Can scientific productivity impact the economic complexity of countries?

Henry Laverde-Rojas¹ · Juan C. Correa²

Received: 6 December 2018 / Published online: 10 May 2019
© The Author(s) 2019

Abstract

The so-called index of economic complexity, based on nations' exports, was initially proposed as an alternative to traditional macroeconomic metrics just as the scientific productivity of countries which has also been deemed as a better predictor of economic growth. Adequate scrutiny to the relationship between these two factors, however, remains little explored. This paper aims to examine the relationship between economic complexity and scientific production while identifying which areas of knowledge hold to this relationship best. By applying panel data techniques to a sample of 91 countries between 2003 and 2014, we found that scientific productivity in basic sciences and engineering has a significant positive effect on the economic complexity of countries. This relationship, however, only remains stable for high-income countries, where university-industry-government capabilities interact to stimulate and generate innovation and strategies for economic growth of firms.

Keywords Economic complexity · Scientific productivity · System GMM · Knowledge-based economy

Introduction

The accumulation of knowledge can be a catalyst that drives nations economic productivity. The relationship between them can emerge when workers' human capital increases due to the technical training that accompanies the adoption of new technology for creating new products and services. Consequently, enhancing workers' human capital also generate multiplier effects on innovation, competitiveness, economic development or economic growth (Solarin and Yen 2016; Hatemi-J et al. 2016; Inglesi-Lotz and Pouris 2013; Inglesi-Lotz et al. 2014; Zaccaria et al. 2016). Discussions about the role of scientific productivity in this virtuous circle are already present in the literature that exploits interdisciplinary approaches to understand macroeconomics (Jaffe et al. 2013). Several scholars suggest the use of different indicators to evaluate the impact of scientific research on nations economic

✉ Henry Laverde-Rojas
henry.laverder@konradlorenz.edu.co

¹ Escuela de Negocios, Fundación Universitaria Konrad Lorenz, Bogotá, Colombia

² Facultad de Psicología, Fundación Universitaria Konrad Lorenz, Bogotá, Colombia

performance: the number of published academic papers, the number of academic papers per capita, participation of the papers published by a country in relation to the world publication, or the number of citations that each paper receives (Kumar et al. 2016; Javed and Liu 2018). As a result, empirical evidence shows the adequacy of these indicators to describe scientific productivity and forecast nations economies (Pouris and Pouris 2009; Inglesi-Lotz et al. 2014).

Another non-monetary approach to understand nations production was proposed by Hidalgo and Hausmann (2009) who coined the so-called term “economic complexity index” (ECI). The development of this index relies on the premise that the economic complexity of a country is higher when its exports are more diverse and have fewer competitors (i.e., the authors used the concepts of diversity and ubiquity to differentiate nations’ economies on these two dimensions). This index was intended as an additional measure to traditional ones, such as GDP per capita (Tacchella et al. 2012; Sweet and Maggio 2015; Hausmann and Hidalgo 2011; Felipe et al. 2012). As a matter of fact, after the publication of ECI, new literature emerged to explore the impact of economic complexity on economic outcomes (Hartmann et al. 2017; Brito et al. 2018).

However, the relationship between economic complexity and scientific productivity remains little explored so far. Our goal here is to provide some preliminary analyses to shed lights in filling this gap. In the first place, we direct ourselves to examine the existence of this relationship. Secondly, we carry on to explore which areas of knowledge best explain this link. Finally, we evaluate the stability of this relationship by discriminating nations according to their levels of development. The organization of this paper is as follows: in section “[Methodology and data](#)”, we provide the methodological aspects of our analytic technique and the construction of the necessary data to implement this methodology; in section “[Results](#)”, we present the results; in section “[Discussion](#)” we discuss the implications of our findings and provide further insights for future research.

Methodology and data

In the first place, we performed bivariate analysis to compare the predictive power of the different areas of knowledge on the economic complexity of countries. To this end, we used the following three methods. We analyzed Spearman correlations between different scientific disciplines and economic complexity; then, we used a simple regression model by ordinary least square (OLS); and, with these results, we filtered out those areas that have a moderate to high correlation, have a positive regression coefficient, highly significant and a relatively high coefficient of determination (R^2). We used Clarke’s test for non-nested models (Clarke 2007), which like the F test, is the standard method for selecting models by testing the significance of differences in R^2 .

Then, we used panel data techniques to further examine the relationship between economic complexity and scientific productivity by exploiting the structure of the data, evaluating the unobserved heterogeneity and the endogeneity present in the variables. We express our specification as follows:

$$ECI_{i,t} = \alpha \times ECI_{i,t-1} + \beta \times SP_{i,t} + X_{i,t} \times \gamma + \mu_i + \delta_t + \epsilon_{i,t} \quad (1)$$

where $ECI_{i,t}$ is the economic complexity index for country i in period t ; $SP_{i,t}$ is scientific productivity; $X_{i,t}$ is a vector of control variables; μ_i are the country-specific effects that are not observed and δ_t includes the temporary effects that have an impact on countries. To deal with these temporal effects, we included a set of time dummies for all the regressions.

We employed two estimation techniques, Pooled OLS and System GMM, to account for alternative treatments of information regarding the type of Scientific Productivity. Despite the difficulties presented by Pooled OLS, we ran this model to contrast the robustness of its results with those of the system GMM estimator. The estimation of Eq. 1 through Pooled OLS faces three problems for its identification. The first problem is the specific characteristics of the countries, which could be correlated with other regressors, making it impossible to estimate the parameters consistently. Panel data models offer the possibility to avoid this problem by treating these individual characteristics as time-invariant and eliminating them through transformations.

The second problem is the endogeneity present in some variables. As usual, the instrumental variables approach can be used. The problem here is the difficulty of finding valid and robust instruments. However, a more general strategy to deal with these problems is to use the difference GMM estimator which was developed by Arellano and Bond (1991). This estimator transforms the variables to eliminate the fixed effects and, later, the endogenous variables are instrumented in levels with the lags of the variables. However, the difference GMM estimator suffers from sample bias when the number of years is small and the dependent variable shows a high degree of persistence (Alonso-Borrego and Arellano 1999). Bond et al. (2001) recommended the System GMM estimator developed by Arellano and Bover (1995) and Blundell and Bond (1998) to obtain more consistent estimates. System GMM estimator uses the lagged values in levels as instruments for the transformed variables in the equation of first differences, as Arellano and Bond (1991) does, but adds the lagged differences to instrumenting the levels of the endogenous variables. This improves the efficiency of the estimates and prevents sample bias. However, the gain of asymptotic efficiency comes with a cost. The number of instruments tends to increase exponentially with the number of periods. This proliferation of instruments can lead to different problems such as a matrix of large estimated variances, downward bias in the standard errors in the two-stage estimators, weakens the over-identification test and over-adjustment of the endogenous variables.

A third problem of estimating Eq. 1 through pooled OLS is that the delayed value of the dependent variable is correlated with the fixed effects on the error term. The independence of $ECI_{i,t-1}$ and the error is a necessary condition for the consistency of OLS. If these variables are correlated, then OLS inflates the estimated coefficient of $ECI_{i,t-1}$ by attributing predictive power to it that belongs to the countries' fixed effect. Besides, if the number of lags is small, then problems of endogeneity arise, and the influence of countries fixed effects could increase. The solution of this third problem is to instrument $ECI_{i,t-1}$ as any other similar endogenous variable and use system GMM as mentioned above (Roodman 2009b).

Following Roodman (2009a), we employed System GMM estimator to estimate the parameters of our models, and to prevent instrument proliferation we followed the rule that the number of instruments does not exceed the number of groups. The consistency of the estimations relies on the fulfillment of the conditions of orthogonality (i.e., the residuals are not serially intercorrelated and the regressors are exogenous). We verified these assumptions by using two tests: the Hansen's overidentification test, to check the validity of the instruments, and the test for serial correlation AR(2).

To obtain data of economic complexity, we relied on the index proposed by Hidalgo and Hausmann (2009) from MIT's Observatory of Economic Complexity (<https://atlas.media.mit.edu/en/>) (Simoes and Hidalgo 2011). As we indicated in the previous section, this indicator uses export data to characterize the products of the countries in terms of diversity and ubiquity. Diversity describes the national production through the breadth

of the products spectrum. Ubiquity takes into account the number of competitors of the underlying products. In this sense, the economic complexity of country is higher if the diversity of its products is higher than those of other countries and there are few competitors for these products.

We approximated the scientific productivity of countries through the number of publications in all the scientific disciplines for the period 2003–2014, using data from SCImago Journal and Country Rank (www.scimagojr.com). Initially, we used publications per capita (Jaffe et al. 2013) and the Revealed Comparative Advantages Index (RCA) (Guevara and Mendoza 2013; Guevara et al. 2016). The first is simply the logarithm of the ratio between the number of publications and the population of each country. Meanwhile, RCA is defined as:

$$RCA_{ij} = \frac{x_{ij}}{X_i} / \frac{x_{aj}}{X_a} \quad (2)$$

where x_{ij} are publications of the country i that publishes in the journal of area j ; X_i is the total of publications in country i ; x_{aj} is the total of publications of area j of all the countries analyzed a ; X_a is the total of publications from all countries a .

The control variables that we used to characterize the countries economic complexity are: (i) patent applications (residents) per capita (Pper) taken from World Bank indicators (<https://data.worldbank.org/indicator>); (ii) an index of Human Capital (HC), based on years of schooling and returns to education from the Penn World Table (PWT) (www.rug.nl/ggdc/productivity/pwt), version 9.0 (Feenstra et al. 2015); (iii) GDP per capita (GDPper) from PWT; (iv) an institutional indicator, Corruption Perception Index (CPI) taken from Transparency International (www.transparency.org). CPI ranges from 0 (very corrupt) to 100 (highly transparent countries); (v) population size from PWT. Excepting ECI and CPI, all variables are in logarithms.

Results

Table 1, shows the Spearman correlations between ECI and GDP per capita with the scientific productivity (SP) in each discipline for the period 2003–2014, as well as ECI and GDP per capita with the Revealed Comparative Advantage Index (RCA). The correlations varied ostensibly both in sign and magnitude. For example, we observed a positive correlation between ECI and the scientific productivity per capita in disciplines such as Agriculture and Biological Sciences or Energy, and a negative correlation between ECI and the RCA for these same disciplines. Furthermore, the magnitude of the correlations also differed substantially, SP always showed higher correlations with ECI and GDP per capita than those of RCA. Such variations show that RCA is more sensitive than SP to the annual variability of countries and disciplines, which might be due to the failure of the index as either a reliable cardinal or ordinal measure of a country's revealed comparative advantage (see further details in Yeats 1985).

By taking $\rho >= 0.6$ as a cut-off criterion, we found that biology, computer science, chemistry, engineering, and exact sciences were those that best correlate with the economic complexity of countries. Table 2 shows a bivariate OLS regression model, where the dependent variable is ECI and the independent one is each scientific discipline. The results confirm the divergences between SP and RCA both in sign and magnitude, as well as in their statistical significance. Given the exploratory nature of this work, we decided to take

Table 1 Spearman correlations of scientific productivity and economic complexity index and gross domestic product per capita

Area of knowledge	ECI		GDP per capita	
	SP	RCA	SP	RCA
Agricultural and biological sciences	0.5459	− 0.4825	0.6372	− 0.5389
Arts and humanities	0.5393	0.1656	0.6040	0.1722
Biochemistry, genetics and molecular Biology	0.6777	0.3652	0.6843	0.3028
Business, management and accounting	0.5091	0.0957	0.6225	0.2483
Chemical engineering	0.7035	0.2553	0.7017	0.3573
Chemistry	0.7503	0.2325	0.7168	0.2374
Computer science	0.6389	0.4340	0.6808	0.5277
Decision sciences	0.5719	0.2519	0.6622	0.3987
Dentistry	0.5969	− 0.0428	0.6141	− 0.0392
Earth & planetary sciences	0.5755	− 0.2546	0.6650	− 0.1318
Economics, econometrics & finance	0.5665	− 0.0141	0.6645	0.0091
Energy	0.5788	− 0.0916	0.6610	0.0114
Engineering	0.6964	0.4159	0.7037	0.4476
Environmental science	0.5651	− 0.4311	0.6533	− 0.4557
Health professions	0.5176	0.3256	0.6235	0.4241
Immunology & microbiology	0.6536	− 0.4767	0.6641	− 0.5881
Materials science	0.7359	0.3792	0.6882	0.3464
Mathematics	0.7003	0.3431	0.7185	0.4466
Medicine	0.6324	− 0.2964	0.6749	− 0.3721
Multidisciplinary	0.5495	− 0.1861	0.5877	− 0.0895
Neuroscience	0.6267	0.5364	0.6459	0.5279
Nursing	0.4916	− 0.1313	0.6008	− 0.0950
Pharmacology, toxicology & pharmaceuticals	0.7044	− 0.2092	0.6978	− 0.2505
Physics & astronomy	0.7494	0.3959	0.6983	0.3993
Psychology	0.4974	0.1643	0.6022	0.1712
Social sciences	0.4987	− 0.2425	0.6116	− 0.2557
Veterinary	0.5774	− 0.3698	0.6194	− 0.4228

Spearman correlation coefficients of comparisons between RCA of publications pertaining to a given area

as a reference an R^2 cut-off criterion equal to 0.1 for several reasons. First, we noted that a higher cut-off criterion in both SP and RCA (e.g., 0.2 or higher) would end up leading us concluding that few disciplines act as predictors of ECI. Second, the majority of disciplines showed coefficients of determination below 0.2 for RCA indexes, so we set up the threshold in 0.1 so as to preserve consistency with similar previous findings (Jaffe et al. 2013). Finally, we rely on the statements of Breiman and Friedman (1982) who suggested 0.1 as the minimum and reliable recommended cut-off criterion to identify meaningful predictors. Authors as Falk and Miller (1992) consider that an R^2 less than 0.1 provides very little information, so the relationships between variables have a very low predictive level. Thus, we found that chemistry, computer science, health sciences, engineering, biology and exact sciences proved to be the best predictors of ECI.

We selected the disciplines that best explain the economic complexity of countries by identifying their higher associations with ECI in both Spearman correlations and OLS

Table 2 Bivariate analysis of scientific productivity- OLS

Area of knowledge	SP	R^2	RCA	R^2
Agricultural & biological sciences	0.569***	0.313	– 0.491***	0.238
Arts & humanities	0.577***	0.314	0.173***	0.030
Biochemistry, genetics & molecular Biology	0.693***	0.472	0.374***	0.138
Business, management & accounting	0.532***	0.273	0.100*	0.010
Chemical engineering	0.725***	0.512	0.256***	0.065
Chemistry	0.762***	0.574	0.234***	0.054
Computer science	0.667***	0.429	0.441***	0.193
Decision sciences	0.592***	0.341	0.257***	0.065
Dentistry	0.605***	0.363	– 0.043**	0.002
Earth & planetary sciences	0.589***	0.341	– 0.259***	0.067
Economics, econometrics & finance	0.596***	0.341	– 0.012	0.000
Energy	0.621***	0.363	– 0.097***	0.010
Engineering	0.710***	0.496	0.416***	0.174
Environmental science	0.586***	0.334	– 0.435***	0.189
Health professions	0.532***	0.277	0.334***	0.110
Immunology & microbiology	0.663***	0.435	– 0.480***	0.230
Materials science	0.746***	0.551	0.380***	0.144
Mathematics	0.714***	0.502	0.346***	0.120
Medicine	0.647***	0.411	– 0.297***	0.088
Multidisciplinary	0.588***	0.326	– 0.186***	0.035
Neuroscience	0.639***	0.402	0.540***	0.291
Nursing	0.510***	0.253	– 0.132***	0.018
Pharmacology, toxicology & pharmaceuticals	0.717***	0.507	– 0.210***	0.044
Physics & astronomy	0.760***	0.571	0.397***	0.158
Psychology	0.517***	0.259	0.166***	0.028
Social sciences	0.533***	0.268	– 0.243***	0.060
Veterinary	0.586***	0.340	– 0.371***	0.138

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Beta Coefficients. Dependent variable: Economic Complexity Index (ECI). SP=Scientific Productivity per capita, RCA=Revealed Comparative Advantages

regressions: Biochemistry, genetics and molecular biology, computer science, engineering, environmental sciences, health professions, materials science, mathematics, neuroscience, physics and astronomy, and veterinary. Then, we used Clarke's test for non-nested models with the aim of determining which of the models is preferable to explain ECI. In Table 3, we found that the best models were those involving biochemistry, engineering, materials science, mathematics, and physics and astronomy (based on Clarke's test for a $p \leq 0.05$). In summary, the economic complexity of countries seems to be better explained by the scientific productivity of those areas related to basic and exact sciences and engineering.

However, one would expect that richer countries are likely to publish relatively more in specific disciplines compared to poorer ones. Figure 1 shows an analysis of correlations between ECI and scientific productivity for 2014, by discriminating the income level of countries. This figure clearly shows that poorest countries were less complex and made fewer research efforts in publications for basic and exact sciences and engineering. However, exceptional countries also exist. For example, Australia and

Table 3 Tests for non-nested model selection

Model 1 (m1)	Model 2 (m2)	Clarke's test <i>p</i> value	Model 1 (m1)	Model 2 (m2)	Clarke's test <i>p</i> value
BGM	CS	0.000 (m1 is preferred)	ES	HP	0.000 (m1 is preferred)
BGM	EN	0.076 (Neither model)	ES	MS	0.000 (m2 is preferred)
BGM	ES	0.000 (m1 is preferred)	ES	MTH	0.000 (m2 is preferred)
BGM	HP	0.000 (m1 is preferred)	ES	NEU	0.000 (m2 is preferred)
BGM	MS	0.000 (m2 is preferred)	ES	PA	0.000 (m2 is preferred)
BGM	MTH	0.000 (m1 is preferred)	ES	VET	0.000 (m2 is preferred)
BGM	NEU	0.000 (m1 is preferred)	HP	MS	0.000 (m2 is preferred)
BGM	PA	0.000 (m2 is preferred)	HP	MTH	0.000 (m2 is preferred)
BGM	VET	0.000 (m1 is preferred)	HP	NEU	0.000 (m2 is preferred)
CS	EN	0.000 (m2 is preferred)	HP	PA	0.000 (m2 is preferred)
CS	ES	0.000 (m1 is preferred)	HP	VET	0.000 (m2 is preferred)
CS	HP	0.000 (m1 is preferred)	MS	MTH	0.000 (m1 is preferred)
CS	MS	0.000 (m2 is preferred)	MS	NEU	0.000 (m1 is preferred)
CS	MTH	0.000 (m2 is preferred)	MS	PA	0.000 (m2 is preferred)
CS	NEU	0.000 (m2 is preferred)	MS	VET	0.000 (m1 is preferred)
CS	PA	0.000 (m2 is preferred)	MTH	NEU	0.000 (m1 is preferred)
CS	VET	0.000 (m1 is preferred)	MTH	PA	0.000 (m2 is preferred)
EN	ES	0.000 (m1 is preferred)	MTH	VET	0.000 (m1 is preferred)
EN	HP	0.000 (m1 is preferred)	NEU	PA	0.000 (m2 is preferred)
EN	MS	0.000 (m2 is preferred)	NEU	VET	0.000 (m1 is preferred)
EN	MTH	0.000 (m2 is preferred)	PA	VET	0.000 (m1 is preferred)
EN	NEU	0.000 (m1 is preferred)			
EN	PA	0.000 (m2 is preferred)			
EN	VET	0.000 (m1 is preferred)			

BGM biochemistry, Genetics and Molecular Biology; *CS* computer science; *EN* engineering; *ES* environmental science; *HP* health professions; *MS* materials science; *MTH* mathematics; *NEU* neuroscience; *PA* physics and Astronomy; *VET* veterinary

Japan behaved differently. Australia is highly productive in research outputs but has a relatively low economic complexity. Japan showed the opposite situation with one of the highest indexes of economic complexity and a relative low scientific production (within developed countries). Although Australia has a world-renowned higher education system, a large part of the products traded in this country come from the Asian market (mainly China). Japan has a highly industrialized productive system that allows it to show a relatively high economic complexity index.

To proceed with a thorough analysis, we performed a pooled OLS regression model that allowed us to control other phenomena. Table 4 shows different specifications alternating the disciplines selected in the bivariate analysis. Columns 1–6 include control variables such as human capital (HC), innovation (Pper), socio-economic (GDPper, POP), and institutional (CPI) conditions. In each model, the scientific productivity was positive and highly significant as a predictor of the economic complexity of countries, with an explained variance ranging from $[0.74 \leq R^2 \leq 0.77]$. Human capital was positive and highly significant to explain the economic complexity of nations. Likewise, richer countries were statistically more complex. Innovation, which we approximated through patents per capita, is positively

Fig. 1 **a** Scattergram of total scientific productivity and ECI, **b** scattergram of scientific productivity in biochemistry and ECI, **c** scattergram of scientific productivity in engineering and ECI, **d** scattergram of scientific productivity in material science and ECI, **e** scattergram of scientific productivity in mathematics and ECI, **f** scattergram of scientific productivity in physics and astronomy and ECI

and highly significant in each of these models. Besides, in some models corruption perception index proved to be non-significant for predicting economic complexity. Finally, population size seems to be an important factor to ECI.

Next, in Table 5 we tackled the problems of unobserved heterogeneity and endogeneity through a System GMM estimator. To avoid the proliferation of instruments (Roodman 2009a), we followed the rule of keeping the number of instruments less than or equal to the number of groups. For different models, the lag of ECI is taken as a predetermined variable, and HC, Pper and GDPper as endogenous variables. In addition, we used the Windmeijer's correction (Windmeijer 2005) for standard errors of statistical estimates. When we take the total scientific productivity of countries (SP) we see that it explained relatively well their economic complexity (positive sign, p -value < 0.10, Hansen test > 0.10 and $AR(2) > 0.10$). But, we wonder, which scientific disciplines explain better this relationship? Table 5 shows that Biochemistry, Genetics & Molecular Biology and Materials Science, failed to explain the economic complexity of countries. In addition, surprisingly, control variables ceased to be significant and in some cases presented the wrong sign.

Next, we analyzed if previous results remained unchanged by splitting the sample by countries levels of income and development. In Tables 6 and 7 we used the average per capita income sample as a cut-off criterion to classify countries into low- and high-income nations. Then, we ran a set of Pooled OLS regression models (see Table 6).

For the group of low-income countries, scientific productivity failed to explain ECI, by revealing both positive and negative significant and non-significant relationships. In these nations, as shown in Fig. 1, there was no clear relationship between these two variables.

Table 4 Economic complexity (ECI) and Scientific productivity – Pooled OLS regression models

Dependent variable: ECI						
	(1)	(2)	(3)	(4)	(5)	(6)
Ln(SP)	0.448*** (0.063)					
Ln(BGM)		0.472*** (0.064)				
Ln(EN)			0.547*** (0.052)			
Ln(MS)				0.918*** (0.058)		
Ln(MTH)					0.698*** (0.073)	
Ln(PA)						0.730*** (0.049)
Ln(Pper)	0.227*** (0.043)	0.218*** (0.043)	0.180*** (0.044)	0.129*** (0.043)	0.237*** (0.043)	0.147*** (0.042)
Ln(HC)	1.979*** (0.167)	1.977*** (0.170)	2.065*** (0.163)	1.972*** (0.160)	2.016*** (0.164)	1.786*** (0.164)
Ln(GDPper)	0.112*** (0.029)	0.132*** (0.029)	0.071** (0.029)	0.078*** (0.027)	0.099*** (0.029)	0.093*** (0.027)
CPI	0.003** (0.002)	0.003** (0.001)	0.003* (0.002)	0.002 (0.001)	0.004** (0.002)	0.002 (0.002)
Ln(POP)	0.111*** (0.013)	0.110*** (0.013)	0.110*** (0.013)	0.111*** (0.013)	0.113*** (0.013)	0.107*** (0.013)
Constant	−6.051*** (0.320)	−6.182*** (0.319)	−5.784*** (0.307)	−5.730*** (0.299)	−6.012*** (0.315)	−5.538*** (0.302)
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1047	1047	1047	1047	1047	1047
R2	0,7379	0,7376	0,7494	0,7671	0,7444	0,7647
F-statistic	185.74***	191.27***	196.28***	241.07***	194.34***	250.50***

* $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$. Robust errors between parentheses.

BGM biochemistry, genetics and molecular biology; *CS* computer science; *EN* engineering; *ES* environmental science; *HP* health professions; *MS* materials science; *MTH* mathematics; *NEU* neuroscience; *PA* physics and astronomy; *VET* veterinary

In contrast, for the high-income countries, these relationships remained positive and highly significant, regardless of the discipline. We confirmed these results in a dynamic analysis using system GMM (Table 7). These results showed how the level of income and development of nations associates with notable differences in their economic complexity. Poorest countries need a minimum threshold of scientific productivity to impact their economic complexity significantly. Rich nations, however, enjoy increases in the economic complexity of their production due to increases in scientific productivity of basic sciences and engineering.

Table 5 Economic complexity (ECI) and Scientific productivity – System GMM

Dependent variable ECI						
	(1)	(2)	(3)	(4)	(5)	(6)
Lag(ECI)	0.930*** (0.054)	0.925*** (0.063)	0.897*** (0.069)	0.909*** (0.068)	0.920*** (0.061)	0.907*** (0.064)
Ln(SP)	0.135* (0.077)					
Ln(BGM)		0.165 (0.094)				
Ln(EN)			0.198** (0.084)			
Ln(MS)				0.200 (0.131)		
Ln(MTH)					0.274* (0.140)	
Ln(PA)						0.157* (0.083)
Ln(Pper)	−0.074 (0.089)	−0.082 (0.101)	−0.122 (0.094)	−0.108 (0.113)	−0.100 (0.104)	−0.111 (0.101)
Ln(HC)	0.043 (0.452)	0.491 (0.503)	0.588 (0.511)	0.650 (0.553)	0.492 (0.532)	0.505 (0.497)
Ln(GDPper)	−0.077 (0.051)	−0.071 (0.045)	−0.111 (0.066)	−0.084 (0.057)	−0.085 (0.060)	−0.075 (0.054)
CPI	0.000 (0.001)	0.000 (0.001)	0.001 (0.001)	0.001 (0.001)	0.000 (0.001)	0.001 (0.001)
Ln(POP)	0.022 (0.019)	0.023 (0.021)	0.030 (0.021)	0.029 (0.023)	0.028 (0.023)	0.029 (0.023)
Instruments	85	85	85	85	85	85
Hansen test (p-value)	[0.131]	[0.097]	[0.141]	[0.094]	[0.098]	[0.087]
AR(1)	[0.011]	[0.011]	[0.011]	[0.011]	[0.011]	[0.011]
AR(2)	[0.464]	[0.469]	[0.442]	[0.468]	[0.442]	[0.474]
Observations	960	960	960	960	960	960
Groups	91	91	91	91	91	91

* $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$. Between parenthesis standard errors, we used two-step and used the Windmeijer correction (2005).

BGM biochemistry, genetics and molecular biology; *CS* computer science; *EN* engineering; *ES* environmental science; *HP* health professions; *MS* materials science; *MTH* mathematics; *NEU* neuroscience; *PA* physics and astronomy; *VET* veterinary

Discussion

An examination of the relationship between economic complexity (Hidalgo and Hausmann 2009) and scientific productivity (Jaffe et al. 2013) was the aim of this work. As science and wealth have always been connected (Jaffe 2009; Erfanian and Neto 2017), those analyses that are intended to unveil how this relationship works are not only a

Table 6 Economic complexity (ECI) and Scientific productivity—Pooled OLS by Income

Dependent variable: ECI

	Low-income countries					
	(1)	(2)	(3)	(4)	(5)	(6)
Ln(SP)	−0.178 (0.470)					
Ln(BGM)		−0.658 (0.777)				
Ln(EN)			0.486 (0.331)			
Ln(MS)				1.205*** (0.416)		
Ln(MTH)					0.495 (0.616)	
Ln(PA)						0.914*** (0.304)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Observations	375	375	375	375	375	1047
R ²	0.5558	0.5564	0.5576	0.5619	0.5562	0.7647
F-statistic	43.67***	42.16***	48.49***	54.43***	44.52***	250.50***
	High-income countries					
	(1)	(2)	(3)	(4)	(5)	(6)
Ln(SP)	0.749*** (0.072)					
Ln(BGM)		0.817*** (0.071)				
Ln(EN)			0.759*** (0.057)			
Ln(MS)				1.161*** (0.061)		
Ln(MTH)					1.023*** (0.083)	
Ln(PA)						1.010*** (0.057)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Observations	672	672	672	672	672	672
R ²	0.6051	0.6089	0.6283	0.6730	0.6185	0.6773
F-statistic	69.01***	78.09***	78.10***	96.87***	72.08***	98.76***

* $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$. Robust errors between parentheses.

BGM biochemistry, genetics and molecular biology; *CS* computer science; *EN* engineering; *ES* environmental science; *HP* health professions; *MS* materials science; *MTH* mathematics; *NEU* neuroscience; *PA* physics and astronomy; *VET* veterinary

Table 7 Economic complexity (ECI) and Scientific productivity—System GMM by Income

Dependent variable ECI						
	Low-income countries					
	(1)	(2)	(3)	(4)	(5)	(6)
Ln(SP)	1.153 (1.096)					
Ln(BGM)		1.909 (1.584)				
Ln(EN)			1.313 (1.121)			
Ln(MS)				1.397 (1.089)		
Ln(MTH)					1.651 (1.489)	
Ln(PA)						0.567 (1.139)
Instruments	43	43	43	43	43	43
Hansen test (p-value)	[0.122]	[0.139]	[0.126]	[0.214]	[0.150]	[0.254]
AR(1)	[0.132]	[0.121]	[0.126]	[0.141]	[0.140]	[0.138]
AR(2)	[0.271]	[0.283]	[0.327]	[0.324]	[0.299]	[0.312]
Observations	336	336	336	336	336	336
Groups	43	43	43	43	43	43
	High-income countries					
	(1)	(2)	(3)	(4)	(5)	(6)
Ln(SP)	0.294** (0.123)					
Ln(BGM)		0.236** (0.105)				
Ln(EN)			0.614*** (0.183)			
Ln(MS)				0.809*** (0.248)		
Ln(MTH)					0.779*** (0.232)	
Ln(PA)						0.575*** (0.180)
Instruments	63	63	63	63	63	63
Hansen test (p-value)	[0.251]	[0.360]	[0.129]	[0.214]	[0.132]	[0.143]
AR(1)	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
AR(2)	[0.356]	[0.308]	[0.780]	[0.549]	[0.503]	[0.384]
Observations	624	624	624	624	624	624
Groups	64	64	64	64	64	64

* $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$. Between parenthesis standard errors, we used two-step and used the Windmeijer correction (2005).

BGM biochemistry, genetics and molecular biology; *CS* computer science; *EN* engineering; *ES* environmental science; *HP* health professions; *MS* materials science; *MTH* mathematics; *NEU* neuroscience; *PA* physics and astronomy; *VET* veterinary

well-deserved intellectual motivation but a matter of policy-makers' interest (Inglesi-Lotz and Pouris 2013; Inglesi-Lotz et al. 2014; Babić et al. 2016; Kutlača et al. 2015; Wong and Fung 2017).

We believe that the relationship between economic complexity and scientific productivity is one of the core elements of the so-called “knowledge-based economy” that uses knowledge to create tangible and intangible values (Nguyen and Pham 2011). However, as this last term is missing in the approaches of economic complexity (Hidalgo and Hausmann 2009) and scientific productivity (Jaffe et al. 2013), few studies discuss its implications in the literature. Nguyen and Pham (2011), for example, have shown the critical role of scientific outputs (i.e., scientific productivity) in knowledge-based economies such as the ones of South East Asian (ASEAN) countries. In these nations, economic and social stability were regarded as potential factors that explain why some countries are more productive than others. By including corruption perception index as a variable that mirrors economic and social (in)stability (Correa and Jaffe 2015), in our analyses, we have shown that these factors are neither related to economic complexity nor the scientific productivity of countries.

What seems to be important for explaining economic complexity is the scientific disciplines where countries compete (Inglesi-Lotz and Pouris 2013). Our results showed that productivity in basic sciences and engineering are the ones with stronger correlations with economic complexity (Jaffe et al. 2013). Arguably, this fact might reflect the so-called “triple helix model” where university-industry-government capabilities interact. This interaction, in turn, stimulate and generate innovation and strategies for economic growth of firms that exploit this infrastructure to create synergistic economic effects (Chung and Park 2014). Here, an illustrative classification of research types is in order. On the one hand, scientific research products are those intended to generate new knowledge, especially the theoretical knowledge that allows the advance of scientific disciplines. We might say that these products are fundamentally produced and funded by universities. On the other hand, research and development (R&D) products aim to solve real-world problems based on well-known scientific principles. These R&D products are produced and funded by industries and government agencies [e.g., projects targeting technology transfer and commercialization purposes (Baek et al. 2018)]. One way of understanding the idea of a triple helix is as a form of general research policies that regulate how scientific and R&D products coexist. Their coexistence, however, not only rely on scientific requirements per se, but necessarily by social and economic factors, and needs and possibilities of each country (Vinkler 2018).

A well-known case that illustrates the idea above is the set of European post-communist countries whose research policy was to increase the number of their research products rather than enhancing the quality of them. According to Jurajda et al. (2017) such policy distracted their limited resources away from internationally more competitive research. Increasing the number of scientific papers in different disciplines is undoubtedly a necessary step for achieving the required diversification that leads nations to scientific competitiveness worldwide, but this diversification works with a large number of citations and a higher effort on R&D projects that require a higher percentage expense concerning countries GDP (Suarez 2014; Cimini et al. 2014; Jaffe 2011).

Our findings are linked with the ideas above when we try to understand the relationship between the economic complexity of nations and their scientific productivity. However, such connection proved to be clear and stable only for high-income countries while low-income economies did not exhibit this same pattern. Arguably, developing countries need to overcome several institutional hurdles (e.g., lack of public and private funding agencies,

underdeveloped scientific institutions, insufficient investment on research and development projects) before they can reach the benefits of the triple helix that exists in wealthier nations. Our results should be seen in perspective with those reported by Solarin and Yen (2016) who found that research outputs had a positive impact on economic growth, in both developing and developed countries. It is important to recall that economic growth does not measure the productive knowledge required to develop the capacity to make a larger variety of products of increasing complexity (Hausmann et al. 2012). The scarce scientific productivity of low-income countries can be deemed as a systematic indicator of weak or non-existent productive knowledge.

We believe that future research should focus on the following four interrelated aspects. Firstly, it is important to quantify the proximity of the industrial sector with basic sciences and engineering and compare its proximity to other areas of knowledge. As patents are more common in basic sciences and engineering but rare in disciplines such as economics, psychology, social sciences or liberal arts, it might be possible that ECI is more sensitive to this proximity. In contrast, the productivity of social sciences might impact ECI in a non-trivial manner through, for example, public policies. Examining the indirect effects of social sciences and other disciplines on ECI is a well-deserved venture. Second, it is worth diving deep on the relationship between economic complexity and scientific productivity in low-income economies, by emphasizing the impact of public and private investment on R&D projects. Third, our work should be replicated with a more extended temporal series. For example, by merging the series analyzed by Jaffe and colleagues between 1982 and 2013 (Jaffe et al. 2013), with a more recent series (between 2014 and 2018) further studies could test the robustness and reliability of our results. Finally, in order to deepen our findings, an analysis might be done on the control variables, such as the prevailing human capital of nations, as captured by modern techniques (Laverde-Rojas et al. 2019), or including another pool of institutional metrics such as public spending in R&D (González and Pazó 2008), economic freedom index, innovation index (Jaffe et al. 2013), among others.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

- Alonso-Borrego, C., & Arellano, M. (1999). Symmetrically normalized instrumental-variable estimation using panel data. *Journal of Business & Economic Statistics*, 17(1), 36–49.
- Arellano, M., & Bond, S. (1991). Some tests of specification for panel data: Monte carlo evidence and an application to employment equations. *The review of economic studies*, 58(2), 277–297.
- Arellano, M., & Bover, O. (1995). Another look at the instrumental variable estimation of error-components models. *Journal of econometrics*, 68(1), 29–51.
- Babić, D., Kutlača, Đ., Živković, L., Štrbac, D., & Semenčenko, D. (2016). Evaluation of the quality of scientific performance of the selected countries of Southeast Europe. *Scientometrics*, 106(1), 405–434.
- Baek, S., Hwang, S., & Park, Y. I. (2018). Determinants of technology transfer and commercialization in national research and development: Focusing on Korea railroad research projects. *Asian Journal of Innovation & Policy*, 7(3), 438–456.
- Blundell, R., & Bond, S. (1998). Initial conditions and moment restrictions in dynamic panel data models. *Journal of Econometrics*, 87(1), 115–143.
- Bond, S., Hoeffler, A., & Temple, J. (2001). GMM estimation of empirical growth models.

- Breiman, L., & Friedman, J. (1982). Estimating optimal correlations for multiple regression and correlation. Technical Reports. Stanford University Technical Reports Orion 010.
- Brito, S., Magud, M. N. E., & Sosa, M. S. (2018). *Real exchange rates, economic complexity, and investment*. Washington D.C.: International Monetary Fund.
- Chung, C. J., & Park, H. W. (2014). Mapping triple helix innovation in developing and transitional economies: Webometrics, scientometrics, and informetrics. *Scientometrics*, 99(1), 1–4.
- Cimini, G., Gabrielli, A., & Labini, F. S. (2014). The scientific competitiveness of nations. *PLoS ONE*, 9(12), e113470.
- Clarke, K. A. (2007). A simple distribution-free test for nonnested model selection. *Political Analysis*, 15(3), 347–363.
- Correa, J. C., & Jaffe, K. (2015). Corruption and wealth: Unveiling a national prosperity syndrome in Europe. *Journal of Economics and Development Studies*, 3(3), 43–59.
- Erfanian, E., & Neto, A. B. F. (2017). Scientific output: Labor or capital intensive? an analysis for selected countries. *Scientometrics*, 112(1), 461–482.
- Falk, R. F., & Miller, N. B. (1992). *A primer for soft modeling*. Ohio: University of Akron Press.
- Feenstra, R. C., Inklaar, R., & Timmer, M. P. (2015). The next generation of the penn world table. *American Economic Review*, 105(10), 3150–82.
- Felipe, J., Kumar, U., Abdon, A., & Bacate, M. (2012). Product complexity and economic development. *Structural Change and Economic Dynamics*, 23(1), 36–68.
- González, X., & Pazó, C. (2008). Do public subsidies stimulate private R&D spending? *Research Policy*, 37(3), 371–389.
- Guevara, M., & Mendoza, M. (2013). Revealing comparative advantages in the backbone of science. In *Proceedings of the 2013 workshop on computational scientometrics: theory & applications*, ACM, (pp. 31–36).
- Guevara, M. R., Hartmann, D., Aristarán, M., Mendoza, M., & Hidalgo, C. A. (2016). The research space: using career paths to predict the evolution of the research output of individuals, institutions, and nations. *Scientometrics*, 109(3), 1695–1709.
- Hartmann, D., Guevara, M. R., Jara-Figueroa, C., Aristarán, M., & Hidalgo, C. A. (2017). Linking economic complexity, institutions, and income inequality. *World Development*, 93, 75–93.
- Hatemi-J, A., Ajmi, A. N., El Montasser, G., Inglesi-Lotz, R., & Gupta, R. (2016). Research output and economic growth in g7 countries: New evidence from asymmetric panel causality testing. *Applied Economics*, 48(24), 2301–2308.
- Hausmann, R., & Hidalgo, C. A. (2011). The network structure of economic output. *Journal of Economic Growth*, 16(4), 309–342.
- Hausmann, R., Hidalgo, C. A., Bustos, S., Coscia, M., Chung, S., Jimenez, J., Simoes, A. & Yildirim, M. A. (2012). *The atlas of economic complexity: Mapping paths to prosperity*. Cambridge, MA: Center for International Development/MIT Media Lab.
- Hidalgo, C. A., & Hausmann, R. (2009). The building blocks of economic complexity. *Proceedings of the National Academy of Sciences*, 106(26), 10570–10575.
- Inglesi-Lotz, R., Balcilar, M., & Gupta, R. (2014). Time-varying causality between research output and economic growth in US. *Scientometrics*, 100(1), 203–216.
- Inglesi-Lotz, R., & Pouris, A. (2013). The influence of scientific research output of academics on economic growth in South Africa: An autoregressive distributed lag (ARDL) application. *Scientometrics*, 95(1), 129–139.
- Jaffe, K. (2009). *What is science? An interdisciplinary perspective*. New York: University Press of America.
- Jaffe, K. (2011). Do countries with lower self-citation rates produce higher impact papers? Or, does humility pay? *Interciencia*, 36(9), 694–698.
- Jaffe, K., Caicedo, M., Manzanares, M., Gil, M., Rios, A., Florez, A., et al. (2013). Productivity in physical and chemical science predicts the future economic growth of developing countries better than other popular indices. *PLoS ONE*, 8(6), e66239.
- Javed, S. A., & Liu, S. (2018). Predicting the research output/growth of selected countries: Application of even gm (1, 1) and ndgm models. *Scientometrics*, 115(1), 395–413.
- Jurajda, Š., Kozubek, S., Münich, D., & Škoda, S. (2017). Scientific publication performance in post-communist countries: Still lagging far behind. *Scientometrics*, 112(1), 315–328.
- Kumar, R. R., Stauvermann, P. J., & Patel, A. (2016). Exploring the link between research and economic growth: An empirical study of china and usa. *Quality & Quantity*, 50(3), 1073–1091.
- Kutlača, D., Babić, D., Živković, L., & Štrbac, D. (2015). Analysis of quantitative and qualitative indicators of see countries scientific output. *Scientometrics*, 102(1), 247–265.

- Laverde-Rojas, H., Correa, J. C., Jaffe, K., & Caicedo, M. I. (2019). Are average years of education losing predictive power for economic growth? An alternative measure through structural equations modeling. *PLoS ONE*, 14(3), e0213651.
- Nguyen, T. V., & Pham, L. T. (2011). Scientific output and its relationship to knowledge economy: An analysis of ASEAN countries. *Scientometrics*, 89(1), 107–117.
- Pouris, A., & Pouris, A. (2009). The state of science and technology in africa (2000–2004): A scientometric assessment. *Scientometrics*, 79(2), 297–309.
- Roodman, D. (2009a). A note on the theme of too many instruments. *Oxford Bulletin of Economics and Statistics*, 71(1), 135–158.
- Roodman, D. (2009b). How to do xtabond2: An introduction to difference and system gmm in stata. *The Stata Journal*, 9(1), 86–136.
- Simoes, A. J. G., & Hidalgo, C. A. (2011). The economic complexity observatory: An analytical tool for understanding the dynamics of economic development. AAAI Workshops, North America, Aug. 2011. Available at: <https://www.aaai.org/ocs/index.php/WS/AAAIW11/paper/view/3948/4325>. Accessed 1 May 2019.
- Solarin, S. A., & Yen, Y. Y. (2016). A global analysis of the impact of research output on economic growth. *Scientometrics*, 108(2), 855–874.
- Suarez, R. K. (2014). Precious papers from 'non-research-intensive' countries. *The Journal of Experimental Biology*, 217, 818–819.
- Sweet, C. M., & Maggio, D. S. E. (2015). Do stronger intellectual property rights increase innovation? *World Development*, 66, 665–677.
- Tacchella, A., Cristelli, M., Caldarelli, G., Gabrielli, A., & Pietronero, L. (2012). A new metrics for countries' fitness and products' complexity. *Scientific Reports*, 2, 723.
- Vinkler, P. (2018). Structure of the scientific research and science policy. *Scientometrics*, 114(2), 737–756.
- Windmeijer, F. (2005). A finite sample correction for the variance of linear efficient two-step gmm estimators. *Journal of Econometrics*, 126(1), 25–51.
- Wong, C. Y., & Fung, H. N. (2017). Science-technology-industry correlative indicators for policy targeting on emerging technologies: Exploring the core competencies and promising industries of aspirant economies. *Scientometrics*, 111(2), 841–867.
- Yeats, A. J. (1985). On the appropriate interpretation of the revealed comparative advantage index: Implications of a methodology based on industry sector analysis. *Weltwirtschaftliches Archiv*, 121(1), 61–73.
- Zaccaria, A., Cristelli, M., Kupers, R., Tacchella, A., & Pietronero, L. (2016). A case study for a new metrics for economic complexity: The netherlands. *Journal of Economic Interaction and Coordination*, 11(1), 151–169.