al, 2001), while not all researchers involved necessarily belong to the same sub-domains, nor that they publish in journals belonging to the same field(s) or that these journals belong to only one field. When using the citation indexes as produced by Thomson Reuters Scientific as the main sources for both publication as well as citation data, the basis for the field normalization can be the classification by Thomson Scientific of the journals covered in the citation indexes into so-called Journal Subject Categories, one of the classification approaches discussed in detail by Zitt et al. (2005). This journal subject classification system contains for example for chemistry several sub-domain descriptions, such as *Chemistry, analytical*, *Chemistry, Physical*, and so on as well as for Physics, with fields like *Physics, condensed matter*, *Physics, applied*, etc. Although not a perfect system, it is currently the only classification system fitting the multidisciplinary nature of the citation indexes. However, in some domains, such a sub-domain description is missing, which can be an obstacle in proper comparisons of research performance, both in terms of output measurement as well as in terms of impact analysis. One of these fields is the field of economics research within the domain of the social sciences. One can distinguish various attempts to apply bibliometrics in this field, most of them using journal impact factors for either journal ranking comparisons (Laband D et al. (1985), Kalaitzidakis P et al. (1999)), department comparisons (Graves et al. (1982), Conroy and Dusansky (1995), Dusansky and Vernon (1998)), or even country comparisons (Harzing AW (2005), Donovan C and Butler L (2008)). An important problem of these approaches is the dependence on journal impact measures for ranking and evaluation.

As journal impact factors are an invalid tool for serious evaluation purposes, one should better avoid applying this indicator for this specific purpose (van Leeuwen and Moed, 2002). An extensive study for the EU, describing the European landscape of the field of economics research applied bibliometric techniques beyond journal impact measures (EUR 20889, 2004), thereby showing the actual impact of European researchers in the field of economics, compared to expected journal and field impact values. The progress of this study compared to previous work consisted of the merging of EconLit information with that of the Social Sciences Citation Index (SSCI), thereby allowing comparisons across the various disciplines in economics based on the EconLit classification of scientific research papers. So although this extensive study succeeded in defining economics through specialties identified in EconLit (the output dimension), it still suffered from the problem of comparing the impact against the classification system applied in the SSCI (the impact dimension). By merging high impact and low impact research specialties together into one field, Economics, one runs the risk of comparing apples with oranges, and although all fruit, a better perspective on research activity and journal performance across specialties in a field is necessary, especially in the light of possible applications for research assessment procedures, which become more and more important in terms of accountability, and consequently, in terms of funding.

Methodology

In this study, we used two sources of information. The first source consists of the citations indexes, nowadays better known as the Web of Science (WoS), the internet version of the combined Science Citation Index, the Social Sciences Citation Index, and the Arts & Humanities Citation Index. Most important features of the citation indexes are the multidisciplinary nature of the system, the international orientation of thejournals processed for the citation indexes, the inclusion of references/citations, and the processing of all the addresses of publishing research institutions. Roughly, the citation indexes covered about 9.000 scholarly journals annually. These journals are classified in a system known as Journal Subject Categories. As stated in the Introduction, if one tries to identify economics research through the journal subject classification system of the citation indexes, one ends up with three main fields (*Economics*, *Business*, and *Business, Finance*), and probably a number of related fields (like *Management*, *Planning & Development*, etc.). The data from the citation indexes were retrieved from an in-house version of the citation indexes. The benefits of the citation indexes, and especially the Social Sciences Citation Index (which covers the domain of economics research and related research areas, such as business, finance, public administration, international relations, etc) like citation data, multiple addresses, international scope, was combined with the benefits of the disciplinary oriented database EconLit, in particular the classification of individual publications to 19 different specialties within the field of economics research in its broad sense. EconLit is a data system that covers in depth the field of economics research and related areas, the document types that are processed include journal publications, as well as monographs, book chapters, conference proceedings papers, etc. The analysis described in this paper covers the period 1993-2002. We collected the journal publications from EconLit, in order to combine them with the publications covered by the citation indexes, comparable to previous studies, for example in medicine (van Leeuwen et al, 2001). This focus on journal papers is not meant to be ignorant of the research output not covered by the citation indexes, and well-covered by EconLit, like books, book chapters, etc. but in this study we want to focus on journal publications, because in the field of economics, journal publications play an important role, and Journal Impact Factors, as produced by Thomson Reuters Scientific play an important role in various systems of rewarding researchers, and rankings based on journal outputs.

Our first objective in this study is that we want to focus on the general coverage of journal literature in EconLit and the citation indexes, and the overlap of both sources of bibliographic and bibliometric data. Hopefully, this should lead to improved field impact scores, thereby making more accurate descriptions and comparisons within the field of economics and related areas possible..

After matching the downloaded publication of EconLit over the period 1993-2002 with the publications covered in the citation indexes, we can apply the field identifications attached to EconLit publications. In EconLit, these field identifiers are attached on a paper to paper basis, and a paper can carry more than one identifier. Therefore, the numbers presented below are the result of a whole counting of paper-field identifier combinations, and are not an exact representation of the total number of publications matched between the citation indexes and EconLit. Likewise, for every publication matched, we can find more than one ISI Journal Subject Category attached. Papers that are the result from combining both databases are indicated below as 'CI/EconLit publications'

The indicators we present in this study are **P**, which stands for the number of journal publications, **C+sc** is the number of received citations, including self citations. **CPP** is the citation per publication ratio, or the mean impact per article. **Pnc** stands for the number of publications not cited within the period of analysis. **CPP/FCSm** is the comparison of the actual impact of set of publications, compared to the mean field impact of that set of papers. The field is determined by the journals in which the papers are published, as journals are classified to journal subject categories (in the case of the citation index data), or the field is determined by the field classification. **JCSm/FCSm** is the score indicative for the impact level of the journal compared to its field(s). Finally, the percentage of self citations (**% Self citations**) is the share of the total number of received citations by the first to n^{th} author of the citing paper to source papers in the citation index.

Results

Table 1 displays the ISI Journal Subject Categories related to the CI/EconLit publications, in a descending order of number of papers involved. Clearly, *Economics* stands out as the largest field, followed by *Business, Finance*. Next we find a number of fields related with economics research, like *Environmental Studies*, *Planning & Development*, etc. These fields appear due to the double (or more) assignments of Journal Subject Categories to a journal, thereby indicating the strong relations between economics and other fields (such as *Political Science*, in political economy, and *Social Sciences, Mathematical Applications* and *Statistics & Probability*, for the quantitatively oriented parts of economics, econometrics).

Table 1 does not contain all Thomson Scientific Journal Subject Categories, only those fields that contain more than 1% of all papers in the set are displayed.

Table 1: The field of economics in the citation indexes, as covered by EconLit, and expressed through Thomson Scientific Journal Subject categories, 1993-2002.

Journal Subject Category	P	C+sc	CPP	% Pnc	CPP/FCSm	JCSm/FCSm	% Self Citations
* ECONOMICS	31,142	158,849	4.44	27%	1.17	1.16	13%
* BUSINESS, FINA	5,393	37,302	6.26	20%	1.88	1.87	10%
* ENVIRONM STUD	2,989	13,289	3.76	22%	1.11	1.09	16%
* PLANNING & DEV	1,956	8,872	3.88	20%	1.65	1.62	14%
* INDUSTR RELAT	1,931	9,305	4.33	23%	1.31	1.29	10%
* SOC SCI,MATH M	1,901	11,463	5.07	23%	1.14	1.11	16%
* BUSINESS	1,363	10,664	6.91	20%	1.48	1.47	12%

* INTERNAT RELAT	1,298	7,191	5.03	25%	1.87	1.94	9%
* POLITICAL SC	1,097	8,128	6.64	24%	3.34	2.88	10%
AGRIC ECON&POL	1,056	3,574	2.78	26%	1.11	1.09	18%
STATISTICS&PROBA	1,051	10,679	8.67	20%	2.70	2.42	15%
* URBAN STUDIES	971	4,521	3.94	18%	1.22	1.20	16%
* DEMOGRAPHY	921	6,843	6.46	15%	1.54	1.48	13%
* LAW	862	7,134	7.51	15%	2.00	1.92	9%
* SOC SCI,INTERD	795	2,711	2.96	33%	1.38	1.27	13%
* PUBLIC ADMIN	771	2,408	2.64	33%	1.21	1.15	16%
* HISTORY SOC SC	619	1,536	2.04	32%	1.19	1.13	18%

* The asterisk relates to the inclusion of this field in the Social Sciences Citation Index.

As Table 1 shows, and this is further underlined in Figure 1, *Economics* as a Journal Subject Category covers over 50% of the CI/EconLit publications. These 31,142 CI/EconLit publications get cited overall 158,849 times in total, which leads to a mean impact score for a paper in that Journal Subject Category of 4.44 (which is corrected for self citations). Roughly one quarter of all published journal papers in that JSC is not cited at all in a ten year timeframe, and 13% of all citations are self citations. The comparison of the mean impact of publications in EconLit indexed as *Economics* with the field impact (CPP/FCSm) results in a score of 1.17, which is 17% above worldwide field average.

The JSC in which CI/EconLit publications reach the highest level of mean impact score is *Statistics & probability*, which is a field that covers journals that are processed for the Science Citation Index. The mean impact level of those 1,051 publications is 8.67, nearly twice as high as the mean impact of the largest JSC, Economics. The field-normalized impact of the scientists publishing in those journals covered in EconLit, and belonging to the much larger domain of *Statistics & probability*, in which scientists from other disciplines publish as well, is 2.70 (which is 170% above worldwide field impact level).

Figure 1 shows the relative distribution of EconLit publications over the various Journal Subject Categories. Here it immediately becomes clear that Economics as a JSC is covering most of the CI/EconLit publications, hence this is also an important drawback, as this is now transformed into one large mass of publications without any distinctions among this set of papers. So on the basis of Journal Subject Categories attached to EconLit publications, one finds large differences in impact level, but, more importantly, one gets stuck with one dominant field (Economics), covering 50% off all publications in the set without any details, while the other 50% of the publications is distributed over 51 other Journal Subject Categories. This clearly asks for some clarification. The results below will provide this clarification.
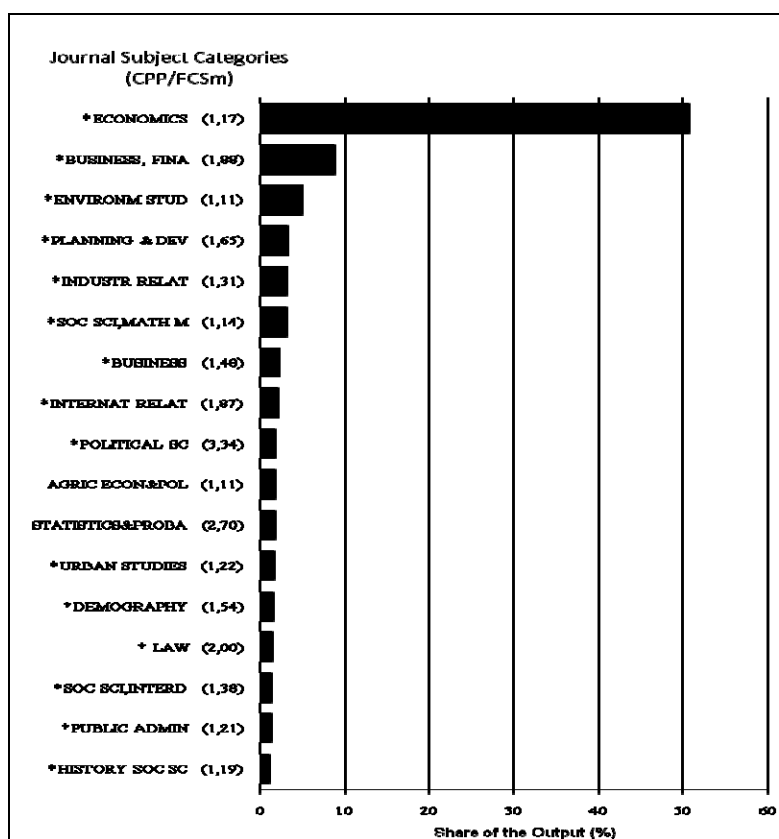
Figure 1: Distribution of EconLit publications of Journal Subject Categories, 1993-2002 (an asterisk * indicates the SSCI origin of the JSC)

In Table 2, the CI/EconLit output is shown, here based on the descriptive elements attached to EconLit publications. Again, we calculated the full range of bibliometric indicators. As the number of EconLit fields is smaller than the number of Thomson Scientific Journal Subject Categories, Table 2 contains all publications in journals processed for both EconLit as the citation indexes.

Table 2: The field of economics in the citation indexes, as covered by EconLit, and expressed through EconLit descriptive elements, 1993-2002.

EconLit Field Indication.	P	C+sc	CPPc	%Pnc	CPP/ FCSm	JCSm/ FCSm	% Self Citations
Agr. & Nat. Res. Econ.	6,818	29,517	3.68	31%	1.08	1.06	15%
Bus. Admin. & Bus. Econ.	2,305	18,365	7.13	20%	1.83	1.66	11%
Econ. Dev., Techn. Change & Growth	6,802	37,996	4.94	31%	1.44	1.13	12%
Economic History	1,978	7,032	2.99	34%	0.97	1.18	16%
Economic Systems	3,735	15,525	3.65	34%	1.16	1.11	12%
Financial Economics	8,132	50,990	5.60	29%	1.61	1.61	11%
Gen. Econ. & Teaching	1,304	5,052	3.47	36%	0.98	1.30	10%
Health, Educ., & Welfare	4,111	25,369	5.32	25%	1.51	1.49	14%
Industrial Organization	7,287	32,347	3.86	32%	1.12	1.11	13%
International Economics	6,540	29,663	3.98	35%	1.15	1.11	12%
Labor & Demogr. Econ.	7,815	42,550	4.78	28%	1.43	1.38	12%
Law & Economics	1,500	9,067	5.35	25%	1.54	1.55	11%
Macroecon.s & Mon. Econ.	6,084	27,731	3.98	38%	1.12	1.15	13%

Math. & Quantitative Meth.	6,724	45,422	5.78	30%	1.52	1.38	14%
Microeconomics	11,698	65,180	4.81	30%	1.36	1.36	14%
Public Economics	4,000	16,396	3.50	35%	1.08	1.23	15%
Schools. Econ. Thought & Methodology	1,391	3,688	2.31	45%	0.62	0.91	13%
Urban, Rural, & Reg. Econ.	4,897	21,290	3.63	28%	1.12	1.11	16%
Other Special Topics	269	892	2.89	40%	1.34	1.28	13%

A first observation learns that the relative distribution of output over fields is much more balanced as was the case in the distribution over Journal Subject Categories. As Figure 2 shows, the largest field, Microeconomics, covers 11,698 publications, which is slightly more than 12% of the journal publications in the analysis, while the next two largest fields (Financial economics and Labor & demographic economics) cover a slightly more than 8% of the total number of journal publications. Next to this difference, we also find a resemblance with the distribution over Journal Subject Categories, namely a strong variation of the mean impact scores across the EconLit fields. While the mean impact level of the largest field (Microeconomics) is 4.81 citations per publication, the highest mean impact score is 7.13 (Business administration & business economics). The field that relates most strongly to the Journal Subject Category Statistics & probability, the EconLit field Mathematics & Quantitative methods, is the seventh ranking field in terms of output numbers covered, with an impact level that is the second highest, closely followed by Financial economics and Law & economics.

Another interesting phenomenon is the strong variation we observe for the impact indicator CPP/FCSm, where we compare the mean impact of the journal publications in the EconLit fields with 'their' field impact level. Here the field impact level is no longer based on one Journal Subject Category (as was the case in Table 1), but the underlying EconLit field impact level can be (and actually will be) constituted by a number of Journal Subject Categories. This is caused by the fact that in EconLit, publications are indexed on a paper-to-paper basis, and not, as is the case in the citation indexes, on a journal-to-journal basis. So a major improvement of this distribution of publications over EconLit fields, with the underlying variation in Journal Subject categories for the field impact level determination is the more detailed insight in impact level differences between specialties in the domain of economics research. However, one still remains dependent on the Journal Subject categorization of journals, and as we have shown above, this is strongly dominated by the Journal Subject Category Economics.
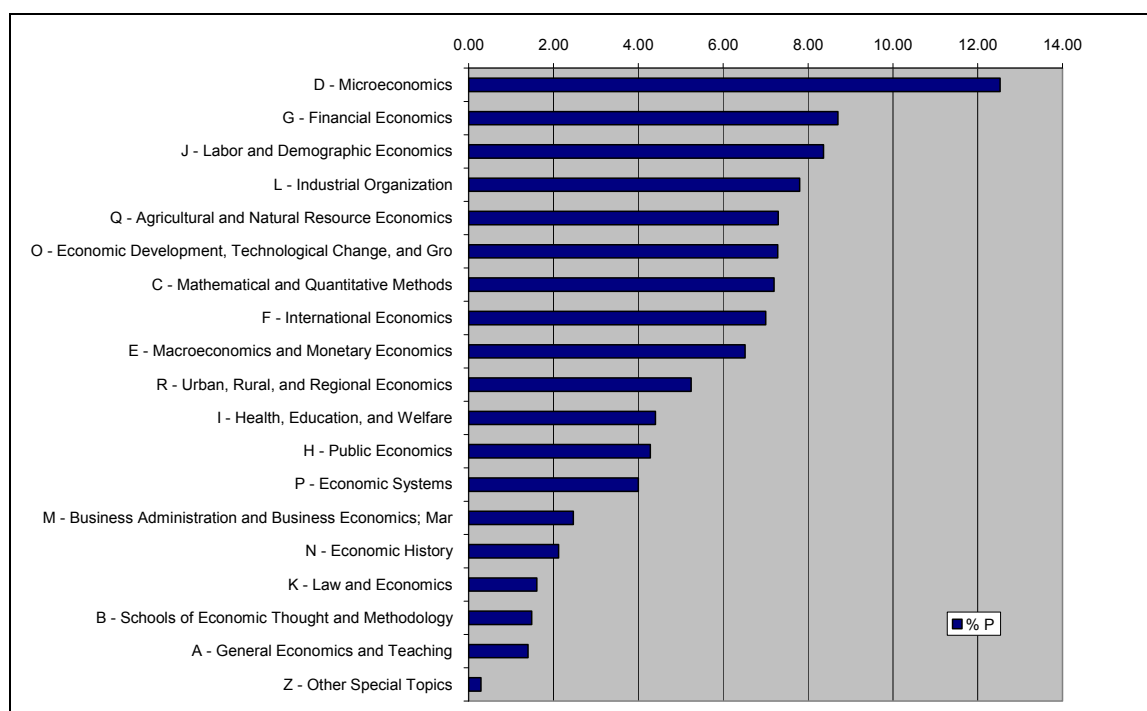
Figure 2: Distribution of EconLit publications, covered in the citation indexes in 1993-2002 over EconLit fields.

As we have observed above, the normalized impact scores across EconLit fields vary strongly. Figure 3 shows that the mean impact scores (CPP) vary strongly as well, within the range of 2.3 (Schools Econ. Thought & Methodology) to 7.1 (Bus. Admin. & Bus. Econ.). Across the EconLit fields, we distinguish four fields with an average impact score within the range of five to six citations per paper Mathematical and Quantitative Methods, Financial Economics, Law and Economics, and Health, Education & Welfare), within the range of four to five citation per publications we find three fields (Economic Development, Technological Change and Growth, Microeconomics, and Labor & Demographic Economics). We then find a sharp distinction between these seven fields, and the next following eight fields: these eight fields have an impact score between three and four citations per publication, on average.
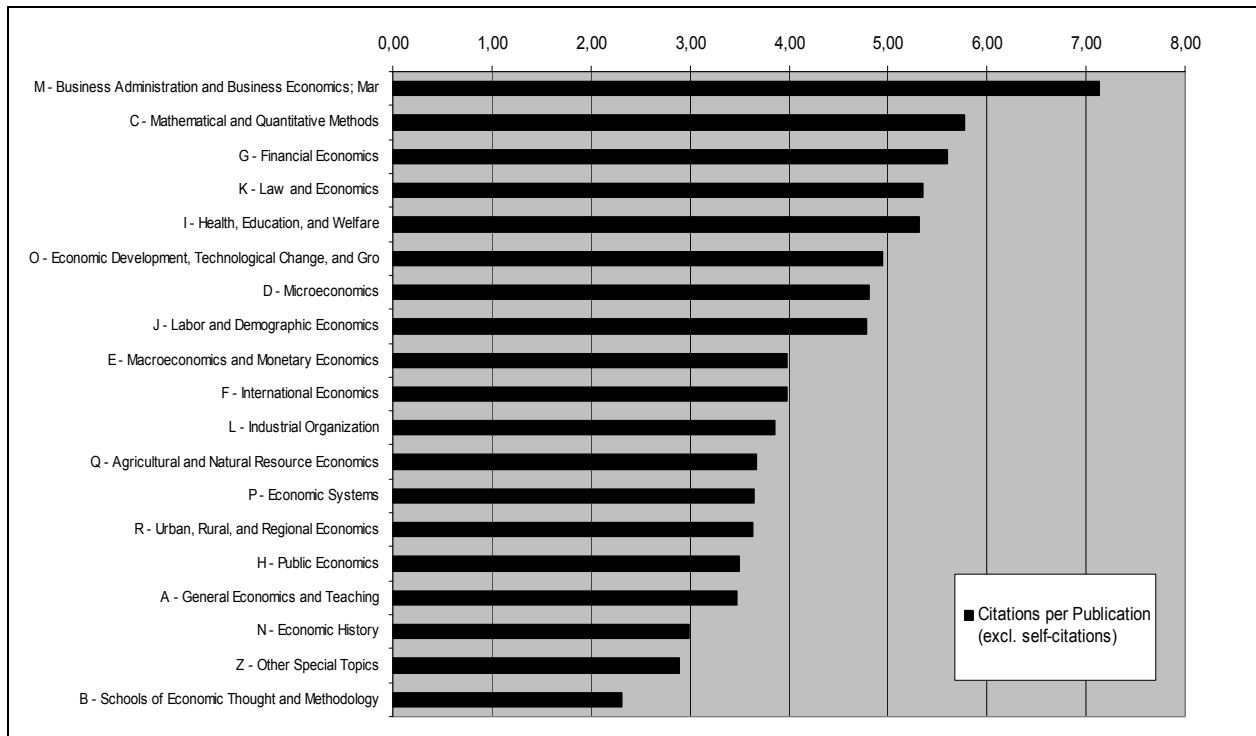
Figure 3: Overall mean impact per EconLit field, 1993-2002

If we focus on the composition of the citation index field *Economics* in terms of the EconLit fields, as displayed in Figure 4, we notice that this field consists of a wide range of economics topics, the most important being Microeconomics (with over 14% of the publications indexed as Economics belonging to that field). The next five fields, each covering roughly 8% of the publications in the citation index field, vary strongly in topic (form Macroeconomics and monetary economics, Industrial organization, International economics, Economics development, technical change and growth, and Mathematical and quantitative methods). So the first six EconLit fields cover roughly 60% of the citation index field *Economics*. On the other end of the range of EconLit fields, we notice some very small specializations in the field of economics research.
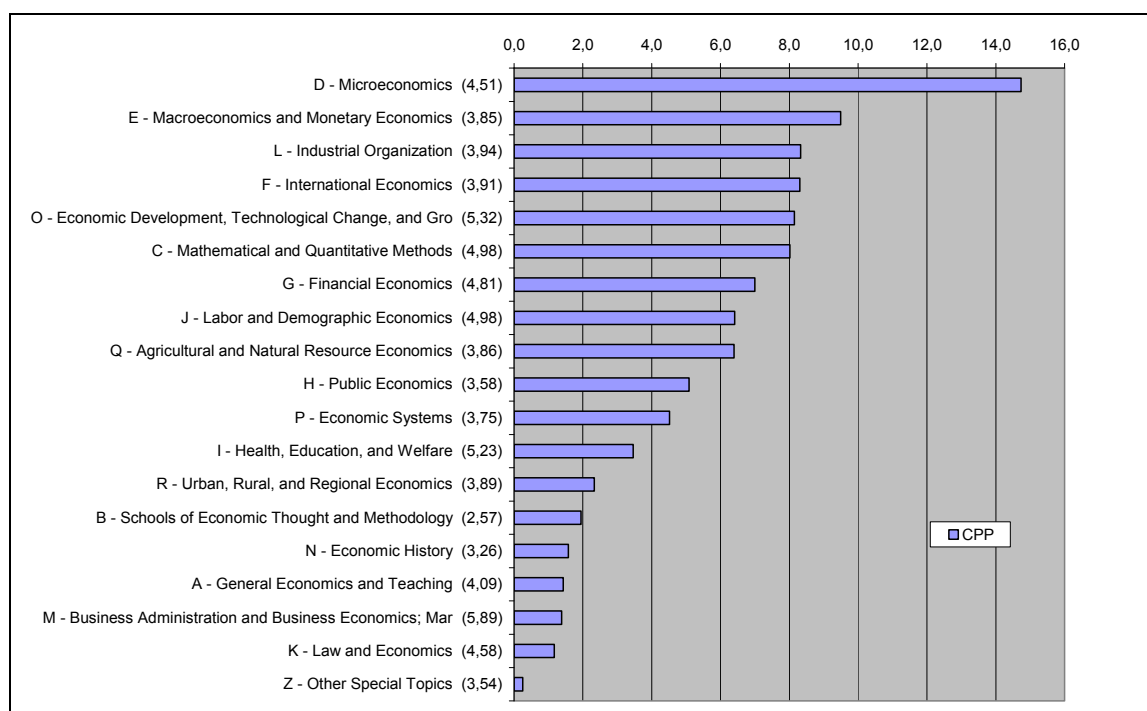
Figure 4: Composition of ISI * Economics by EconLit fields, 1993-2002

From the perspective of the output share in the field of Economics, we shift to the perspective of impact scores. The results of this analysis are shown in Figure 5. This Figure displays the descending mean citation scores for the various EconLit fields as show in Figure 4. The mean impact of the citation index field *Economics* is displayed by the line indicating the mean impact of the overall field at the level of 4.44 (see Table 1). It becomes clear that the order of the fields changes when focusing on impact scores, in comparison with the output-based order in Figure 4. We find Microeconomics at the average of the citation index field Economics, while some of the other somewhat larger fields show mean impact levels that are below the average impact score of the citation index field *Economics*, e.g. Industrial organization, International economics, Macroeconomics and monetary economics. The consequence of this situation is that, in case of a research assessment based on the citation index classification, the researchers working in this specialisation are in terms of field comparison, compared to relatively high field citation impact scores. On the other hand, we also observe some of the smaller fields such as Business administration and business economics and Health, education, and welfare with relatively high impact scores compared to the average impact of the citation impact of the citation index field of Economics. For researchers in these specialisations the current field normalization in the citation index is too low for proper comparisons, or in other words, the mean field impact is underestimated. And although Figures 4 and 5 only display the situation for the EconLit based composition of the citation index field Economics, a similar situation exists for the total set of CI/EconLit publications.
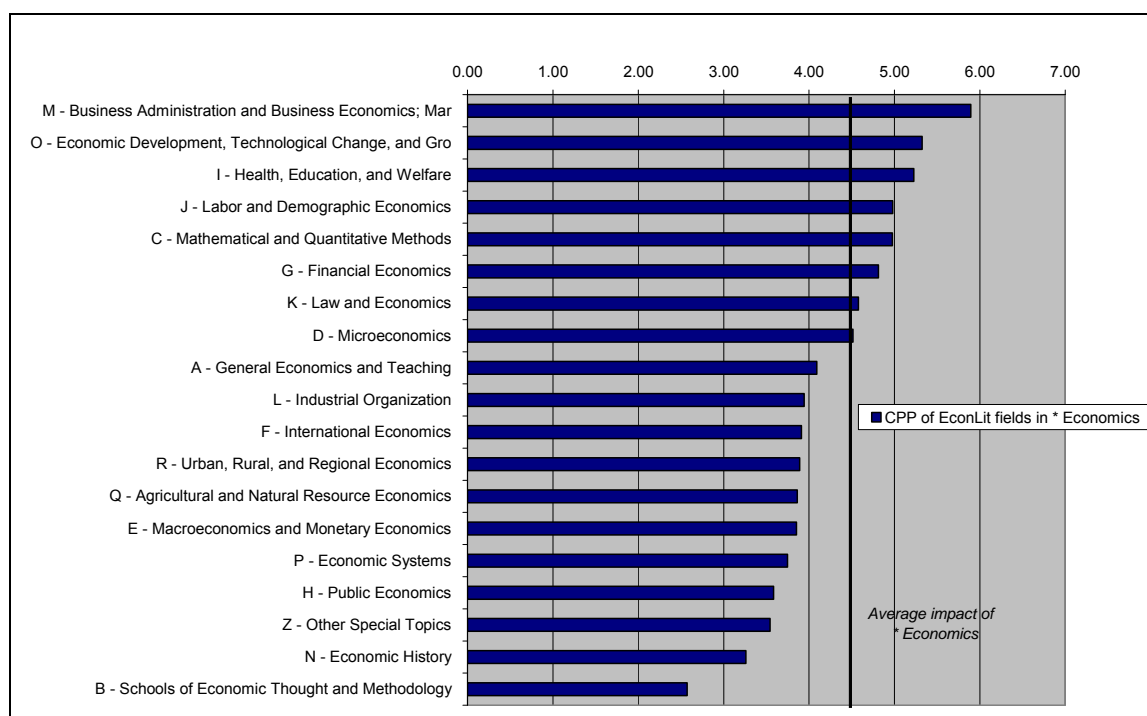
Figure 5: Mean impact values for specialties within the field of economics, based on EconLit fields, 1993-2002.

Conclusions and discussion

This study has shown that it is possible to combine the best of both worlds on a large scale. In this study, we combined the classification system of EconLit (which is attributed on a paper-to-paper level), to the publications in the citation indexes (which are classified to fields through the journal(s) in which the papers were published, on the basis of a journal-to-field system, that is, one journal belongs entirely to one or more Journal Subject Categories). And while this latter system is in itself not a mall functioning system, as it is used across all fields of scientific research, ranging from the humanities to high energy physics, the description of the whole field of economics with only one single Journal Subject Category can be seen as a serious handicap, especially in the light of possible research performance assessment exercises, in which detailed information, and distinction in research activity among groups in the assessment are crucial to make proper comparisons. Without any further distinction in terms of economic specialties, the assessment of groups in the field of economics based on only the journal subject categories in the citation indexes, suffers from the lack of a properly working field-normalization process. As this study has shown, one can observe strong differences across sub-fields in economics (e.g., the fields dealing with business aspects, and those dealing with history and philosophy of the field, where the former has an average impact three times as high as the latter specialtiers). A crucial conclusion is that if one conducts an analysis on the research performance of economics groups, based on the current filed normalization based on o Journal Subject Categories, the main part of the output will be in the field of economics, in journals classified by Thomson Scientific as '*Economics*'. In this case, the mean average impact of this larger field is determining the outcome of the performance assessment, while the fields with relative high numbers of citations per publications will probably reach relative high field-normalized impact levels, while some of the sub-fields within economics research (with relatively low citation numbers) will suffer from the fact that the mean impact within '*Economics*' is partially determined by the impact of specialties with relatively higher impact levels. In future studies we will test the application

of field-normalized impact scores, based on EconLit classification in an analysis of the research performance of academic economists. Yet another important improvement would be to create journal impact measures that are not only field normalized (contrary to the current practice in economics and related fields of using journal impact factors), but that will also be field normalized based on EconLit classification rather than that of Thomson Scientific.

References

EUR 20889 (2004), *Mapping of Excellence in Economics*, European Commission, DG Research Brussels, pages 1-227.

Bassecoulard. E and Zitt, M, (1999) *Indicators in a research institute: a multi-level classification of scientific journals*, Scientometrics, 44, 323-345

Conroy ME and R Dusansky (1995), *The productivity of economics departments in the U.S.: publications in the core journals*, Journal of Economics Literature, 33, 1966-1971

Donovan C and L Butler (2008), *Testing novel quantitative indicators of research 'quality', esteem and 'user engagement': an economics pilot study*, Research Evaluation, 16, 231-242

Dusansky, R. and CJ Vernon (1998), *Ranking of U.S. Economics departments*, Journal of Economic Perspectives, 12 , 157-170.

Graves P. J Marchand, and R Thompson (1982), *Economics departmental rankings: research incentives, constraints, and efficiency*, American Economic Review, 72, 1131-1141.

Harzing, AW (2005), *Australian research output in Economics & Business: high volume, low impact ?* The Australian Journal of Management, 30, 183-200

Kalaitzidakis P, P Mamuneas-Theofanis P, and T Stengos (1999), *European Economics: an analysis based on publications in the core journals*, European Economics review, 43, 1150-1168.

Laband, D. and MJ Pierre (1994), *The relative impact of economics journals*, Journal of Economic Literature, 32, 650-666.

Moed, HF, de Bruijn, RE, and van Leeuwen, TN, (1995) *New bibliometric tools for the assessment of national reserach performance: database description, overview of indicators and first applications*, Scientometrics, 33, 381-422

Van Leeuwen, TN, Visser MS and Moed HF, (2001), *Bibliometric Profiles of Academic Chemistry Research in the Netherlands. 1991 – 2000.* Research report to the Dutch Association of Universities (VSNU).

Van Leeuwen, TN, van der Wurff, LJ and van Raan, AFJ (2001), *The use of combined bibliometric methods in research funding policy*, Research Evaluation, 10, 195-201

Van Leeuwen, TN and HF Moed (2002), *Development and application of journal impact measures in the Dutch science system*, Scientometrics, 53, 249-266

Zitt, M, Raman-Rahary, S, and Bassecoulard, E (2005), Relativity of citation performance excellence measures: from cross-field to cross-scale of field-normalization, Scientometrics, 63, 373-401