ISSN 2318-2377

TEXTO PARA DISCUSSÃO N°553

CROSS-BORDER CO-AUTHORSHIPS IN SCIENTIFIC ARTICLES AND KNOWLEDGE FLOWS:IMPLICATIONS FOR INVESTIGATING AN EMERGING INTERNATIONAL SYSTEM OF INNOVATION

Leonardo Costa Ribeiro
Márcia Siqueira Rapini
Leandro Alves Silva
Eduardo da Motta e Albuquerque

Maio de 2017

http://www.cedeplar.ufmg.br

Textos para Discussão

A série de Textos para Discussão divulga resultados preliminares de estudos desenvolvidos no âmbito do Cedeplar, com o objetivo de compartilhar ideias e obter comentários e críticas da comunidade científica antes de seu envio para publicação final. Os Textos para Discussão do Cedeplar começaram a ser publicados em 1974 e têm se destacado pela diversidade de temas e áreas de pesquisa.

Ficha catalográfica

C951 2017 Cross-border co-authorships in scientific articles and knowledge flows: implications for investigating an emerging international system of innovation / Leonardo Costa Ribeiro ... [et al.]. - Belo Horizonte : UFMG/CEDEPLAR, 2017.
43 p. : il. - (Texto para discussão, 553)

Inclui bibliografia (p. 35-38)
ISSN 2318-2377

1. Gestão do conhecimento 2. Ciência - periódicos - publicação. 3. Redes de informação. I. Ribeiro, Leonardo Costa. II. Universidade Federal de Minas Gerais. Centro de Desenvolvimento e Planejamento Regional. III. Título. IV. Série.

CDD: 658.4038

Ficha catalográfica elaborada pela Biblioteca da FACE/UFMG - JN 043/2017

UNIVERSIDADE FEDERAL DE MINAS GERAIS
FACULDADE DE CIÊNCIAS ECONÔMICAS
CENTRO DE DESENVOLVIMENTO E PLANEJAMENTO REGIONAL

CROSS-BORDER CO-AUTHORSHIPS IN SCIENTIFIC ARTICLES AND KNOWLEDGE FLOWS:IMPLICATIONS FOR INVESTIGATING AN EMERGING INTERNATIONAL SYSTEM OF INNOVATION

Leonardo Costa Ribeiro
Inmetro, Rio de Janeiro, Brazil

Márcia Siqueira Rapini
Cedeplar-UFMG, Belo Horizonte, Brazil

Leandro Alves Silva
Cedeplar-UFMG, Belo Horizonte, Brazil

Eduardo da Motta e Albuquerque
Cedeplar-UFMG, Belo Horizonte, Brazil

CEDEPLAR/FACE/UFMG
BELO HORIZONTE
2017

SUMÁRIO

ABSTRACT

Size matters: the total of internationally co-authored scientific articles in 2015 corresponds to the global scientific production in 1993. The steady and systematic growth in international collaboration in science provides a strong basis for an emerging GIS. Therefore, it is important to map international flows that connect different national systems of innovation. This paper tracks knowledge flows through cross-border co-authorships in scientific publications, through a database with 10 million papers published in 2000, 2003, 2006 2009, 2012 and 2015. The data show an increase in international co-authorships from 10.7% in 2000 to 21.3% in 2015. However, this growth has network properties, since the number of international flows has grown from 545,372 in 2000 to 7,083,075 in 2015. Those international co-authorships signal networks of universities and research institutes, providing international connections to firms that eventually interact only locally with those universities and research institutes. The growth in the size, dimension and quality of those scientific flows strengthens a broad and variegated mosaic of interconnections can be grasped by the size of the network of cross-border co-authorships, a network that might be supporting an emerging and rudimentary global system of innovation.

Key words: knowledge flows, international co-authorships, science, innovation systems

JEL CLASSIFICATION: O30

RESUMO

Tamanho importa: o total de artigos científicos com co-autoria internacional em 2015 corresponde à produção científica global em 1993. O crescimento constante e sistemático da colaboração internacional em ciência fornece uma base sistema internacional de inovação emergente. Portanto, é importante mapear os fluxos internacionais que conectam diferentes sistemas nacionais de inovação. O presente trabalho aborda os fluxos de conhecimento através de co-autorias internacionais em publicações científicas indexadas, através de uma base de dados com 10 milhões de artigos publicados em 2000, 2003, 2006, 2009, 2012 e 2015. Os dados mostram um aumento das co-autorias internacionais de 10,7% 2000 a 21,3% em 2015. No entanto, esse crescimento tem propriedades de rede, já que o número de fluxos internacionais passou de 545.372 em 2000 para 7.083.075 em 2015. Essas co-autorias internacionais sinalizam redes de universidades e institutos de pesquisa, fornecendo conexões internacionais a empresas que eventualmente interagem apenas localmente com essas universidades e institutos de pesquisa. O crescimento do tamanho, da dimensão e da qualidade desses fluxos científicos fortalece um amplo e variado mosaico de interconexões que pode ser apreendido pelo tamanho da rede de co-autorias internacionais, uma rede que pode estar apoiando a emergência de um sistema global de inovação ainda rudimentar.

Palavras-chave: fluxos de conhecimento, co-autorias internacionais, ciência, sistemas de inovação

INTRODUCTION

Empirical evidence and theoretical reflection have been accumulating on the growth of international knowledge flows that connect different systems of innovation. A report from OECD summarizes data and evidence on a mosaic of forms that connections to knowledge spill over national boundaries (OECD, 2015, chapter 3). Therefore, there is an increasing evidence on the role of those knowledge flows tensioning the national dimension of innovation systems. It’s time to begin to evaluate how far those changes are preparing the basis for an emerging international system of innovation.

The literature on national systems of innovation (NSIs) has been aware of the role of international flows since its earlier works. More than twenty years ago, Nelson and Rosenberg (1995, pp. 17-18) have highlighted this tension - a growing tension since then - between the international dimension of technology and the national boundaries of innovation systems. Recently, in a conclusion of a broad review of the literature on NSIs - a section entitled "from national to international systems of innovation" -, Soete, Verspangen and Weel (2010, p. 1176) mention that "a feature which has increasingly challenged the notion of NSI is of course the rapid growth in international research and knowledge flows". Soete et al (2010, p. 1176) conclude their review stressing how "globalization of knowledge flows represents a real challenge for systems of innovation policies".

This paper focuses on science to stress how far and how consolidated are those international networks, which may constitute another solid basis for corroding national boundaries between systems of innovation. After all, science can be understood as an international endeavor*per se*, always avoiding to be restrained by national boundaries.

This paper investigates global knowledge flows that are related to scientific production and may be empirically investigated through statistics of ISI-indexed papers and the institutional location of their authors. A data base with 10,021,195 papers and documents published in 2000, 2003, 2006, 2009, 2012 and 2015 was prepared, with an initial effort to identify international co-authorship. Table 1 shows the increase in international co-authorship - the clue for another international knowledge flow in this paper - from 10.7% in 2000 to 21.3% in 2015.

TABLE 1
Articles, total and with international co-authorship, also in percentage of international co-authorship
(2000 - 2015)

Year	Total Articles	Articles With Inter Flow	%
2000	1,274,329	136,483	10.7%
2003	1,360,275	166,672	12.3%
2006	1,517,189	197,940	13.0%
2009	1,885,092	265,460	14.1%
2012	2,019,563	329,190	16.3%
2015	1,964,747	418,866	21.3%

Source: Web of Knowledge, authors' elaboration (Database A)

What does mean those 418,000 papers with international collaboration in 2015? The total of scientific papers published in 1993, according to data from the National Science Foundation (NSF, 1996, p. 5-31).

This rise is a trend, identified by other sources - "scientific papers with authors from multiple countries increased by 16% in top natural science journals between 2012 and 2015, according to data from the Nature Index" (Nature, 2016b). According to the National Science Foundation, "between 2000 and 2013, collaboration has been increasing, with higher shares of scientific publications with institutional and international co-authorships" (NSF, 2016, p. 5-102).

For our quantitative analysis, throughout this paper the benchmark of 21.3% (the global average for international co-authorship in 2015) will be used to investigate diverse dimensions of scientific production: location, S&E fields, and institutions. An important question in this preliminary analysis is whether or not international co-authorships and their links have network properties.

The focus of this paper is scientific production, using data from the Web of Science. This focus leads this paper to underestimate the real dimension of global scientific knowledge flows, because it does not capture migration, visits, study abroad, research efforts that did not become formalized in papers.

With this underestimation in mind, the focus of this paper helps to interpret the specific contribution of the scientific dimension for the global knowledge flows and their consolidation as a structural feature of contemporary world.

This paper is organized in seven sections.

1. TENTATIVE FRAMEWORK: CROSS-BORDER KNOWLEDGE FLOWS CONNECTING DIFFERENT NATIONAL SYSTEMS OF INNOVATION

The elaboration on NSIs organizes the role of different institutions (firms, universities, research institutes, etc) for technological progress. This approach focused mainly in national spaces, but since the beginning put forward the role of international relationships and related tensions created by the transnational nature of technology and the international propensity of science (see, as an example, Nelson and Rosenberg, 1995, pp. 17-18). Even the literature that tried to question impressionist visions about the rhythm of internationalization (see the CJE special issue 1995) illustrates international connections and flows (Patel, 1995).

There are flows that connect firms from different countries. There is a huge and growing literature on the internationalization of R&D. This huge literature may be summarized by Dunning and Lundan (2009), Patel (2011), Cantwell (2013) and Laurens et al (2015).

Dunning and Lundan (2009) focus on "patterns of the internationalization of the knowledge-creating and knowledge-sourcing activities of MNEs" (p. 13). They collect evidence on "three main trends": 1) "the internationalization of the innovative activities of MNEs has lagged behind that of their productive activities"; 2) "foreign affiliates have gained substantially more autonomy and now play a far more important role in the knowledge-creating activities of the MNE as a whole", furthermore, Dunning and Lundan stress that those foreign affiliates "link the internal network of the MNE with national or regional innovation systems within which they are embedded"; 3) given new players in the global economy, "the innovative activities of MNEs have become more geographically dispersed than has been the case before" (p. 13). Their paper discusses the "organization of affiliate R&D", "evidences on the internationalization of corporate R&D", "external technology sourcing by multinational enterprises", and "policy implications for home countries". In the conclusion, Dunning and Lundan (2009, pp. 27-28) suggest a correlation between the growth of internationalization of "knowledge-creating activities of MNEs" "over the past three decades" (p. 27) and an evolution towards "more autonomy to foreign affiliates" (p. 28). However, they highlight that the persistence of the role of home country in those networks and that the internationalization of R&D has been "directed to the traditional host countries" (p. 28). They also highlight a co-evolutionary pattern between the "innovative activity of MNEs" and the "home country NIS" (p. 28).

Patel (2011) focuses on the activities of European Union large firms - an opportunity to revisit his own work in this field, especially his influential 1995 paper (Patel, 1995), searching for changes in patterns of geographical concentration of R&D activities of MNEs. Patel organizes a database with 963 worldwide technologically actives large firms, reviews available statistics to measure the location of knowledge-creating activities, presents his data and methods and discusses the main results. According to his measures, "all technologically active firms have become much more globalized since the early 1990s" (Patel, 2011, p. 12). Differences between different sectors are also presented by Patel (pp. 16-19), that indicates "pharmaceuticals", "food, drink and tobacco", "mining and petroleum" and "ICT" as the most internationalized sectors. In his "assessment and conclusion", Patel evaluate globalization of technology in two dimensions ("spread vs. volume"). Focusing in EU companies, Patel suggests that given the difference between the spread and volume of those international flows, those foreign centers may be "involved in adapting products developed elsewhere... or may be 'listening posts' aimed at monitoring developments in science and technology in foreign locations" (p. 21). Finally, Patel's results "are also consistent with the notion that companies are increasingly involved in different foreign locations in order to tap into local S&T resources rather than simple adapt their products for local markets" (p. 23).

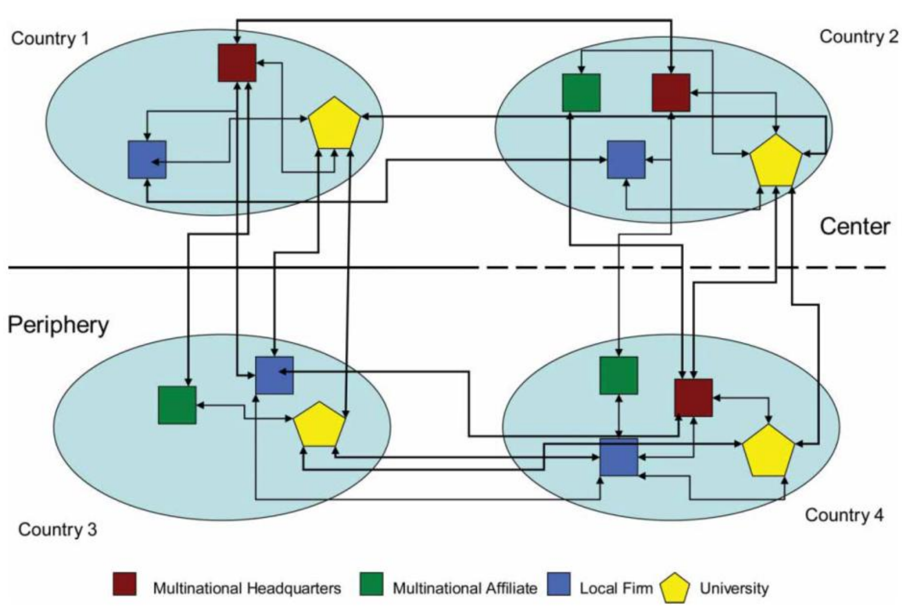
Cantwell (2013) integrates changes in internal knowledge flows within the MNEs (that depend upon ICT revolution) and their new relationships with external knowledge flows with changes that have been blurring the frontiers between science and technology. Those changes may signal the "changing nature of knowledge creation" (p. 4), with implications leading to a more complex set of knowledge flows, both within MNEs and between them and other firms and other institutions. The aim of Cantwell's paper is to investigate "the linkages between intra-firm and inter-firm networks for knowledge creation and exchange" (p. 4). As Dunning and Lundan (2009), Cantwell (2013, p. 13) points to a greater role for subsidiaries, stressing an evolutionary process within which they "may evolve to become competence-creating on their own right". It would be necessary now to go beyond the "conventional parent-driven view of the MNC" (p. 14). With subsidiaries more autonomous actors, their "external business networks" "may well also become international" (p. 14). This evolution of "subsidiaries or sub-contractors towards competence-creating activities, partly reflects the more general dispersion of innovative efforts across countries since 1992" (p. 15). This whole process and the ICT resources generates a key transformation: "the shift from the MNC as an institution for technology transfer between established activities frequently organized along miniature replica lines in different locations, and towards the MNC as a developer of international networks for technology creation, which combine formerly unconnected streams of innovation" (p. 18). Cantwell suggests an evolutionary patterns for multinationals (p. 21), suggesting that "the competitive advantage of established or mature MNCs increasingly stems instead from their ability to build and control a network of global flows of information, resources and people" (p. 21). All those developments indicate that "the role of the multinational corporate group as a whole continues to remain critical in the story of an internationally distributed system of innovation" (p. 4).

Laurens et al (2015) address two questions: is there a "growing trend in the internationalization of technology creation" and does "the 'home-base-augmenting' strategy still dominate? To answer those questions they organize a broad review of the literature, prepare a database of 946 firms (using ORBIS database to consolidate names and GUOs), and 3,251,515 applications of priority patents (from Patstat). Laurens et al (2015, pp. 769-770) find that 7.1% of 880,460 patents from 2003-2005 are international - 17.34% among US firms, 30.37% among EU firms and 2.54% among Asian firms. The internationalization rate for 1994-1996 was 5.2% (p. 769), with increase in the rates for US and Asia, but a decrease for EU firms. Laurens et al (2015, p. 770-771) conclude that there is "no deepening of the global internationalization rate" and identify "very contrasted situations and opposite dynamics by continent firms". Regarding the strategies, Laurens et al (2015, p. 773) do not find that "is internationalization driven mostly by home-base-augmenting motives (p. 773). In the conclusion, Laurens et al (p. 773) raise the issue of a plateau in the internationalization rate, "beyond which, as suggested by Gammeltoft, organizational issues become too complex to be efficiently managed". Laurens et al (2015, p. 773) caution against "generalizations previously made from case study analyses and by quantitative analyses done one decade ago".

Other tools to capture firms' involvement in those flows are patent citations of patents: they may indicate - *inter allia*, "knowledge spill-over ... between countries" (Jaffe, Trajtenberg and Fogarty, 2000, p. 218). Ribeiro et al (2014) suggest a methodology to uncover international knowledge flows through patents citations of ISI-indexed scientific papers - basically tracking those firms citing papers to support technological innovations formalized in patents.

Those references and a large literature on innovation focusing in international flows reviewed by Britto et al (2013, pp. 79-83) support a suggestion of a tentative framework presented in Figure 1.

FIGURE 1
Multinationals, their R&D activities and the scientific institutions as connectors between different national systems of innovation - a tentative framework



Source: Britto et al (2013, p. 80)

Each flow presented is Figure 1 is supported by this literature review, by case studies, surveys of firms and/or university research (Britto et al, 2013, pp. 79-83). The tentative framework presented in Figure 1 shows how different national systems of innovation are connected through those multifarious flows. TNCs are key organizers of several of those international flows (Fuchs, 2014; Cantwell, 2009), but they are not alone. As Figure 1 suggests, universities and research institutes have their own international connections, and those connections transforms the interactions between universities and firms. A very simple example is a firm in country 3 that has interaction with one university in the same country, but this specific university has connections with universities in countries 1 and 4 - this is an apparently local interaction, but in reality it is an international interaction, since firm in country 3 may be using knowledge developed either internationally or in a foreign country.

Those flows that connect different NSIs have received systematic and continuous attention by researchers and institutions, as the literature shows. OECD (2015, chapter 3) is useful to summarize the role of broader flows such as "mobility of highly skilled individuals", "scientists on the move", "excellence in scientific collaboration", "research across borders", "inventions across borders", "international markets for knowledge", "open innovation" and "collaboration on innovation".

Those references put forward two issues, one theoretical and another empirical.

Theoretically, the size and scope of all those cross-border flows put forward a question: where are we in the transition from national systems of innovation to an emerging global system of innovation? Of course it is too early to talk about a global system of innovation, but we are in a point where the tension between the national dimension of systems of innovation and the international proclivity of science and technology has increased strongly.

Empirically, those broad flows and connections might suggest that our tentative framework simplifies and underestimate the scale and scope of international knowledge flows.

Between those diverse flows, the flows through science that connect different NSIs have a very peculiar and important role. A small set of those scientific flows are investigated in this paper: science, narrowly reduced to ISI-indexed papers, more specifically to international co-authorship in those papers, may be represented by sets of arrows in Figure 1 that connect different NSIs. As such, the analysis that is organized in this paper completes another feature of Figure 1, adding another piece to that framework. To investigate how important those connections are - and how they can be measured - is the subject of this paper.

2. LITERATURE REVIEW: GLOBAL COLLABORATION IN SCIENCE

Knowledge flows in science are increasingly acknowledged in the literature as an important issue. Beyond the OECD's (2015, chapter 3) references to connections in knowledge, other benchmark publications stress the importance of those flows. NSB (2016, pp. 5/105-5/109), for instance, evaluates carefully those international connections, presenting general data (Figure 5-26, p. 5/103), co-authorship by country (Figure 5-28, p. 5/106), disaggregating international collaboration by field (Figure 5-27, p. 5/101). NSB (2016, p. 5/110) also analysis the international flows through cross-national citations, another "evidence that S&E research is increasingly international in scope". OECD (2016, p. 28) includes this growing importance among other shifts in science: "[t]he production of scientific research is progressively shifting from individuals to groups, from single to multiple institutions, and from national to international level".

Academically, there is a long list of important works dealing with internationalization of science (a brief review is presented by Wagner et al, 2005, pp. 1609-1610).

This literature is one of the supports of a strong argument presented by Wagner et al (2015): given the size of those networks, new dynamics might be being created. Beyond this literature, Wagner et al (2015) present previous research of their group (for instance, Leydesdorff et al, 2008) and focus on "global networks of science", analysing their formation and stabilization, to finally collect evidence to support their suggestion that international collaboration "represent a self-organizing phenomena" (2015, p. 8). The growth of the global network of research "is an emerging organization added (and possibly superseding) the national model". More than the stabilization and self-organization of those scientific networks, Wagner et al (2015, p. 12) suggest that they impact other levels of innovation systems: "[t]he global network is arguably now a more stable system that serves as a source of vitality and direction to R&D at all lower levels". As a evidence of a transition to a more global system, Wagner et al (2015, p. 12) evaluate that those changes create new governance problems, as the "global cooperation networks", a " dynamic system", operate "orthogonally to national systems".

The growing importance of international collaboration may also be grasped by the attention given by leading scientific journals such as *Nature*.

In a "Comment", Adams (2013, p. 557) suggests that "we are entering a fourth age of research, driven by international collaborations between elite research groups".

Nature organize an Index (see http://www.natureindex.com/), that is very useful for ranking scientific institutions and firms in their production and involvement with science. Beyond data on scientific institutions, *Nature Index* organizes statistics from "articles... from 68 natural science journals" (Nature, 2015, p. S34).

As Rosenberg (1990) has long time ago highlighted, firms produce basic research with their own money, and as Hicks (1995) has shown, firms publish the results of their research. *Nature*

Index corporate institutions ranking show this, with data for the XXI Century (see http://www.natureindex.com/annual-tables/2016/institution/corporate/all). It is important to highlight the general position of those firms in the Nature Index with all institutions: IBM is at the 237th position, F. Hoffmann-La Roche at 244th, and Merck, at 456th position is the last firm to be shown in the Top 500 institution's in Nature Index ranking. Those positions are in line with Hicks (1995) evaluation that leading firms published as much as medium size universities: IBM is one position above the University of Ottawa (Canada) and the University of Sussex, for instance, is in the 366th position.

But this Index also shows a further point in the logic of firms' involvement with basic research: they have international collaboration. Table 2 shows both the leading firms in scientific production, according to the *Nature Index* and the percentage of those articles that was produced in international collaboration.

TABLE 2

***Nature* Index, leading corporations according to total of articles and the percentage of articles with international collaborations**

(2016)

Rank	Corporate Institution	Articles	Inter. Colab.
1	IBM Corporation, USA	129	0.7053
2	F. Hoffman-La Roche AG, Switzerland	136	0.9396
3	Novartis International AG, Switzerland	92	0.8851
4	Merck KGaA, Germany	76	0.4899
5	AstraZeneca plc, United Kingdom (UK)	59	0.5584
6	Samsung Group, South Korea	76	0.54
7	GlaxoSmithKline plc. (GSK), United Kingdom (UK)	71	0.614
8	Interuniversity Microelectronics Centre (IMEC), Belgium	45	0.6169
9	Pfizer Inc., USA	50	0.5196
10	NTT Group, Japan	34	0.3154
11	Amgen Inc., USA	39	0.7874
12	BGI, China	53	0.7919
13	Bristol-Myers Squibb (BMS), USA	21	0.2478
14	Toyota Group, Japan	23	0.6537
15	GNS Science, New Zealand	54	0.6397
16	Sanofi, France	31	0.7977
17	Eli Lilly and Company, USA	28	0.4811
18	BASF SE, Germany	45	0.7243
19	Hitachi, Ltd., Japan	27	0.798
20	Bruker Corporation, USA	44	0.8531
21	Microsoft Corporation, USA	27	0.5638

Source: Nature Index (http://www.natureindex.com/annual-tables/2016/institution/corporate/all), authors' elaboration

Table 2 shows that, according to the *Nature Index*, the scientific production of those leading firms has strong international collaboration: 17 out of 21 firms have international collaborations in more than 50% of their articles.[1]

Those data help to fill another set of arrows in Figure 1, as they present interactions in scientific knowledge flows connecting firms and other institutions in other countries. The presence of firms in those networks may be a consequence of the consolidation of global coperation networks in science that might be influencing and orienting R&D in other dimensions, as Wagner et all (2015, p. 12) suggested. At least, this high quality participation of firms in networks of international cooperation in science indicates a more complex set of flows connecting - and tensioning - different national systems of innovation.

As our data show in Table 1, there are 418,000 papers with international collaboration in 2015, as many papers as the global scientific production in 1993. So, size matters (Anderson, 1972). Therefore an increase may be generating, as suggested by Wagner et al (2005 and 2015), a new dynamics that may affect the overall system of innovation.

How to trace and measure this tip of the iceberg of global knowledge flows from science is the task of the next section.

3. DATA AND METHODOLOGY

One contribution of this paper is a methodology to analyze data of ISI-indexed papers. To analyze those data, a matching strategy connecting information from other three databases (Orbis, USPTO and PATSTAT) was proposed and implemented for this paper.

3.1.The Database

Our starting point is a database generated through information provided by Thomson Reuters' WebOfScience (www.webofknowledge.com). From this source- WebOfScience - on papers published from 2000, 2003, 2006, 2009, 2012 and 2015, totaled more than 10 million articles, according to Table 1. The basic search in the WebOfScience involved: a) publications that included articles, meeting abstracts, book reviews etc (see Appendix Table A1); b) publications from all fields;[2]c) sources such as journals, conference proceedings, etc (our database included

---

[1] Articles from firms may be of high scientific quality. This feature may be hinted by Nobel Prizes awarded to scientists working in firms (Schlagberger et al, 2016). A good case study of a transnational corporation generating basic science through a large international network is presented by a Nobel Prize winner that worked in Merck (Campbel, 2015).

[2] WebOfScience has 252 different S&E disciplines. To process those data by more aggregated areas we applied Braun et all (1995a, 1995b) suggestion - as close as possible: 170 S&E disciplines were aggregated into 27 fields (see Table

information from 105,840 different sources[3] - this is an important difference vis-à-via *Nature Index*, that includes 68 top journals); d) the data for our database involved the following topics on each document: title, author(s), institutional location(s) (address: institution, country), science and engineering field, information on funding of the paper.

To analyze those data, we need information about the nature of the institution (university, firm, etc) and we need to process information to homogenize different names and ways of writing the name of one institution. That is why we need those other three databases.

First, a database on firms and institutions - ORBIS (http://www.bvdinfo.com/en-gb/home). This database helps our data processing in two main ways, with information on GUO and NACE sectors. GUO (Global Ultimate Owner) is a key variable that organizes under one name all different units and branches of one firm. The GUO is the entity in the highest level of the ownership chain of a network of entities, i.e., it is the entity that is not a subsidiary of any other one. GUO is the top of a hierarchy, to use Williamson's concept of modern multidivisional firm (Silva, 2014). An entity is considered a subsidiary of another when the owner share in the company is higher than 50.01%. Example: Pfizer US, Pfizer Switzerland, etc, all will be processed under one single GUO. And those GUO are classified under NACE sectors, a contribution to understanding different economic and manufacturing sectors. The Nace sector is the classification of economic activities developed by the European Community (European Commission, 2008). This paper use the first two digits of this classification, presented in Appendix I (Table A1).

Second, a database on USPTO patents. This database - used in other analyses of our basic framework (see Ribeiro et al, 2014) - contains the full text of patents (including claims, description, citations, authors, assignee, IPC, etc.) and covers he grants and applications from 2000 to 2014 at the USPTO, totaling more than 6 million patents.

Third, EPO's Patent database (Patstat) covering about 70 million patents from about 90 patent offices worldwide. Patstat includes a table containing a list of standardized names that reduces the diversity of different spelling of firms' names. This standardization helps our data processing.

3.2.A Matching Strategy and our Methodology

Why do we want to integrate the database on scientific papers and the other three? The reason is very simple: we would like to connect the institution of the first author to a GUO, when possible. To do this, we need standardization and a search in the ORBIS database with a name of the institution as close to the standard name as possible.

A2,Appendix for this list). In this aggregation 82 S&E disciplines were not used, since they are related to human and social sciences and have no correspondence in the aggregation strategy of Braun et al (1995a, 1995b).

[3] We do as Wagner et al (2005, p. 1611) that uses all types of documents in their analysis.

How do we connect the name of the institution of the first author of the paper with a GUO in the ORBIS database?

STEP 1: START THE SEARCH: in this step we get the name of the institution that the first author is affiliated - we start with around 10 million first authors (see Table 1).

STEP 2: STANDARTIZATION OF THE NAME - FIRST ROUTE: in this step we try to organize under a single and a standardized name this institutional name. For this, we begin with a search in the database PATSTAT and its list of standardized names. If the names are exactly the same, we successfully found an association between the first author's institution with the standardized firm ID from Patstat.

STEP 3: STANDARTIZATION OF THE NAME - SECOND ROUTE: If the names are not the same, a second search is processed: we search the longest common substring (LCS) of each word of the first author's institution in the list of standardized names of Patstat and search the LCSs of the first author's institution. If successful, this search associates the full name of the institution in Patstat with different abbreviations of the name in the scientific papers database.

STEP 4: STANDARTIZATION OF THE NAME - THIRD ROUTE: if no associations have been found in the two previous steps, our algorithm implements a Google Search with the first author's institution and that gets the first result obtained. Associated the first author institution name with the standardized firm ID from Patstat, when the first Google result matches to standardized name using the same LCS procedure above. In this step we have associated misspelled or using more complex abbreviation institutions names with the standardized Patstat name.

STEP 5: PREPARING THE SEARCH IN THE ORBIS DATABASE: After associating the institutional names from ISI papers to the standardized Patstat names we search the USPTO database - a connection between the information of the patents applied by these institutions with the standardized names associated with papers of our first database.

STEP 6: CONNECTING THE INSTITUTION OF THE PAPER'S AUTHOR AND THE ORBIS DATABASE: From patent information we get the publication number (USPTO identification), and this patent number is searched in the list of patents available in a database prepared from Orbis for Firms that applied patent in UPTO from 2000 to 2014. In this step, as we find the patent identification in the Orbis list, we finally have associated the first author institution with the firm ID in Orbis database. Therefore, we have information regarding the firms/institutions in the ORBIS database associated with an institution that published at least one ISI indexed publication - a link between an author and a GUO.

The outcome of this data processing based on our matching strategy is shown in Table 3. For all years in our database of scientific publication, we got the GUO of at least 31% of the first author's institution.

TABLE 3
Total of Web of Science articles and the percentage of papers with first author identified in ORBIS (2000 - 2015)

Year	1st Author Institution	% of Total Papers
2000	1098286	37.6
2003	1203816	37.2
2006	1335445	36.5
2009	1718450	33.2
2012	1886181	31.9
2015	1850945	31.2

Source: WebOfScience, ORBIS, authors' elaboration (Database A, first column; Database B, second column)

The average of those six years is 34.6%. Table 3 informs the analysis that we will do in the next sections - it is important to note the limitations of our methodology, the scope of the data that we processed for some tables, and avenues for improvement in the methodology.

The information that we are missing beyond our 34.6% of identified GUOs might be mainly related to universities and research institutions. ORBIS is a database focused on firms, therefore information on universities - a sector with important share of public and non-profit organizations - seems not to be systematically collected and organized. But, what might be missing in relation to universities may be gaining in relation to firms - the possibility of a more careful and informed analysis of their scientific production. This is the reason why we want to organize this connection between those two databases.

Keeping in mind those biases, how our databases and our methodology differentiate from other available in the literature? *Nature Index* presents data from a more selected number of publications - 68 top journals - whileour database uses 105,840 different publications. Beyond this broad scope, the main contribution of our methodology is the connection with ORBIS, which allows our analysis to investigate the sectoral distribution of scientific production from firms' side.

For this paper, we may explore those two different databases: 1) the full database, with only WefOfScience data (we will call it Database A); 2) a database prepared by our methodology, using WebOfScience and ORBIS data (we will call it Database B). Depending on the subject to be analyzed, we can choose which database is adequate.

4. AN OVERALL PICTURE: INTERNATIONAL CO-AUTORSHIP IN SCIENTIFIC PRODUCTION

This section organizes a preliminary presentation of our databases, providing basic statistics about them. Three topics are described for this preliminary mapping of international co-authorship: countries, S&E disciplines and types of interactions.

4.1.Countries

There are 206 countries in our database for 2015. Table 4 shows the scientific production ranked by leading countries, showing the total production and the international collaboration (co-authorship) in 2015. USA, China, England, German and Japan are the leaders in the scientific production of 2015, representing 47.9% of total production. These same countries, with the exception of Japan, also lead the international co-authorship articles. These five countries in 2015 represented 18.78% of total international collaborative scientific production, being less concentrated than the total production. So, an important observation is that collaborative scientific production is more spread around the Global Knowledge System.

TABLE 4
Ranking of leading countries according to total of articles and the percentage of their international co-authorships (2015)

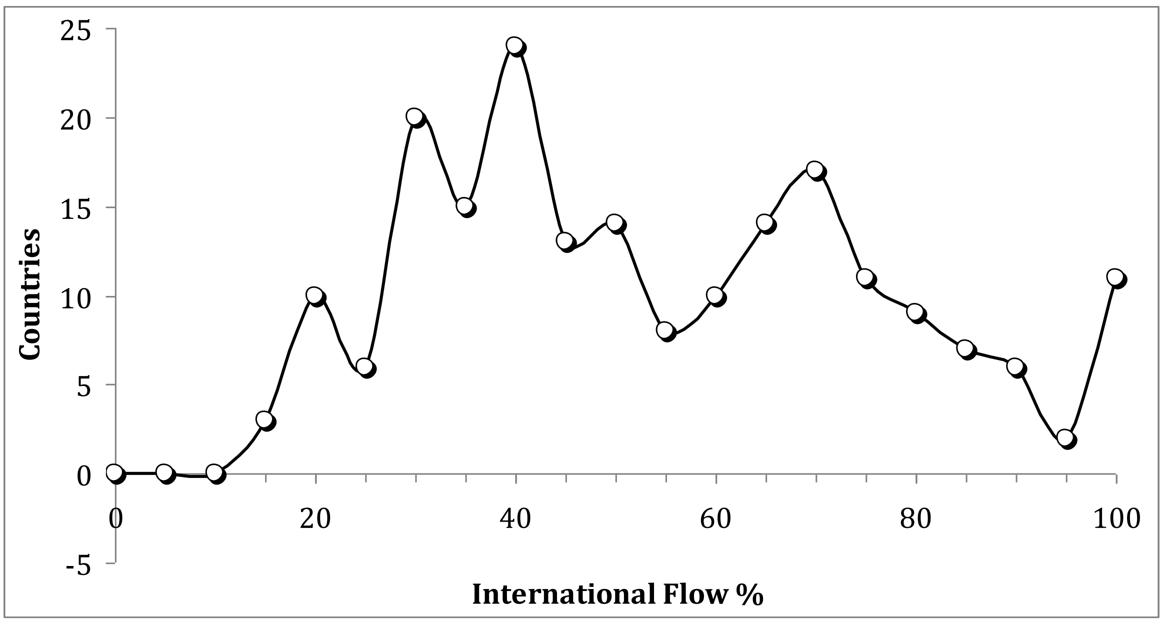
Rank	Country	Articles	% Inter
1	USA	415421	15.8%
2	PEOPLES R CHINA	278613	17.9%
3	ENGLAND	91111	28.4%
4	GERMANY	87891	29.6%
5	JAPAN	72330	13.3%
6	INDIA	59732	12.7%
7	ITALY	59671	27.8%
8	CANADA	55496	26.3%
9	FRANCE	55150	31.0%
10	SOUTH KOREA	53446	15.3%
11	AUSTRALIA	51590	27.2%
12	SPAIN	48872	28.4%
13	BRAZIL	38650	20.7%
14	RUSSIA	31422	16.7%
15	NETHERLANDS	29840	35.1%
16	TURKEY	28721	10.3%
17	IRAN	27443	16.5%
18	POLAND	24110	17.1%
19	TAIWAN	23836	15.2%
20	SWITZERLAND	19152	42.9%

Source: WebOfScience, authors' elaboration (Database A)

Since we have data since 2000, to be further evaluated in other works, we may follow China's evolution: in 2000 China was at the 8th position (12th in international flow), in 2006 it reached the 2nd position (5th in international flow), and in 2012, she was ranked at the 2nd position in both.

The distribution of countries according to their percentage of international co-authorship is summarized in Graph 1 - a histogram that shows the number of countries in each range of percentage of international co-authorship (5% is the range chosen).

GRAPH 1
HISTOGRAM - distribution of 206 countries according to the percentage of their articles with international co-autorship
(percentage of articles with international co-autorship X number of countries)
(2015)



Source: WebOfScience, authors' elaboration (Database A)

Three peaks may be shown in Graph 1.

The first, around 30%, (20 countries) - is composed basically by countries that their national systems of innovation are not completely formed - countries that are in "regime of interaction" 2 according to Chaves et al (2017): examples are Mexico (25.11%), Philippines (27.49%), South Africa (28.08%) and Thailand (29.90%). In this peak there are also countries in the "regime of interaction" 3: Canada (26.27%), England (28.42%) and Germany (29.58%) - not so small "regime 3" countries.

The second peak, around 40%, is predominantly from small countries located in the 3rd "regime of interaction": Sweden (36.38%), Netherlands (35.12%) are examples of this set of countries. There are also smaller countries from "regime of interaction" 2: Chile (39.70%).

Finally, the third peak (around 70%, with 21 countries) is composed only by countries located at "regime of interaction" 1: Uganda, Ecuador, and Kenya. It seems that beyond 60% of international co-authorship, there are only countries located at "regime of interaction" 1 or countries in a worst position than that - countries without USPTO applications.[4]

Those correlations are to be further evaluated, but, a preliminary analysis may suggest some patterns. First, least developed countries, countries with rudiments of national systems of innovation, depend strongly of international cooperation to start (Kruss et al, 2015, p. 15), therefore their high levels of international co-authorship. Second, dynamic innovation systems of small countries are more internationalized than the average. Thirdly, larger countries with strong national scientific base are proportionally less internationalized than the average, although they are leaders in absolute terms - see the positions of USA and China in Table 4.

4.2. WebOfScience S&E Disciplines

There are 252 WebOfScience S&E disciplines in our database. Table 5 shows the scientific production ranked by leading S&E disciplines, displaying the total production and the percentage of international collaboration (co-authorship) in 2015.

[4] Exceptions: Qatar ("regime of interaction" 2, with 63.47% of international co-authorship); and very small European countries - Andorra, Liechtenstein, Monaco, San Marino.

TABLE 5
Leading Science and Engineering disciplines and international co-autorship
(total and percentage of articles with international co-autorship)
(2015)

Web of Science	Articles	% Inter Flow
CHEMISTRY, MULTIDISCIPLINARY	64,525	21.27
BIOCHEMISTRY E MOLECULAR BIOLOGY	64,342	23.16
MULTIDISCIPLINARY SCIENCES	62,161	29.88
ONCOLOGY	56,323	18.73
CHEMISTRY, PHYSICAL	45,922	26.59
ENGINEERING, ELECTRICAL E ELECTRONIC	45,203	20.87
CLINICAL NEUROLOGY	40,966	18.45
CARDIAC E CARDIOVASCULAR SYSTEMS	36,588	17.65
MEDICINE, GENERAL E INTERNAL	34,394	14.05
GASTROENTEROLOGY E HEPATOLOGY	32,063	13.31
MATERIALS SCIENCE, MULTIDISCIPLINARY	31,277	21.33
ENVIRONMENTAL SCIENCES	28,788	27.92
NEUROSCIENCES	23,633	25.46
PHARMACOLOGY E PHARMACY	23,286	19.27
PUBLIC, ENVIRONMENTAL E OCCUPATIONAL HEALTH	23,069	25.3
COMPUTER SCIENCE, ARTIFICIAL INTELLIGENCE	22,884	24.36
BIOTECHNOLOGY E APPLIED MICROBIOLOGY	22,878	23.52
MATHEMATICS, APPLIED	22,369	27.89
PLANT SCIENCES	20,861	28.77
PHYSICS, MULTIDISCIPLINARY	20,532	25.38
ENDOCRINOLOGY E METABOLISM	20,481	21.32

Source: WebOfScience, authors' elaboration (Database A)

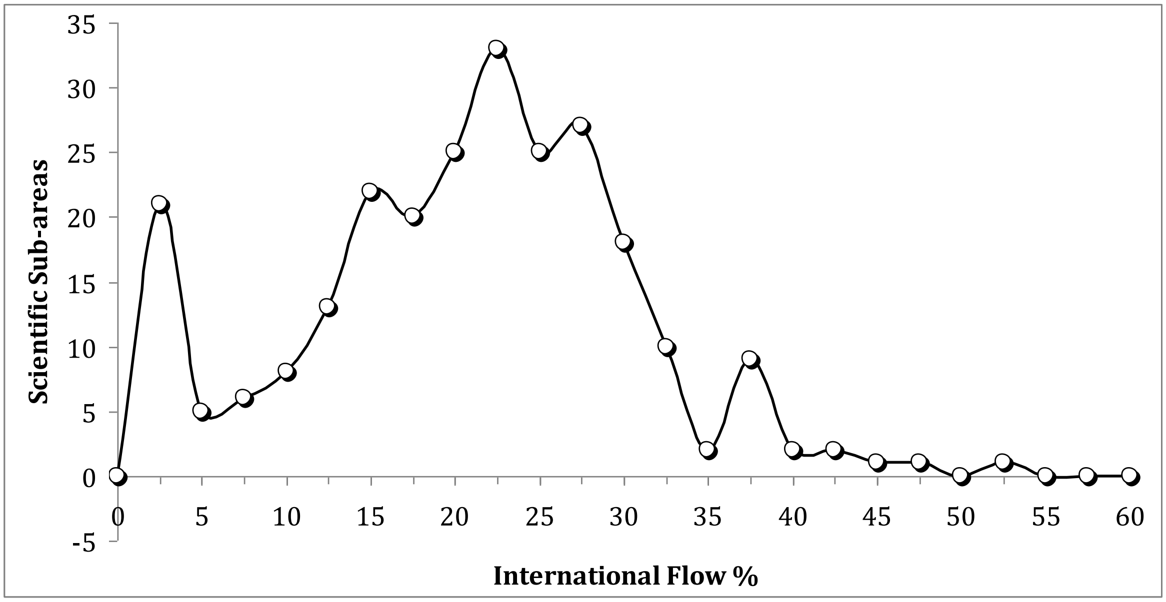
Table 5 presents the leading S&E disciplines as organized by the WebOfScience, their total and their percentage of international co-authorship. As in the distribution for countries, the S&E disciplines with more articles are not the disciplines with more international co-authorship. The two leading S&E disciplines in scientific production (Chemistry and Biochemistry) are not the leading disciplines in percentage of international co-authorship (Astronomy and Physics, particles). However, if in the case of countries the highest levels in the percentage of international co-authorship were correlated to least developed countries, in the case of S&E disciplines other logic might be operating - international nature of S&E fields might be an explanation.

The distribution of S&E disciplines according to their percentage of international co-authorship is summarized in Graph 2 - a histogram that shows the number of S&E disciplines in each range of percentage of international co-authorship (2.5% is the range chosen).

GRAPH 2

HISTOGRAM - distribution of 252 S&E disciplines according to the percentage of their articles with international co-autorship

(percentage of articles with international co-autorship X number of S&E disciplines (2015)



Source: WebOfScience, authors' elaboration (Database A)

Graph 2 shows three peaks.

The first peak, around 2.5% of international co-authorship, involves 21 S&E disciplines - all are not related to the S&E fields organized by Braun (1995a, 1995b). In fact, until 5% of international co-authorship the disciplines are not of those S&E fields.

The second peak, around 22.5% (close to the international average of 21.3%), involves 33 S&E disciplines, that together with neighbor ranges 15% to 30% there are 180 S&E fields (the leading fields in Table 6 are in this broad region).

The third peak, around 37.5%, involves 9 S&E disciplines, highly internationalized - examples are Geochemistry, Parasitology, and Meteorology.

4.3.International Co-Authorship by Types of Interaction

The data on international co-authorship have been presented so far as numbers of publications with international co-authorships and percentages. Those data, for the purposes of this paper, present the scope of this international phenomenon (21.3% of total scientific production, as big as the global scientific production in 1993). However, those data might underestimate the importance of those international co-authorship flows. An article is produced by a research effort, generally a collaborative effort. Articles with international co-authorship are produced by international research efforts that involve more than two different institutions. Therefore, one article with international co-authorship may involve more than one international flow linking two institutions in two different countries - one article may generate various international flows. International flows may grow more than international co-authorship.

Table 6 shows those important data.

TABLE 6
Size of the network of international co-authorships, according to number of pairs (links) by interaction types
(2000 - 2015)

Interaction Type	2000	2003	2006	2009	2012	2015
RI-RI	530424	566816	968723	1613794	7257669	7019906
RI-Firm	14154	17983	20717	31668	56129	60940
Firm-Firm	794	856	1085	1324	3309	2229
Total	545372	585655	990525	1646786	7317107	7083075

Source: WebOfScience, authors' elaboration (Database A)

Table 6 shows that, although international co-authorship has grown 3.06 times between 2000 and 2015 - reaching 418,866 articles with international co-authorship -, the international flows between institutional authors have grown 12.99 times - reaching 7,083,075 cross-border pairs, or individual international flows, in 2015. Those very simple calculations based on data from Table 6 are informative enough to suggest important network properties of those scientific international collaborations, already identified by Wagner et al (2005, 2015).[5]

Those international flows are dominated by research institutions. The first line shows the academic interactions between Research Institutes (RI-RI). The flows among them were 530,424 in 2000 and exponentially grew to 7,0199,966 pairs in 2015. The second line of Table 7 shows interactions between Research Institutes and Firms, which grew 4.31 times between 2000 and 2015, from 14,154 pairs in 2000 to 60,940 in 2015. The third line shows Firms-Firms pairs that grew only 2.81 times, from 794 to 2,229 pairs.

Table 6 puts forward the dimension of the phenomenon that this paper investigates: given the networks properties displayed by our analysis, international co-authorships involve 7,083,075

[5] Wagner et al (2015, p. 6) analyze data from 1990 to 2011, and find a growth in the "number of coauthor relationships (links)" that is "disproportionately large compared to the growth in the number of addresses".

cross-border pairs or individual international flows. A phenomenon that goes far beyond what the figures for articles with international co-authorships may show. Just to stress how important are those data from Table 6 for our tentative framework presented in Figure 1, the reader could imagine a World Map with 7,083,075 arrows connecting universities and firms across planet Earth.

5. A PRELIMINARY ANALYSIS OF THE NETWORK OF INTERNATIONAL CO-AUTHORSHIP

The main findings of section IV are the exponential growth of the international co-authorship in scientific papers and the size of the pairs of co-authorship. In other words, the growth of knowledge flows is greater than the growth of articles with international co-authorships. This section focuses in the 7,083,075 pairs of international co-authorships to grasp the meaning of this network. Since this network is an important backbone of cross-border interactions that may be supporting an emerging global innovation system, a more detailed analysis on its specific features is important. This section investigates this network and its connections through two different approaches, each one with a different node in its structure: 1) institutions; 2) countries.

5.1.Institutions as Nodes of the Network

Table 7 shows each institution as a node of the network and the total connections in each node. Since each of the 7.08 million links shown in Table 6 have two participants, the total of column Links in Table 7 is 14.16 million. The most connected institutions and their links are displayed in Table 7.

TABLE 7
NETWORK ANALYSIS:
Leading institutions (network nodes) according to the number of links with international co-authors (2015)

International Coauthorship Links per Node - 2015	
Node	Links
UNIV OXFORD (ENGLAND)	54813
UNIV CAMBRIDGE (ENGLAND)	50481
UNIV TORONTO (CANADA)	47843
HARVARD UNIV (USA)	47100
UCL (ENGLAND)	45773
UNIV COPENHAGEN (DENMARK)	45395
CHARLES UNIV PRAGUE (CZECH REPUBLIC)	45074
UNIV ATHENS (GREECE)	45073
CNRS (FRANCE)	44920
UNIV BOLOGNA (ITALY)	44447
UNIV EDINBURGH (SCOTLAND)	44087
CENTRE ETUDES RECHERCHE (SWITZERLAND)	42351
CHINESE ACAD SCI (PEOPLES R CHINA)	41786
IST NAZL FIS NUCL (ITALY)	41767
UNIV LONDON IMPERIAL COLL SCI TECHNOL E MED (ENGLAND)	41477
UNIV BELGRADE (SERBIA)	40466
UNIV MELBOURNE (AUSTRALIA)	40418
HEIDELBERG UNIV (GERMANY)	40299
UNIV MALAYA (MALAYSIA)	40103
MOSCOW MV LOMONOSOV STATE UNIV (RUSSIA)	40099
UNIV PISA (ITALY)	40025
UNIV PARIS 11 (FRANCE)	39977
OHIO STATE UNIV (USA)	39735
UNIV MANCHESTER (ENGLAND)	39203
MIT (USA)	39166
UNIV NAPLES FEDERICO II (ITALY)	38489
UNIV GENOA (ITALY)	38433
UNIV SYDNEY (AUSTRALIA)	38248
UNIV AUTONOMA MADRID (SPAIN)	38085
UNIV GLASGOW (SCOTLAND)	37886
UNIV BRITISH COLUMBIA (CANADA)	37257
LUND UNIV (SWEDEN)	37198

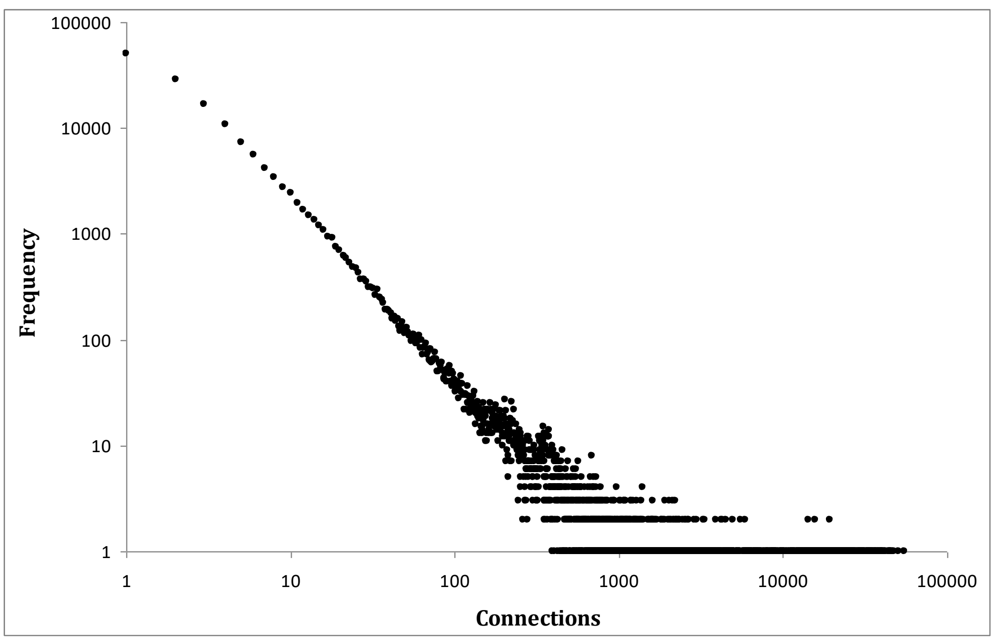
Source: WebOfScience, authors' elaboration (Database A)

Table 7 shows universities leading the ranking of connections (8 first positions). The logic of connections is different than the logic of total of articles and of the logic of total international co-authorships: the Chinese Academy of Sciences leads both rankings (Table 4), but is in the 13th position in the ranking of connections (Table 7).

Location of most connected firms would be around the 545th position (Novartis, with more than 5,000 connections), and the 7,200th position (Robert Bosch, with 80 connections).[6]

Graph 3 presents the distribution of the 163,410 institutions according to the number of connections that they have presented - a histogram that shows the number of institutions (frequency) according to the number of connections that they have. There are 50,889 institutions with one link, and one institution with 54,813 links.

GRAPH 3
HISTOGRAM: distribution of institutions with international co-authorships according to the number of connections
(2015)



Source: WebOfScience, authors' elaboration (Database A)

[6] A comparison with patent citation of ISI-indexed paper as a knowledge flow for firms and co-authorship of scientific papers: in 2009, the leading firm in patents with citation of ISI-indexed paper was Microsoft, with 333 patents - or links (Ribeiro et al, 2014, Table 2, p. 72). Microsoft has more than 650 connections (links) through international co-authorship. The leading firm in connections through international co-authorship in our preliminary calculations is Novartis, with more than 5,000 links.

The power law (the straight line in a log x log scale) behavior of the distribution of connections (links) indicates that the co-authorship network is a free-scale one. So, it shows a small number of nodes with a very high connection (called hubs) and a huge number of nodes with not so many connections. The hubs represent the institutions that attract and concentrate a large number of international collaborations translated into co-authorship in papers. Their behavior influence the global features of the network (they might define the mainly research interests). In general, this sort for network are generated by systems that show a self-organization property, therefore, the power law in Graph 3 may indicate that the system that generated it is in a self-organization state.

The identification in the network of international co-authored scientific papers of properties of a self-organized system, already highlighted by Wagner (2005, 2015), is a finding that supports an evaluation of the contribution of science and its flows for the emergence of a global system of innovation.

5.2.Countries As Nodes Of The Network

Table 8 shows 2.31 million links. Since one paper may have more than one co-author institution, their country will count only once in this paper. Therefore, there are less links in Table 8 than in Table 7.

Once more, the ranking of countries according to connections in the international co-authorships network is different from conventional rankings (total papers and papers with international co-authorship). The USA keep the leading position in those three rankings (see Tables 4 and 8), while China is at the second position in Table 4 but in the 4th position in Table 8.

TABLE 8

NETWORK ANALYSIS: Ranking of leading countries as nodes according to total of links in their international co-authorships, and the number of countries that those links connect (2015)

Node	Links	Countries
USA	273628	202
ENGLAND	150343	193
GERMANY	139864	185
PEOPLES R CHINA	108608	173
FRANCE	107849	188
ITALY	91429	187
SPAIN	78042	178
CANADA	76588	182
AUSTRALIA	73851	186
NETHERLANDS	69379	179
SWITZERLAND	61305	181
JAPAN	50182	171
SWEDEN	48931	173
BELGIUM	43943	174
BRAZIL	37141	171
DENMARK	36190	162
AUSTRIA	34259	158
SCOTLAND	32850	165
SOUTH KOREA	32695	147
RUSSIA	32656	153
POLAND	32382	136
INDIA	30708	167
NORWAY	25934	158
FINLAND	25291	147
PORTUGAL	24833	156
CZECH REPUBLIC	24631	143
GREECE	23737	140
TAIWAN	21663	144
TURKEY	20962	148
SOUTH AFRICA	20851	171

Source: WebOfScience, authors' elaboration (Database A)

The distribution of connections per country is also very uneven. The United States of America have connections with 202 other countries, China with other 173 countries, Netherlands with 179 countries and Brazil with 171 countries. Least developed countries are connected with fewer countries: Mozambique with 101 countries, Bolivia with 78 countries, Liberia with 42 countries, and Angola with 32 countries.

This network connects all 206 countries. As discussed in sub-section IV.1, least developed countries or countries with rudiments of an innovation system, begin their science and technology activities with strong support from more developed countries.

6. SCIENTIFIC PRODUCTION AND INTERNATIONAL CO-AUTHORSHIP ACCORDING TO ECONOMIC AND MANUFACTURING SECTORS

This section focuses the distribution of scientific production across economic and manufacturing sectors. For this analysis, data from Orbis is necessary - therefore we use our Database B in this section. It is important to remember that we were able to find the first author of 31.2% of papers published in 2015 (see Table 3).

6.1 A Ranking of Nace Sectors

Table 9 shows the leading NACE sectors (see Appendix Table A2 for a list of sectors) in scientific publications, data for 2015.

TABLE 9
Leading NACE sectors in articles published and percentage of articles with internacional co-authorship
(2015)

Rank	NACE	Article	% Inter Co-authorship
1	85	325606	20.87%
2	86	40824	24.07%
3	21	36788	23.10%
4	72	17428	26.38%
5	84	14714	21.03%
6	68	5739	28.82%
7	30	2988	25.37%
8	77	2808	33.51%
9	58	2661	29.73%
10	46	2529	22.42%
11	55	2522	32.04%
12	82	2450	26.98%
13	94	2260	21.37%
14	64	2141	28.44%
15	26	1670	23.59%
16	65	1558	12.90%
17	20	1465	32.70%
18	96	957	34.59%
19	56	884	35.29%
20	71	847	23.61%
21	47	795	22.89%
22	74	685	28.61%
23	28	503	20.08%
24	32	285	12.63%
25	62	229	21.40%
26	22	209	14.83%
27	27	170	11.18%
27	29	170	19.41%
28	24	163	11.04%
29	61	160	20.00%
30	79	146	14.38%
31	10	108	13.89%
32	19	87	8.05%
33	23	47	14.89%
34	70	40	42.50%

Source: WebOfScience, ORBIS, authors' elaboration (Database B)

The leading NACE sector is 85 (Education). Although, as mentioned in section III, Orbis is not focused in the type of institutions that prevail in this sector (non-profit or governmental), there are enough data to preserve their leading position. This position is not a surprise, given the information displayed in Table 7. The percentage of international co-authorship is near the general average.

The second position is sector 86 (Human health services). This is an important indication for further research, given the position of the two leading institutions of this sector in the ranking of links shown in Table 7: Mayo Clinic is in the 402nd position and Massachusetts General Hospital in the 462nd position - above the first firm, Novartis.[7] The percentage of international co-authorship of this sector is slightly greater than the average - 24.07%.

The third position is NACE sector 21 (Manufacture of pharmaceuticals, medicinal chemical and botanical products) - the first manufacturing sector in this list. The percentage of international co-authorship is also slightly above the average: 23.10%.

Since 2000 the ranking of those three sectors is the same.

The other manufacturing industry in the list of top ten NACE sectors is "Manufacture of transport equipment" (NACE 30), with a percentage of international co-authorship of 25.39%.

It is interesting to note that there are 14 manufacturing sectors among the top 35 listed in Table 9. Only 4 have percentages of international co-authorship greater that 21.3% (20, 21, 26 and 30).

6.2.Nace Sectors and S&E Fields

A closer look to manufacturingsectors is worthwhile. This subsection, on the one hand, focuses on NACE sectors related to those sectors - from NACE sectors 01 to 33, plus sectors 61 to 63 (telecommunications and IT). On the other hand, this section investigates the S&E fields that firms from those sectors publish.

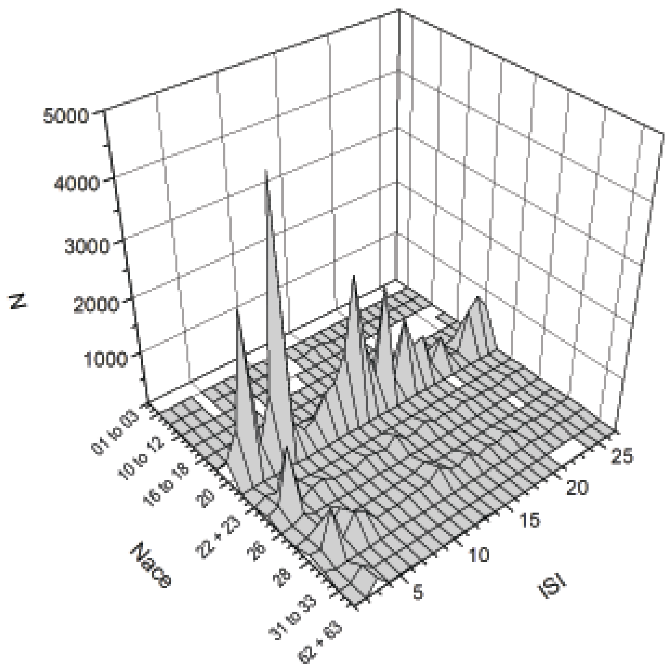
The outcome of this investigation is in Graph 4, which presents a Matrix with total publications by each cell from a NACE sector and an S&E field.[8]

---

[7] The importance of hospitals (included in NACE 86) for innovation has been stressed by many works. Gulbrandsen et al (2016) evaluate this role, as for them "[h]ospitals are, in other words, organizations that play many essential roles in medical innovation: production and diffusion of knowledge, linking of practice with science and technology, use of and feedback on prototypes and concepts, and implementation of new medical routines, devices and procedures. All hospital outputs and services may be tied to innovation" (Gulbrandsen et al, 2016, p. 1493),

[8] Those matrices technology sectors X scientific sectors were prepared in Ribeiro et al (2010).

GRAPH 4:
Matrix Selected Nace Sectors X 27 S&E Fields
(2015, all papers)

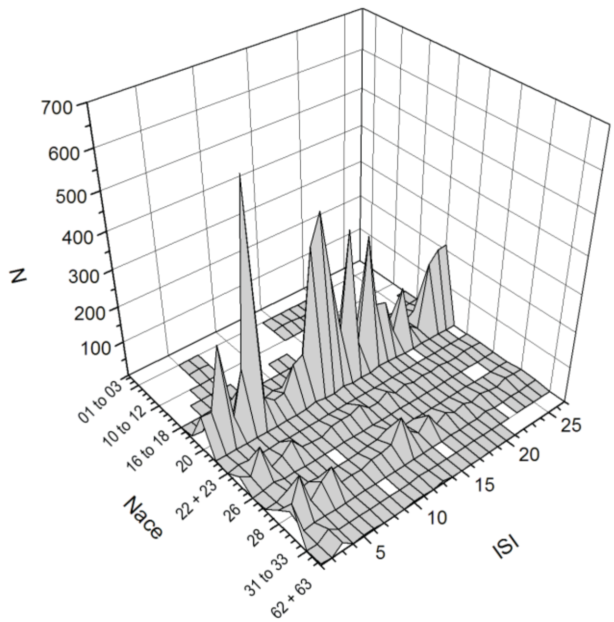


Source: WebOfScience, ORBIS, authors' elaboration (Database B)

Graph 4 shows empty cells in NACE sector 01 to 09 (agriculture and mining). The peaks are in NACE sector 21: first the cell NACE 21 and ISI 6 (Inorganic Chemistry and Eng.), second NACE 21 and ISI 3 (Electronic Eng.). Another peak is in NACE 26 and ISI 3. NACE 29+30 also has a peak at ISI 3. A predicted pattern of scientific production can be found in NACE 10-12, with its peak at ISI 15 (Food Science and Agriculture) and in NACE 31-33, with a peak at ISI 1 (Mathematics).

Graph 5 was built using only internationally co-authored papers.

GRAPH 5:
Matrix Selected Nace Sectors X 27 S&E Fields
(2015, papers with international co-authorship)



Source: WebOfScience, ORBIS, authors' elaboration (Database B)

It is a less compact matrix, with more empty cells. However, the general structure shown in Graph 4 is found in Graph 5: the peaks are in the same cells, with fewer articles, as Table 9 shows.

7. CONCLUDING REMARKS: INTERNATIONALIZATION OF SCIENTIFIC PRODUCTION AND ITS CONTRIBUTION TO THE EMERGENCE OF AN GSI

International knowledge flows in science, captured through international co-authorship in scientific papers, have grown steadily in the last decades. In 2015 there were 418,866 articles with international co-authorship. The investigation of those flows is important to support, statistically and empirically, our tentative framework presented in Figure 1. This manuscript, therefore, provides another piece of our puzzle. How does it articulate with other works? In a previous article we investigated another set of knowledge flows: flows through patent citations of ISI-Indexed papers, as clues for ways that firms (with their patents) used knowledge generated in research institutions. In that investigation, with data for 2009, we found 70,000 links - a patent on one side, a paper in the other side, predominantly a connection between a citing firm and a cited university (Ribeiro et al, 2014, p. 77). A rough quantitative comparison juxtaposes 418,866 international co-authorships and 70,000 cross-border patent citations of scientific papers. Cross-border flows from international co-authorships in articles are broader than flows from patent citations of ISI-indexed papers: the USA, for instance, in 2015 have connections with 202 countries through international co-authorships (see Table 8), while have only 81 countries connected through patent citations in 2009 (Ribeiro et al, 2014, p. 77).

The analysis of data on international co-authorship on how institutions, countries, firms, economic sectors are distributed around the global average (21.3% in 2015) revealed how consistent is this phenomenon – and those differences in relative international co-authorships help the analysis of those variables.

However, those rough data on total of articles with international co-authorship apparently underestimate the nature of this phenomenon: the links between co-authors grow faster than the growth of scientific papers and international co-authorships. Properties of networks were unveiled by our analysis, which may be illustrated by a comparison between 418,866 articles with international co-authorships and 7,083,075 links between those co-authors. A network that has power law properties and that is a free scale network.

The nature of this network evaluated by our analysis also revealed elements of a self-organized system. This is important to highlight a contribution of the network built from the scientific production to the process of emergence of rudiments of a global system of innovation. Science and the knowledge flows are important components in our tentative framework, not only quantitatively – size of the flows -, but also qualitatively – the nature of those flows and the properties of their network.

Those findings are important, but we need to stress the exploratory nature of this manuscript - first round of processing and analyzing a rich database. A specific methodological contribution is presented, because the articulation between a database built from the WebOfScience and ORBIS was a precondition for an evaluation of the distribution of scientific papers and international co-authorship across different economic and manufacturing sectors.

Further research may follow three broad lines. First, to improve our methodology and matching strategy, basically to increase the percentages of Table 3. Second, to explore the data to present analysis of how firms and hospitals participate in knowledge flows captured by international co-authorships. And, third, to elaborate and re-elaborate our tentative framework, to include those findings related to networks and self-organizing systems – in science and in technology – in the interpretation of the emergence of a global innovation system - or at least to understand newer and stronger tensions between the international nature of S&T and the national dimension of innovation systems.

REFERENCES

ADAMS, J. (2013) The fourth age of research. *Nature*, v. 497, 30 May 2013, n. xxx, pp. 557-559.

ANDERSON, P. W. 1972. More is different: broken symmetry and the nature of the hierarchical structure. *Science*, 177(4047), 393-396. doi: 10.1126/science.177.4047.393

BRAUN, T.; GLANZEL, W.; GRUPP., H. (1995a) The Weight of 50 Nations in 27 Science Areas, 1989-1993 (Part I: All fields combined, Mathematics, Engineering, Chemistry and Physics). *Scientometrics*, v. 33, n. 3, pp. 263-293.

BRAUN, T.; GLANZEL, W.; GRUPP., H. (1995b) The Weight of 50 Nations in 27 Science Areas, 1989-1993 (Part II: Life Sciences). *Scientometrics*, v. 34, n. 2, pp. 207-237.

BRITTO, G.; CAMARGO, O. S.; KRUSS, G.; ALBUQUERQUE, E. M. (2013) Global interactions between firms and universities: global innovation networks as first steps towards a Global Innovation System. *Innovation and Development*, v. 3, n. 1, pp. 71-88.

CAMPBELL, W. C. (2015) Ivermectin: a reflection on simplicity. *Nobel Lecture*. Stockolm: Nobel Foundation (https://www.nobelprize.org/nobel_prizes/medicine/laureates/2015/campbell-lecture.pdf).

CANTWELL, J. (2009) Innovation and information technology in the MNE. In: RUGMAN, A. M. (ed.), *The Oxford Handbook of International Business*. Oxford University Press: Oxford, Second Edition, pp. 417-446.

CANTWELL, J. (2013) Blurred boundaries between firms, and new boundaries within (large multinational) firms: the impact of decentralized networks for innovation. *Seoul Journal of Economics*, v. 26, n. 1, pp. 1-32.

CHAVES, C. V., RIBEIRO, L. C.; SANTOS, U. P.; ALBUQUERQUE, E. (2017) Sistemas de inovação e mudanças na divisão centro-periferia: notas sobre uma metodologia para identificar trajetórias de países a partir de estatísticas de ciência e tecnologia. Belo Horizonte: Cedeplar-UFMG (Unpublished manuscript).

CHEN, C.J., HUANG, Y.F., LIN, B.W. (2012) How firms innovate through R&D internationalisation? An S-curve hypothesis. *Research Policy* 41, 1544–1554.

DUNNING, J.; LUNDAN, S. (2008) *Multinational enterprises and the global economy*. Cheltenham: Edward Elgar (Second edition).

DUNNING, J.H., LUNDAN, S.M., (2009) The internationalisation of corporate R&D: a review of the evidence and some policy implications for home countries. *Review of Policy Research*, 26, 13–33.

EUROPEAN COMMISSION (2008) *EUROSTAT*: Methodologies and Workingpapers, NACE Rev. 2: Statistical classification of economic activities in the European Community, Luxembourg: Office for Official Publications of the European Communities.

ERNST, D. (2006) Innovation offshoring: Asia's emerging role in Global Innovation Networks. East-West Center: Honolulu.

FUCHS, M. (2014) *Worldwide knowledge?* Global firms, local labour and the region. Surrey: Ashgate.

GEORGHIOU, L. (1998) Global cooperation in research. *Research Policy*, v. 27, n. 6, pp. 611-626.

GILLIES, J.; CAILLIAU, R. (2000) *How the web was born*: the story of the World Wide Web. Oxford/New York: Oxford University Press.

GLÄNZEL, W.; SCHUBERT, A. (2005). Domesticity and internationality in co-authorship, references and citations. *Scientometrics*, v.65, n. 3, pp. 323–342.

GUELLEC, D.; POTTERIE, B. (2004) Measuring internationalization of the generation of technology: an approach based on patent data. In: MOED, H.; GLÄNZEL, W.; SCHMOCH, U. (eds). *Handbook of quantitative science and technology research*: the use of publication and patent statistics in studies of S&T systems. Dordrecht: Kluwer Academic Publishers, pp. 645-662.

HICKS, D. (1995) Published papers, tacit competencies, and corporate management of public/private character of knowledge. *Industrial and Corporate Change*, v. 4, n. 2, pp. 401-424.

HYMER, S. (1976) *The International Operations of National Firms*: A study of direct foreign investiment. Cambridge: MIT Press.

JAFFE, A.; TRAJTENBERG, M.; FOGARTY, M. (2000) Knowledge Spillovers and Patent Citations: Evidence from a Survey of Inventors. *American Economir Review*, v. 90, n. 2, pp. 215-218.

KERR, S. P.; KERR, W.; OZDEN, C.; PARSONS, C. (2016) Global talent flows. *Journal of Economic Perspectives*, v. 30, n. 4, pp. 83-106.

KRUSS, G.; LEE, K.; SUZIGAN, W.; ALBUQUERQUE, E. (2015) Introduction. In: ALBUQUERQUE, E.; SUZIGAN, W.; KRUSS, G.; LEE, K. *Developing National Systems of Innovation*: University-Industry interactions in the Global South. Cheltenham/Ottawa: Edward Elgar/IDRC, pp. 1-28 (available at http://www.elgaronline.com/view/9781784711092.00007.xml)

LAURENS, P.; Le BAS, C.; SCHOEN, A.; VILLARD, L.; LARÉDO, P. (2015) The rate and motives of the internationalization of large firm R&D (1994-2005): towards a turning point? *Research Policy*, v. 44, n. 3, pp. 765-776.

LEYDESDORFF, L.; WAGNER, C. (2008) International collaboration in science and the formation of core group. *Informetrics*, v. 2, n. 4, pp. 317-325.

LINDEN, G.; KRAEMER, K. L.; DENDRICK, J. (2007), Who captures value in a Global Innovation System? The case of Apple`s iPod. Personal Computer Industry Center/University of California: Irvine Irvine.

LOS, B.; TIMMER, M.; de VRIES, G. J. (2015) How global are global value chains? A new approach to measure international fragmentation. *Journal of Regional Science*, v. 55, n. 1, pp. 66-92.

NATIONAL SCIENCE BOARD. 1996. *Science and Engineering Indicators 1996*.Arlington, VA: National Science Foundation

NATIONAL SCIENCE BOARD. 2016. *Science and Engineering Indicators 2016*. Arlington, VA: National Science Foundation

NATURE (2016a) Nature Index (http://www.natureindex.com/) (accessed 19 november 2016).

NATURE (2016b) International research collaborations on the rise. *Nature Index* (http://www.natureindex.com/news-blog/international-research-collaborations-on-the-rise - accessed 19 november 2016).

NELSON, R.; ROSENBERG, N. (1995). Introduction. In: NELSON, R. (ed). *National innovation systems:* a comparative analysis. New York, Oxford: Oxford University.

OECD (1998), L'internationalisation de la R&D industrielle: structures et tendances. OECD: Paris.

OECD (2008), The internationalization of business R&D: evidence, impacts and implications. OECD: Paris.

OECD (2013) Interconnected economies: benefiting from global value chains. Paris: OECD.

OECD (2015) *OECD Science, Technology and Industry Scoreboard 2015*: Innovation for growth and society, OECD Publishing, Paris. http://dx.doi.org/10.1787/sti_scoreboard-2015-en

OECD and SCImago Research Group (CSIC) (2016) *Compendium of Bibliometric Science Indicators*. OECD, Paris. Accessed from http://oe.cd/scientometrics.

PATEL, P. (2011) Location of innovative activities of EU large firms. SPRU Electronic Working Paper Number 190, downloaded on 17 May 2015 from http://www.sussex.ac.uk/spru/documents/sewp190.pdf

PATEL, P.; PAVITT, K. (1998), National systems of innovation under strain: the internationalization of corporate R&D. SPRU: Brighton.

RIBEIRO, L. C.; RUIZ, R. M.; BERNARDES, A. T.; ALBUQUERQUE, E. M. (2010) Matrices of science and technology interactions and patterns of structured growth: implications for development. *Scientometrics*, v. 83, pp. 55-75.

RIBEIRO, L. C. ; KRUSS, G. ; BRITTO, G. ; RUIZ, R. M. ; BERNARDES, A. T. ; ALBUQUERQUE, E. M. (2014) A methodology for unveiling global innovation networks: patent citations as clues to cross border knowledge flows. *Scientometrics*, v. 101, pp. 61-83.

ROSENBERG, N. (1990). Why do firms do basic research (with their money)? *Research Policy*, v. 19, pp.165-174.

SCHLAGBERGER, E. M.; BORNMANN, L.; BAUER, J. (2016) At what institutions did Nobel laureates do their prizewinning work? An analysis of biographical information on Nobel laureates from 1994 to 2014. *Scientometrics*, v. 109, n. 2, pp. 723-767.

SILVA, L. (2014) Tensões e conexões: um estudo sobre multinacionais e sistemas nacionais de inovação. Belo Horizonte: Cedeplar-UFMG (Tese de Doutorado).

SOETE, L.; VERSPAGEN, B.; WEEL, B. (2010) Systems of innovation. In: HALL, B.; ROSENBERG, N. (eds) *Handbook of the economics of innovation*. Volume II. Amsterdam: North Holland.

TEECE, D. (2006) Reflections on the Hymer thesis and the multinational enterprise. *International Business Review*, v. 15, n. 2, pp. 124-139.

TIMMER, M. P.; ERUMBAN, A.; LOS, B.; STEHRER, R.; de VRIES, G. (2014) Slicing up global value chains. *Journal of Economic Perspectives*, v. 28, n. 2, pp. 99-118.

UNCTAD (2005), World Investment Report 2005 – Transnational corporations and the internationalization of R&D. United Nations Conference on Trade and Development: Geneva.

UNCTAD (2006), World Investment Report 2006 – FDI from developing and transition economies: implications for development. United Nations Conference on Trade and Development: Geneva.

UNCTAD (2011), World Investment Report 2011 – Non-equity modes of international production and development. United Nations Conference on Trade and Development: Geneva.

UNCTAD. (2011) World Investment Report 2011 – Non-equity modes of international production and development. Geneva: United Nations Conference on Trade and Development.

UNCTAD. (2013) World Investment Report 2013 – Global value chains: investment and trade for development. Geneva: United Nations Conference on Trade and Development.

WAGNER, C.; LEYDESDORFF, L. (2005) Network structure, self-organization, and the growth of international collaboration in science. *Research Policy*, v. 34, n. 10, pp. 1608-1618.

WAGNER, C.; PARK, H. W.; LEYDESDORFF, L. (2015) The continuing growth of global cooperation networks in research: a conundrum for national governments. *PLOS one*, July 21, 2015

APPENDIX TABLES

Appendix TABLE A1

Document Type	Articles
ARTICLE	5.688.001
MEETING ABSTRACT	1.113.336
PROCEEDINGS PAPER	1.107.212
EDITORIAL MATERIAL	439.781
BOOK REVIEW	410.670
ARTICLE; PROCEEDINGS PAPER	400.577
REVIEW	311.397
LETTER	232.354
NEWS ITEM	127.357
CORRECTION	61.917
POETRY	31.037
BIOGRAPHICAL-ITEM	30.376
ART EXHIBIT REVIEW	14.230
FILM REVIEW	9.707
RECORD REVIEW	8.676
REVIEW; BOOK CHAPTER	7.614
FICTION, CREATIVE PROSE	4.440
ARTICLE; BOOK CHAPTER	4.158
THEATER REVIEW	2.995
DANCE PERFORMANCE REVIEW	2.886
REPRINT	2.756
BIBLIOGRAPHY	1.889
MUSIC SCORE REVIEW	1.887
TV REVIEW, RADIO REVIEW	1.788
MUSIC PERFORMANCE REVIEW	1.737
SOFTWARE REVIEW	1.386
EXCERPT	586
EDITORIAL MATERIAL; BOOK CHAPT	522
SCRIPT	241
DATABASE REVIEW	183
MUSIC SCORE	92
HARDWARE REVIEW	63
BIOGRAPHICAL-ITEM; BOOK CHAPTI	16
MEETING SUMMARY	8
CORRECTION; BOOK CHAPTER	7
LETTER; BOOK CHAPTER	5
MAIN CITE	5
CHRONOLOGY	5
REPRINT; BOOK CHAPTER	4
REVIEW; BOOK	3
ABSTRACT OF PUBLISHED ITEM	3
MAIN CITE; BOOK CHAPTER	2
ARTICLE; BOOK	2
BOOK	1

Source: WebOfScience, authors' elaboration, Database A

Appendix TABLE A2

NACE	NACE Divisions Names
01 to 03	Agriculture, forestry and fishing
05 to 09	Mining and quarrying
10 to 12	Manufacture of food products, beverages and tobacco products
13 to 15	Manufacture of textiles, apparel, leather and related products
16 to 18	Manufacture of wood and paper products, and printing
19	Manufacture of coke, and refined petroleum products
20	Manufacture of chemicals and chemical products
21	Manufacture of pharmaceuticals, medicinal chemical and botanical products
22 + 23	Manuf. of rubber and plastics products, and other non-metallic mineral prod.
24 + 25	Manuf. of basic metals and fabricated metal products, exc. mach. and equip.
26	Manufacture of computer, electronic and optical products
27	Manufacture of electrical equipment
28	Manufacture of machinery and equipment n.e.c.
29 + 30	Manufacture of transport equipment
31 to 33	Other manufacturing, and repair and installation of machinery and equip.
35	Electricity, gas, steam and air-conditioning supply
36 to 39	Water supply, sewerage, waste management and remediation
41 to 43	Construction
45 to 47	Wholesale and retail trade, repair of motor vehicles and motorcycles
49 to 53	Transportation and storage
55 + 56	Accommodation and food service activities
58 to 60	Publishing, audiovisual and broadcasting activities
61	Telecommunications
62 + 63	IT and other information services
64 to 66	Financial and insurance activities
68	Real estate activities
69 to 71	Legal, accounting, manag., architec., engin., tech. testing and analysis activit.
72	Scientific research and development
73 to 75	Other professional, scientific and technical activities
77 to 82	Administrative and support service activities
84	Public administration and defence, compulsory social security
85	Education
86	Human health services
87 + 88	Residential care and social work activities
90 to 93	Arts, entertainment and recreation
94 to 96	Other services
97 + 98	Activities of households as employers; undif. goods- and services-producing activities of households for own use
99	Activities of extra-territorial organisations and bodies

Source: European Commission (2008)

Appendix Table A3

LIST OF ISI S&E FIELDS

1	Mathematics
2	Materials Science
3	Electronic Engineering
4	Nuclear Sciences
5	Mechanical, Civil and Other Engineering
6	Inorganic Chemistry and Engineering
7	Analytical Chemistry
8	Physical Chemistry
9	Organic Chemistry
10	Applied Physics
11	Solid State Physics
12	Geosciences
13	Other Physics
14	Ecology
15	Food Science and Agriculture
16	Biotechnology
17	Microbiology
18	General Biology
19	Pharmacology and Pharmacy
20	Public Health
21	Pathology
22	Neuroscience
23	Reproduction Medicine and Geriatrics
24	General Medicine
25	Internal Medicine
26	Research Medicine
27	Immunology

Source: Braun et al (1995a, 1995b)

APPENDIX MATRIX 1
NACE SECTORS X 27 S&E FIELDS (2015, all papers)

		Agriculture, forestry and fishing	Mining and quarrying	Manufacture of food products, beverages and tobacco products	Manufacture of textiles, apparel, leather and related products	Manufacture of wood and paper products, and printing	Manufacture of coke, and refined petroleum products	Manufacture of chemicals and chemical products	Manufacture of pharmaceuticals, medicinal chemical and botanical products	Manufacture of rubber and plastics products, and other non-metallic mineral products	Manufacture of basic metals and fabricated metal products, except machinery and equipment	Manufacture of computer, electronic and optical products	Manufacture of electrical equipment	Manufacture of machinery and equipment n.e.c.	Manufacture of transport equipment	Other manufacturing, and repair and installation of machinery and equipment	Telecommunications	IT and other information services
		01 to 03	05 to 09	10 to 12	13 to 15	16 to 18	19	20	21	22 + 23	24 + 25	26	27	28	29 + 30	31 to 33	61	62 + 63
1	Mathematics	0	0	0	0	0	0	11	420	4	0	85	1	16	143	96	3	11
2	Materials Science	0	0	2	2	1	2	28	893	11	101	155	15	48	97	0	5	9
3	Electronic Engineering	0	0	2	2	1	8	70	2872	14	39	1160	143	170	746	6	207	245
4	Nuclear Sciences	0	0	0	0	0	22	18	428	10	6	20	22	45	69	14	18	3
5	Mechanical, Civil and Other Engineering	0	0	0	0	6	24	72	1201	13	36	136	28	236	421	1	28	21
6	Inorganic Chemistry and Engineering	0	0	16	0	2	26	433	4745	19	15	180	10	31	291	16	12	22
7	Analytical Chemistry	0	0	2	0	0	0	22	462	1	2	52	1	1	30	1	0	0
8	Physical Chemistry	0	0	0	0	0	0	6	213	9	1	50	7	5	45	18	0	7
9	Organic Chemistry	0	0	0	4	0	3	38	320	0	0	17	1	3	18	0	0	0
10	Applied Physics	0	0	1	0	0	3	19	669	9	9	125	33	29	62	1	1	17
11	Solid State Physics	0	0	0	0	0	0	18	797	22	6	80	6	16	66	1	1	8
12	Geosciences	0	0	0	0	0	12	24	1255	3	0	29	1	41	83	1	11	1
13	Other Physics	0	0	0	0	3	3	95	2544	7	7	195	28	126	267	0	9	33
14	Ecology	0	0	1	0	2	1	30	1200	19	0	21	1	12	113	6	4	5
15	Food Science and Agriculture	0	0	73	1	0	1	96	555	3	0	24	4	7	37	0	0	0
16	Biotechnology	0	0	59	1	0	4	153	2071	3	3	120	4	1	163	1	2	3
17	Microbiology	0	0	3	0	0	0	16	270	1	0	6	0	0	3	20	0	0
18	General Biology	0	0	9	0	1	0	73	1277	19	1	46	2	4	78	0	2	4
19	Pharmacology and Pharmacy	0	0	8	0	0	0	38	693	11	2	11	1	1	23	11	0	0
20	Public Health	0	0	3	0	0	0	74	764	11	0	75	1	7	114	4	5	3
21	Pathology	0	0	0	0	0	0	10	148	0	0	22	2	1	11	45	0	0
22	Neuroscience	0	0	2	0	0	0	66	553	0	0	28	1	2	63	2	0	0
23	Reproduction Medicine and Geriatrics	0	0	0	0	0	0	9	128	3	0	9	0	0	24	5	0	0
24	General Medicine	0	0	1	0	0	0	2	45	0	0	1	0	0	2	64	0	0
25	Internal Medicine	0	0	6	0	0	0	44	764	1	0	52	0	2	34	1	1	1
26	Research Medicine	0	0	1	0	0	0	32	977	1	0	46	0	4	38	36	1	1
27	Immunology	0	0	4	0	0	2	39	693	0	1	19	1	1	29	28	0	0

Source: WebOfScience, ORBIS, authors' elaboration (Database B)

APPENDIX MATRIX 2

NACE SECTORS X 27 S&E FIELDS (2015, papers with international co-authorship)

		Agriculture, forestry and fishing	Mining and quarrying	Manufacture of food products, beverages and tobacco products	Manufacture of textiles, apparel, leather and related products	Manufacture of wood and paper products, and printing	Manufacture of coke, and refined petroleum products	Manufacture of chemicals and chemical products	Manufacture of pharmaceuticals, medicinal chemical and botanical products	Manufacture of rubber and plastics products, and other non-metallic mineral products	Manufacture of basic metals and fabricated metal products, except machinery and equipment	Manufacture of computer, electronic and optical products	Manufacture of electrical equipment	Manufacture of machinery and equipment n.e.c.	Manufacture of transport equipment	Other manufacturing, and repair and installation of machinery and equipment	Telecommunications	IT and other information services
		01 to 03	05 to 09	10 to 12	13 to 15	16 to 18	19	20	21	22 + 23	24 + 25	26	27	28	29 + 30	31 to 33	61	62 + 63
Mathematics	1	0	0	0	0	0	0	7	95	2	0	25	0	4	40	96	1	4
Materials Science	2	0	0	0	0	0	0	3	93	0	7	9	2	6	15	0	0	1
Electronic Engineering	3	0	0	0	0	0	0	15	261	1	2	75	10	20	101	3	23	24
Nuclear Sciences	4	0	0	0	0	0	2	4	47	0	1	1	1	1	11	2	2	0
Mechanical, Civil and Other Engineering	5	0	0	0	0	1	1	14	166	1	2	15	2	8	46	1	2	4
Inorganic Chemistry and Engineering	6	0	0	3	0	0	2	75	632	0	3	42	0	3	70	1	1	4
Analytical Chemistry	7	0	0	0	0	0	0	3	55	0	0	6	0	1	7	0	0	0
Physical Chemistry	8	0	0	0	0	0	0	0	29	1	0	2	0	0	9	1	0	1
Organic Chemistry	9	0	0	0	0	0	0	12	35	0	0	0	0	0	6	0	0	0
Applied Physics	10	0	0	0	0	0	0	4	102	1	0	15	1	2	17	0	0	4
Solid State Physics	11	0	0	0	0	0	0	6	116	8	1	18	0	5	19	1	0	3
Geosciences	12	0	0	0	0	0	1	13	392	3	0	15	0	8	27	0	4	0
Other Physics	13	0	0	0	0	0	0	43	469	0	0	28	2	20	87	0	2	7
Ecology	14	0	0	0	0	0	0	17	291	7	0	4	0	5	32	1	1	0
Food Science and Agriculture	15	0	0	5	0	0	0	32	131	0	0	3	0	1	9	0	0	0
Biotechnology	16	0	0	4	0	0	0	39	387	1	0	21	2	1	45	0	0	1
Microbiology	17	0	0	1	0	0	0	7	66	0	0	0	0	0	2	2	0	0
General Biology	18	0	0	1	0	0	0	30	347	7	1	10	0	1	19	0	1	0
Pharmacology and Pharmacy	19	0	0	0	0	0	0	10	147	5	0	0	0	0	4	1	0	0
Public Health	20	0	0	0	0	0	0	21	138	3	0	6	0	1	13	0	0	1
Pathology	21	0	0	0	0	0	0	3	41	0	0	1	0	1	1	6	0	0
Neuroscience	22	0	0	0	0	0	0	26	153	0	0	6	0	1	11	0	0	0
Reproduction Medicine and Geriatrics	23	0	0	0	0	0	0	4	27	2	0	1	0	0	4	0	0	0
General Medicine	24	0	0	0	0	0	0	2	5	0	0	0	0	0	0	2	0	0
Internal Medicine	25	0	0	0	0	0	0	17	182	1	0	5	0	0	11	0	0	0
Research Medicine	26	0	0	0	0	0	0	15	213	0	0	4	0	3	8	3	0	0
Immunology	27	0	0	0	0	0	2	18	214	0	0	1	0	0	7	4	0	0

Source: WebOfScience, ORBIS, authors' elaboration (Database B)