Revista de Análisis Económico, Vol. 37, Nº 2, pp. 49-67 (Octubre 2022)

OPPORTUNITIES TO UPGRADE THE SCIENTIFIC DISCIPLINES SPACE*

OPORTUNIDADES PARA LA MEJORA DEL ESPACIO CIENTIFICO

NESTOR GANDELMAN**
Universidad ORT Uruguay

OSIRIS J. PARCERO***
Kazakh-British Technical University

FLAVIA ROLDAN****
Universidad ORT Uruguay

Abstract

Research resources in a given scientific domain may spill over into other close scientific disciplines, thereby improving performance. Using bibliometric data from the SCImago database drawn from a sample of 174 countries, we implement a measure of proximity-based on revealed comparative advantage (RCA) as a specialization or activity index. Our estimates show that proximity between disciplines positively and significantly related to the publication growth rate.

Keywords: *Revealed comparative advantage, revealed proximity, scientific production.*

JEL Classification: *L3, L38, O3, O5.*

* The authors acknowledge the financial support of the Agencia Nacional de Investigación e Innovación (ANII), Uruguay, under grant FMV_1_2019_1_156201.
We would like to thank Matilde Pereira for her excellent research assistance. We also thank Almat Kenen, Yerkezhan Kenzheali, and Zhibek Kassymkanova for their invaluable help with data processing.
** Universidad ORT Uruguay (gandelman@ort.edu.uy)
*** Kazakh-British Technical University (osiris.jorge.parcero@gmail.com)
**** Universidad ORT Uruguay (roldan@ort.edu.uy)

Resumen

Los recursos de investigación en un dominio científico determinado pueden usarse en disciplinas científicas cercanas, mejorando así su producción. Utilizando datos bibliométricos de la base de datos SCImago extraídos de una muestra de 174 países, implementamos una medida de proximidad basada en el concepto de ventajas comparativas reveladas (RCA) como índice de especialización o de actividad. Nuestras estimaciones muestran que la proximidad entre disciplinas científicas se relaciona de manera positiva y significativa con la tasa de crecimiento de producción de las publicaciones científicas.

Palabras clave: *Ventaja comparativa revelada, proximidad revelada, producción científica.*

Clasificación JEL: *L3, L38, O3, O5.*

I. INTRODUCTION

"…science allows us to build taller and taller ladders to reach ever-higher-hanging fruit". In *Building Taller Ladders.* Joel Mokyr, 2018.

In idea-based growth models, economic growth arises from people creating ideas. Ideas are the heart of science, and science, as a source of knowledge, is an engine that drives growth and productivity. This notion can be traced back to Adam Smith (1776) but has also sparked more recent studies aimed at explaining the relationship between scientific research and its economic growth effects.[1] Accordingly, the following questions become relevant. How can a country increase its scientific production? How might a country upgrade its performance in scientific research production? This article seeks to address these questions to the extent it explores the relationship between scientific production and the proximity between scientific disciplines.

In the trade literature, Hidalgo, Klinger, Barabási & Hausmann (2007) (HKBH hereafter) have developed a framework where a country's existing industrial structure determines its potential for technological upgrades. They find that the proximity between goods in which the country is already specialized and those not yet specialized plays a crucial role in the success of industrial upgrade processes.

---

1 Mansfieled (1972 & 1995); Rosenberg (1990); Jaffe in (1989); Adams (1990); Rosenberg and Nelson (1994); Partha and David (1994); Stephan (1996); Griliches (1998); Henderson *et al.* (1998), and even more recently Bloom *et al.* (2020).

Building on HKBH's (2007) framework, this article examines whether proximity between scientific disciplines affects their production growth rate. Informally, two scientific domains will be close to the extent that they require similar inputs. Our main idea is that some capabilities, institutions, knowledge, and other inputs necessary to produce science in certain disciplines might be used in other scientific domains. More precisely, suppose that a country is specialized in a particular domain that requires certain inputs. To the extent that these inputs are not completely specific, they might generate spillover in producing in other "close" scientific domains, i.e., those requiring similar resources.

Consequently, the ability to upgrade a country's science production in a scientific domain will depend on how similar (how close) this scientific domain is to those in which a country is already specialized. We approach science production as the number of papers produced by a country at a particular point in time and in a particular scientific discipline.[2] We further discuss the limitation of this approach since the difficulty quantifying the output of scientific research.

Our main methodological challenge is to define closeness between disciplines. A way to address this problem is to observe scientific production patterns. In the spirit of Balassa (1965) for trade patterns and following HKBH (2007), we compute a measure where bibliometric data (publications) reveal proximity among scientific disciplines. The revealed proximity measure indicates the likelihood of taking advantage of existing knowledge or other non-specific inputs with public goods' characteristics that will allow countries to advance in scientific domains.

To test how proximity between scientific disciplines affects a country's scientific performance, we use the SCImago dataset of world scientific production between 1996 and 2019. Our research offers three main contributions. First, inspired by the trade literature, we propose and implement a measure of proximity between scientific disciplines based on revealed comparative advantage (RCA) as an index of specialization or activity. Second, we provide evidence that proximity between disciplines positively and significantly affects a country's publication growth rate. Third, we include a wide range of countries and disciplines from all scientific areas, including Social Science, which has been excluded from most bibliometric studies (Harzing, 2013).

Place in the literature

Our article is placed in the economics of science literature (Partha & David, 1994 and Stephan, 1996). A large part of this literature concentrates on studying the linkages between basic science, technological innovation, and economic growth.[3]

---

[2] In some robustness exercises, we approach production using the number of citations instead of papers produced.

[3] To cite some relevant studies, Adams (1990) tests the effects of accumulated academic science on productivity in manufacturing industries. Academic science has spillover across industries, and, with

Assuming that science impacts economic growth, we investigate factors that boost scientific production and stimulate performance in scientific domains. To do that, we use bibliometric data for 307 disciplines rather than only basic research. We consider disciplines from Life Sciences, Health Sciences, Physical Sciences, Social Sciences and Humanities, and multidisciplinary fields. At this point, we are close to a line of research focused on studying the determinants of research output. Some research documents the relationship between scientific publications and individual scientists' demographic characteristics (Cole and Cole, 1974; Levin and Stephan,1991). Moreover, other articles analyze the impact of the institutional factors in the productivity of scientists (Stephan, 1996), and others the effects of funding on research output.[4]

The current article focuses on proximity between scientific disciplines to stimulate publication growth rates. We claim that there are spillovers between scientific disciplines and that such spillovers occur between close disciplines. Close to that idea, Crespi & Geuna (2008) study the relationship between investment in science and research outputs, quantified by publications and citations. They devote particular attention to studying the role of knowledge spillovers. Using panel data for a sample of OECD countries, authors quantify spillover of science investment from one country to another using a measure of proximity based on international scientific co-authorship. The higher the level of co-authorship, the more likely the existence of scientific investment spillover between countries. In the same spirit, we measure the proximity between scientific disciplines based on their level of co-production, i.e., when a country simultaneously produces a pair of disciplines. The higher the level of co-production in two scientific domains, the greater the probability of research resources spillover between disciplines.

This article also relates to international trade literature, from which we borrow two key and interrelated concepts: revealed comparative advantage and revealed proximity. In particular, by using the RCA measure from Balassa (1965), we compute the revealed proximity index from HKBH (2007). They argue that the more a given product has nearby products, the faster a producing nation can transform its productive profile. In the same vein, articles like Hausmann & Klinger (2006) and Thomson & Athukorala (2020), among others, explore the underlying idea that a country's existing industrial structure determines its industrial upgrade opportunities. All these studies use data on

---

long lags (roughly 30 years), it affects industrial productivity. Fleming & Sorenson (2004) study the mechanism through which science accelerates the rate of invention. They argue that science changes inventors' search processes, and they support this claim with an empirical test using patent data. Sorenson & Fleming (2004) consider that the act of publication, which makes information publicly available and encourages the rapid diffusion of knowledge, accounts for the linkage between science and economic growth. They find evidence that publication is an important mechanism in accelerating the rate of technological innovation.

4 Irwin and Klenow (1996), Lerner (1999), Klette *et al.* (2000) study the effects of public sponsored commercial R&D. For a review, see Hall and Van Reenen (2000). Moreover, Adams and Griliches (1998), Payne and Siow (1999), Bonoccorsi and Daraio (2003) among others, focus on the impact of funding on research output proxy by publications or citatitons at university or department level.

exports from the manufacturing sector. Our paper explores this concept but focuses on scientific knowledge production, proxied by bibliometric data.

The RCA index has been adopted for a wide variety of industries[5] and contexts, including scientometrics (Chuang *et al.*, 2010). The RCA index is usually called the activity index in scientometric literature. It was introduced by Frame (1977) and computed equal as the RCA index. We interpreted it as a measure of specialization rather than a measure of scientific capability.

The paper is organized as follows. In Section II, we present data and discuss methodology. Section III describes the econometrics and results. We conclude in Section IV.

II. DATA AND METHODOLOGY

This section describes the database and two fundamental concepts underlying the estimated equation presented in Section III. The key concepts are: (i) revealed comparative advantage, and (ii) revealed proximity between disciplines. We use both notions to compute the main variables in the econometric model.

II.a. Data source

We use public data from the SCImago Journal & Country Rank (SJCR) website based on Scopus.

This source covers citable documents per country, including those in over 23,452 peer-reviewed journals (and other serial publications). Our scientific production measure includes citable documents from journals and trade journals (articles, reviews of scientific relevance, and conference papers published in journals). We do not include book series or book reviews, letters, conference meeting abstracts, or non-serial sources. Articles are classified in 307 *non-exclusive* disciplines.[6] In addition, each paper is *non-exclusively* classified by the countries of the authors' affiliations. We access data from 1996 to 2019 and include 174 countries that have had at least 100 documents published in 2019.

[5] For example, forestry (Dieter and Englert, 2007), the manufacture of pharmaceutical products (Cai, 2018), or agriculture and food (Jambor and Babu, 2016). It has also been used in patent analysis (Soete and Wyatt, 1983; Zheng *et al.*, 2011), electronic commerce in the tourism sector and prevalence of the internet (Ruiz Gómez *et al.*, 2018), and start-ups and venture capital (Guerini and Tenca, 2018).

[6] It means that an article could be classified in more than one scientific field among the classification of the 307 scientific domains.

II.b. Revealed comparative advantage (RCA) or the activity index

The direct measure of comparative advantage is a complex calculation because it requires measuring the opportunity costs of production factors. Balassa (1965) developed an indicator that demonstrates how trade patterns reveal which products a country has a comparative advantage.

This indicator has been adopted in a wide variety of contexts, including scientometrics. Nonetheless, rather than a measure of strength (or revealed advantage), we interpret it, in our context, as a measure of specialization. Formally, the RCA (or specialization or activity) index is defined as the ratio between a discipline's participation in a country's scientific production and the participation of this same discipline in world scientific production. A country is specialized (or has RCA) in a particular discipline if within-country participation is larger than expected based on the participation of the discipline in the world scientific production. As proxies for scientific production, we use the number of published documents.[7] This approach involves some drawbacks. We are aware that the scientific process delivers several outputs. We can mention, among others: (i) new knowledge; (ii) highly qualified human resources; and (iii) new technology or knowledge with socio-economic impact.[8] In this article, we restrict our attention to the first kind of output, and we measure it by publications produced.

Formally, we calculate RCA for each of the 307 scientific disciplines identified for each country at a given time. The RCA of discipline *i* in country *c* at time *t* is computed as:

$$RCA_{i,t}^{c} = \frac{x_{i,t}^{c} / x_{t}^{c}}{X_{i,t}^{*} / X_{t}^{*}}$$

where $x_{i,t}^{c}$ is the number of published documents f discipline *i* in country *c* at time *t*, x_{t}^{c} is the number of documents published in *all* disciplines in country *c* at time *t,* $X_{i,t}^{*}$ is the number of documents published by discipline *i* in the world at time *t*, and X_{t}^{*} is the number of documents published by *all* disciplines in the world at time *t*.

When $RCA_{i,t}^{c}$ exceeds unity, country *c* at time *t* reveal specialization in the discipline *i*. Conversely, if $RCA_{i,t}^{c}$ is less than unity, country *c* is not specialized in the discipline *i* at time *t*.

---

7 In some exercises, we approach the output of scientific research by the citations generated by the published documents.

8 See Crespi & Geuna (2008).

It is worth noting that the RCA index allows for comparison across scientific domains (within a specific country) and among countries (within a specific discipline).

II.c. Revealed proximity and characterization of the space of scientific disciplines

In this paper, we study whether existing research resources to produce science positively contribute to producing science in close scientific domains. The main idea is that if a country is specialized in a discipline, some research resources like particular capabilities, institutional resources, environment, knowledge, and other inputs make this possible. Thus, if two disciplines are similar in the sense that they would require similar capabilities and similar other research resources, then it is likely that if we observe an RCA (or specialization) in one of them, we will also observe it in the other.

We follow a revealed approach based on the RCA (or specialization or activity) index to measure the proximity between two disciplines. This approach has the same inspiration as RCA: we let ex-post data "reveal" how similar scientific disciplines are without ex-ante considerations. In this way, we remain agnostic about factors determining proximity between different scientific disciplines.

Following HKBH (2007), revealed proximity is based on conditional probabilities. Looking at world data, we measure the probability of having RCA (specialization) in discipline *i* conditional on having RCA (specialization) in area *j*. Since conditional probabilities are not symmetric, we should also consider the converse: the probability of having RCA in area *j* conditional on having RCA in discipline *i*. The two probabilities are not necessarily equal. Formally, the revealed proximity measure is the minimum between these two statistics:

$$\varphi_{ij} = min\left\{Pr\left(RCA_{it} \mid RCA_{jt}\right), Pr\left(RCA_{jt} \mid RCA_{it}\right)\right\}$$

As the minimum of two conditional probabilities, this measure lies between 0 and 1; the larger the value, the closer the two disciplines are.[9]

Conditional probability is computed for each year using all the countries studied. Given that, φ_{ijt} has no country subscript, and by the definition of conditional probability, we have:

[9] The symmetric imposed solves a technical problem that arises when few countries have RCA in certain disciplines. As an extreme case, suppose that discipline *j* is only produced with RCA (or specialized in) by country c. Then, for every other discipline in which the country c is specialized, $Pr\left(RCA_{it} \mid RCA_{jt}\right)$ will be equal to 1. This fact would reflect the particular characteristic of the scientific profile of country c rather than similarity between disciplines. By taking the minimum, we overcome such a problem. See, for example, Hausmann & Kinger (2006) and Thomson & Athukorala (2020).

$$\varphi_{ijt} = min\left\{Pr\left(RCA_{it} \mid RCA_{jt}\right), Pr\left(RCA_{jt} \mid RCA_{it}\right)\right\}$$

$$= min\left\{\frac{\Pr\left(RCA_{it} \cap RCA_{jt}\right)}{\Pr\left(RCA_{it}\right)}, \frac{\Pr\left(RCA_{it} \cap RCA_{jt}\right)}{\Pr\left(RCA_{jt}\right)}\right\}$$

$$= \frac{Pr\left(RCA_{it} \cap RCA_{jt}\right)}{max\left\{Pr\left(RCA_{it}\right), Pr\left(RCA_{jt}\right)\right\}}$$

$$= \frac{number\ of\ countries\ that\ have\ RCA\ indisciplines\ i\ and\ j\ at\ time\ t}{\max\{number\ of\ countries\ with\ RCA\ indiscipline\ i\ at\ time\ t,\ number\ of\ countries\ with\ RCA\ indiscipline\ j\ at\ time\ t\}}$$

Proceeding in this way, we obtain the yearly revealed proximity matrix that accounts for each discipline's proximity to each remaining discipline. Table 1 lists disciplines that result in greater and lesser proximity for two scientific domains selected for illustration purposes from the 2019 proximity matrix.

TABLE 1

PROXIMITY MEASURES BETWEEN DISCIPLINES

By Publications

Business and International Management				Surgery			
High proximity		Low proximity		High proximity		Low proximity	
Business, Management and Accounting (miscellaneous)	65.50%	Medical Terminology	1.92%	Pathology and Forensic Medicine	51.80%	Computer Science Applications	2.40%
Strategy and Management	63.50%	Reviews and References (medical)	1.92%	Clinical Psychology	50.00%	Critical Care Nursing	0.00%
Management, Monitoring, Policy and Law	55.80%	Aerospace Engineering	0.00%	Otorhino-laryngology	50.00%	Environ-mental Chemistry	0.00%

III. ECONOMETRIC ANALYSES

We claim that increases in the output of scientific discipline can be brought about by the performance of nearby disciplines. For instance, if a country is specialized at period t in medicine and agriculture (i.e., both produced with RCA greater than one), it would be expected that a close scientific discipline such as veterinary medicine would have a potential advantage to take from medicine and agriculture toward itself. The following regression tests our hypothesis:

$$\begin{aligned} &Growth\ Publication_{j,c,t+1} \\ &= \alpha_1 AvgProximity_{j,c,t} + \alpha_2 Publication_{j,c,t} + \eta_j + \eta_t + \varepsilon_{j,c,t} \end{aligned} \tag{3}$$

The dependent variable, $Growth\ Publication_{j,c,t+1}$ is the annualized publication growth rate in scientific discipline *j,* in country *c*, between period *t* and *t+1*. More precisely, using the publication level for each discipline *j* in each country *c*, we compute geometric average growth for periods 2000-1996, 2004-2000, 2008-2004, 2012-2008, 2016-2012, and 2019-2016 as follows:

$$Growth\ Publication_{j,c,t+1} = \sqrt[n]{\frac{Publication_{j,c,t+1}}{Publication_{j,c,t}}} - 1,$$

where *n* is the number of years in the considered period.

Average Proximity $\left(AvgProximity_{j,c,t}\right)$ measures the proximity of discipline *j* to those disciplines in which country *c* at time *t* is specialized.[10] This variable is computed from the revealed proximity matrix illustrated in Table 1. Formally, $AvgProximity_{j,c,t}$ is calculated as follows:

$$AvgProximity_{j,c,t} = \frac{\sum_i \varphi_{ijt} I_{RCA_{i,t}}}{\sum_i \varphi_{ijt}} \in [0,1],$$

where $I_{RCA_{i,t}}$ is an indicator variable equal to one when country *c* has RCA greater than the unity in scientific domain *i* at time *t,* and zero otherwise. Thus, the numerator is the sum of the proximity of scientific domain *j* to all other disciplines where country *c* is specialized at time *t*. The denominator sums the proximity of domain *j* to all other disciplines at time *t*. Therefore, $AvgProximity_{j,c,t}$ is interpreted as the percentage of scientific space around discipline *j* in which country *c* is already specialized. For example, $AvgProximity_{j,c,t} = 0.2$ means the country *c* is specialized in the 20% of scientific space around *j* at time t. The larger the $AvgProximity_{j,c,t}$ the greater the probability of knowledge spillover toward *j* from disciplines in which country *c* is specialized.

---

[10] HKBH (2007) call this measure as *density*.

As in the growth literature, there may be convergence between disciplines. Those that start with lower values might experience more rapid growth rates than those with larger initial values. In order to control for this possible convergence in publication levels, we control for the level of publication in the domain j in country c in the previous period. If conditional convergence exists, a discipline with a higher level of publication in a particular period should experience lower growth in publication level in the following one (negative α_2). We take this variable in logs.

In some specifications, we include in equation (3) an interaction term to test whether the effect of *AvgProximity* depends on the level of publication.

Finally, to control the potential existence, at the country-discipline level, of unobserved characteristics that are correlated with both the growth of scientific production and proximity, we include a country-discipline fixed effect, η_j. In this way, we control for any country-discipline time-invariant unobservable variable that may lead to a spurious relationship between the growth of scientific production and proximity. Additionally, η_t is a time-fixed effect to control for possible common trends that may affect both the growth of scientific production and proximity across countries, and $\varepsilon_{j,c,t}$ corresponds to the error term.

It is natural to conjecture that *AvgProximity* might have different effects for disciplines in which country c is (or not) specialized. Consequently, we estimate equation (3) in two subsamples. The first group includes scientific disciplines-countries with an RCA of less than one at the beginning of a period, time t. The second group comprises scientific disciplines-countries with an RCA greater than or equal to one at the beginning of a period, time t.

Conditional on the inclusion of these control variables, our underlying identification assumption is that shocks affecting the proximity measure are uncorrelated with shocks affecting average publication growth. This assumption is plausible. Let us consider, for example, a shock that occurs in a particular discipline-country that affects the production of documents at a specific moment.[11] This shock will affect average publication growth in such a discipline, but not the *AvgProximity* variable. This fact is because that variable is a function of φ_{ijt}, which considers all countries' information beyond the country where the shock has taken place, as well as a function of indicator variable, I_{RCA}, which refers to all other scientific domains different from such a discipline.

[11] For example, in the last two years, as a consequence of the COVID-19 pandemic, the scientific production in domains related to health has seen an explosion. In the same way, one can think that some particular discovery or fact that could occur in a country can generate an unexpected shock in the country's scientific production in a specific discipline.

III.a. Descriptive evidence

Table 2 reports descriptive statistics for the variables detailed above. From the table, three observations are in order.

First, Publication growth differs depending on whether a scientific domain has an RCA (or specialization or activity) index greater or smaller than the unity. In particular, when the specialization index is less than unity, the mean growth in publications is 9.5%. In contrast, when the RCA index is greater than or equal to one, the mean growth in publication –6.9%.

Second, the *AvgProximity* variable has a smaller mean for disciplines in which a country is not specialized (RCA <1) than for disciplines that do (RCA ≥ 1).

Third, the distribution of publications exhibits a larger dispersion when disciplines have a RCA index greater than one than in the opposite case.

TABLE 2

SUMMARY STATISTICS

At time *t*: RCA < 1

	Count	Mean	Sd	Min	Max
GrowthPublicat	88,320	0.095	0.364	-1.000	4.848
AvgProximity	88,320	0.341	0.098	.0056	0.808
Publication	88,320	176	932	0.000	77,459

At time *t*: RCA≥ 1

	Count	Mean	Sd	Min	Max
GrowthPublicat	83,614	-0.069	0.378	-1.000	2.986
AvgProximity	83,614	0.403	0.122	0.022	1.000
Publication	83,614	273	1,399	0.000	90,875

Let us now examine in more detail the behavior of the *AvgProximity* variable. Figure 1 presents the distribution of *AvgProximity* variable. Panel (A) draws the *AvgProximity* for disciplines that initially (at time *t*) country *c* is not specialized. This group of disciplines identifies those in which the country will achieve specialization in the next period (solid green line) and those that remained with an RCA index less than the unity (red dashed line). As we can see, those disciplines that made a jump and gained a specialization (RCA) index greater than the unity between periods had a higher value according to our proximity indicator than those that failed to make such a jump. This fact is supportive of our claim that knowledge or other research resources in nearby disciplines in which a country is specialized might spill over and positively affect other scientific domains.

FIGURE 1

AVERAGE PROXIMITY DISTRIBUTION. DISCIPLINES IN WHICH AT TIME *t*,
THE COUNTRY c IS NOT SPECIALIZED

Panel A Panel B

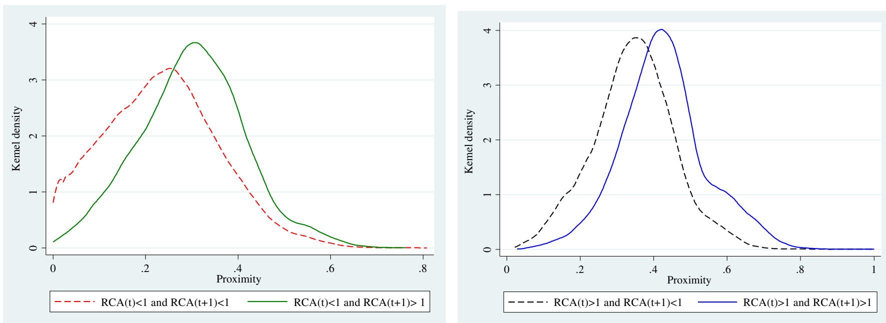

Figure 1. The distribution of the *AvgProximity* variable. The graph plots the Kernel density function of the *Avgproximity*$_{j,t}$ variable. In Panel (A), the solid green line plots the *Avgproximity* of those disciplines in which a country c is not specialized, but it does that in the next period. The red dashed line plots the *AvgProximity* of those disciplines in which a country c is specialized neither at time *t* nor at the next period. In Panel (B), the solid blue line plots the *Avgproximity* of those disciplines in which country *c* is specialized at time *t* and maintains it in the next period. The black dashed line plots e *AvgProximity* of those disciplines in which country *c* is specialized at time *t,* but stopped it in the next period.

Panel (B) also describes the distribution of the *AvgProximity* variable, but this time for disciplines that initially (at time *t*) country *c* is already specialized (RCA≥1). We are also able to distinguish two groups: those that maintained RCA index greater than or equal to one in the next period, $t + 1$ (solid blue line), and those that reduced the specialization index to less than one (black dashed line). The figures show that *AvgProximity* takes larger values when disciplines maintain their specialization than when disciplines stop having an RCA index greater than unity. This evidence is consistent with the idea that knowledge or other research resources might spill over from scientific domains where a country is already specialized towards those nearby disciplines.

III.b. Baseline results

We begin with baseline estimates from (3), asking whether the proximity between a given discipline *j* and the set of disciplines in which country *c* is specialized (RCA ≥1) will affect the rate at which the country's publications grow. Thus, the critical parameter is the coefficient of *AvgProximity*, coefficient α_1.

TABLE 3

AVERAGE PROXIMITY BETWEEN DISCIPLINES AND PUBLICATIONS GROWTH RATES

VARIABLES	PUBLICATIONS					
	At t:RCA < 1	At t:RCA≥ 1	At t:RCA < 1	At t:RCA≥ 1	At t:RCA < 1	At t:RCA≥ 1
AvgProximity	0.166	0.793***	0.288**	0.771***	1.355***	1.388***
	(0.105)	(0.081)	(0.136)	(0.098)	(0.153)	(0.097)
L.Publications			–0.187***	–0.121***	–0.065***	–0.000
			(0.008)	(0.008)	(0.011)	(0.013)
AvgProximity x L.Publications					–0.392***	–0.320***
					(0.030)	(0.023)
Constant	–0.060*	–0.416***	0.174***	–0.169***	–0.141***	–0.353***
	(0.035)	(0.032)	(0.041)	(0.036)	(0.045)	(0.039)
Marginal effects						
AvgProximity					0.383***	0.458***
					(0.130)	(0.083)
L.Publications					–0.199***	–0.130***
					(0.008)	(0.008)
Observations	88,320	83,614	88,320	83,614	88,320	83,614
R-squared	0.035	0.038	0.154	0.085	0.172	0.107
Number of Country-Discipline units	27,486	28,800	27,486	28,800	27,486	28,800
Fixed Effects	YES	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES	YES

Notes: Cluster robust standard errors in parentheses. This table presents the estimate for the average proximity effect on average publications annual growth in periods 2000-1996, 2004-2000, 2008-2004, 2012-2008, 2016-2012, and 2019-2016. Regressions in the first, third, and fifth columns consider country-discipline observations such that at the beginning of the period, the RCA (specialization or activity) index is less than 1, while the second, fourth, and sixth columns consider observations such that at the beginning of the period, the index is at least equal to 1. Each regression includes country-discipline fixed effects and year dummies. Regressions control for Publications at the country-discipline-year level at the beginning of the period. *** Significant at the 1 percent level. **Significant at the 5 percent level. *Significant at the 10 percent level.

Table 3 reports the estimate of the coefficients. The dependent variable is *Publication Growth* in discipline *j* in country *c* at time *t+1*. As indicated earlier, all specifications include time and country-discipline fixed effects.

From the table, we observe that the impact of the proximity between discipline *j* and the set of disciplines in which country *c* is specialized (the *AvgProximity* variable) has different impacts depending on whether the matter discipline reveals not to have (Columns 1 and 3) or to have (Columns 2 and 4) an RCA (or specialization or activity) index greater than 1.

The first four columns of Table 3 indicate that the *Avg Proximity* variable positively impacts the growth rate of the discipline's publications. Nonetheless, this result hides differences in the magnitude according to the level of publication a discipline has.

Columns 5 and 6 in Table 3 report these heterogeneous effects in detail. Accounting for the marginal effect of *AvgProximity*, if a country is not specialized in discipline *j*, an increase in one standard deviation in *AvgProximity* will push up growth in publications by 3.76 percentage points.[12] Moreover, when the country is already specialized in discipline *j*, the publication growth rate increases by 5.62 percentage points.

The above estimates confirm the intuition that the publication growth rate in scientific domains depends on the average proximity that each discipline has to other scientific domains in which a country is currently specialized.

We now illustrate our results using the estimation reported in Table 3. We project the publication growth rate of each discipline. Figure 2 plots publication growth projected for the following years against 2019's average proximity for a small set of countries chosen only to illuminate the intuition of our results.

For China and South Korea, disciplines in the Physical Sciences are closer to other domains in which these countries are currently specialized and have the largest expected growth rates. In contrast, for Uruguay and Ghana, disciplines in the Physical Sciences, which appear closer to the origin of the graphs, pointing to more isolation of disciplines (small average proximity), have lower expected publication growth rates.

From the pictures, we can conclude that the more developed a nearby scientific space (larger average proximity), the higher the publication growth rate of a specific discipline.

The estimations in Table 3 and the above examples suggest that greater dynamism will be observed in scientific disciplines that exhibit higher average proximity to scientific space that has already been developed.

III.c. Robustness

We now run the following exercises for robustness,[13] where we use an alternative measure of average proximity and alternative measures of scientific production.

An alternative measure of average proximity. The average proximity variable captures the idea that knowledge and skills, among other variables, accumulated by a country in the set of disciplines that it is currently specialized in, can spill over to nearby disciplines and boost its growth rate.

It is worth noting that given how we compute the *AvgProximity* variable, it correlates with countries' attributes or factors that simultaneously produce pairs of scientific disciplines. A natural concern is whether all these factors are transferable

[12] This impact is computed as follows: the marginal effect of *AvgProximity* for publications when the activity index is less than unity (0.383) times one standard deviation of *AvgProximity* in such a case (0.098). The remaining impacts on the growth rates are calculated in the same way.

[13] We also run various other exercises (not reported) where we include additional controls, for instance Gross Domestic Product per capita, to capture differences in resources available. The results remain almost identical.

FIGURE 2

PROJECTION OF PUBLICATION GROWTH RATE

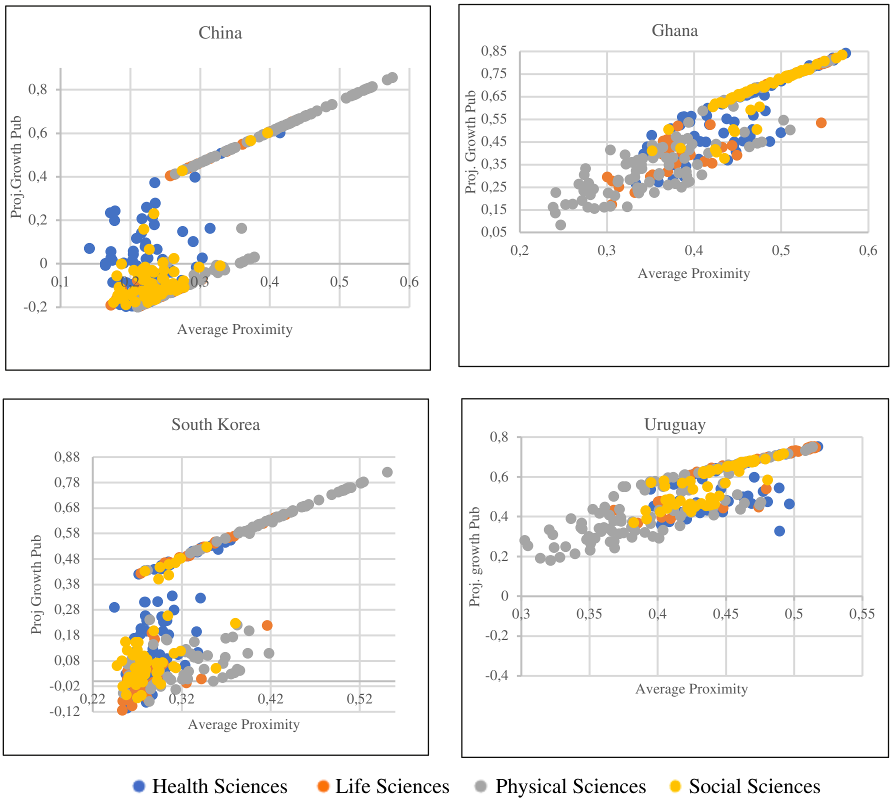

Figure 3. Projection Publication growth rates. The graph plots projected publication growth rates for the following years using estimates from Table 3. Countries are chosen only to illustrate the results.

in such a way that they contribute to producing science in nearby disciplines. This concern exists because some attributes might play only minor roles as determinants of scientific research output. Among these factors, we can distinguish (i) factors or attributes like language and natural resources that do not change over time; (ii) factors such as institutional frameworks or policies (related to education and support for science and innovation) that are somewhat stable over time; and (iii) factors like transferable skills, shared resources not exhausted in producing one of the disciplines, knowledge, technologies, and techniques transferable between areas. We claim that this last category is more likely to be directly linked to the production process in scientific research.

Although it is difficult to distinguish among the three categories above, we follow Thomson & Athukorala (2020) and explore the impact of factors that change over time to focus on the third set. We measure the change in the variable *AvgProximity* to isolate these factors, eliminating the impact of factors that do not vary over time. The underlying assumption in this estimation is that noise in the *AvgProximity* variable at time $t+1$ is not correlated with its error at time t, so it plays no systematic errors in the estimate. In the proximity matrix, given a large number of observations at each period (a matrix of 307 x 307) and a large number of countries in our sample (174), we can safely assume that distortions or shocks affect a single country or discipline can be smoothed out. Moreover, as we explained earlier, noise in the *AvgProximity* variable is uncorrelated with $\varepsilon_{j,c,t}$ in equation (3).

TABLE 4

CHANGE IN AVERAGE PROXIMITY AND PUBLICATION GROWTH RATES

VARIABLES	PUBLICATIONS					
	At t:RCA < 1	At t:RCA≥ 1	At t:RCA < 1	At t:RCA≥ 1	At t:RCA < 1	At t:RCA≥ 1
ΔAvgProximity	0.612***	0.824***	0.615***	0.768***	0.993***	1.692***
	(0.095)	(0.076)	(0.099)	(0.075)	(0.151)	(0.120)
L.Publications			–0.186***	–0.119***	–0.184***	–0.115***
			(0.008)	(0.008)	(0.008)	(0.008)
ΔAvgProximity x L.Publications					–0.153***	–0.332***
					(0.035)	(0.036)
Constant	–0.010	–0.129***	0.261***	0.112***	0.256***	0.107***
	(0.011)	(0.008)	(0.014)	(0.019)	(0.014)	(0.018)
Marginal effects						
ΔAvgProximity					0.614***	0.729***
					(0.099)	(0.065)
L.Publications					–0.186***	–0.117***
					(0.008)	(0.008)
Observations	88,320	83,614	88,320	83,614	88,320	83,614
R-squared	0.043	0.043	0.160	0.089	0.162	0.103
Number of Country-Discipline units	27,486	28,800	27,486	28,800	27,486	28,800
Fixed Effects	YES	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES	YES

Notes: Robust clustered standard errors in parentheses. This table presents the estimate for variation in the average proximity effect on average publications annual growth in periods 2000-1996, 2004-2000, 2008-2004, 2012-2008, 2016-2012, and 2019-2016. Regressions in the first, third, and fifth columns consider country-discipline observations such that at the beginning of the period, the RCA (specialization or activity) index is less than 1, while the second, fourth, and sixth columns consider observations such that at the beginning of the period, the index is at least equal to 1. Each regression includes country-discipline fixed effects and year dummies. Regressions control for Publications at the country-discipline-year level at the beginning of the period. *** Significant at the 1 percent level. **Significant at the 5 percent level. *Significant at the 10 percent level.

Table 4 reports our results. As can be seen, they are similar and confirm that *AvgProximity* plays a significant role in determining the publication growth rates.

Alternative measures of scientific performance. We compute the annualized growth for each discipline country using information regarding citations. Table 5 reports the impact of average proximity on this alternative measure of scientific production. We find that the main results of our baseline specification still hold despite some loss of precision.

TABLE 5

AVERAGE PROXIMITY BETWEEN DISCIPLINES AND GROWTH RATES ON CITATIONS

VARIABLES	CITATIONS					
	At t:RCA < 1	At t:RC≥ 1	At t:RCA < 1	At t:RCA≥ 1	At t:RCA < 1	At t:RCA≥ 1
AvgProximity	0.039 (0.142)	0.613*** (0.103)	0.119 (0.202)	0.623*** (0.140)	1.923*** (0.308)	1.937*** (0.230)
L.Citations			–0.288*** (0.006)	–0.227*** (0.006)	–0.191*** (0.013)	–0.134*** (0.013)
AvgProximity x L.Citations					–0.324*** (0.040)	–0.270*** (0.031)
Constant	0.120*** (0.045)	–0.303*** (0.040)	1.325*** (0.061)	0.847*** (0.053)	0.797*** (0.091)	0.441*** (0.082)
Marginal effects						
AvgProximity					0.226 (0.199)	0.422*** (0.127)
L.Citations					–0.301*** (0.007)	–0.242*** (0.006)
Observations	66,634	65,182	66,634	65,182	66,634	65,182
R-squared	0.018	0.032	0.378	0.275	0.386	0.286
Number of Country-Discipline units	23,656	25,283	23,656	25,283	23,656	25,283
Fixed Effects	YES	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES	YES

Notes: Cluster robust standard errors in parentheses. This table presents the estimate for the average proximity effect on average citation annual growth in periods 2000-1996, 2004-2000, 2008-2004, 2012-2008, 2016-2012, and 2019-2016. Regressions in the first, third and fifth columns consider country-discipline observations such that at the beginning of the period, the RCA (specialization or activity) index is less than 1, while the second, fourth, and sixth columns consider observations such that at the beginning of the period, the index is at least equal to 1. Each regression includes country-discipline fixed effects and year dummies. Regressions control for Citations at the country-discipline-year level at the beginning of the period. *** Significant at the 1 percent level. **Significant at the 5 percent level. *Significant at the 10 percent level.

IV. DISCUSSION AND CONCLUSIONS

This paper has provided evidence of how proximity between scientific disciplines plays a crucial role in the process of scientific upgrades.

The underlying idea is that there is likely to be knowledge or research resource spillover between scientific disciplines and that such spillover is larger between close scientific disciplines. We have implemented a definition of closeness according to what the data reveals. Using the SCImago dataset for publications, empirical findings support the notion that the publication growth rate in scientific domains depends on each discipline's average proximity to those scientific domains where a country currently is specialized.

These average effects conceal a diverse range of impact across disciplines. In particular, greater proximity will have a larger effect on scientific domains that are not yet specialized.

The finding that the initial level of publication has a negative and statistically significant effect on the growth rate of scientific research output can be interpreted through the lens of convergence.

These results may have significant policy implications. Considering the existence of factors like skills, knowledge, technologies, and techniques transferable between scientific domains, public and private efforts to support bundled disciplines will have increasing returns due to positive externalities over proximate scientific domains.

REFERENCES

ADAMS, J. D. (1990). Fundamental stocks of knowledge and productivity growth. *Journal of Political Economy*, *98*(4), 673-702.

ADAMS, J. D., & GRILICHES, Z. (2000). Research productivity in a system of universities. In *The economics and econometrics of innovation* (pp. 105-140). Springer, Boston, MA.

BLOOM, N., JONES, C. I., VAN REENEN, J., & WEBB, M. (2020). Are ideas getting harder to find? *American Economic Review*, *110*(4), 1104-44.

BALASSA, B. (1965). Trade Liberalisation and Revealed Comparative Advantage. The Manchester School, 33, 99-123.

BONACCORSI, A., & DARAIO, C. (2003). A robust nonparametric approach to the analysis of scientific productivity. *Research evaluation*, *12*(1), 47-69.

CAI, J., ZHAO, H., & COYTE, P. C. (2018). The effect of intellectual property rights protection on the international competitiveness of the pharmaceutical manufacturing industry in China. *Engineering Economics*, *29*(1), 62-71.

CHUANG, Y. W., LEE, L. C., HUNG, W. C., & LIN, P. H. (2010). Forging into the innovation lead—A comparative analysis of scientific capacity. *International Journal of Innovation Management*, *14*(03), 511-529.

COLE, J. R., & COLE, S. (1974). Social stratification in science. *American Journal of Physics*, *42*(10), 923-924.

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

DIETER, M., & ENGLERT, H. (2007). Competitiveness in the global forest industry sector: an empirical study with special emphasis on Germany. *European Journal of Forest Research*, *126*(3), 401-412.

FLEMING, L., & SORENSON, O. (2004). Science as a map in technological search. *Strategic Management Journal*, *25*(8-9), 909-928.

FRAME, J. D. (1977). Mainstream research in Latin America and the Caribbean. *Interciencia*, *2*(3), 143-148.

GRILICHES, Z. (1998). R & D and productivity: the unfinished business. *Estudios de Economía*, *25*(2), 145-160.

GUERINI, M., & TENCA, F. (2018). The geography of technology-intensive start-ups and venture capital: European evidence. *Economia e Politica Industriale*, *45*(3), 361-386.

HALL, B., & VAN REENEN, J. (2000). How effective are fiscal incentives for R&D? A review of the evidence. *Research Policy*, *29*(4-5), 449-469.

HARZING, A. W. (2013). A preliminary test of Google Scholar as a source for citation data: a longitudinal study of Nobel prize winners. *Scientometrics*, *94*(3), 1057-1075.

HAUSMANN, R., & KLINGER, B. (2006). Structural transformation and patterns of comparative advantage in the product space.

HENDERSON, R., JAFFE, A. B., & TRAJTENBERG, M. (1998). Universities as a source of commercial technology: a detailed analysis of university patenting, 1965–1988. *Review of Economics and Statistics*, *80*(1), 119-127

HIDALGO, C. A., KLINGER, B., BARABÁSI, A. L., & HAUSMANN, R. (2007). The product space conditions the development of nations. *Science*, *317*(5837), 482-487.

IRWIN, D. A., & KLENOW, P. J. (1996). High-tech R&D subsidies Estimating the effects of Sematech. *Journal of International Economics*, *40*(3-4), 323-344.

JAFFE, A. B. (1989). Real effects of academic research. *The American Economic Review*, 957-970.

JAMBOR, A., & BABU, S. C. (2017). *Competitiveness of global agriculture: Policy lessons for food security: Synopsis*. Intl Food Policy Res Inst.

KLETTE, T. J., & JOHANSEN, F. (2000). Accumulation of R&D capital and dynamic firm performance: a not-so-fixed effect model. In *The Economics and Econometrics of Innovation* (pp. 367-397). Springer, Boston, MA.

LERNER J. THE GOVERNMENT AS VENTURE CAPITALIST: THE LONG-RUN IMPACT OF THE SBIR PROGRAM. JOURNAL OF BUSINESS. 1999;72(3):285–318.

LEVIN, S. G., & STEPHAN, P. E. (1991). Research productivity over the life cycle: Evidence for academic scientists. *The American Economic Review*, 114-132.

MANSFIELD, E. (1972). Contribution of R&D to economic growth in the United States. *Science*, *175*(4021), 477-486.

MANSFIELD, E. (1995). Academic research underlying industrial innovations: sources, characteristics, and financing. *The Review of Economics and Statistics*, 55-65.

PARTHA, D., & DAVID, P. A. (1994). Toward a new economics of science. *Research Policy*, *23*(5), 487-521.

PAYNE, A. A., & SIOW, A. (1999). *Does federal research funding increase university research output?*. Institute of Government and Public Affairs, University of Illinois.

ROSENBERG, N. (1990). Why firms do basic research (with their own money)? *Research Policy*, v. 19, n. 2.

ROSENBERG, N., & NELSON, R. R. (1994). American universities and technical advance in industry. *Research Policy*, *23*(3), 323-348.

RUIZ GÓMEZ, L. M., RODRÍGUEZ FERNÁNDEZ, L., & NAVIO-MARCO, J. (2018). Application of communication technologies (ICT) within the tourism industry in the European Union. *Tourism: An International Interdisciplinary Journal*, *66*(2), 239-245.

SOETE, L., & WYATT, S. (1983). The use of foreign patenting as an internationally comparable science and technology output indicator. *Scientometrics*, *5*(1), 31-54.

SORENSON, O., & FLEMING, L. (2004). Science and the diffusion of knowledge. *Research Policy*, *33*(10), 1615-1634.

STEPHAN, P. E. (1996). The economics of science. *Journal of Economic Literature*, *34*(3), 1199-1235.

THOMSON, R., & ATHUKORALA, P. C. (2020). Global production networks and the evolution of industrial capabilities: does production sharing warp the product space?. *Oxford Economic Papers*, *72*(3), 731-747.