SCIENTIFIC NETWORKS AND COLLABORATIONS WITH INDUSTRY IN LATIN AMERICA

Hugo Confraria[a]
Fernando Vargas[b]

May 2015

Abstract

In this paper, we use a combination of bibliometric, social network and econometric techniques to increase understanding on how research institutions are interacting with the private sector in Latin America (LAC). We first study recent trends in the scientific outcome and specialization. On average, LAC countries have been reducing the gap with the world leading regions. They also have tended to specialize in fields related to economic activities based on natural resources, such as Agricultural and Plant & Animal Sciences. However, collaborations with the private sector remain scarce. We build scientific networks composed by what we define as Research Departments (RD), which can belong to universities, research institutes or government agencies, in different scientific fields, at the country and LAC level. We model the intensity of collaboration of a RD with the industry as a function of its characteristics and its position in the LAC and local scientific network. Our results show that collaborations with industry are influenced by the intensity of previous interactions with the private sector. Additionally, RDs that have a higher degree and betweenness values in their local scientific network are more likely to show higher rates of collaboration with industry. Centrality values at the LAC level do not seem to play a significant role in this regard.

JEL Classification Code: O30, O39, O54

Keywords: Bibliometrics, Science, Networks, Technology transfer, Industry-university interaction

---

[a] UNU-MERIT, Keizer Karelplein 19 6211TC Maastricht, The Netherlands. Email: confraria@merit.unu.edu
[b] UNU-MERIT, Keizer Karelplein 19 6211TC Maastricht, The Netherlands. Email: vargas@merit.unu.edu

1. Introduction

Latin American economic growth of recent years has been remarkably based on the exploitation of natural resources, but besides this trend there has been no sustained productivity growth in other sectors of the economy (IADB, 2010a). This stagnation harms the abilities of the country to successfully catch-up with industrialized economies. It has been argue that low productivity is a consequence of lack of innovative capacities of incumbents and the absence of new economic activities more based on scientific knowledge, both related to the capacity to transform new knowledge into technological innovations (Katz, 2006).

In Latin America, investing in knowledge is an activity with high returns. According to Lederman & Maloney (2003), social returns of R&D investments are between 60% in middle-income countries to nearly 100% in low-income countries. Despite this evidence, LAC countries still lag significantly behind industrialized countries (IADB, 2010b). Furthermore, the literature on innovation systems illustrates that it is not only a matter of knowledge production, but also on the linkages and relationship between scientific institutions and the private sector (Lundvall, 1992).

Evidence from the USA in Agrawal & Henderson (2002), shows that scientific institutions are often an important source of information for innovation, but a preferred technology transfer channel is far to be determined. In particular, authors point out the relevance of joint research projects. In LAC, evidence from innovation surveys show that universities and research centers tends to be less relevant partners for technological innovation (IADB, 2010a). On the other hand, LAC countries have been doing efforts to promote universities-industry linkages, but the size of the interventions does not correspond to the dimension of the problem faced (Cimoli, 2010). However, positives effect of linkages between industry and research institutions has been found. Crespi (2012) finds that public policies that promote collaboration activities for innovation between industry and universities regularly show positive impacts.

In this paper, we are going to explore the industry-science collaborations, in terms of co-authored scientific publications, from the supply of knowledge side. Our main goal is to increase understanding of the characteristics of the research institutions that work more closely with the private sector. Our approach will consider quantity and quality of scientific publications, but also a network analysis view that will allow estimating the relevance of cooperation patterns in this regard.

In what follows, we will first do a revision of relevant literature; then section three describes data sources that we use. Section four shows the evolution of scientific production process and major trends on the specialization of LAC countries. Section five presents the conceptual framework for the estimation of the relations between science and industry and the role of scientific networks. Section six describes LAC scientific networks in five selected disciplines: Agriculture, Engineering, Environmental, Geosciences and Plant and Animal Science. Section seven presents the estimation strategy of the econometric model and main descriptive statistics. Section 8 present and analyze main results and finally, conclusions will be put forward.

2. Literature Review

The role of science and knowledge creation is becoming even more important for economic development in the twenty-first century than ever before. Endogenous economic growth models remark that investments and accumulation of knowledge are fundamental for economic prosperity (Romer, 1990). Thus, the creation of science and technology (S&T) knowledge that allows for improvements in how physical and human capital, labor, and natural resources are combined translates in high long-term economic growth rates. Griliches (1979) shows that, when considering externalities, differences in the amount of investments in Research & Development (R&D) accounts for nearly 75% of divergences in economic growth rates. Furthermore, Rouvinen (2002) finds that it is not the case that only rich and more productive countries show higher R&D investments levels because they can afford it. On the contrary, it is knowledge production what causes productivity growth. On top of that, Griffith et al. (2004) shows that returns to investments in R&D tend to grow in line with the distance to the technological frontier.

In Latin America, not only high impacts on productivity are observed as a consequence of investments in R&D. Social returns are also considerable and sometimes can be even higher than private returns (Lederman & Maloney, 2003). This is valid for both, medium income countries like Argentina, Chile or Mexico where social returns are close to 60%, to low income countries, like Nicaragua where estimated returns are nearly 100% (Lederman & Maloney, 2003). Despite this evidence, LAC countries still lag significantly behind in terms of R&D investments. Brazil is the only country of the continent that spends more than 1% of its gross domestic product (GDP) on R&D. Even so, its investment are just nearly two-thirds of the average of OECD countries and sits far below R&D expenditures of countries with accelerated economics growth rate, like China or Israel (Catanzaro et al., 2014).

However, the mechanisms through investments in R&D and S&T impacts economic growth are not that straightforward. Since the introduction of the National Innovation Systems framework (Lundvall, 1992), the proposition that innovation flows directly from S&T activities to the economy has lost importance. Instead, the predominant view is that innovation is the result of a process in which different types of actors and institutions interact, contributing with resources, information or knowledge, thus fostering ideas with potential for the innovative process.

In this framework, universities and research institutions play a pivotal role. Not only by providing highly skilled labor supply that could increase the absorptive capacity of the economy (Goode, 1959; Cohen & Levinthal, 1990), but also as a source of knowledge and information on which other entities can rely to build market-valued goods and services. These institutions produce and diffuse codified knowledge, such as scientific publications, licenses or patents that are used as a source of information for innovation activities by the private sector. Nevertheless, there are other channels through which also tacit knowledge, key component of successful technology transfer, can be transmitted, such as meetings and conferences, informal exchanges of information, joint ventures, consulting and research projects (Cohen et al., 2000). Although these authors found that across productive sectors, universities and firms use a variety of channels to establish relationships depending on current objectives, some main tendencies can be identified: industries with a strong investment in R&D tend to go for collaborative research, whereas service firms lean towards personal mobility and training. These results are in line with findings from Agrawal & Henderson (2002), where it shows that consultancies and collaborative research are perceived as far more important for knowledge transfer, than patents and licenses in USA.

There is some evidence that these positive linkages are also present in Latin American countries. Particularly, Marotta et al. (2007) find that linkages with universities increase the likeliness to innovate by 29% in Chilean firms and 44% in the case of Colombian firms. Furthermore, La Paz et al. (2012) shows that start-ups that partner with universities have better performance indicators than those working with consultancies firms. Despite this evidence, cooperation with universities happens far less frequently than with other agents of the national innovation systems (Anllo & Suarez, 2009). Understanding the determinants of the existence and intensity of these mechanisms of technology transfer and collaborations remains highly relevant for policy design aimed to foster industry-university linkages.

In this regard, studying the knowledge production process of scientific systems facilitates the understanding of a significant share of the stock of knowledge in an economy. This can be addressed focusing on the understanding of knowledge network structures. It is commonly argued (e.g. Shilling & Phelps, 2005) that a dense network enables richer and greater amounts of knowledge to be exchanged and integrated more readily. On the other hand, the diversity of knowledge distributed across different clusters in a network provides the requisite variety for recombination. This apparent tension between knowledge diffusion and knowledge creation is solved in network structures which are at the same time highly clustered (easing knowledge diffusion) and with short paths between clusters (fostering knowledge creation)[c].

Besides the structure of the networks, characteristics of the agents (nodes) of the system and its position within the network also play a role in determining performance and evolution of a system. For example, Gulati and Gargiulo (1999) found that organizations tend to form new alliances with partners that have complementary capabilities, and that the information from previous alliances plays a significant role in the decision of choosing a new partner. Even more, previous alliances tend to remain or to be repeated because routines decrease asymmetries of information among partners and facilitate the estimation of future returns of joint activities (Gulati, 1995). Still in the management literature we can find also that organizational units can produce more innovations and enjoy better performance if they occupy central network positions that provide access to new knowledge developed by other units (Tsai, 2001).

Taking into consideration the specificities of knowledge creation and diffusion, and network structures, in this study we will use a heterogeneous set of methodologies and metrics to analyse the Latin American scientific system, its interactions and the proximity with the industry. Our contribution is, firstly, to update trends and specializations patterns of scientific knowledge in LAC

[c] Systems that show these types of network characteristics are known as "small-worlds" (Watts & Strogatz, 1998).

countries using bibliometric analysis and descriptive statistics, and secondly to assess to what extent scientific networks structures affect collaborations between science and the industry.

3. Data

We will use the *InCites* tool proposed by *Thomson Reuters*, which is a web-based research evaluation tool that facilitates national and institutional comparisons across long time periods using publication output, productivity, specialization and normalized citation impact values. *InCites* provides output and citation metrics from the WoS (Web of Science, *Thomson Reuters*), which in turn will access data and metrics from a dataset (SCI-EXPANDED, SSCI and A&HCI) of 22 million WoS papers from 1981 to 2013. The metrics for comparisons are created based on address criteria, using the whole-counting method, i.e. counts are not weighted by number of authors or addresses.

It is important to notice that Latin America's research output may be underestimated because its researchers often publish in journals that are not indexed in major citation databases, such as WoS or Elsevier's Scopus. However, SciELO Citation Index, a collection of roughly 750 journals from research in Latin America, Spain, Portugal, the Caribbean and South Africa, has become available from the WoS platform as of January 2014. This initiative allows citations to SciELO articles in the WoS database to be counted, thereby increasing their visibility and impact.

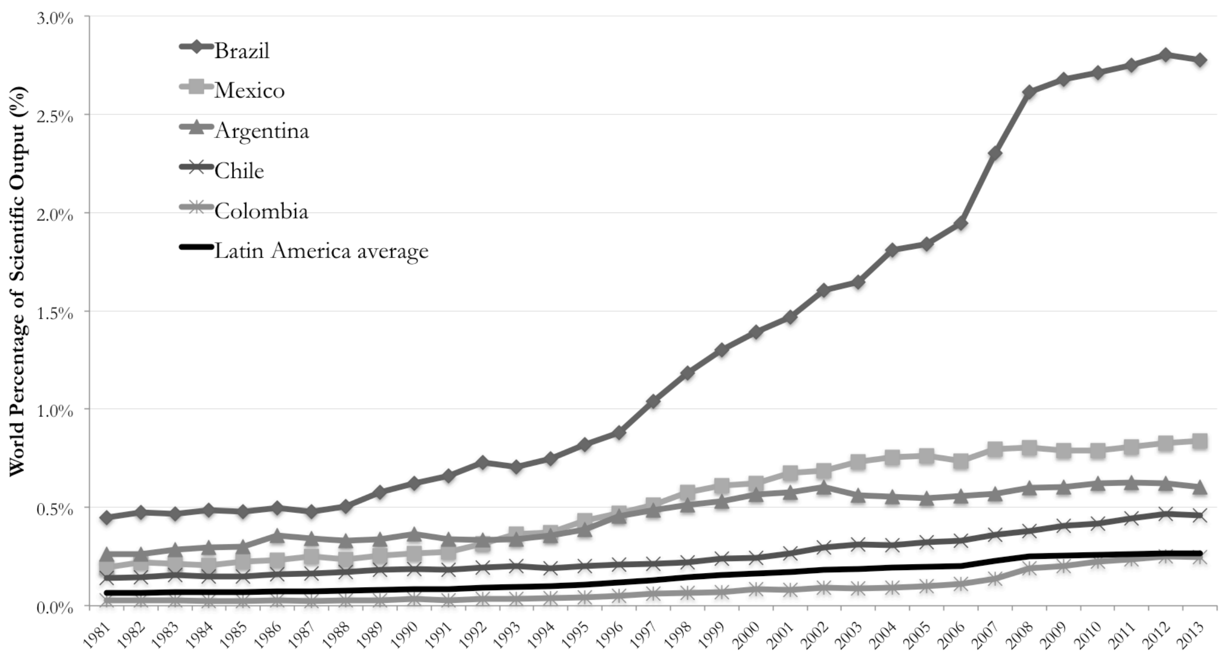
In our study, we will use bibliometric indicators of scientific impact or quality. These are normalized citation impact measures, which evaluate the scientific influence or visibility of a set of publications in a specific period of time. For the *Quality Citation Index* a value of 1.2 for a specific country, for example, indicates that the citation impact of papers published by scientists in this country has, on average, 20 percentage points above the worldwide average. For the *Quality Top 10% Index* a value of "10" for a set of documents represents that ten percent of the publications in that set are in the top ten percent of the world regardless of subject, year and document type and would therefore be considered to be performing at the same level as world average. A higher value is considered to be higher performance.

An important caveat in our analysis is related to the definition of organization type in *InCites*. An industry collaborative publication is one that lists its organization type as "corporate" for one or more of the co-author's affiliations. However, not all single affiliations of all publications in *InCites* are unified as "university", "research institute", "corporate", etc. There are corporate affiliations that have not been unified yet not having an organization type assigned and, therefore, are not identified as industrial collaborations. Large multinational corporations (MNE) have a higher probability of being identified and unified. Therefore, publications listed as industry collaborations are a lower boundary of the real co-publications activities. We would expect that countries with lower presence of MNEs have larger differences between the number of publications authored by the industry captured by *InCites* and the real activity.

4. Science in LAC: Trends and Specialization

Latin America's long-term world percentage of publication output has increased from 1.32% in 1981 to 5.03% in 2013. In 2013, all Latin America countries accounted for 71391 publications in WoS. In Fig. 1 we can observe the evolution of publication output in world percentage of a group of LAC.

Fig. 1 Trends in publication output of top5 countries (shares as a world percentage) ordered by publication output in 2013 (1981-2013)



Source: Own calculations based on InCites ™

Brazil´s share is particularly high when compared with other countries of the region, replicating differences in the size of the economies. The share of world scientific output rate increased at a constant rate from 1993 to 2006, where it exploded to the levels that Brazil showed in 2013. Vargas et al. (2014) argue that, in areas such as Agricultural Sciences this increase since 2006 was due to the expansion of Brazilian journals in WoS and an increase in the number of issues published by these journals. Other countries that show higher than LAC average shares of scientific output in the last decade are Mexico, Argentina, and Chile.

Table 1 provides data adjusting scientific output by other characteristics of the countries, thus allowing for an assessment of the scientific "productivity", per billion USD and per million inhabitants.

Table 1 Research performance of LAC: Summary statistics (2004-2013)

Country	Total publications output			Productivity				Quality Citation Index	
				Docs/GDP [a]		Docs/POP [b]			
	04-08	09-13	Variation	04-08	09-13	04-08	09-13	04-08	09-13
BRAZIL	106692	178313	67%	23	32	113	181	0.61	0.65
MEXICO	38729	52715	36%	9	11	69	88	0.70	0.79
ARGENTINA	28440	39992	41%	24	26	146	196	0.77	0.93
CHILE	17185	28632	67%	26	37	208	330	0.90	0.92
COLOMBIA	6436	15169	136%	8	16	29	64	0.67	0.80
VENEZUELA	5948	5801	-2%	8	6	44	39	0.62	0.71
CUBA	3630	4093	13%	16	14	64	73	0.59	0.72
URUGUAY	2388	3553	49%	26	29	143	210	0.83	0.95
PERU	2100	3562	70%	5	6	15	24	0.94	1.32
COSTA RICA	1679	2160	29%	16	17	76	91	0.99	1.15
ECUADOR	1136	1875	65%	5	7	16	25	1.08	1.17
PANAMA	1007	1576	57%	12	12	59	84	1.48	1.48
BOLIVIA	768	1028	34%	15	16	16	20	0.81	1.17
GUATEMALA	326	583	79%	2	3	5	8	0.81	0.97
NICARAGUA	242	342	41%	7	9	9	12	0.78	0.99
PARAGUAY	175	307	75%	4	5	6	9	1.00	1.07
DOMINICAN REPUBLIC	145	258	78%	1	1	3	5	1.00	0.80
HONDURAS	143	254	78%	3	4	4	7	1.23	0.96
HAITI	94	208	121%	4	9	2	4	0.94	1.32
EL SALVADOR	102	196	92%	1	2	3	6	0.57	1.11
LATIN AMERICA	209756	325069	55%	14	18	74	108	0.66	0.70
WORLD	5017977	6493637	29%	21	24	153	186	1.00	1.00

Source: Own calculations based on InCites ™ & World Bank. [a] Docs/GDP - Publications per billion constant 2005 US$ of GDP (yearly average). [b] Docs/Population - Publications per million inhabitants (yearly average).

LAC countries are ranked in Table 2 by aggregate scientific production from 2004-2013. Although Brazil has the highest number of publications, it also has the lower scientific impact. This may happen due to a significant percentage of articles are being published in national journals that had recently been included in the databases. In general Latin America's scientific impact, although growing, remains relatively low when compared to the world average. Despite their low productivity and scientific output, Peru's and Panama's do best in these terms but mainly because more than 90% of their publications are co-authored with researchers outside their country (Van Raan, 1998). Chile, thanks to its research into the genetics of food cultivation and also successful international collaboration through its astronomical observatories, has increased its research output and maintained a medium level of scientific impact.

We also provide in appendix (Figs. 2 & 3) information about long-term trends in scientific productivity. It is possible to observe a slightly convergence, both in terms of publications per GDP and publications per population, from Latin America towards the world average. However, a significant gap still remains. The only countries that have a relative good performance since late 90s are Chile, Uruguay, Argentina and Brazil.

Countries often try to invest strategically in research areas important to their economic development. Creation of applied specific local knowledge may increase innovation capacities of incumbents, but also promote the birth of start-ups or spin-offs that could operate as suppliers of established firms. These trends run in parallel with others that not necessarily operates in the same direction. Historical and cultural influences, strengths of scientific establishments, as well as incentives and government funding for scientific research plays a relevant role in defining the revealed scientific specialization of a country. Also important is the scientific system size, as larger science systems have the capacity for more diversity and more coverage of the full scope of sciences while smaller systems may be limited in their ability to invest in specific domains.

The specialization analysis that is next provided is based on the 22 Essential Science Indicators (ESI) areas[d]. Table 2 contains the five subject areas of higher specialization for the 9 countries in Latin America with more than 1% of Latin American total scientific output over the 2009-2013 period. Table 2 also provides information on aggregate specialization level (given by the SII index) for each of these 9 countries

Research specialization is quite similar across these LAC countries. In aggregate terms, the top5 areas with the largest output from Latin America, relative to the world are: Agricultural sciences (15.7%), Plant & Animal science (12.3%), Space science (9.3%), Environment/Ecology (7.7%) and Microbiology (7.3%). The higher Latin American specializations are in Agricultural sciences and Plant & Animal science, which is in line with the high importance of agricultural and livestock activities in the region, and the outstanding productivity increases in these areas in the last decade.

Chile, Peru and Uruguay cases are interesting because they revealed more specialized[e] in specific subject areas than others countries of the sample. The specialization of Peru is related to issues in public health (prevention of HIV, tuberculosis and lupus) that they also have a high scientific impact (Van Noorden, 2014). Uruguay, on the other hand, has more scientific publications in fields associated to its main economic products: bovine meat, rice and soya beans[f]. Chile's high specialization in Space Science is certainly related to its extraordinary infrastructure of giants telescopes housed in the Atacama Desert. According to Catanzaro et al. (2014), funding for astrophysics has grown from $2 million in 2006 to $6.8 million in 2010. Over the same period, the number of faculty positions has almost doubled. This has led not only to an increase in the number of publications in this field but also to an increase in quality. The country has also found scientific success working on food crops, such as highly cited collaboration on the genome of the potato (Van

---

[d] The Essential Science Indicators schema (*Thomson Reuters*) comprises 22 subject areas in science and social sciences and is based on journal assignments. Arts & Humanities journals are not included. Each indexed journal (11,000+) is found in only one of the 22 subject areas and there is no overlap between categories.

[e] If a country has a scientific output structure exactly similar to the World, the value of the indicator will be zero. The size of SII is an indication of how strongly each country is specialized.

[f] http://atlas.cid.harvard.edu/country/ury/

Noorden, 2014). These specializations can reveal themselves as potential sources for knowledge-based economic growth in the near future (See Box 1). On the other hand, Economics & Business, Materials science, Computer science, Psychiatry/Psychology and at a certain level Engineering seem to be neglected disciplines across countries.

Table. 2 Top 5 subject areas, in the 9 LAC with higher scientific output (2009-2013)

Country	Relative Specialization Intensity (Rank)										SII [c]
	1		2		3		4		5		
	RSI [b]	Docs [a]	RSI [b]	Docs [a]	RSI [b]	Docs [a]	RSI [b]	Docs [a]	RSI [b]	Docs [a]	
BRAZIL	Agricultural sciences		Plant & Animal science		Microbiology		Pharmacology & Toxicology		Environment & Ecology		0.43
	4.00	20033	2.42	21557	1.36	3428	1.18	5557	1.17	6082	
MEXICO	Space science		Plant & Animal science		Environment & Ecology		Agricultural sciences		Microbiology		0.28
	2.51	1366	2.50	6600	2.12	3258	1.92	2851	1.47	1095	
ARGENTINA	Plant & Animal science		Microbiology		Space science		Agricultural sciences		Geosciences		0.39
	2.89	5791	2.25	1273	2.09	861	2.06	2318	1.90	2176	
CHILE	Space science		Plant & Animal science		Agricultural sciences		Environment & Ecology		Mathematics		0.95
	9.70	2863	1.86	2669	1.82	1467	1.71	1430	1.45	1206	
COLOMBIA	Agricultural sciences		Plant & Animal science		Engineering		Immunology		Physics		0.21
	2.23	949	1.88	1425	1.73	2048	1.48	390	1.35	1706	
VENEZUELA	Agricultural sciences		Immunology		Plant & Animal science		Environment & Ecology		Economics & Business		0.35
	2.44	397	2.41	244	2.41	700	2.24	379	1.85	192	
CUBA	Immunology		Pharmacology & Toxicology		Agricultural sciences		Plant & Animal science		Microbiology		0.34
	2.58	184	2.43	263	2.24	258	2.13	437	1.59	92	
URUGUAY	Plant & Animal science		Agricultural sciences		Microbiology		Environment & Ecology		Immunology		0.64
	3.37	600	3.32	331	2.71	136	1.77	183	1.54	95	
PERU	Immunology		Plant & Animal science		Microbiology		Environment & Ecology		Social sciences		0.96
	5.64	350	2.52	450	2.33	117	2.30	239	2.21	463	

Source: Own calculations; InCites™

[a] Docs = Scientific publications.

[b] RSI = Share of a country's papers in a given field, relative to the share of world papers in that field.

[c] SII = Specialization Intensity Index. This measure provides a ratio to assess whether a country is "specialized" or "not specialized". It grows with the specialization intensity of a country.

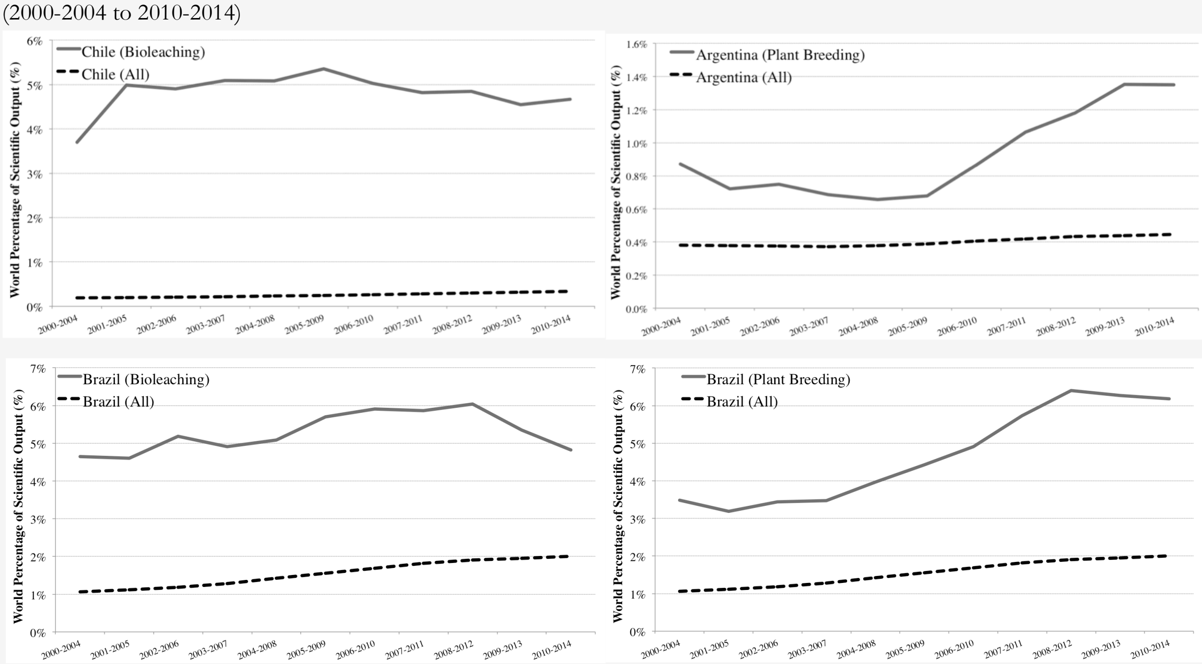
Box 1 Plant Breeding and Bioleaching as opportunities for knowledge based growth

Natural resource based industries are commonly categorized as having low technological dynamism. However, it has been argued that there are certain specificities in the current context that are different from those which prompted ideas such as the "curse of Natural Resources" and that open a temporary window of opportunity for resource-based growth (Pérez, 2010). For example, Marin and Stubrin (2015) show that new seed improvement techniques, based on genetic, agricultural, biotechnology and/or transgenic knowledge has played a critical role on the evolution of the local seed industry in Argentina. We can observe that the development of these techniques has been correlated with the increasing specialization of the scientific system of the country in these technologies. This trend is also observed in Brazil, another LAC country with a large seed industry.

The same correlations between increasing use of a technology and the production of local scientific knowledge are observed in the case of bioleaching (the extraction of metals from their ores through the use of bacteria) in Chile and Brazil, which are countries with large mining industries.

In Fig. 4 we show that in these two specific technologies, Brazil, Argentina and particularly Chile show a high level of scientific specialization. Although the accumulation of this type of knowledge is promising for promoting knowledge-intensive growth in natural resources exploitation, it is a matter of further research to understand if it is the production of scientific knowledge that is fostering the application in the industry or the other way around.

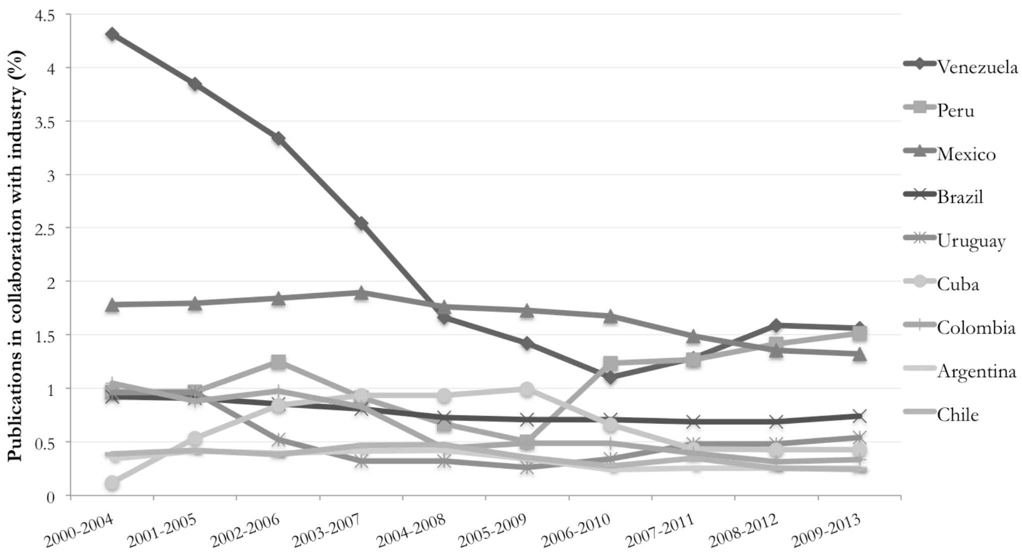
Fig. 4 Evolution of the world percentage of publications in "Bioleaching" and "Plant Breeding" of Brazil, Chile and Argentina (2000-2004 to 2010-2014)



In what follows, we will focus on the study of five main scientific fields: Agricultural, Engineering, Environmental, Geosciences, and Plant and Animals sciences. We choose these fields in part because in the case of Agricultural, Geosciences and Plant and Animals, are closely related to natural resources-based economic activities, in which Latin American countries are more intensive. We also include engineering and environmental sciences, because we expect that this type of knowledge is consistently applied across different economic activities.

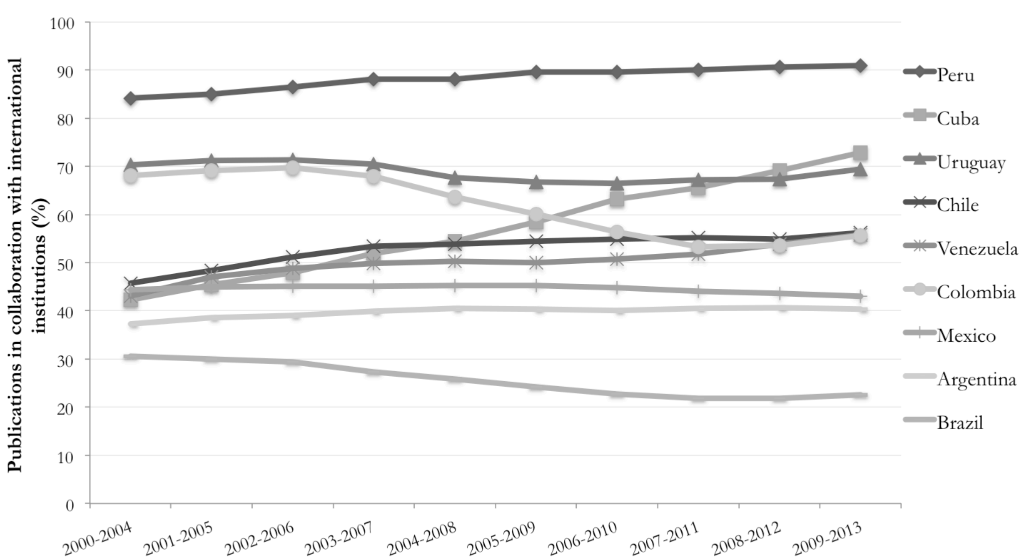
While looking at the aggregation of these five topics, we can appreciate a relatively low percentage of collaboration with industries, comparing with countries like USA or Germany (>2%). This situation is clear in Fig. 5. Besides Venezuela, Peru and Mexico all the other LAC countries have an average industry collaboration percentage below 1%. Regarding the percentage of international collaborations, as expected, that number is negatively correlated with the size of the local scientific output. Fig. 6 shows that countries as Peru, Cuba and Uruguay have the highest levels of international co-authorship mostly due to the activity of one local institution engaging in research with many international partners.

Fig. 5 Evolution of the percentage of collaborations between research institutions in LAC and industry from 2004-2008 to 2009-2013



Source: InCites™

Fig. 6 Evolution of the percentage of collaborations between research institutions in LAC and international institutions from 2004-2008 to 2009-2013



Source: InCites™

In summary, scientific activity has been growing in LAC countries during the last decade but still is not enough to catch-up with the rest of the world. Only 5 or 6 countries show productivity and quality levels near to the world averages. Research specialization seems influenced by economic specialization[g]. In our selected scientific fields, co-publications with international institutions are frequent and highly relevant for the smaller scientific systems. On the other hand, collaborations with the industry are scarce.

5. Conceptual framework

The aim of the following analysis is to gather new evidence on research organizations characteristics that facilitate technology transfer, specifically in the form of co-publication with the industry. We focus our analysis on what we are going to call from now the Research Department level (RD)[h]. This unit of analysis is defined by an output measure, i.e. we assume that all publications from an institution in a specific scientific field were done by a particular RD. Therefore, we treat an institution publishing in, for example, two scientific fields as two different RDs (but we keep track that both are part of the same institution). Besides this, we assume that research performed in a particular field face its own conditions, and it is embedded in a particular scientific network, separately from other scientific topics. We are aware that this is debatable approach, but RDs tend to be highly specialized because each scientific field demands highly specific knowledge, making very costly to get involved in research in other disciplines (Jeffrey, 2003). Thus we expect that this definition may include some errors, but not a consistent bias.

We model the intensity of research collaborations between companies and RDs as a direct function of characteristics of the latter and its position in the relevant scientific network. The underlying assumption is that firms are actively searching for partners, and those that conduct scientific and R&D activities have the capacity to screen RDs in order to choose the best source of knowledge or partner for collaboration. Thus, the quantity and quality of the research performed by the RDs are one of the signals that would impact on the level of involvement in research projects with the industry. In addition, RDs could also be access points to other sources of information and knowledge through their work and ties with third parties. Under this perspective, RDs that are better connected to other institutions would be preferred to work with by the industry because they could lower the cost of screening other RDs and also decrease the risk of knowledge lock-in (Menzel & Fornahl, 2010). On the other hand, working with more central RDs increases the potential damages of eventual leakages of knowledge relevant to the firm. The latter characteristics are studied through social network analysis, obtaining the centrality features of each RD. These families of indicators

---

[h] The number of publications of each institution is retrieved based on address criteria. RDs can belong to universities, research institutes or governmental agencies

[h] The number of publications of each institution is retrieved based on address criteria. RDs can belong to universities, research institutes or governmental agencies

assess the importance of the position of a node in a network. In this paper, we rely on three commonly used measures of centrality (Freeman, 1978):

i. Degree:

This measure of centrality accounts for the total number of links that a node have in a network. In the case of the networks that we are studying it will account for the total number of different research partners with whom each RD collaborates. RDs with higher degree number could be considered popular among their peers, enjoying benefits from reputation. Furthermore, they also hold what could be considered a more diversified set of research partners. However, regularly, maintaining links is a costly endeavor, and then we would expect to find limits on the utility of getting extra linkages. We use the normalized version of the indicator implemented by the *igraph* package of the R software. Formally:

$$C_D(i) = \frac{1}{(n-1)} \sum_{j=1}^{n} l(i,j) \quad (1)$$

Where $l(i,j) = \begin{cases} 1 & if\ there\ is\ an\ edge\ between\ i\ and\ j \\ 0 & otherwise \end{cases}$, and n is the number of nodes of the network

ii. Betweenness:

This index accounts for the total number of shortest paths[i] in which a node is involved. Under the assumption that shortest paths are preferred in diffusion of knowledge in a network, RT with higher betweenness values may be connecting knowledge from two very distant RD, broadening the scope of potential sources of information and allowing them to play a role of broker of knowledge. We use the normalized version of the indicator implemented by the *igraph* package of the R software. Formally:

$$C_B(i) = \frac{1}{(n(n-3)(n+2))} \sum_{j \neq k} \frac{g_{jk}(i)}{g_{jk}} \quad (2)$$

Where n is the number of nodes of the network, $g_{jk}(i)$ is the number of shortest paths that pass through node i, and g_{jk} is the total number of shortest paths.

iii. Closeness:

This index is defined by the inverse of the average shortest path to all other nodes in the network. An RD with higher values of closeness would require less effort to reach any other source of information. At the same time, at least theoretically, it could access new knowledge more quickly than others. We use the normalized version of the indicator implemented by the *igraph* package of the R software. Formally:

---

[i] The shortest path is the minimum distance, accounted by links, between two nodes of a network.

$$C_C(i) = (n-1)\left[\sum_{j=1}^{n} d(i,j)\right]^{-1} \quad (3)$$

Where n is the number of nodes of the network, $d(i,j)$ is the length of the shortest path between nodes i and j.

Our main variable of interest is the percentage of publications of RDs coauthored with the private sector. These activities are far to be common representing a case of corner outcomes, with a corner at zero and a continuous distribution for strictly positive values (upper-censored at 100). Wooldridge (2002) suggest addressing this cases implementing "hurdle" or "two-tiered" models, allowing explanatory variables to differently affect the participation decision, i.e. the co-authorship of at least one publication, and the intensity of those collaborations. In this paper, we firstly follow the specification of the two-tiered model developed by Cragg (1971). In the called "first-tier" of the model, the probability of participation in co-publication with the industry is estimated through a probit model. In the "second-tier" a truncated normal model is used to estimate the intensity of the collaboration with the industry, formally:

$$f(w,y|x_1x_2) = \{1-\Phi(x_1\gamma)\}^{1(w=0)}\left[\Phi(x_1\gamma)(2\pi)^{-\frac{1}{2}\sigma^{-1}} exp\left\{-\frac{(y-x_2\beta)^2}{2\sigma^2}\right\} / \, \Phi\left(\frac{x_2\beta}{\sigma}\right)\right]^{1(w=1)} \quad (4)$$

Where w is a dichotomous variable equals to 1 if a RD has at least one publication with the industry and 0 otherwise, and y is the percentage of publications coauthored with the private sector. When w is equal to 0, then y also takes the value of 0. While $w = 1$, then $y > 0$
x_1 and x_2 are a set of RD characteristics that affect the likeliness to co-publish with the industry and the intensity of these activities, respectively. Thus, γ captures the effects on the participation, and β those associated with the intensity of co-publication. This specification assumes conditional independence between the two tiers of the model. In this case, after controlling for observable characteristics and network features of RDs, there is no correlation between the decision to participate and the intensity of co-publications. We are aware that the latter assumption is debatable. For that reason, and for checking the consistency of the estimations, we also use the approach of Heckman (1979). Although this model is aimed to address the selectivity problem that arise when an interval of the outcome variable is not observable, statistically is very similar to Cragg's model and its flexibility allows for correlation between the participation and intensity equations. However, a variable that affect the participation but not the intensity needs to be included in order to identify the model. Regularly, the selection of this variable is also a matter of discussion.

6. Network Analysis

As it was mentioned above, we define that institution conducting research in more than one field have different RDs operating separately in each one of them. To build the networks we create a threshold to select the most prolific RDs in LAC. Those could be part of universities, governmental agencies or research institutes (private or public). For each one of the five fields studied, we selected RDs with more than 50 publications in each of the two five-year's periods analyzed. Afterward, to create the network of collaborations, for each "elite" RD we gather all partners with 5 or more collaborations in the same field, in the same period. Thus, two RDs are going to be linked if and only if they have 5 or more co-authorships in the field and period. It is worth mentioning that collaboration partners are not necessarily part of the "elite" RDs group, given that they only need to satisfy the minimum of 5 co-publications with one "elite" RD. We include non-LAC RD that collaborate on publications, but we drop of the network all of them that are linked with only one LAC RD. All the network structures can be seen in the appendix (Figs. 7). The thicknesses of the edges reflect the number of collaborations between institutions. Table 3 gives some network summary statistics from the five scientific fields that we are analyzing.

Table. 3 Summary statistics of scientific networks

Subject Area	Number nodes		Number Communities		Average Path Length	
	04-08	09-13	04-08	09-13	04-08	09-13
Agricultural sciences	96	138	11	25	3	2.86
Engineering	77	139	10	12	3.41	3.01
Environmental sciences	119	224	8	10	2.96	2.53
Geosciences	98	165	4	8	2.62	2.52
Plant & Animal sciences	164	253	8	8	2.74	2.51

Source: Own calculations. InCites™

The 55% growth of Latin America scientific production between 2004-2008 and 2009-2013 is also more or less proportionally reflected in the growth of number of nodes (RDs) in all subject areas networks. In both periods, the average path length is relatively low, meaning that knowledge that is created in one node has the potential to be diffused quite fast to the rest of the network.

Interestingly, the change in the number of communities does not follow a common trend. Engineering, Plant and Animal, and to some extent, Environmental sciences shows a remarkable increment in the number of RDs in the LAC network, but marginal changes in the number of communities, suggesting that newcomers were rapidly absorbed by well-established groups of collaborations. On the other hand, Agricultural and Geosciences at least double the number of knowledge communities, what can be interpreted that evolving networks are creating new niches of knowledge, either with new local actors or increasing diversification of knowledge sources through new international collaborations.

Shortest average paths together with an increasing number of communities are signals that a network structure is evolving towards a structure that facilitate both knowledge creation and knowledge diffusion. However, attention needs to be paid to the fact that in almost all scientific fields studied is common to observe that two neighboring countries, regularly with very similar economic structures, are only connected with each other through an institution from a third country. Even when this situation gives potential brokerage power to the external RD, it is not clear what the impact is for the performance of LAC networks. Clearly, this is a topic that requires further research.

As we mentioned before, for the econometric implementation we also estimate centrality measures for local/national networks. For each country, these networks are formed by all elite national RD (same threshold defined before) and its research partners. Foreign institutions are also included in the network, but those that have collaborations with only one local RD are considered peripheral, and then dropped from the network. After application of these filters, we are left with data only from Argentina, Brazil, Chile, and Mexico[j].

7. Econometric implementation and descriptive statistics

We firstly estimate the Cragg (1971) model described above, using the user-developed craggit routine in the *Stata* software. Thereafter we estimate a two-step Heckman selection model, using the same mentioned software. We use pooled data from LAC scientific networks, at the RD level, in five scientific fields: Agricultural, Engineering, Environmental, Geosciences, and Plant and Animals sciences.

For some economic activities, and particularly the case of natural resources based industries, higher degrees of local knowledge are required for technological innovations (Katz, 2004) At the same time, the tacit component of knowledge is key to success in technology transfer activities and it is highly correlated with geographic proximity (Pinch et al., 2003). Thus, firms should tend to favor collaborations with local or national RD. We try to capture these differences by also including in our model network features of the RDs relative to the national scientific networks. We control for the type of organizational structure where the RDs are operating, accounting for differences in objectives of the research performed in each one of them. We also control for unobservable effects of the scientific field on the degree of closeness to the industry and for idiosyncratic characteristics of national innovation systems including a set of field and country dummies. We cluster error at the institution level, allowing errors to be correlated among RDs that are part of the same institution.

Usually, quantitative studies assessing causality based on statistics and data from networks are subject to endogeneity biases. In our case, it would be in the causal direction in the relationship between the position of a RD in the relevant scientific network and the degree of the intensity of the collaboration with the industry. We try to address this potential problem using information from two separate periods of time, allowing RDs characteristics and network position of one period to impact on the collaborations with the private sector in the following period. We choose 5-years time span in order to account for both, the duration of the publication process, and the expected time in which a publication would impact in the scientific communities (Crespi & Geuna, 2008).

[j] We could get the national network of Venezuela, but data requirements for the econometric estimations left these observations out of the final dataset.

After the application of these data requirements to the nodes gathered from the social network analysis and dropping outliers on the outcome variable[k], we finish with a database that includes 324 observations from four LAC countries in the five selected scientific topics. It is worth mentioning that our estimation of collaborations between RD and industry suffers from some measurement shortcomings, mostly derived from the fact that we are only relying on scientific publications data. According to Cohen et al. (2000), joint publications are just one type of different channels of technology transfer available. Research funded by industry, co-patenting, or even research collaborations that, perhaps because of confidentiality issues, does not involve the publication of scientific papers could be happening and are not being captured by the data that we are using. Subsequently, our results should be considered as a lower bound on the relation between science and industry. However, Agrawal & Henderson (2002) show that collaboration in publications is, for example, far more common as a mechanism of technology transfer than patenting.

Table 4, in the appendix, summarizes descriptive statistics of the main variables used in the econometric analysis. Overall, 0.36% of the publications of a typical RD in LAC were coauthored with the industry in the period 2009-2013, a lower share than the 0.45% of period 2004-2008. Normalized degree shows that the average node in the LAC network has linkages to the 6% of the total number of members of its network. When considering the local network, that indicator goes up 16%. In the period 2004-2008, the typical LAC RD was involved in an equivalent of the 3% of the number of shortest paths that pass through the node with the maximum betweenness value of that network, a third of the percentage when considering national networks. Normalized closeness shows, as expected, that members of national networks are at shortest distances than when considering all member of the LAC network. Regarding number of publications, the average RD published approximately 179 papers in the field, during the period 2004-2008. In the same period, the quality of publications indexes considered here show that on average publications of LAC RT underperform in relation to the rest of the world, having 21% fewer citations (citation index of 0.79) than the average paper in the same field, and only 6.1% of publications in the field are in the top 10% of the discipline.

Brazil concentrates almost 50% of the RD considered in this study, mirroring the regional importance of the country relative to total number of publications and in terms of GDP. Argentina, Chile, and Mexico account for the other half of the observations in our sample. Plant and Animal sciences account for 27% of the RDs here considered. Agricultural, Engineering, and Environmental sciences represents roughly 20% each while Geosciences accounts for the remaining 14% of the cases. The majority of the RDs in the sample are part of universities (86%) and the remaining 14% belong to research institutes and governmental agencies.

---

[k] There was only one observation that showed an outcome above 3 standards deviations above the mean. In fact, the geological sciences of the "Universidade do Vale do Rio dos Sino" from Brazil, was 8 standard deviations above the mean.

8. Results

Table 5 shows findings regarding determinants of the participation of co-publication with the industry and its intensity. Unsurprisingly, past collaborations with the industry reveal as a strong predictor of participation in new collaborations. Across different implementations of the model the sign and statistical significance of these coefficients remains. However, it is quite surprising that the statistical significance of the positive impact does not hold for the intensity of collaborations equation.

Interestingly, the average academic quality of the publications in the relevant field does not impact neither on the likeliness of collaboration with the industry nor the intensity in none of the different specifications of our model. On Table 5 we are reporting results associated to the citation index. We also run estimations including the percentage of publications in the world top10% of the field, and results were extremely similar. One possible interpretation of these results is that quality in academy and quality in the industry means totally different things. Therefore, private firms would be relying on other sources of information to assess the quality of potential partners. It also could be the case that scientific fields defined in this analysis are too broad to capture the specificity required by the industry; therefore aggregate measures of quality could be hiding the quality of the valuable knowledge of the RD.

Centrality in the LAC scientific network does not seem to be important regarding the relationship with the private sector. Indeed, neither degree (estimations (2) and (3)), betweenness (estimations (4) and (5)), nor closeness (estimations (6) and (7)) of the LAC networks show statistically significant coefficients. If any, effects seem to be positive for participation but negative for intensity, suggesting that RDs more central in the LAC context are going to prefer to work with by the private sector, but the focus of its publications is strictly academic.

The main finding in our results is the fact that two of the centrality measures of the national scientific networks (estimations (2) and (4)), as opposed to the LAC level, indeed are positively correlated, and are statistically significant, with the intensity of the collaboration with the industry. RDs with higher values of local degree and betweenness engage more intensively in research activities with the private sector, but it does not seem to affect in the participation equation. These findings suggest that the access to a more diversified set of knowledge sources at the local level makes RDs more suitable to perform research also valuable for the local industries. Somehow puzzling, closeness centrality measure do not show a statistical significant relation neither participation nor intensity of collaborations with the industry. Furthermore, as opposite to the other two centrality measures studied, the sign of local closeness seems negatively correlated with the intensity of co-publications with the private sector. However, the latter correlation does not hold in the two-steps Heckman estimation, suggesting that it is only a statistical issue.

Table. 5 Craggit estimation of intensity of collaboration with the industry

Industry collaboration intensity (2009-2013)	No Centrality		Degree			
	(1)		(2)		(3)	
	Tier1	Tier2	Tier1	Tier2	Tier1	Tier2
Industry collaboration intensity (2004-2008)	0.4458***	0.3579	0.4563***	0.3259	0.4561***	0.2802
	(0.1165)	(0.2484)	(0.1200)	(0.2530)	(0.1199)	(0.2025)
Number of publications (log) (2004-2008)	0.7133***	-1.3882**	0.5874***	-0.7842	0.5877***	-0.3753
	(0.1122)	(0.5444)	(0.1813)	(0.6362)	(0.1811)	(0.6986)
Quality (2004-2008)	0.0089	-0.1093	0.0076	-0.0797	0.0075	-0.0438
	(0.0231)	(0.1186)	(0.0236)	(0.1142)	(0.0237)	(0.1011)
Brazil	0.4351*	4.0722**	0.4063	4.1814**	0.4014	3.4978**
	(0.2496)	(1.8118)	(0.2535)	(1.8458)	(0.2557)	(1.5916)
Chile	-0.1100	-0.1071	-0.1391	0.2295	-0.1373	0.4643
	(0.3198)	(1.7456)	(0.3101)	(1.8384)	(0.3105)	(1.6820)
Mexico	0.1539	2.6570	0.1092	2.7322*	0.1014	2.0342
	(0.3484)	(1.6268)	(0.3522)	(1.6432)	(0.3568)	(1.5811)
Agricultural	0.6438**	2.0409	0.5804*	2.1106	0.5738*	1.5893
	(0.3053)	(3.2737)	(0.3154)	(3.3053)	(0.3139)	(2.9091)
Engineering	1.3783***	9.8622**	1.3203***	9.6726**	1.3133***	8.1659**
	(0.3138)	(4.5963)	(0.3249)	(4.5390)	(0.3232)	(3.6974)
Environmental	0.5307*	5.6520	0.4536	5.7809	0.4492	4.7709
	(0.2940)	(3.9083)	(0.3191)	(3.9234)	(0.3198)	(3.3091)
Geosciences	0.7910**	8.9886**	0.6281	9.2261**	0.6270	8.1466**
	(0.3977)	(4.4545)	(0.4755)	(4.4834)	(0.4747)	(3.6893)
University	0.5258**	-1.3041	0.5694**	-1.3348	0.5643**	-1.0294
	(0.2332)	(1.6462)	(0.2288)	(1.5327)	(0.2287)	(1.3925)
Normalized degree (2004-2008)			2.2307	-9.6152	2.1664	-11.7398
			(2.2971)	(7.6304)	(2.2777)	(7.1588)
Normalized local degree (2004-2008)					0.1155	2.7692**
					(0.4597)	(1.2593)
Constant	-5.4481***	-4.1878	-4.9338***	-6.5906	-4.9380***	-7.2720
	(0.8585)	(6.7205)	(1.0633)	(7.2772)	(1.0650)	(6.5955)
Sigma	1.8431***		1.8317***		1.6888***	
	(0.3091)		(0.3028)		(0.2678)	
N	324		324		324	

Notes: Clustered errors at institution level. Standard errors in parentheses. * Coefficient is statistically significant at the 10 percent level; ** at the 5 percent level; *** at the 1 percent level; no asterisk means the coefficient is not different from zero with statistical significance.

Table. 5 (cont.) Craggit estimation of intensity of collaboration with the industry

Industry collaboration intensity (2009-2013)	Betweenness				Closeness			
	(4)		(5)		(6)		(7)	
	Tier1	Tier2	Tier1	Tier2	Tier1	Tier2	Tier1	Tier2
Industry collaboration intensity (2004-2008)	0.4547***	0.3620	0.4543***	0.3286*	0.4747***	0.3283	0.4651***	0.3285
	(0.1195)	(0.2493)	(0.1193)	(0.1964)	(0.1201)	(0.2756)	(0.1199)	(0.2761)
Number of publications (log) (2004-2008)	0.6213***	-1.6080**	0.6213***	-1.1141	0.5950***	-1.0672*	0.6152***	-1.1270
	(0.1436)	(0.6502)	(0.1437)	(0.7083)	(0.1283)	(0.6322)	(0.1311)	(0.7439)
Quality (2004-2008)	0.0068	-0.1238	0.0067	-0.0628	0.0125	-0.0901	0.0116	-0.0914
	(0.0242)	(0.1282)	(0.0243)	(0.1155)	(0.0240)	(0.1258)	(0.0248)	(0.1260)
Brazil	0.4261*	3.9814**	0.4242*	3.2640**	0.3280	4.3527**	0.2728	4.4197**
	(0.2538)	(1.7920)	(0.2555)	(1.5256)	(0.2498)	(1.7759)	(0.2660)	(1.8297)
Chile	-0.1557	-0.2540	-0.1540	-0.1488	-0.2142	0.3227	-0.2593	0.3378
	(0.3134)	(1.7903)	(0.3139)	(1.6503)	(0.3024)	(1.8479)	(0.3078)	(1.8600)
Mexico	0.1037	2.5558	0.0976	1.8905	0.1463	2.4304	0.0985	2.5040
	(0.3549)	(1.6259)	(0.3595)	(1.5557)	(0.3475)	(1.6192)	(0.3636)	(1.7214)
Agricultural	0.5909*	1.9563	0.5875*	1.6884	0.7687**	1.5155	0.7614**	1.5376
	(0.3092)	(3.1977)	(0.3081)	(2.8596)	(0.3171)	(3.2933)	(0.3166)	(3.2565)
Engineering	1.2916***	9.6009**	1.2859***	8.0995**	1.6320***	8.7661*	1.6008***	8.8415*
	(0.3322)	(4.4741)	(0.3315)	(3.8140)	(0.3806)	(5.0669)	(0.3745)	(5.0038)
Environmental	0.4138	5.2310	0.4085	4.3991	0.5016*	5.7184	0.4992*	5.6818
	(0.3210)	(3.8456)	(0.3229)	(3.4247)	(0.2946)	(3.8693)	(0.3005)	(3.8683)
Geosciences	0.6367	8.6376*	0.6302	7.6083**	0.7193*	8.8090**	0.7683*	8.8182**
	(0.4544)	(4.4138)	(0.4531)	(3.7590)	(0.4157)	(4.4590)	(0.4186)	(4.4389)
University	0.5589**	-1.3637	0.5536**	-1.2150	0.5463**	-1.0862	0.5246**	-1.1351
	(0.2328)	(1.6437)	(0.2333)	(1.4357)	(0.2252)	(1.4468)	(0.2227)	(1.4860)
Normalized beteweenness (2004-2008)	2.5164	3.9438	2.4597	-1.4633				
	(2.2906)	(9.1984)	(2.3112)	(7.5636)				
Normalized local beteweenness (2004-2008)			0.1604	3.2971**				
			(0.5978)	(1.4762)				
Normalized closeness (2004-2008)					23.2792	-59.7806	20.9869	-57.3279
					(15.6393)	(92.7442)	(15.2277)	(93.0904)
Normalized local closeness (2004-2008)							6.6198	-4.4030
							(6.3685)	(23.1923)
Constant	-5.0130***	-2.7913	-5.0142***	-3.7046	-5.9531***	-3.0014	-6.2486***	-2.6051
	(0.9727)	(6.4384)	(0.9723)	(6.2620)	(0.9233)	(7.4359)	(0.9928)	(7.9800)
Sigma	1.8343***		1.6925***		1.8249***		1.8282***	
	(0.3038)		(0.2636)		(0.3175)		(0.3178)	
N	324		324		324		324	

Notes: Clustered errors at institution level. Standard errors in parentheses. * Coefficient is statistically significant at the 10 percent level; ** at the 5 percent level; *** at the 1 percent level; no asterisk means the coefficient is not different from zero with statistical significance.

Table. 6 Two-step Heckman of intensity of collaboration with the industry

Industry collaboration intensity (2009-2013)	No Centrality (8)		Degree (9)		Betweenness (10)		Closeness (11)	
	Participation	Intensity	Participation	Intensity	Participation	Intensity	Participation	Intensity
Industry collaboration intensity (2004-2008)	0.4458***	0.3541**	0.4561***	0.3358*	0.4543***	0.3361**	0.4651***	0.3452*
	(0.1696)	(0.1476)	(0.1449)	(0.1788)	(0.1664)	(0.1661)	(0.1630)	(0.2035)
Number of publications (log) (2004-2008)	0.7133***	--	0.5877***	--	0.6213***	--	0.6152***	--
	(0.1440)	--	(0.2073)	--	(0.1600)	--	(0.1620)	--
Quality (2004-2008)	0.0089	-0.0303	0.0075	-0.0239	0.0067	-0.0223	0.0116	-0.0277
	(0.0286)	(0.0427)	(0.0295)	(0.0414)	(0.0279)	(0.0447)	(0.0283)	(0.0430)
Brazil	0.4351	0.7875**	0.4014	0.7260**	0.4242	0.7159**	0.2728	0.7313**
	(0.3095)	(0.3311)	(0.3009)	(0.3467)	(0.2990)	(0.3154)	(0.3819)	(0.3047)
Chile	-0.1100	-0.2022	-0.1373	-0.1445	-0.1540	-0.2121	-0.2593	-0.2942
	(0.3629)	(0.4267)	(0.3637)	(0.4296)	(0.4596)	(0.4693)	(0.4360)	(0.5341)
Mexico	0.1539	0.2080	0.1014	-0.0575	0.0976	-0.0574	0.0985	0.0851
	(0.4529)	(0.5426)	(0.4657)	(0.5920)	(0.4310)	(0.4742)	(0.4906)	(0.4725)
Agricultural	0.6438**	0.1663	0.5738*	0.0272	0.5875	0.1271	0.7614**	0.1090
	(0.3241)	(0.1977)	(0.3031)	(0.2647)	(0.3590)	(0.2287)	(0.3181)	(0.4931)
Engineering	1.3783***	1.5513***	1.3133***	1.3846***	1.2859***	1.3639***	1.6008***	1.4349
	(0.3606)	(0.3078)	(0.3033)	(0.4184)	(0.3678)	(0.3449)	(0.4287)	(0.9541)
Environmental	0.5307	0.3684	0.4492	0.2370	0.4085	0.2347	0.4992	0.3665
	(0.3376)	(0.2536)	(0.3295)	(0.2591)	(0.3745)	(0.2437)	(0.3130)	(0.2797)
Geosciences	0.7910	1.4312*	0.6270	1.4080**	0.6302	1.4013**	0.7683	1.3811**
	(0.5017)	(0.7351)	(0.4670)	(0.6332)	(0.5233)	(0.6258)	(0.4904)	(0.6776)
University	0.5258	-0.1325	0.5643*	-0.1951	0.5536*	-0.1885	0.5246*	-0.1599
	(0.3260)	(0.5253)	(0.3079)	(0.6095)	(0.2936)	(0.4921)	(0.2819)	(0.6050)
Normalized degree (2004-2008)			2.1664	-0.3799				
			(2.8682)	(2.5330)				
Normalized local degree (2004-2008)			0.1155	1.3424**				
			(0.5662)	(0.6475)				
Normalized beteweenness (2004-2008)					2.4597	-0.4082		
					(2.8869)	(2.1994)		
Normalized local beteweenness (2004-2008)					0.1604	1.8189**		
					(0.7469)	(0.7842)		
Normalized closeness (2004-2008)							20.9869	-2.2580
							(20.9892)	(39.7345)
Normalized local closeness (2004-2008)							6.6198	9.1075
							(7.5071)	(8.1165)
Constant	-5.4481***	-0.8432	-4.9380***	-0.8648	-5.0142***	-0.8211	-6.2486***	-1.0981
	(1.1140)	(0.7192)	(1.1832)	(1.0783)	(1.0536)	(0.7752)	(1.1215)	(2.8895)
lambda	0.8369***		0.8634*		0.8539**		0.8087	
	(0.2314)		(0.4931)		(0.3790)		(0.5537)	
N	324		324		324		324	
sigma	1.1598		1.1307		1.1195		1.1389	
chi2	61.8592		70.3405		67.2307		68.6761	
p	0.0000		0.0000		0.0000		0.0000	
rho	0.7215		0.7636		0.7628		0.7100	

Notes: Coefficients reported are marginal effects. Bootstrapped clustered errors at the institution level (100 repetitions). *Coefficient is statistically significant at the 1 percent level; ** at the 5 percent level; *** at the 1 percent level; no asterisk means the coefficient is not different from zero with statistical significance.

Country dummies shows that there is no significant difference on the likeliness of participation with the private sector, but in Brazil RDs that do participate, are doing it more intensively than their counterparts in Argentina, Chile and Mexico. Finally, engineering sciences are consistently the research field with more collaboration with the industry in both participation and intensity, a result that we would expect given the nature of the engineering activities (less "basic" and more "applied" activities). The result in the intensity equation also holds for Geosciences, probably due to the importance of mining activities in the sample of countries included in our analysis. However, as it was shown in section 4, it is worth to remark that these two disciplines are far to be one the most preferred research topics of scientific institutions in LAC.

Finally, even when controlling for networks centrality features, the size of the RD reveals as a strong predictor of the likeliness of performing research with the private sector, i.e. larger RDs tend to collaborate with the industry. On the other hand, the intensity of these collaborations is lower when comparing with smaller RD that participates in publications with the private sector. We are aware that there is a risk of multicollinearity between the size of the RD and the centrality measures at the LAC level, but we are confident that theoretically both types of variable measure different things then it must be included in the estimations. Excluding either size or centrality at the LAC level will give rise to a problem of omitted variables. However, this potential problem needs to be taken into account when interpreting results.

In table 6 we present estimations of a two-step Heckman model for specifications (3), (5) and (7) of the model. Based on results of the Craggit estimations we use the size of the firm as the exclusion variable, i.e., that affect the participation decision, but no in the intensity equation. Most of the results of the previous estimations hold. However, in this set of estimations the previous collaboration with the industry not only increases the likeliness of participation in co-publication, but also in the intensity of these collaborations. An extra 1% of co-publications with the industry in one period increase these activities in 0.35-0.45% in the next period. Despite this change, the degree and betweenness values of RDs in their local scientific networks positively affect the level of collaborations with the industry.

9. Conclusions

In this paper, we use a combination of bibliometric, social network and econometric techniques to increase the understanding of scientific systems and its relationship with the private sector in LAC. We studied recent trends on scientific outcome, measured as publications, by LAC countries and the linkages that exist between RDs within and between these countries, with a particular focus on collaboration activities with the industry.

We found that LAC share of global scientific publications started to increase at a higher rate since 1993, thus revealing a trend for convergence with the world leading regions. This increase has been mainly caused by Brazil and most notably in subject areas such as Agricultural, and Plant & Animal sciences. Moreover, when analyzing the relative scientific output normalized by GDP (Docs/GDP) and population (Docs/Pop), the results show that in the most recent years Chile, Uruguay, Argentina, and Brazil have levels of scientific productivity higher than the world average. Furthermore, specialization of scientific systems in LAC tends to follow economic specialization, focusing on scientific fields related to natural resources. There exist examples where the intensive use of certain technologies in natural resources exploitation activities is accompanied by a notable increment in research and publications related to these technologies. However, most LAC countries have an average industry collaboration percentage in the last decade below 1%. This is a very low number when compared to the rest of the world. There are differences between fields (Engineering and Geosciences show higher levels) but, in general, collaborations between science and industry, measured as co-publications, are scarce.

The growth of scientific production can also be appreciated in the increasing number of RDs embedded in LAC scientific networks. Regardless scientific fields, international collaborations of LAC RDs have been increasing, enhancing diversification of knowledge sources and improving channels for diffusion of new knowledge. However, collaboration within LAC RDs remains low. In most of the fields studied in this paper, linkages between LAC RDs are scarce, even when these countries tend to specialize in similar scientific fields and economic activities. Understanding if this lack of integration between LAC scientific institutions is harming potential gains of complementary knowledge is clearly a matter of further research.

The main finding is that the access to more diversified sources of knowledge matter for the intensity of the collaboration between science institutions and the industry, suggesting that there is a space for promoting industry-science collaboration through increasing linkages of local RDs with other local and foreign scientific institutions. However, other public policies such grants that encourage joint research projects between science institutions and the private sector should remain or increase relevance, because dependency on previous collaborations is revealed as key predictor of futures relations between RD and the industry.

Complementing this analysis with qualitative approaches and with primary data that accounts for other types of technology transfer activities and sources of funding for research, would certainly improve the understanding of LAC knowledge production, transfer and diffusion systems. Focusing on science-industry linkages through examining scientific publications strictly related to the most relevant economic activities in LAC are suggestions for further research.

10. References

Abramovitz, M. (1986). Catching Up, Forging Ahead, and Falling Behind. *The Journal of Economic History*, 46(2), 385–406.

Agrawal, A., & Henderson, R. (2002). Putting Patents in Context: Exploring Knowledge Transfer from MIT. *Management Science*, 48(1), 44–60.

Anlló, G., and D. Suárez. (2009). "Innovación: Algo más que I+D. Evidencias Ibeoramericanas a partir de las encuestas de innovación: Construyendo las estrategias empresarias competitivas". CEPAL-REDES, Buenos Aires, Unpublished.

Catanzaro, M., Miranda, G., Palmer, L., & Bajak, A. (2014). South American science: Big players. *Nature*, 510(7504), 204–206.

Cimoli, M. (2010). Vínculos entre universidades y empresas para el desarrollo tecnológico. CEPAL. Secretaría General Iberoamericana. Santiago de Chile.

Cohen, W. M., & Levinthal, D. A. (1990). Absorptive Capacity: A New Perspective on Learning and Innovation. *Administrative Science Quarterly*, 35(1), 128–152.

Cohen, W. M., Nelson, R. R., & Walsh, J. P. (2000). Protecting their intellectual assets: Appropriability conditions and why US manufacturing firms patent (or not). National Bureau of Economic Research.

Cragg, J. G. (1971). Some Statistical Models for Limited Dependent Variables with Application to the Demand for Durable Goods. *Econometrica*, 39(5), 829–844.

Crespi, G. (2012). "Fiscal Incentives for Business Innovation". In . Corbacho, coord., The Fiscal Institutions of Tomorrow. Institutions for People series. Washington, DC: Inter-American Development Bank.

Crespi, G. A., & Geuna, A. (2008). An empirical study of scientific production: A cross country analysis, 1981–2002. *Research Policy*, 37(4), 565–579.

Freeman, L. C. (1978). Centrality in social networks conceptual clarification. *Social Networks*, 1(3), 215–239.

Goode, R. B. (1959). Adding to the Stock of Physical and Human Capital. *The American Economic Review*, 49(2), 147–155.

Griffith, R., Redding, S., & Reenen, J. V. (2004). Mapping the Two Faces of R&D: Productivity Growth in a Panel of OECD Industries. *The Review of Economics and Statistics*, 86(4), 883–895.

Griliches, Z. (1979). Issues in Assessing the Contribution of Research and Development to Productivity Growth. *Bell Journal of Economics*, 10(1), 92–116.

Gulati, R., and Martin G. (1999). "Where do interorganizational networks come from? 1." *American journal of sociology* 104, no. 5: 1439-1493.

Gulati, R. (1995). "Social structure and alliance formation patterns: A longitudinal analysis." *Administrative science quarterly*: 619-652.

Heckman, J. J. (1979). Sample Selection Bias as a Specification Error. *Econometrica*, 47(1), 153–61.

IADB. (2010a). The Age of Productivity: Transforming Economies from the Bottom Up.

IADB. (2010b). Science, Technology, and Innovation in Latin America and the Caribbean: A Statistical Compendium of Indicators. Washington, DC: Inter-American Development Bank.

InCites™ Thomson Reuters. (2014). Data processed 04 December, 2014. Data source: Web of Science.

Jeffrey, P. (2003). Smoothing the waters: Observations on the process of cross-disciplinary research collaboration. *Social Studies of Science*, 33, (4): 539-562.

Katz, J. (2004). Economic, institutional and technological forces inducing the successful inception of salmon farming in Chile. In Vandana Chandra ed. Technology, adaption and exports: how some countries got it right, Washington DC, World Bank; 32pp.

Katz, J. (2006). Structural change and domestic technological capabilities. CEPAL Review , 89, 55-68, August.

La Paz, A., Cancino, C., & Miranda, J. (2012). Start-ups success using public funds: university versus industry sponsorship. *Multidisciplinary Business Review*. Vol. 5 N°1 (Junio 2012).

Lederman, D., & Maloney, W. F. (2003). R&D and Development (SSRN Scholarly Paper No. ID 402480). Rochester, NY: Social Science Research Network.

Lundvall, B.-Å. (1992). National Systems of Innovation: Towards a Theory of Innovation and Interactive Learning. Pinter Publishers.

Marin, Anabel, and Lilia Stubrin. "Innovation in natural resources: New opportunities and new challenges The case of the Argentinian seed industry." UNU-MERIT Working Paper 2015-015.

Marotta, D., Mark, M., Blom, A., & Thorn, K. (2007). Human capital and university-industry linkages‟ role in fostering firm innovation: an empirical study of Chile and Colombia. World Bank Policy Research Working Paper 4443, World Bank.

Menzel, M.-P., & Fornahl, D. (2010). Cluster life cycles—dimensions and rationales of cluster evolution. *Industrial and Corporate Change*, 19(1), 205–238.

Pérez, C. (2010). Technological dynamism and social inclusion in Latin America: a resource-based production development strategy. Cepal Review.

Pinch, S., Henry, N., Jenkins, M., & Tallman, S. (2003). From "industrial districts" to "knowledge clusters": a model of knowledge dissemination and competitive advantage in industrial agglomerations. *Journal of Economic Geography*, 3(4), 373–388.

Romer, P. M. (1990). Endogenous technological change. *Journal of Political Economy*, S71–S102.

Rouvinen, P. (2002). Characteristics of product and process innovators: some evidence from the Finnish innovation survey. *Applied Economics Letters*, 9(9), 575–580.

Schilling, M. A., & Phelps, C. C. (2005). Interfirm Collaboration Networks: The Impact of Small World Connectivity on Firm Innovation (SSRN Scholarly Paper No. ID 564422). Rochester, NY: Social Science Research Network.

Tsai, W. (2001). Knowledge Transfer in Intraorganizational Networks: Effects of Network Position and Absorptive Capacity on Business Unit Innovation and Performance. *Academy of Management Journal*, 44(5), 996–1004.

Van Noorden, R. (2014). The impact gap: South America by the numbers. *Nature*, 510(7504), 202–203.

Van Raan, A. F. (1998). The influence of international collaboration on the impact of research results. *Scientometrics*, 42(3), 423–428.

Vargas, R. A., de Souza Vanz, S. A., Stumpf, I. R., & others. (2014). The role of National journals on the rise in Brazilian Agricultural Science Publications in Web of Science. *Journal of Scientometric Research*, 3(1), 28.

Watts, D. J., & Strogatz, S. H. (1998). Collective dynamics of "small-world" networks. *Nature*, 393(6684), 440–442.

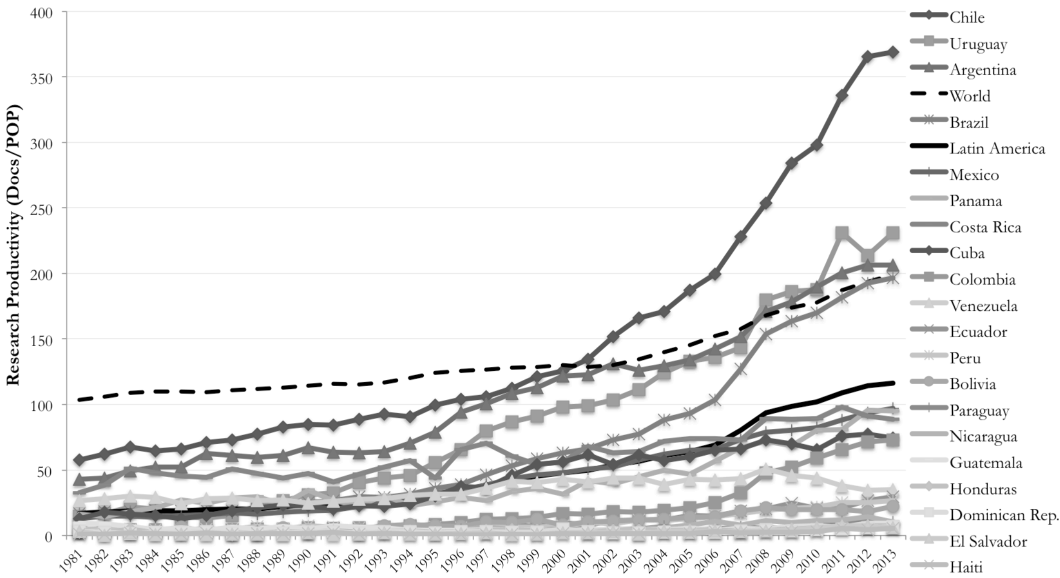
Wooldridge, J. M. (2002). Econometric Analysis of Cross Section and Panel Data. MIT Press.

Appendix

Table. 4 Descriptive statistics of variables used in the econometric analysis

Variable	Observations	Mean	Standard Deviation	Min	Max
Industry collaboration (0/1) (2004-2008)	324	0.22	0.41	0	1
Industry collaboration (0/1) (2009-2013)	324	0.30	0.46	0	1
Industry collaboration intensity (2004-2008)	324	0.45	1.42	0	10.87
Industry collaboration intensity (2009-2013)	324	0.36	0.98	0	6.44
Normalized degree (2004-2008)	324	0.06	0.08	0.01	0.47
Normalized local degree (2004-2008)	324	0.16	0.20	0.01	1
Normalized beteweenness (2004-2008)	324	0.03	0.06	0	0.37
Normalized local beteweenness (2004-2008)	324	0.09	0.17	0	1
Normalized closeness (2004-2008)	324	0.04	0.01	0.02	0.07
Normalized local closeness (2004-2008)	324	0.06	0.02	0.02	0.11
Number of publications (2004-2008)	324	178.73	284.35	5	2368
Number of publications (log) (2004-2008)	324	4.45	1.19	1.61	7.77
Quality Citation Index (2004-2008)	324	0.79	0.32	0.21	2.49
Quality Top 10% Index (2004-2008)	324	6.10	5.43	0	36.36
Argentina	324	0.15	0.35	0	1
Brazil	324	0.47	0.50	0	1
Chile	324	0.15	0.36	0	1
Mexico	324	0.23	0.42	0	1
Agricultural	324	0.20	0.40	0	1
Engineering	324	0.18	0.39	0	1
Environmental	324	0.20	0.40	0	1
Geosciences	324	0.14	0.35	0	1
Plant and Animal	324	0.27	0.44	0	1
Universities	324	0.86	0.35	0	1

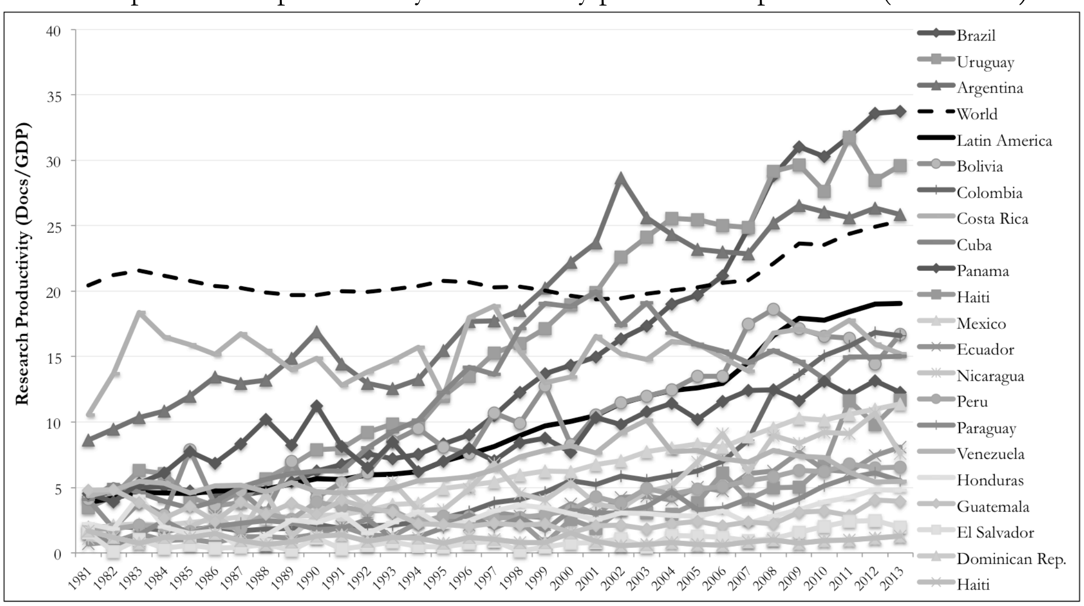
Fig. 2 Trends in publication productivity measured by publications per population (1981-2013)



Source: Own calculations based on InCites ™

[a] Docs/Population - Publications per million inhabitants.

Fig. 3 Trends in publication productivity measured by publications per GDP (1981-2013)

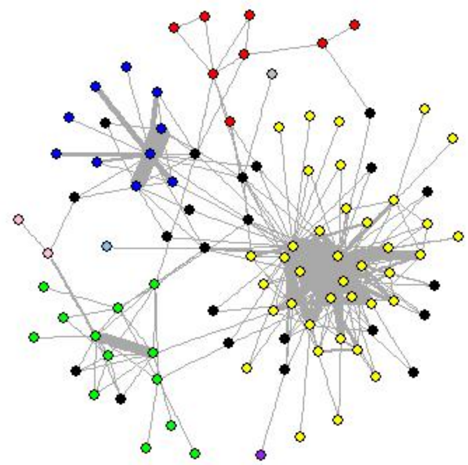


Source: Own calculations based on InCites ™

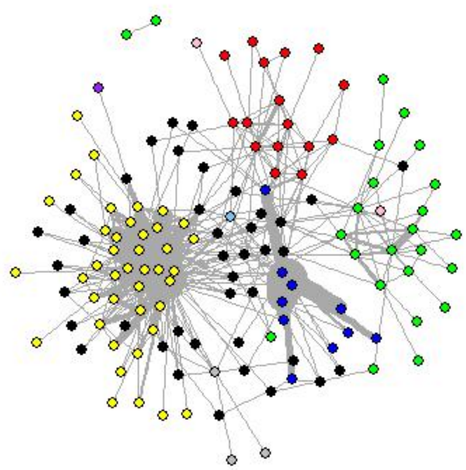
[a] Docs/GDP - Publications per billion constant 2005 US$ of GDP.

Figs. 7 Network structures of all subject areas in two periods (2004-2008 & 2009-2013)[12]

i. Agricultural sciences (2004-2008)



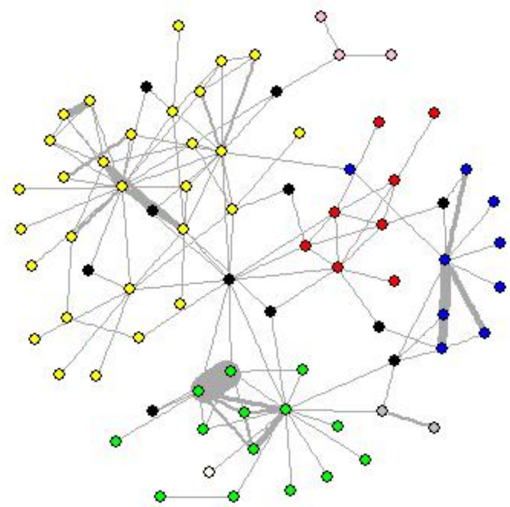
ii. Agricultural sciences (2009-2013)



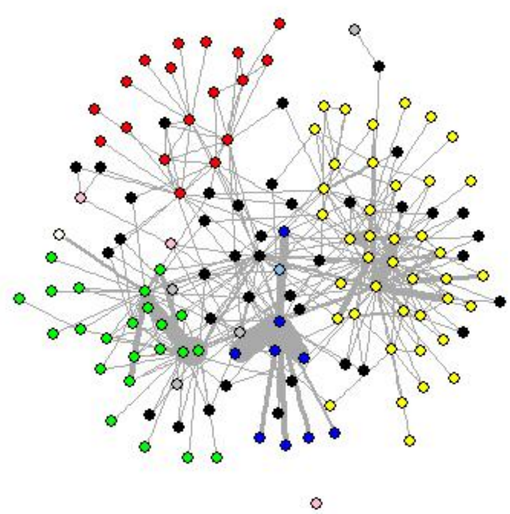
---

[12] Green: Mexico; Blue: Argentina; Red: Chile; Yellow: Brazil; Black: Non-LAC institutions; Purple: Peru; Pink: Venezuela; Light Blue: Uruguay; Grey: Colombia; Purple: Costa Rica; Ivory: Cuba; Orange: Panama.

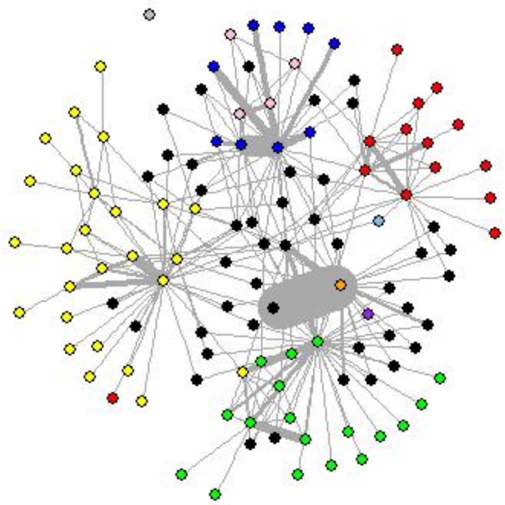
iii. Engineering (2004-2008)



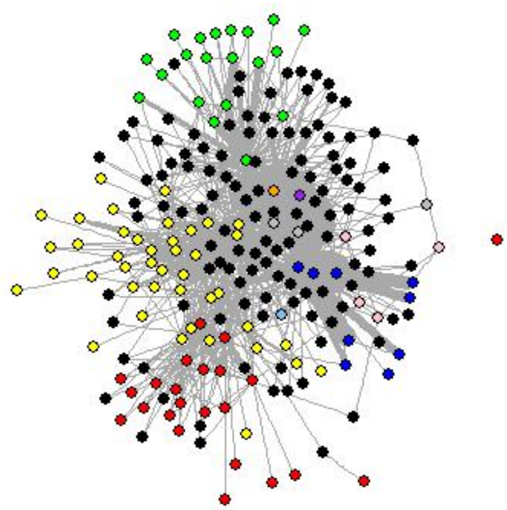
iv. Engineering (2009-2013)



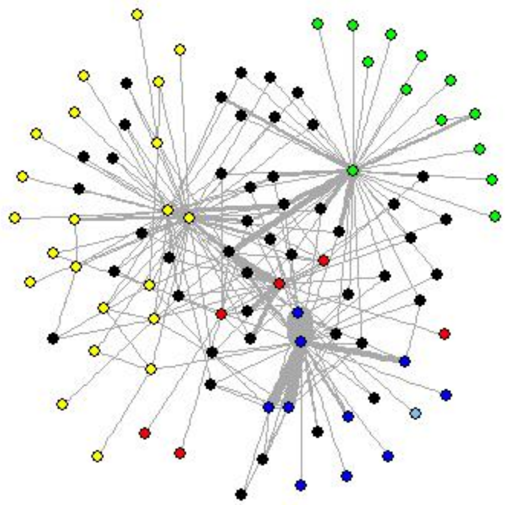
v. Environmental sciences (2004-2008)



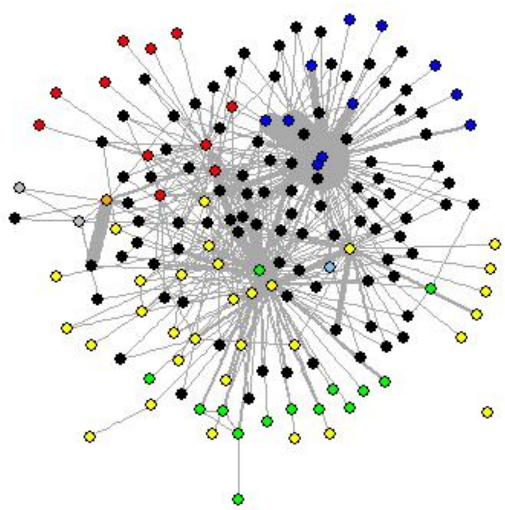
vi. Environmental sciences (2009-2013)



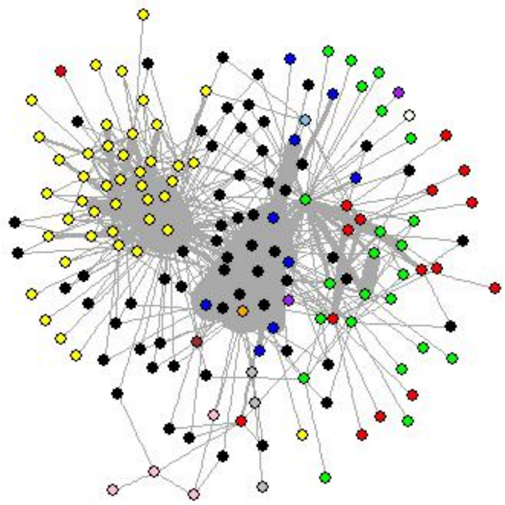
vii. Geosciences (2004-2008)



viii. Geosciences (2009-2013)



ix. Plant & Animal sciences (2004-2008)



x. Plant & Animal sciences (2009-2013)