

Citation Flows in the Zones of Influence of Scientific Collaborations

Bárbara S. Lancho Barrantes ¹

Vicente P. Guerrero Bote ¹

Zaida Chinchilla Rodríguez ²

Félix de Moya Anegón ²

1. *Departamento de Información y Comunicación, Grupo Scimago, Universidad de Extremadura, Plazuela IbnMarwan, 06071 Badajoz, Spain.*
2. *Grupo Scimago, Consejo Superior de Investigaciones Científicas, Centro de Ciencias Humanas y Sociales, Instituto de Políticas y Bienes Públicos, C/Albasanz, 26-28. 28037 Madrid, Spain.*

Abstract

Domestic citation to papers from the same country and the greater citation impact of documents involving international collaboration are two phenomena that have been extensively studied and contrasted. Here, however, we show that it is not so much a national bias, but that papers have a greater impact on their immediate environments, an impact that is diluted as that environment grows. For this reason, the greatest biases are observed in countries with a limited production. Papers that involve international collaboration have a greater impact in general, on the one hand, because they have multiple "immediate environments," and on the other because of their greater quality or prestige. In short, one can say that science knows no frontiers. Certainly there is a greater impact on the authors' immediate environment, but this does not necessarily have to coincide with their national environments, which fade in importance as the collaborative environment expands.

Introduction

Some workers have analyzed the degree of collaboration using the coauthorship of publications (Carpintero & Peiró, 1983; Alcaín & Gálvez, 1998). There is indeed some correlation between the coauthorship index and the impact of papers, so that scientific communities gain in visibility as their networks grow and bring together a greater number of associates (Moya et al., 2008). In general, high levels of collaboration lead to high levels of impact, greater quality of the papers published, and greater productivity of the authors in their particular scientific fields (Lewinson & Cunningham, 1991; Narin, Stevens, & Whitlow, 1991; Glänzel, 2001; Glänzel, 2002; Leimu & Koricheva, 2005; Katz & Hicks, 1997; Persson, Glanzel, & Danell, 2004; Hsu & Huang, 2010; Aksnes, 2003; Moya-Anegón et al., 2008; Chinchilla, Vargas, Hassan, González, & Moya, 2010)

The potential benefits of scientific collaboration may depend on the discipline. The effect of collaboration on scientific impact appears to be more positive in the "hard" sciences such as physics and astronomy, than in the "soft" sciences such as sociology or social sciences (Stack, 2002; Bandyopadhyay, 2001; Moed, Bruin, Nederhof, & Tijssen, 1991; Bridgstock, 1991), with citation behaviour sometimes differing considerably from one field to another (Lancho, Guerrero, & Moya, 2010a, 2010b). The benefits will also depend on the different types of collaboration (Leimu & Koricheva, 2005; Katz & Hicks, 1997): (a) domestic in-house collaboration (all authors from the same institution); (b) domestic institutional collaboration (all authors from the same country but from more than one institution); and (c) international collaboration (authors from more than one country; (Leimu & Koricheva). Although institutional collaboration is more important than domestic in-house collaboration, international collaboration is even more so in the sense that it increases the citation rates far above those of domestic national collaboration (Narin et al., 1991; Katz & Hicks; Goldfinch, Dale, & De Roue, 2003; Sooryamoorthy, 2009). Gómez, Fernandez, and Sebastian (1999) consider that international collaboration increases the visibility of research papers since they are published in journals of greater impact than those of national collaborations. Narin and Whitlow (1990) find evidence that papers with a multiple international authorship have double the citation frequency of those without such collaboration. Schmoch and Schubert (2008) suggest that international papers are more highly cited because their potential community is larger. In other words, international papers may be more highly

cited simply because more people know about them. This and the overlap of the communities involved could explain the difference in impact between collaborations in different fields, because the overlap of communities in fields that are close to each other is greater than in fields that are farther apart. To this one must add a possible national bias in the various countries involved in a collaboration that could work to the benefit of papers with international collaboration.

Are these suggestions and theoretical expectations of a greater impact and citation rate of international collaborations actually borne out in reality, however? If so, where does that impact come from? Does it come mainly from the countries included in the collaboration? Could the impact be because of the international partners in the collaboration? Do these papers also have a greater national impact? Do countries have a citation bias towards papers from their own country? Is the greater impact of such international collaboration papers because the authors come from more countries? The main hypothesis of the present study was that science knows no frontiers—that there is no national citation bias. If there is a certain bias towards what is nearby, this simply reflects the density of the authors' personal relationships and therefore of the information channels (both incoming and outgoing) that they use, a density that becomes diluted with increasing size of the area considered. This means that collaborative papers may gain in citations because of the increase in these informal channels of information dissemination, but above all, there exists a general gain in impact.

In the present work, we inquire into the provenance of the citations received by the different countries and the different types of papers, and into the targets of the citations made by the different countries. We shall take a country's production into account in two ways: first, that clearly the greater the production, the more citations the country has the potential to receive, but also the more domestic citations (or references) it will introduce into the arena.

Data and Methods

We used Scopus as the data source for the computation of the indicators because it best represents the overall structure of world science at a global scale. Scopus is the world's largest scientific database. It covers most of journals included in Thomson Reuters ScientificWeb of Science (WoS) and more (Moya-Anegón et al., 2007), and its coverage is statistically balanced in terms of subjects, countries, languages, and publishers. The greater part of the data was retrieved from the SCImago Journal & Country Rank (<http://www.scimagojr.com/>) database, limiting the search to citations to papers published in 2004 from papers published in 2005, 2006, and 2007. The data were downloaded in December 2009.

We chose the 20 countries with the greatest production in 2004, which are the countries with the greatest statistical significance, after which the following countries show a sharp drop in production. It is not the purpose of this study to make inferences from a random sample, but to study what happens in these 20 countries with the greatest production. They account for nearly 90% of overall scientific production, and nearly 97% of the scientific production involving international collaboration.

The main indicators used are as follows:

- Citations per paper: Average citations received by the papers of 2004 from the papers of 2005–2007, calculated by dividing the number of citations/references by the number of target papers.
- References per paper: Average references in papers of 2005–2007 to papers of 2004, calculated by dividing the number of citations/references by the number of source papers.
- Impact on papers per paper: The impact (or influence) of the papers of 2004 on the papers of 2005–2007, calculated by dividing the number of citations/references by the number of target papers and by the number of source papers.

$$I_{pp} = \text{citations (references)} / \text{num.target papers.num. source papers}$$

For example, to calculate the impact on nondomestic papers per collaboration paper for Germany, in the numerator, we put the citations from the nondomestic papers corresponding to the case of Germany to Germany's international collaboration papers. In this case, the number of target papers would be the number of Germany's collaboration papers of 2004, and the number of source papers

would be the number of nondomestic papers for the case of Germany from 2005 to 2007. To calculate the impact on collaboration papers per nondomestic paper for Germany, in the numerator, we put the citations from Germany's international collaboration papers to the nondomestic papers corresponding to the case of Germany. In this case, the number of target papers would be the number of nondomestic papers for the case of Germany of 2004, and the number of source papers would be the number of Germany's collaboration papers from 2005 to 2007.

Unlike the previous two indicators, which are simple ratios, this last indicator takes into account the size of both the target set of papers (obviously, the larger this set, the greater the likelihood of attracting more citations) and the source set (again of course, the larger this set, the more likely that references will be made to the target set). This takes into account not only the average citations per paper, but also the maximum that this ratio could reach. The size of the set of source papers has normally not been taken into account because the sources are usually taken to be the set of the global production of papers. However, grouping citations in accordance with their provenance can have a major effect if source sets of different sizes are generated. This type of indicator was first used by Pichappan (1995) for the citation analysis of journals of a particular subject field. In all cases, both the target and the source papers are classified by source (domestic, i.e., of the same country, or nondomestic), by collaboration type (with or without international collaboration, only distinguishing collaboration with institutions of other countries), and specific classifications for citations to or from the collaborating countries.

TABLE 1. The 20 countries ranked by total number of papers, with the corresponding percentage of papers in collaboration, percentage of citations from collaborating countries, and percentage of references to collaborating countries.

Country	Papers 2004	Papers 2005 to 2007	% Collaboration papers 2004	% Collaboration papers 2005 to 2007	% Citations from collaborators	% References to collaborators
United States	410,521	1,359,565	23.28	25.29	15.57	14.36
China	113,292	552,902	16.09	13.99	24.18	32.65
Japan	110,622	347,570	19.79	20.94	35.30	34.89
United Kingdom	107,143	362,408	37.46	38.99	30.86	30.24
Germany	98,949	326,678	39.65	41.47	30.65	29.49
France	69,591	232,091	41.64	43.33	32.51	31.07
Canada	55,929	196,960	39.27	40.78	39.06	40.04
Italy	54,112	183,950	34.81	36.64	33.39	33.67
Spain	39,776	144,564	32.88	34.89	33.29	32.92
Australia	35,886	127,554	38.03	38.85	36.18	35.26
Russian Federation	35,000	100,774	30.93	33.20	34.89	32.77
India	33,169	128,505	18.22	18.28	33.61	33.08
Republic Of Korea	31,332	120,173	25.00	25.45	38.68	42.53
Netherlands	30,168	104,253	45.14	45.90	35.05	35.36
Switzerland	22,049	75,615	55.05	56.95	35.58	35.59
Brazil	21,658	88,335	27.75	25.35	36.16	35.50
Sweden	21,280	69,274	45.75	48.90	34.38	32.30
Taiwan	21,071	81,275	17.49	18.53	38.25	42.70
Poland	20,378	67,225	31.32	31.18	36.80	34.40
Turkey	18,170	65,901	16.75	15.39	32.61	36.60

Results and Discussion

By way of context, one observes in Table 1 that the United States gets only a small percentage of citations from its collaborators and in turn provides them with only a small percentage of references. We understand this to be because of its large production and hence the large number of national references involved, i.e., because U.S. production is so large, its domestic citations reduce the percentage from other countries (and analogously with regard to references). But China, which went from 27% of U.S. production in 2004 to 40% in 2005–2007, despite having only a small percentage of collaboration (which even decreased in the second period), devotes a major proportion of its references to collaborating countries. Turkey, which is the country with the smallest production of the 20 countries considered and with the fewest collaboration papers, receives a high percentage of citations from collaborators, and references many of its collaborating countries, far more in relative terms than the United States or China. Although Switzerland has the highest percentage of collaboration papers, its percentage of citations received from its collaborating countries is not the highest.

The countries with the highest percentages of citations received from their collaborating countries and the highest percentages of references given to those countries are Canada, Republic of Korea, Australia, and Taiwan. These are not countries characterized by a great volume of scientific production. The Pearson correlation coefficient between the two percentages (citations received from collaborating countries and references to those collaborating countries) is 0.88. This implies a major degree of reciprocity between references and citations to and from collaborators. The correlation coefficient between the percentage of citations received from collaborators and the papers published in 2005–2007 is negative, -0.9 . And that between the percentage of references to collaborators and papers published in 2004 is also negative, -0.84 , i.e., the percentage traffic of citations with collaborating countries decreases as production increases. This makes sense simply in mathematical terms: the lower the production, the fewer domestic papers there will be to reference, and, vice versa, the fewer domestic papers there will be from which to receive citations, and hence the greater the respective percentages that will correspond to the traffic to and from nondomestic papers.

One can say in general that for the three citation indicators shown in Figure 1 (of the collaboration, no-collaboration, and overall papers), the nondomestic average exceeds the domestic average. The value of citations per collaboration paper is considerably greater than those of the citations per no-collaboration paper and citations per paper overall, with this difference originating mainly from nondomestic papers. The case is similar for the indicators relating to references, although the differences are less marked. The average of nondomestic references is far greater than that of the domestic references in all cases, and the average references per collaboration paper is considerably greater than those of the references per no-collaboration paper and per paper overall.

As one can see in Figure 2, the greater part of these high values of citations and references per collaboration paper comes from nondomestic papers. The average domestic citation obtained in collaborative papers is 1.9, while the average citation per paper coming from domestic papers is 6.85, three times greater. With respect to references, the average of domestic references per collaborative paper targeted at collaborative papers is 0.52, while the average of references to nondomestic papers per collaboration paper is 1.62, i.e., triple.

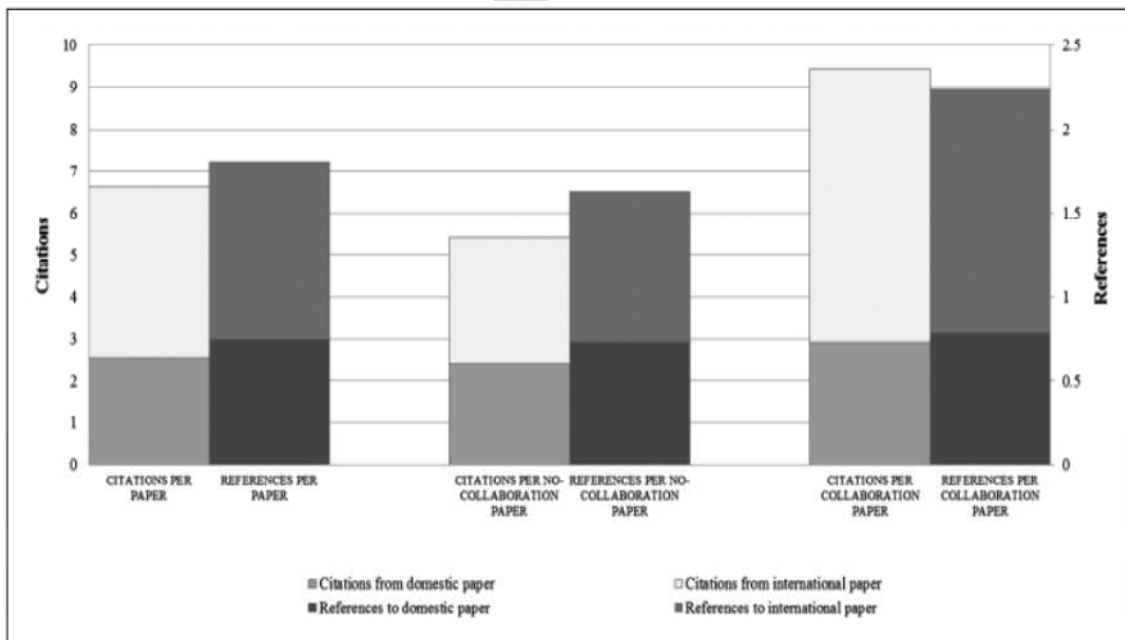


FIG. 1. Comparison of weighted average citations (per paper overall, per paper without collaboration, and per paper with collaboration) and references (per paper overall, per paper without collaboration, and per paper with collaboration) of the 20 countries with the highest production in 2004, distinguishing in both cases between domestic and nondomestic papers.

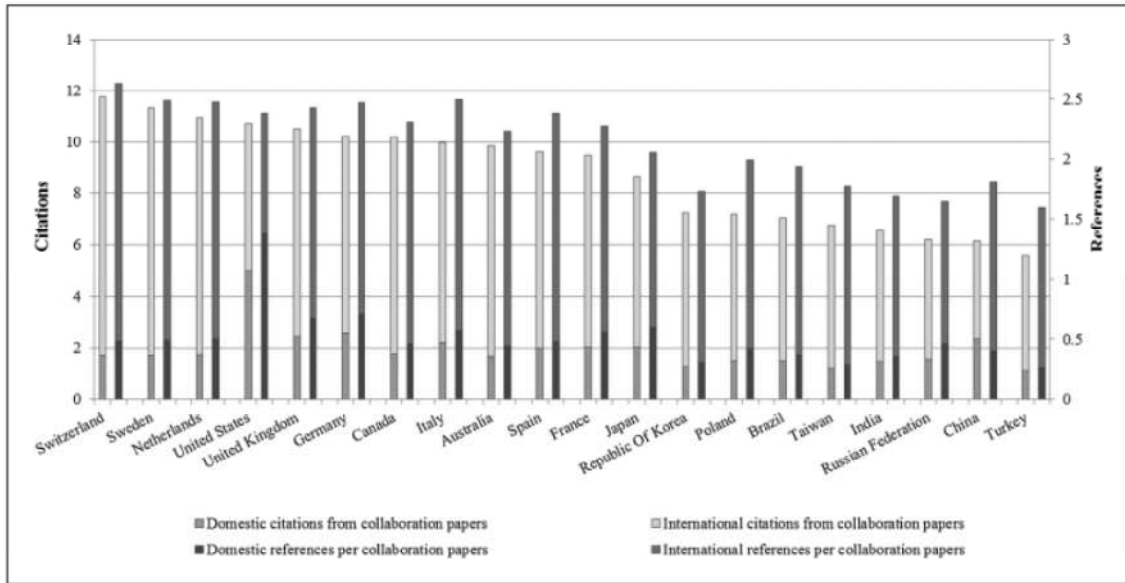


FIG. 2. Citations per collaboration paper (domestic and nondomestic) and references per collaboration paper (domestic and nondomestic), ordered from greater to lesser values of the former of these two parameters.

But it has to be said that there are some cases that deviate somewhat from these averages. For example, for the United States and China, almost 50% (to a lesser extent China) come from their domestic citations (recall that they were the countries with the greatest production of documents in 2004, but with the lowest percentage of citations from their international collaborators). The countries receiving the fewest domestic citations, Taiwan and Turkey, were those that received the greatest percentage of citations from their international collaborators. The same is the case with respect to the references. Although at first glance there is some correlation between citations and references, in countries with the fewest citations there is a stagnation in the number of references. As one observes in Figure 3, the greater part of the nondomestic citations per international collaboration paper comes from other countries. The overall average received from the collaborators is 2.9, and the average received from other countries is 3.9. The Pearson correlation coefficient between the two is 0.84.

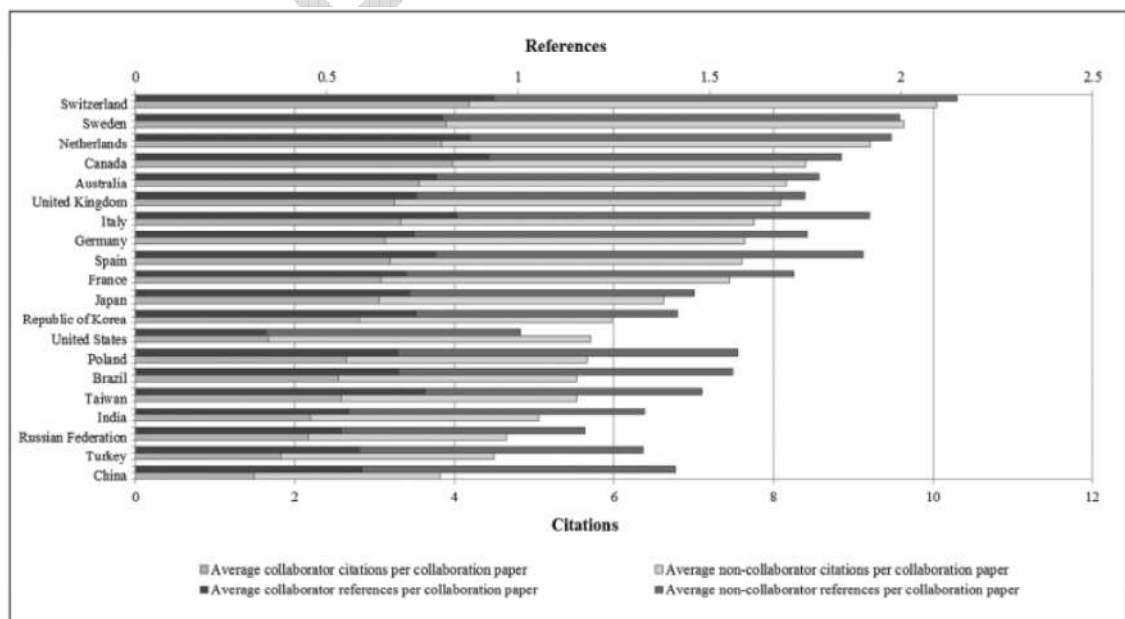


FIG. 3. Comparison between nondomestic citations per collaboration paper and references to nondomestic papers per collaboration paper, showing the average citations from/to collaborator and non collaborator countries, ordered from greater to lesser values of nondomestic citations per collaboration paper.

Again, there are countries that deviate a little from this pattern, as is the case of the United States and China, which, compared with other countries, receive less from their international collaborators and in turn give them less. The case of China may be slightly different because although it captures relatively few citations, it does make a greater number of nondomestic references. The average of references given to collaborating countries is 0.71 and to other countries 0.91. The Pearson correlation coefficient between the two sets of data is 0.71. The Pearson correlation coefficient between the average collaborator citations and references is 0.81, and between the same two sets of data for the other countries is 0.90. The value of the average nondomestic citations of the collaboration papers is greater than that of the no-collaboration papers, even excluding the citations of the collaborating countries (see Figure 4). The same is the case with the references. There are countries for which the difference in the averages is quite noticeable, as is the case of the Russian Federation, China, and Poland, and other countries for which the indicators are quite similar, as is the case of Switzerland, the Netherlands, Canada, and the United States. Switzerland, Sweden, and the Netherlands are the countries that make most references to nondomestic papers in their no-collaboration papers, and are also the countries with the highest nondomestic citation per paper in their international collaborations, although their productions are quite small.

Naturally, the impact on domestic papers per paper of the country is greater than both the impact on nondomestic papers per paper of the country and the impact on papers of the country per nondomestic paper, as shown in Figure 5 in which two different scales have had to be used. Coincidentally, the first three countries are the three with the greatest values of citations per collaboration paper, and the greatest national impact. These are Sweden, Switzerland, and the Netherlands, which are also those with the greatest nondomestic impact, probably because of their high proportion of papers with both national and international collaboration. At the opposite end are the countries with the greatest production, China and the United States. Coincidentally, they are also countries with high values of domestic citations per collaboration paper. This high value of domestic citations is thus due to the large domestic production and the great number of references contributed by this production, because if this volume is taken into account, then these countries turn out to be the least biased towards their own production. There is a negative correlation of -0.6 between production and the impact on domestic papers per paper of the country, while the correlation between domestic and nondomestic impacts is less than 0.38.

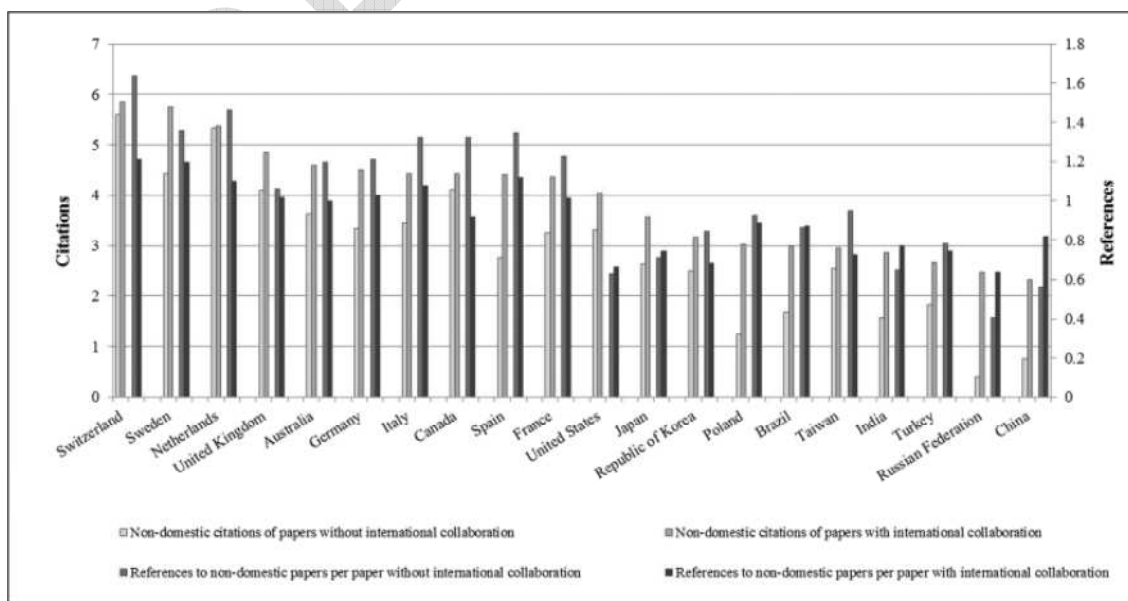


FIG. 4. Average nondomestic citations of papers without international collaboration, average nondomestic citations of papers with international collaboration (excluding citations from the collaborating countries), references to nondomestic papers per paper without international collaboration, and references to the nondomestic papers per collaboration paper (excluding those corresponding to the collaborators).

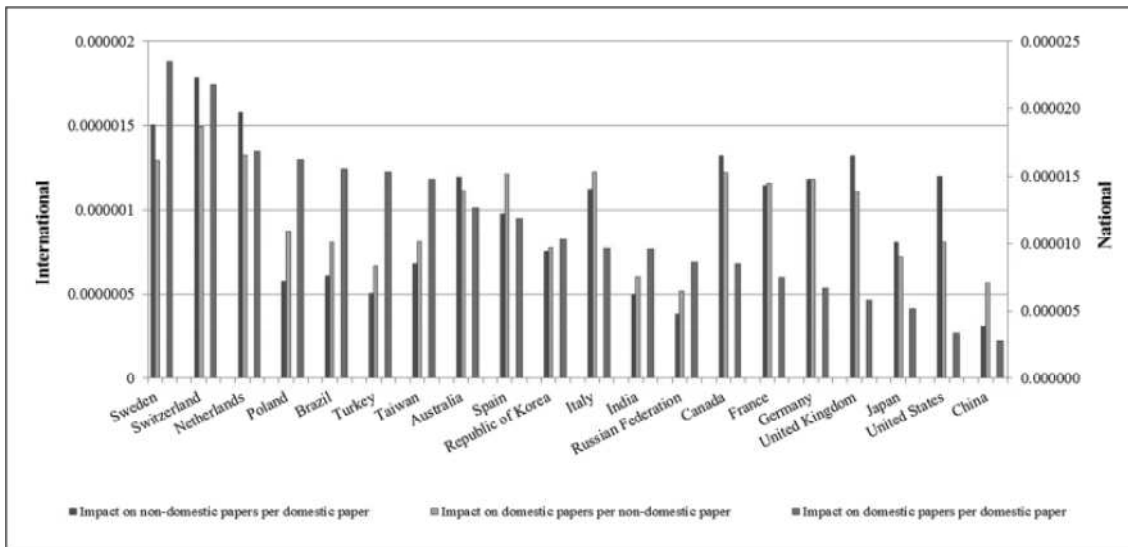


FIG. 5. Impact on nondomestic papers per domestic paper, impact on domestic papers per nondomestic paper, and impact on domestic papers per domestic paper, ordered from greater to lesser values of this last parameter.

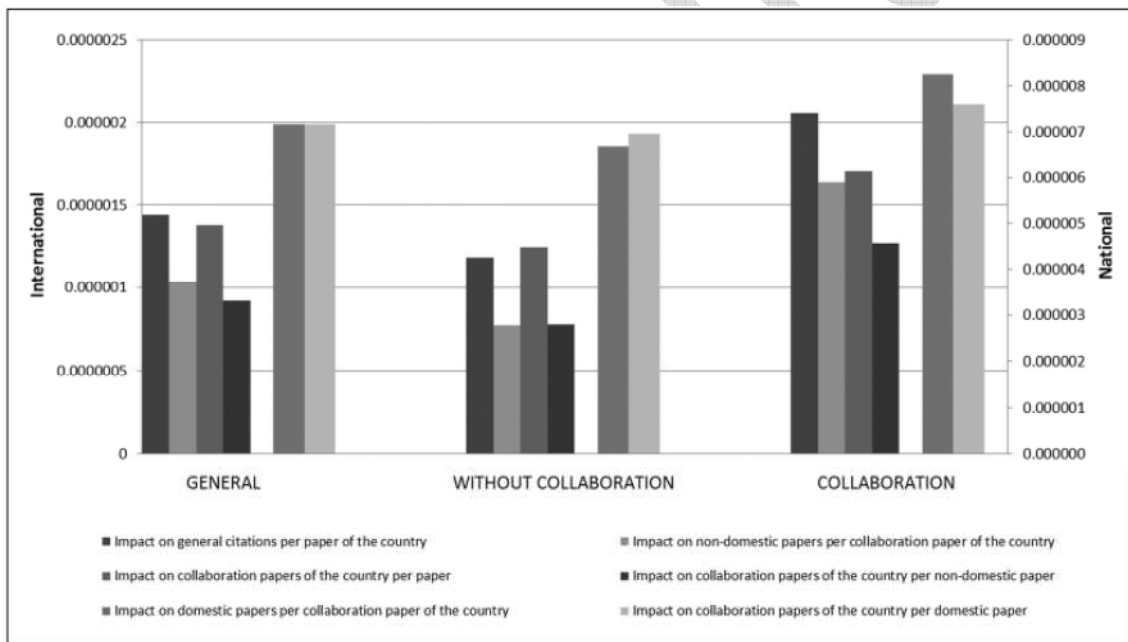


FIG. 6. Weighted averages for the production of 2004 of the impact on general papers (1st), on domestic papers (5th), and on nondomestic papers (2nd) per domestic paper, per domestic paper, and per collaboration paper, and impact on domestic papers, on papers without collaboration, and on collaboration papers per general paper (3rd), per domestic paper (6th), and per nondomestic paper (4th).

Disaggregating these influences by documents with or without international collaboration, one finds that the average national impact is the greatest. This means that, as a rule, all countries tend to have more impact on their own production (Figure 6). Collaborative papers also have a greater impact overall (both domestic and nondomestic), with the nondomestic growing more percentage-wise, so that in them there is less difference between the domestic and nondomestic impacts.

Disaggregating by country the impact of papers with international collaboration, one finds the same general features. As in Figure 5, one observes groups of countries with similar behavior:

- Great production, with little domestic impact (United States, China, and Japan).
- Major domestic and nondomestic impact, but with a not very prominent production (Switzerland, Sweden, and the Netherlands).
- Little impact, whether domestic or nondomestic (India, Russia, Taiwan, and Korea).
- Major nondomestic impact, but little domestic impact (U.S., U.K., Germany, and Japan).
- Major domestic impact, but little nondomestic impact (Turkey, Poland, and Brazil – with Russia coming close to this group in Figure 7).

It has to be said that simple observation of the figures shows that for all the countries, the impact on the collaboration papers per domestic paper is fairly large, although the collaboration papers have a greater impact on the domestic papers than vice versa.

Interpretation

Although some countries with a large production, such as China and the United States, do not have a large percentage of international collaboration papers, there is no significant correlation between scientific production and percentage of collaboration. However, there is a significant negative correlation between production and the percentage traffic of citations to/from the collaborating countries, as was indeed to be expected mathematically. In general, the average nondomestic citation per paper exceeds the domestic, and the average citation per paper of collaboration papers exceeds that of the no-collaboration papers (by more than 70%). Although there is an increase in the average domestic citation per paper of collaboration papers, there is comparatively a greater increase in the nondomestic citation per paper (by more than 110%).

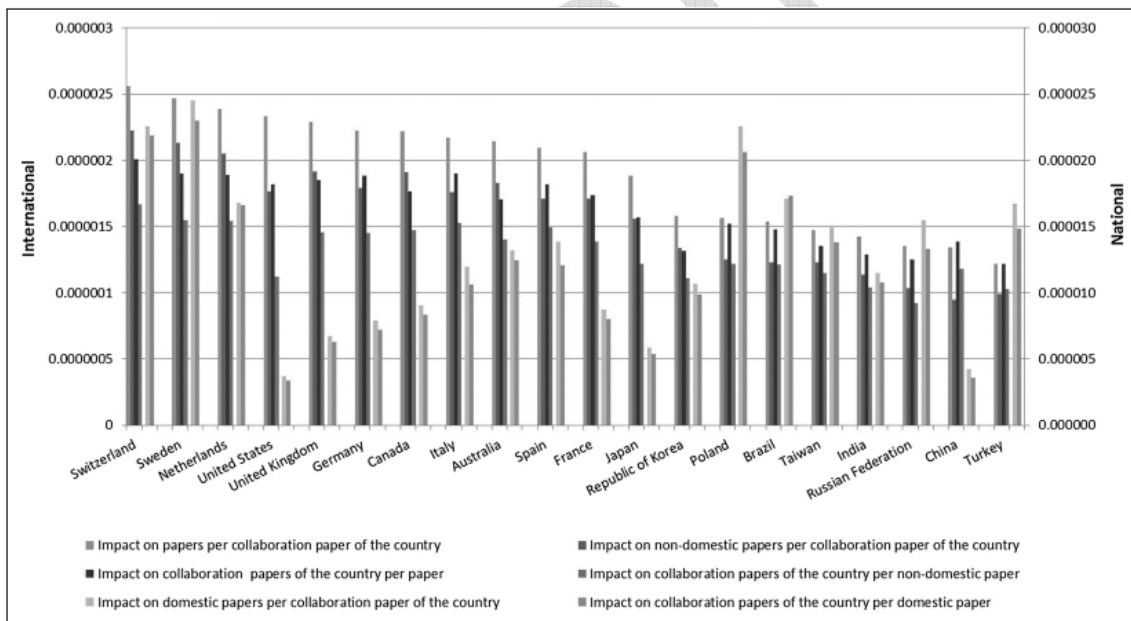


FIG. 7. Impact on papers per collaboration paper of the country, impact on nondomestic papers per collaboration paper of the country, impact on collaboration papers of the country per paper, impact on collaboration papers of the country per nondomestic paper, impact on domestic papers per collaboration paper of the country, and impact on collaboration paper of the country per domestic paper, ordered from greater to lesser values of the first of these parameters.

Some 40% of this nondomestic citation per paper of collaboration papers comes from the collaborating countries, and there is a strong correlation between the collaborator citations and citations from the rest of the countries. The average nondomestic citation per paper of collaboration papers is greater than that of the no-collaboration papers, even when excluding that coming from the collaborating countries.

The impact on domestic production is greater than on nondomestic production. As expected, collaboration papers have a greater impact and there is less difference between the domestic and nondomestic impact. There is a negative correlation between a country's production and its impact on domestic papers per paper. The countries with the largest productions tend to have the least impact on their own production, as is especially the case for the United States and China. On the contrary, the countries of small sizes in terms of scientific production may be favored in this sense, because either their international frontiers with other countries form more a part of their authors' immediate environments or, for some of them (e.g., Netherlands, Switzerland), they are also small in area with high population densities. Both these factors would naturally enhance the dissemination of ideas between countries. However, no significant correlation is observed between domestic and nondomestic impact. Some countries (United States, United Kingdom, Germany, and Japan) have a great impact on nondomestic papers but little impact on domestic papers. At the opposite end of the spectrum, there are countries with a major domestic impact but little nondomestic impact, as is the case for Turkey and Poland, and Brazil. Emerging countries have, understandably, less impact outside their country as most of their researchers do not yet perform at the scientific frontier.

In general, one may conclude that science knows no national frontiers. Certainly there is a greater impact on authors' immediate environments, which do not necessarily coincide with national boundaries. This impact fades as the environment grows in size, while a small environment maximizes the bias produced by self-citation. Indeed, the greatest biases towards domestic production occur in small and developing countries.

Acknowledgments

This work was financed by the Junta de Extremadura. Consejería de Educación, Ciencia & Tecnología and the Fondo Social Europeo as part of the pre-doctoral studentship PRE07052 and the research group grant GR10019, and by the Plan Nacional de Investigación Científica, Desarrollo e Innovación Tecnológica 2008e2011 and the Fondo Europeo de Desarrollo Regional (FEDER) as part of research projects TIN2008-06514-C02-01 and TIN2008-06514-C0202.

References

- Aksnes, D. (2003). Characteristics of highly cited papers. *Research Evaluation*, 12(3), 159–170.
- Alcaín, M., & Gálvez, M. (1998). Evolución de las revistas españolas de psicología. *Papeles del Psicólogo*, 70, 35–42.
- Bandyopadhyay, A.K. (2001). Authorship pattern in different disciplines. *Annals of Library and Information Studies*, 48(4), 139–147.
- Bridgstock, M. (1991). The quality of single and multiple authored papers—An unresolved problem. *Scientometrics*, 21, 37–48.
- Carpintero, H., & Peiró, J. (1983). The significance of the bibliometric methodology to the studies of the history of psychology. *Revista de Historia de la Psicología*, 4(1), 21–32.
- Chinchilla, Z., Vargas, B., Hassan, Y., González, A., & Moya, F. (2010). New approach to the visualization of international scientific collaboration. *Information Visualization*, 9(4), 277–287.
- Glänzel, W. (2001). Domestic characteristics in international scientific coauthorship relations. *Scientometrics*, 51(1), 69–115.
- Glänzel, W. (2002). Coauthorship patterns and trends in the sciences: A bibliometric study with implications for database indexing and search strategies, 1980-1998. *Library Trends*, 50(3), 461–473.
- Goldfinch, S., Dale, T., & De Roue, K. (2003). Science from the periphery: Collaboration network and 'periphery effects' in the citation of New Zealand Crown Research Institutes articles, 1992–2000. *Scientometrics*, 57(3), 321–337.

- Gómez, I., Fernández, M.T., & Sebastián, J. (1999). Analysis of the structure of international scientific cooperation networks through bibliometric indicators. *Scientometrics*, 44(3), 441–457.
- Hsu, J.W., & Huang, D.W. (2010). Correlation between impact and collaboration. *Scientometrics*, 86(2), 317–324.
- Katz, J., & Hicks, D. (1997). How much is a collaboration worth? A calibrated bibliometric model. *Scientometrics*, 40(3), 541–554.
- Lancho, B., Guerrero, V., & Moya, F. (2010a). The iceberg hypothesis revisited. *Scientometrics*, 85(2), 443–461.
- Lancho, B., Guerrero, V., & Moya, F. (2010b). What lies behind the averages and significance of citation indicators in different disciplines? *Journal of Information Science*, 36(3), 371–382.
- Leimu, R., & Koricheva, J. (2005). Does scientific collaboration increase the impact of ecological articles? *Bio Science*, 55, 438–443.
- Lewinson, G., & Cunningham, P. (1991). Bibliometric studies for the evaluation of trans-domestic research. *Scientometrics*, 21(2), 223–244.
- Moed, H., Bruin, R., Nederhof, A., & Tijssen, R. (1991). International scientific co-operation and awareness within the European Community: Problems and perspectives. *Scientometrics*, 21(3), 291–311.
- Moya, F., Chinchilla, Z., Corera, E., González, A., Hassan, Y., & Vargas, B. (2008). *Indicadores bibliométricos de la actividad científica española: 2002–2006*. Madrid: Fecyt.
- Moya, F., Chinchilla, Z., Vargas, B., Corera, E., Muñoz, F., González, A., & Herrero, V. (2007). Coverage analysis of Scopus: A journal metric approach. *Scientometrics*, 73(1), 53–78.
- Narin, F., Stevens, K., & Whitlow, E. (1991). Scientific cooperation in Europe and the citation of multidomestically authored papers. *Scientometrics*, 21(3), 313–323.
- Narin, F., & Whitlow, E. (1990). *Measurement of scientific cooperation and coauthorship in CEC-related areas of science*. Luxembourg: European Community.
- Persson, O., Glanzel, W., & Danell, R. (2004). Inflationary bibliometric values: The role of scientific collaboration and the need for relative indicators in evaluative studies. *Scientometrics*, 60(3), 421–432.
- Pichappan, P. (1995). A dual refinement of journal self-citation measures. *Scientometrics*, 33(1), 13–21.
- Schmoch, U., & Schubert, T. (2008). Are international co-publications an indicator for quality of scientific research?. *Scientometrics*, 74(3), 361–77.
- Sooryamoorthy, R. (2009). Do types of collaboration change citation? Collaboration and citation patterns of South African science publications. *Scientometrics*, 81(1), 177–193.
- Stack, S. (2002). Gender and scholarly productivity: The case of criminal justice. *Journal of Criminal Justice*, 30(3), 175–182.