Research Policy 46 (2017) 265–279

Contents lists available at ScienceDirect

Research Policy

journal homepage: www.elsevier.com/locate/respol

Determinants of citation impact: A comparative analysis of the Global South versus the Global North

Hugo Confraria[a,*], Manuel Mira Godinho[b,c], Lili Wang[a]

[a] UNU-MERIT, Maastricht University, Boschstraat 24, 6211AX Maastricht, The Netherlands
[b] ISEG (Lisbon School of Economics and Management), Universidade de Lisboa,Rua do Quelhas 6, 1200-781 Lisboa, Portugal
[c] UECE (Research Unit on Complexity and Economics), Rua Miguel Lupi 20, 1249-078 Lisboa, Portugal

ARTICLE INFO

Article history:
Received 21 June 2016
Received in revised form
17 November 2016
Accepted 21 November 2016
Available online 1 December 2016

JEL classification:
O38
O39
O57

Keywords:
Science policy
Development
Global South
Bibliometrics
Scientific impact
Citation impact

ABSTRACT

The impact of the scientific output produced by different nations in different fields varies extensively. In this article, we apply bibliometric and econometric analysis to study how citation impact varies across countries. This paper differs from previous research in that a cross-section model is put forward to account for such variation. A special focus is given to the Global South, as countries in this group have been converging with the Global North recently. We find that previous citation impact, level of international collaboration and total publications in a specific scientific field are important determinants of citation impact among all nations. However, specialization in particular scientific fields seems significantly more important in the Global South than in the Global North. These findings imply that most lower- and middle-income countries would better concentrate their resources in generating higher critical masses in specific fields, in addition to pursuing long-lasting international collaboration partnerships, as these actions may lead to higher impact research.

1. Introduction

There is a widely held assumption that scientific research has positive effects on economic development, namely by increasing human capital, by driving productivity growth, or by providing evidence to inform policies and practice (DFID, 2014; Salter and Martin, 2001). However, the process by which this happens is complex, and there has been extensive debate about the extent to which development funders and governments in the Global South, or more generally in the peripheries, should invest in research.

A crucial aspect for analysing the scientific performance of countries is to understand whether their scientific output is having an international impact or influence. The impact of published articles can be regarded as being one crucial aspect of scientific quality, and is thus a "proxy" for quality, as follows from the bibliometrics literature (Moed, 2005). Studies that focus on measuring the scientific impact of countries usually use citation analysis, as this arguably enables international comparisons to be more objective (Garfield, 1979).

There are numerous studies in this field that assess research at the country level, however only a few try to understand what the determinants of citation impact are. This type of analysis can help to understand why some scientific systems are performing better than others. Overcoming this gap in the literature can be particularly helpful to provide relevant insights for science policy, for furthering the policy learning cycle and ultimately for increasing the accountability of public policies.

Using the *InCites*™ tool of Web of Science/Thomson Reuters (*WoS*™), this article applies bibliometric and econometric analysis to evaluate which countries in the world are producing research with higher research citation impact, and to account for those factors that lead to higher results. The ability to estimate the expected number of citations of countries, by taking country characteristics and other variables at the subject category level, can be helpful for policy-makers in low-income and middle-income countries (the

* Corresponding author at: UNU-MERIT, Maastricht University, Boschstraat 24, 6211AX Maastricht, The Netherlands.
E-mail addresses: confraria@merit.unu.edu, hugoconfraria@gmail.com (H. Confraria), mgodinho@iseg.ulisboa.pt (M. Mira Godinho), wang@merit.unu.edu (L. Wang).

http://dx.doi.org/10.1016/j.respol.2016.11.004
0048-7333/

Global South), where public funds for financing the research system are scarce.

Our main objectives are: first, to create a comprehensive framework that can be used in the interpretation of different countries' citation impact, particularly in the Global South; second, to contribute to citation theory by understanding how the citation impact indicators commonly used in high-income countries can be used in lower income contexts, and; third, to provide assistance to policy-makers by identifying those independent variables that significantly influence the citation impact of countries.

In what follows, we will first focus on the framework aspects of our analysis, then we will describe the data and methodology used, and afterwards we will discuss the results obtained. Finally, conclusions will be put forward.

2. Background

2.1. Science in the Global South

The North–South divide is generally considered based on its political and socio-economic dimensions. Commonly, definitions of the Global North include North America, Western Europe, and developed parts of East Asia, while the Global South is perceived as being made up of Africa, Latin America, and developing Asia, including the Middle East. In this study we define Global North and Global South in two ways: firstly, by using the World Bank definition of low & lower-middle-income countries versus upper-middle & high-income countries[1]; and secondly, by dividing the world between OECD countries[2] and non-OECD countries. This possible division of the world into Global South and Global North has been perceived to be not only in terms of wealth or human development, but also in terms of scientific development.

In this context, the understanding of the links between research investment and development has attracted an increasing attention. Although it has been recognized that there is no unique path to successful economic development which should be emulated by every country, scholars such as Bernardes and Albuquerque (2003), Fagerberg and Godinho (2004) and Lall (2000) have stated that, in recent decades, countries that have caught up rapidly have tend to invest in their higher education system and have developed indigenous research efforts. According to Mazzoleni and Nelson (2007), the research programmes that effectively contributed to catch-up did not operate within "ivory towers", but were rather oriented towards an actual, or potential user-community. These programmes were projected to help solve problems, and to advance technology, being applicable to a particular economic area.

There are several ways in which research carried out within national borders can help provide both effective and focused responses to domestic problems, namely by being an enabler for providing up-to-date and qualified training for the new generations of university graduates, and also by helping to attract qualified people to the country, whilst improving the quality of local advice to government and industry (Goldemberg, 1998). Investments in science can not only provide knowledge and skills for increasingly knowledge-intensive industries, but they also generate a "domestic base of good scientists, which can break into the international networks where new technologies are being hatched" (Nelson, 2005). These scientists can act as important conduits of frontier knowledge into the local academic research community (Barnard et al., 2012), which can potentially diffuse that knowledge to students, the economy, and the general public.

Hence, the "scientific culture" of nations (Godin and Gingras, 2000) has been recognised as being a relevant dimension which is achieved through countries investing in science. As stated in the latest *UNESCO Science Report,* "the critical thinking that comes with science education is vital to train the mind to understand the world in which we live, make choices, and solve problems. Science literacy supplies the basis for solutions to everyday problems, reducing the likelihood of misunderstandings by furthering a common understanding. It provides answers that are testable and reproducible and, thus, provides the basis for informed decision-making and effective impact assessments" (UNESCO, 2015).

These arguments reveal the importance of science for international development, although from an economic perspective, one has to take into account the opportunity costs arising from investing in research. Therefore, a necessary and integral part of science policy is to monitor and evaluate the various facets of the scientific enterprise. By measuring the different characteristics of the scientific systems, it is possible to create and manage policies for improving the scientific performance of countries.

2.2. Can the Global South use the same bibliometric indicators as those used in the North?

The use of bibliometric indicators for assessing the impact of scientific publications has been on the rise in recent years. The ability of the use of such indicators to lower costs and time of assessment, without being invasive, and to enlighten political choices by carrying out international comparisons, as well as their perceived objectivity, have all been some of the main forces behind its growing popularity (Moed, 2005). However, the bibliometric assessment of research performance is based on a central assumption: namely that scientists who have to communicate something important, do so by publishing their findings in international peer-reviewed journals. This choice unavoidably introduces a limited view of a complex reality (van Raan, 2004). For instance, regionally focused papers in the Global South (e.g. in Agricultural Sciences) may make particularly important contributions to the local economy, yet remain uncited, as researchers elsewhere are indifferent to those topics. Citation patterns can also differ for other reasons: there are considerable database coverage biases (Moed, 2005; Rafols et al., 2015); the research focus can be locally or more internationally oriented (van Raan, 2003); there is a language bias, as most journals in *WoS*™ are written in English (Leeuwen et al., 2001), and; finally, countries have different levels of access to some journals, due to their financial constraints, selectivity, or publication policies (Lawrence, 2003). This last limitation is particularly relevant in the Global South and may have acted in the past as a stimulus for researchers from those countries to seek publication through other channels, namely through other means that are not registered in *WoS*™, or in other similar databases. This problem was challenged recently by the Research4life[3] partnership, which intends to provide developing countries with easy access to peer-reviewed content. This initiative, which aims to reduce the "e-gap" between rich and poor countries, could contribute to a "normalization" of access to the international circuit in the future. Yet this is still a limitation that we have to keep in mind when interpreting our results.

At the same time, both *WoS*™ and other indexing systems have considerably enlarged the database's coverage of Latin American and Caribbean (LA-C) journals in recent years. According to

[1] See the list of countries here: http://data.worldbank.org/about/country-and-lending-groups.

[2] See the list of countries here: http://www.oecd.org/about/membersandpartners/list-oecd-member-countries.htm.

[3] http://www.research4life.org/.

Collazo-Reyes (2014), the number of LA-C indexed journals in *WoS*™ has increased from 69 to 248 titles in just a period of four years (2006–2009). This unprecedented growth is mainly related to a change in the editorial policy of *WoS*™.

For these reasons, and despite some recognized limitations, the use of bibliometric data and indicators has also been rising in the context of the Global South, where this type of analysis can be particularly relevant to understand successful processes of closing the S&T gap with the most advanced economies (Albuquerque, 2004).

2.3. Factors associated with higher levels of citation impact at the country level

In line with this framework, one way to assess scientific impact is by citation analysis. According to the seminal work of Merton (1973), when a scientist cites a given article, he or she indicates that the article was somehow relevant to their research. The citing author calls attention to some useful information included in an article, be it a method, a statistic, a result, or other[4] information, and then acknowledges intellectual or cognitive influence. Therefore, when a comparable article is cited more times than others, it is considered to have more international scientific influence or impact (Moed, 2005).

Numerous studies assessing research at the individual, institutional, and country level can be found in the literature. Many other studies create and discuss new methods and metrics for evaluating citation impact. However, few try to understand what the determinants of citation impact are. Table A1 in the Appendix A summarises some of the factors that are known to be associated with higher citation rates at the article, author, institutional, and country level. In our study, we focus on those factors that are known to be associated with higher citation impact at country level, namely: level of international collaboration (Glänzel et al., 1995; Katz and Hicks, 1997; Narin et al., 1991; Puuska et al., 2013; van Raan, 1998), Wealth intensity[5] (King, 2004) and having English as an official language (Leimu and Koricheva, 2005). These determinants centre on ad hoc considerations, and the literature has not, to the best of our knowledge, presented a comprehensive framework that could be used to interpret a country's citation impact, particularly in the Global South. By bringing together the main arguments in this literature, this study aims to fill such a gap in the literature.

In our analysis we will also include, as explanatory variables previous citation impact, logarithmic scientific output, the percentage of publications in collaboration with industry, and we will also control for population size. Our argument regarding previous citation performance is that there might be path dependency, or the "Matthew Effect" (Merton, 1968) in science. Research communities, whose work has been highly cited in the past, are more likely to receive citations in the future. Regarding scientific output, the rationale is that a higher scientific production, in the specific subject area, is a sign of higher critical mass and resources applied to the field that tend to foster quality and impact (Shibayama and Baba, 2015). This measure can also be used as a proxy for scientific specialization, as we are controlling for the total number of publications produced by a country. With regards to the percentage of publications in collaboration with industry, we intend to understand whether citation impact is higher when research performed by a country has a higher level of collaboration with industry (Perkmann et al., 2011). This indicator can thus be seen to be a measure of knowledge transfer between industry and academia, and therefore, if a country has a higher percentage of publications with at least one author from a corporation, then we assume that this country is performing more applied research.[6] At the same time, nations have obvious differences in size, and to control for this we add logarithmic population as an independent variable.

3. Methodology

3.1. Data

Publication data were extracted from the InCites™ (2014) platform provided by *Thomson Reuters*, which facilitates national comparisons across time periods. *InCites*™ provides output and citation metrics from *WoS*™, based on a dataset of more than 27 million papers from 1981 to 2014. The metrics for comparisons are created based on address criteria, using the whole-counting method, that is to say, counts are not weighted by number of authors, neither by addresses.

In this study, our main research question is to understand whether there are different determinants of citation impact between the Global South and the Global North across different subject areas. To solve this, we used two different specifications for South and North (GDPpc levels and being an OECD country or not), and we adopted the disciplinary breakdown of the *Essential Science Indicators* (ESI) areas.

The ESI scheme incorporates a selection of journals carried out by Thomson Reuters. Our dataset covers 21 of the 22 ESI categories with a time span of 5 years (2008–2012). The research fields retained are as follows: Agricultural Sciences, Biology & Biochemistry, Chemistry, Clinical Medicine, Computer Science, Economics & Business, Engineering, Environment/Ecology, Geosciences, Immunology, Materials Science, Mathematics, Microbiology, Molecular Biology & Genetics, Multidisciplinary, Neuroscience & Behaviour, Pharmacology & Toxicology, Physics, Plant & Animal Science, Psychiatry & Psychology, Social Sciences (general), and Space Science. The Multidisciplinary area was excluded, as the publications included in this category could not be unambiguously classified into any of the 21 disciplinary areas.

The option to choose the ESI scheme took into account the fact that there are several approaches to define a research field: on the basis of selected concepts (keywords), selected sets of journals, a database of field-specific publications, or any combination of these. The selection of a specific scheme[7] for the division of research fields needs to take into account the trade-off between robustness of results and specificity of the subject category. Bibliometric data are characterised by skewed distributions, and hence robust statistics require considerably large sample sizes. This favours the use of

[4] Authors also write self-citations, cite peers based on personal networks, use flattery (citations of editors and potential referees), and write "negative" citations (contradicting another author). However, it is reasonable to assume that most citations are "positive", that is to say, they are a sign of the fact that the citing author finds something useful in the material that they cite. Deviating citation patterns, such as negative citations, can affect an analysis of an individual article or author, however this adverse effect tends to disappear in an analysis of a larger aggregations of authors, such as departments, universities, or countries (Moed, 2005).

[5] Gross Domestic Product per capita.

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

[7] InCites™ provides six further schemes besides the 21 ESI, based on a conglomerated of journals indexed in the WoS™, e.g. the 251 WoS™ subject categories, or the 6 OECD categories.

fewer categories, with more observations per category. However, articles from different subject categories have different citation propensities. Therefore, the use of very broad categories (e.g. the 6 OECD categories scheme) can lead to differences in citation impact levels, which only reflect differences in the research portfolios of countries, as some countries are more specialized in fields within a given category which have a higher citation propensity.

We believe that the choice of the ESI scheme is the most adequate solution for solving this trade-off in this study. A common, although arbitrary, threshold is often a minimum of 50 full count publications for citation analysis. We use this threshold at a country/category level, and we only consider those countries that have at least 400 publications between 2008 and 2012. By applying these thresholds, we dropped the outliers that have extremely low numbers of publications, which occur particularly in countries/subject areas of the Global South. This is markedly the case for observations before 2003-2007, and therefore we are only able to use two periods in our analysis (2003-2007 and 2008-2012).

A common debate in bibliometric studies is the use of social sciences and humanities for analysis (e.g. Hicks et al., 2015; Marx and Bornmann, 2014). The usefulness of citation impact indicators depends on the extent to which the research outputs are covered in bibliometric databases, and this coverage varies by subject category. The coverage tends to be higher for natural sciences, which gives high priority to journal publications. In the case of social sciences and humanities, where the publication of books, book chapters, monographs, etc. is more traditional, the extent of coverage is reduced. The 21 ESI categories include three categories which are related to social sciences, namely: Economics & Business; Psychiatry & Psychology, and; Social Sciences (general), excluding Humanities. Although the exclusive use of *WoS*™ data might not be appropriate for the analysis of citation impact in the social sciences, we decided that coverage was sufficient enough to include these three categories in our broad, country-level, analyses.

3.2. Approach and metrics

It is well known that different subject areas have different output propensities, and that publications belonging to each field have singular characteristics. Therefore, to be able to explain the different citation performances among countries and subject areas, we compute a multivariate regression analysis (OLS) with fixed effects, at the subject area level.[8]

At the same time, ordinary regression assumes that all observations are independent. However, in our case, each country has 21 subject areas.[9] As these potential 21 observations share specific country characteristics, our observations are not independent of each other, which could potentially lead to a correlation of errors within countries, implying that the findings of statistical significance would be spurious. To tackle this, we had to relax the independence assumption, by clustering the errors at the country level (McCaffrey et al., 2012; Moulton, 1990).

When interpreting the results presented in this study, it should be borne in mind that indicators measuring citation impact capture the influence of journal articles in the scholarly communication system. As a consequence of the partial and one-dimensional nature of these impact indicators, it is recommended to use more than one single indicator in order to obtain more robust conclusions (Bornmann and Leydesdorff, 2013; Waltman, 2016). Consequently, for this study, our dependent variable will be measured by two different indicators: (1) the share of highly cited publications (PPtop10%), which shows the proportion of publications belonging to the top ten percent most cited documents in a given subject category, year and publication type, and; (2) the field normalized citation score (FNCS), which calculates the mean citation rate of a country's set of publications in a specific subject area, period of time, and document type, divided by the mean citation rate of all publications in that subject area/period/document type. Both these variables are normally distributed indexes, with some outliers on the right tale:

$$PPtop10\%(\%) = \frac{PPtop10\%(n)}{P} \tag{1}$$

$$FNCS = \frac{\sum_{i=1}^{P} c_i}{\sum_{i=1}^{P} \left[\mu_f\right]_i} \tag{2}$$

Currently, there are several ways to calculate citation impact indicators. From basic calculations such as: raw citation counts; citations per publication; the h-index; geometric means (Fairclough and Thelwall, 2015); or discretized lognormal and hooked power law distributions (Thelwall, 2016), to normalized methods controlled for research field, publication year, and document type as: the "crown indicator"; field normalized citation score (Waltman et al., 2011); percentile-based approaches (Pudovkin and Garfield, 2009), and; source normalized indicators (Waltman and Eck, 2012), amongst others. As publications belonging to different subject areas have different propensities for being cited (Bornmann et al., 2012; Peters and van Raan, 1994; van Raan, 2003; Waltman et al., 2011), we decided to use normalized indicators. The use of percentile ranking and field normalized citation score can avoid bias toward a large size of country or field. Both these indicators can be computed by using the consistent *InCites*™/*Thomson Reuters* databases, to which we had access.

In our model (3), I is a measure of citation impact in a certain period t, subject area s, and country c. LI is a lag dependent variable from the previous period, O is the number of articles and reviews in *WoS*™, IC is the percentage of publications of a country in international collaboration, and IND is the percentage of publications of a country in collaboration with industry. C is a set of country controls, including total output, gross domestic product per capita (GDPpc), population size, and English as an official language. Finally, α is the constant, and ε is the unobserved residual.

$$I_{cs} = \alpha + \mu LI_{cs} + \lambda O_{cs} + \eta IC_{cs} + \phi IND_{cs} + \beta C_c + \varepsilon_c \tag{3}$$

Some of our independent variables have an exponential distribution (GDPpc, population size and number of articles). We decided to apply logarithms in these cases. As for the multicollinearity problem, none of our independent variables is highly correlated with another one (>60%).

Variables such as R&D intensity, or numbers of researchers, were dismissed as typically these are highly correlated with GDPpc or the number of articles. Furthermore, for many countries, the availability or reliability of this type of data is dubious. In an earlier phase we also included in our model the variable "percentage of individuals using the internet" as a proxy for level of access to scientific journals, which is provided by the *International Telecommunications Union*. However, this indicator is also highly correlated with GDPpc. Finally, we also try to include an indicator of civil liberties, which is provided by the *Freedom House*, in order to account for country's freedom of expression and individual rights. Our argument is that a freer society is more creative, and therefore it is more prone to generate radical ideas. Nonetheless, our preliminary results showed

[8] According to McCaffrey et al. (2012), this method is designed for efficient computation in models with many fixed effects at one level, on the assumption that fixed effects are included as nuisance parameters to control for differences among units that could bias the estimates of interest (our coefficients).

[9] As we use the threshold of at least 50 publications per subject areas, not all countries have 21 observations. This may lead to selection bias, as in those countries that do not fulfil these threshold, the categories that are being computed are potentially those that the country performs better in. Thus this hypothesis needs more research in order to be fully understood.

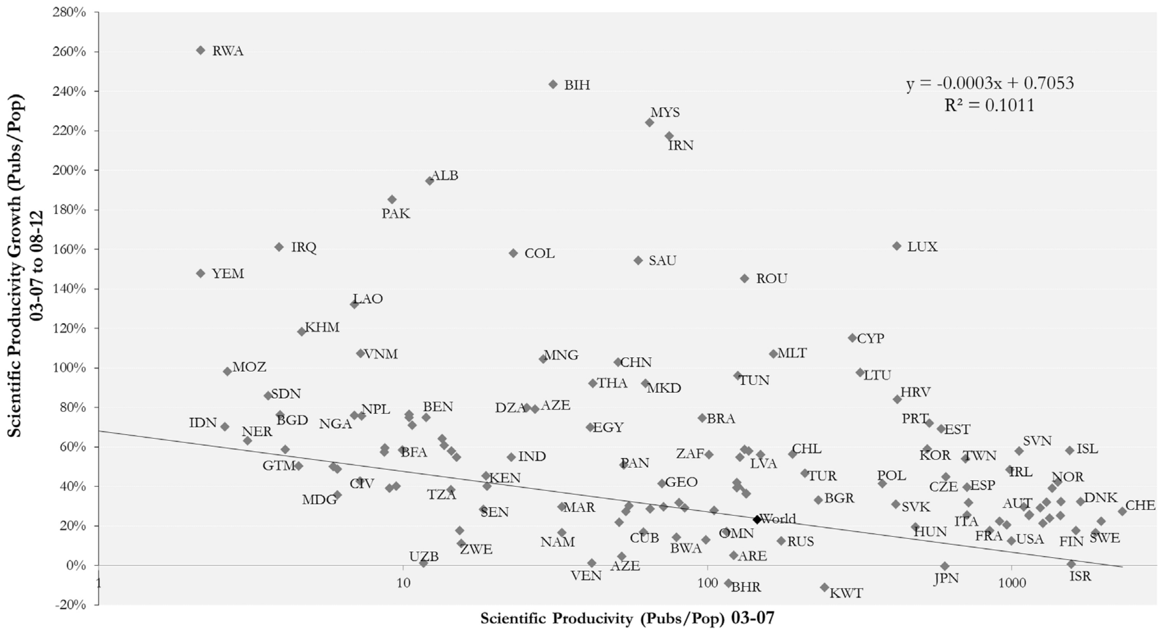

Fig. 1. Scientific productivity (publications per million inhabitants): growth rate versus previous situation.
Source: Own calculations based on *InCites*™. **Note:** In this chart Pubs stands for Publications and Pop for Population (million inhabitants). The vertical axis shows the growth rate of the Pubs/Pop ratio between 2003-2007 and 2008-2012. The horizontal axis shows the number of publications per million people (Pubs/Pop) in 2003-2007 (yearly average).

that this variable was insignificant in all model specifications, and therefore we decided to drop it from our model.

4. Results

4.1. Global trends

The world's long-term publication output in *WoS*™ has increased at an average rate of 3.5% since 1981. This growth rate has increased in the decade between 2004 and 2013 to an average of 5%. In 2013, the EU28 was still the world leader for publications (35%), followed by the US (27%), China (15%), and Japan (6%). Despite these impressive figures, the world shares of the EU28, US, and Japan all fell during the preceding decade. This decline was not due to the reduction of their scientific productivity (number of publications per population size), but rather was due to the higher growth rates of other rising players, such as China or Brazil.

In Fig. 1, by comparing the scientific productivity growth rates of 132 countries between 2003-2007 and 2008-2012 with their scientific productivity in 2003–2007, we observe a modest trend of convergence, denoted by the negative slope of the adjusted line.

The Chinese case is impressive. China's scientific publications have more than doubled over the past ten years. In contrast with the explosive growth of publications, the citation impact of that output has been perceived to still be at a relatively low level (Jin and Rousseau, 2004; Zhou and Leydesdorff, 2006). Using the data for the period between 1997 and 2001, Zhou and Leydesdorff (2006) argue that Chinese citation rates do not match the exponential growth of their scientific production. However, based on our more recent dataset, which refers to the 2008-2012 period, our analysis shows that the average Chinese citation impact is very close to the world average, and that China already is performing considerably better than the world average in some scientific areas, such as "Agricultural Sciences"; "Engineering"; "Mathematics"; "Plant & Animal Science", and; "Social Sciences". A similar finding has been highlighted previously by Wang (2016). This rapid growth reflects the coming of age of the Chinese research system, be it in terms of publications, number of researchers, or investment (UNESCO, 2010).

As for Brazil, its share of world scientific output has increased at a constant rate from 1993 to 2006, followed by a fast rise in 2007 and 2008 to the level shown by Brazil in 2013. Vargas et al. (2014) argue that, in areas such as Agricultural Sciences, Brazil's output increase since 2006 was mainly due to the expansion of Brazilian journals in *WoS*™, and also to an increase in the number of issues published by these journals. This phenomenon may have led to more publications, but fewer citations, as journals edited in Portuguese have less international visibility.

Iran presents another remarkable story. This country more than tripled its number of publications between the two periods analysed. According to Akhondzadeh (2013), "scientific progress in Iran over the past few years was the result of the country's recent policies and programmes to develop knowledge and facilitate researchers' access to the world's top academic resources".

In general, the world figures show a global converging trend in science regarding the quantity of publications from a significant number of world regions. This result may be inflated by changes in the size of the database, although we do not know to what extent this may be the case. *WoS*™ significantly increased between 2005 and 2010, in order to enlarge its regional coverage (Testa, 2011), and also in response to competition from *ScopusTM*, which entered the market in 2004. Despite these relatively recent expansions of *WoS*™ possibly being one reason for the convergence that has been noted in scientific publication worldwide, a similar convergence trend has also been observed for R&D investment by the public sector between the Global North and Global South (UNESCO, 2015). As Radosevic and Yoruk (2014) argue, these trends are possibly associated with a change in the scientific absorptive capacity of countries in the Global South. While 'absorptive capacity' has been generally defined as being "the ability to learn and implement knowledge" (Cohen and Levinthal, 1990), Radosevic and Yoruk (2014) defined absorptive capacity in the context of scientific research as being "the ability to recognize the value of new, external information, assimilate it, and apply it in another context". In accordance with this view, researchers recombine and re-contextualize existing

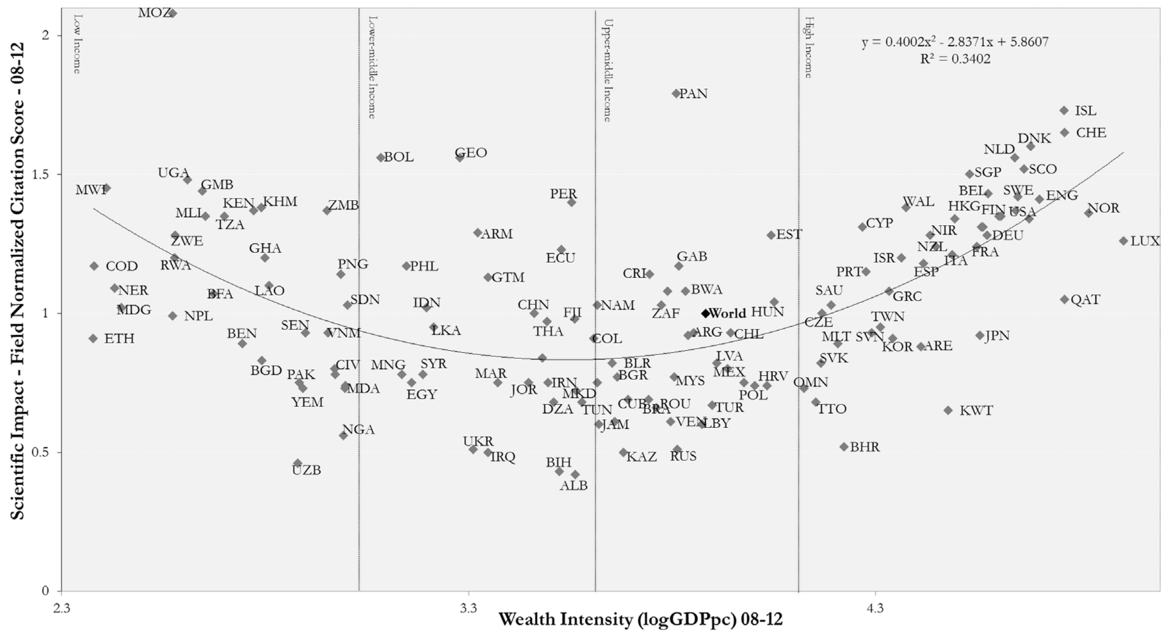

Fig. 2. Citation Impact (FNCS) versus GDPpc (2008-2012).
Source: Own calculations based on *InCites*™ & World Bank. Note: The vertical axis shows normalized citation impact in 2008–2012; the horizontal axis shows the logarithm of GDPpc (constant 2005 USD) in 2008-2012 (yearly average).

scientific knowledge and are able to generate novelty through their new publications.

There is still a huge gap to overcome between higher-income and lower-income nations, however the convergence which has been noted over the most recent years has certainly occurred due to the fact that some countries in the Global South are expanding their scientific capabilities, and are increasing their presence in scientific journals that have high international visibility. Such changing trends provide some support to our quest to understand the determinants of citation impact in the Global South, despite the fact that we analyse this by adopting indicators that are normally used to assess science in the Global North.

4.2. Wealth intensity versus research citation impact

In this study, we are particularly interested in understanding whether there are different determinants of citation impact (scientific influence) at different levels of GDPpc. In order to have a general overview of the relation between wealth intensity and citation impact, in Figs. 2 and 3 we scatter the relation between average GDPpc and citation impact, as measured respectively by FNCS and PPtop10% between 2008 and 2012, for countries that have more than 400 publications.

In both charts we can observe a U-shaped pattern, with the adjusted lines having their inflexion points close to the *World Bank* borderline which divides "low & lower-middle-income" countries from "upper-middle & higher-income" countries. For "low & lower-middle-income" countries, the citation impact performance seems to follow a downward trend, although with substantial deviations from the curve. For "upper-middle & higher-income" countries, there seems to be a positive relation between the two variables. Such an upward trend has already been revealed by King (2004). In contrast, the U-shaped pattern of our data seems to suggest that a nation's wealth only correlates positively with citation impact after a certain level of GDPpc.

One should be aware that the countries shown in these two charts have different size dimensions. If a small country, in terms of publication output, has a set of publications that is very influential, then the citations received by the articles produced by those researchers will improve its citation intensity score significantly. Mozambique, for example, is one of these cases. Despite its total production normalized by population being very low when compared to the world average (5 vs. 179 yearly publications per million people), its FNCS is two times higher than the world average, and its PPtop10% is close to 14%. In Mozambique, from 2008 to 2012, 95% of the country's publications have a foreign author. The high levels of citation impact in Mozambique may stem from the country having only a small group of scientists who produce scientific publications with highly reputed international co-authors (Confraria and Godinho, 2014).

Another outlier in our charts is Panama, and there is also an explanation for this case. Its citation impact (in intensity) is 79% higher than the world average for the FNCS indicator, and 18% of its publications are in the top ten most highly-cited papers. If we take a close look at the most highly cited publications from Panama between 2008 and 2012, we find that most of them come from researchers affiliated to the *Smithsonian Tropical Research Institution.* This organization is a bureau of the *Smithsonian Institution* based outside the United States, which is dedicated to understanding biological diversity. According to its website,[10] its "facilities provide a unique opportunity for long-term ecological studies in the tropics, and are used extensively by some 900 visiting scientists from academic and research institutions in the United States and around the world every year". In a country such as Panama, which had a scientific output close to 1500 publications during the five years analysed, the presence of this research institute certainly makes a difference. The *Smithsonian Institution* functions as a hub, attracting world leading scientists, and it certainly has a huge influence on the high citation impact of Panama.

On the right edge of the U curve, we find high-income countries such as Switzerland, Denmark, Iceland and The Netherlands. These are all relatively small European nations, which have been leading performers in this indicator for quite some time. For example, in

[10] http://www.stri.si.edu/english/about_stri/index.php.

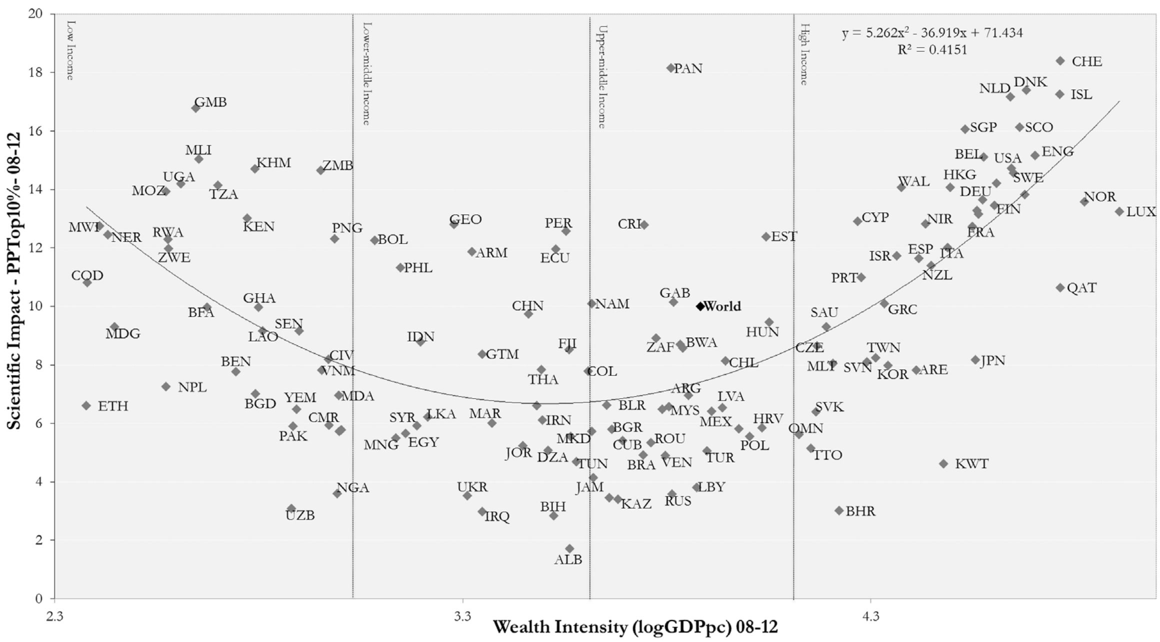

Fig. 3. Share of highly cited publications (PPtop10%) versus GDPpc (2008-2012).
Source: Own calculations based on *InCites*™ & World Bank. Note: The vertical axis shows the share of highly cited publications in 2008–2012; the horizontal axis shows the logarithm of GDPpc (constant 2005 USD) in 2008-2012 (yearly average).

one of the first studies analysing this issue, May (1997) also found that these countries were already leading the world in terms of "citation intensity".

In summary, our descriptive analysis suggests that higher levels of international collaboration may be extremely relevant for countries that may simultaneously have low GDPpc and relatively smaller scientific communities. It is also perceivable that despite middle-income countries may enjoy more resources and larger scientific communities, they do not engage in overseas collaboration so much, and this is reflected in the lower levels of impact on average. As we progress from the left to the right in our chart, the initial downward trend of the U-shaped pattern indicates that the improvement in GDPpc, from low-income to middle-income levels, leads to lower citation impact, on average. Finally, for high-income countries, both higher levels of GDPpc and country size, which is again negatively correlated with higher levels of international collaboration, seem to be critical factors.

4.3. Regression analysis

We used *Stata*™ (StataCorp, 2013) to compute the multilevel regression (OLS) with fixed effects at the subject area level, and errors clustered at the country level. The determinants of citation impact for publications between 2008 and 2012 were examined for 21 subject areas for countries with at least both 50 publications in a subject area and a total of 400 publications. After applying these restrictions, 126 countries and 1686 observations compose our global sample (see Table 1 for descriptive statistics).

Previously, in Figs. 2 and 3 we have seen that for different levels of GDPpc (below and above world average level), there are different patterns of citation impact. To understand whether these differences are substantive for the purpose of our analysis, we split our sample into two groups of countries, following the World Bank's definition of "low & lower middle income" countries (Global South), and "upper middle & higher income" countries (Global North). Furthermore, we also introduce another North-South distinction, namely being, or not being an OECD country. In Table A2, in the Appendix A, we provide descriptive statistics of these four groups. In both North-South specifications, citation impact, number of publications, and level of industry collaboration is significantly higher in the North. Level of international collaboration is substantially higher in the South.

Regressions were carried out for each of these groups separately. Generally, the results for both North and South specifications are robust. Table 2 reports the effect of the predictor variables on citation outcomes, using two dependent variables for the citation rates, which are respectively, PPtop10% and FNCS. The South samples include 54 "low & lower-middle-income" countries (490 observations), and 89 non-OECD countries (928 observations), while the North samples include 72 "upper middle & higher income" countries (1196 observations), and 37 OECD countries (758 observations).[11] Our model not only identifies those variables that are significant in predicting higher levels of citations, but it also identifies the relative contribution of each independent variable to the citation rates of countries.

These results show that, in both groups of countries, previous citation impact, level of international collaboration, and number of publications in the specific area are strongly associated with higher citation rates.

The first of these results indicates that, despite the fast growth of some countries in recent years, globally a strong path-dependency in citation impact still holds. There are specific reasons for this occurrence. It is well known in the literature that better-known scientists tend to receive more credit than less well-known ones, even if their research is similar (Merton, 1968). Frequently cited researchers generally have higher status than those researchers who are cited less frequently. Because status influences perceptions of quality, those with a high reputation are more likely to be cited in the future, which thus further reinforces their status. If we admit similar self-reinforcing mechanisms exist at a more aggregate level, then we can argue that nations performing better are also more likely to attract more tangible resources, such as research funding and outstanding graduate students, which can result in

[11] We count England, Scotland, Wales and Northern Ireland as separate nations.

Table 1
Descriptive Statistics.

Variable	N	Mean	SD	Min	Max	Correlation									
						1	2	3	4	5	6	7	8	9	10
1) PPtop10%_0812	1686	9.39	5.28	0.00	55.51	1.00									
2) PPtop10%_0307	1686	8.45	4.73	0.00	32.81	**0.78**	1.00								
3) FNCS_0812	1686	1.02	0.44	0.09	5.75	**0.90**	**0.69**	1.00							
4) FNCS_0307	1686	0.92	0.34	0.03	3.09	**0.78**	**0.93**	**0.71**	1.00						
5) Pubs Area (log)	1686	3.03	0.72	1.72	5.57	**0.24**	**0.29**	**0.19**	**0.29**	1.00					
6) Internat. Collab	1686	56.07	19.55	9.66	100	**0.34**	**0.23**	**0.36**	**0.25**	**−0.51**	1.00				
7) Industry. Collab	1686	1.75	2.50	0.00	23.38	**0.35**	**0.28**	**0.32**	**0.29**	**0.13**	**0.09**	1.00			
8) Total Pubs (log)	1686	4.40	0.77	2.63	6.24	**0.20**	**0.26**	**0.13**	**0.26**	**0.87**	**−0.52**	**0.17**	1.00		
9) English Official	1686	0.25	0.43	0.00	1.00	**0.20**	**0.17**	**0.18**	**0.19**	**0.09**	0.03	−0.02	**0.08**	1.00	
10) GDPpc (log)	1686	3.95	0.62	2.38	4.91	**0.38**	**0.36**	**0.30**	**0.38**	**0.47**	**−0.21**	**0.31**	**0.55**	−0.02	1.00
11) Popul. (log)	1686	7.27	0.67	5.50	9.13	**−0.20**	**−0.12**	**−0.20**	**−0.14**	**0.42**	**−0.27**	**−0.12**	**0.46**	**0.07**	**−0.39**

Note 1: Correlation with bold numbers significant at $p < 0.05$;
Note 2: The numbers 0307 and 0812 in the variables stand for the time periods 2003-2007 and 2008-2012.

Table 2
Determinants of citation impact in the Global South and the Global North.

	Dependent variables							
	PPtop10%_0812				FNCS_0812			
Independ. Variables	South (GDPpc)	North (GDPpc)	non-OECD	OECD	South (GDPpc)	North (GDPpc)	non-OECD	OECD
PPtop10%_0307	0.581***	0.583***	0.557***	0.692***				
	(0.061)	(0.046)	(0.048)	(0.053)				
FNCS_0307					0.643***	0.561***	0.576***	0.687***
					(0.107)	(0.067)	(0.076)	(0.086)
Pubs Area (log)	2.281***	1.189***	2.049***	0.944**	0.156**	0.015	0.128***	−0.013
	(0.572)	(0.335)	(0.404)	(0.402)	(0.062)	(0.035)	(0.040)	(0.043)
Int. Collab	0.082***	0.090***	0.091***	0.087***	0.007***	0.007***	0.008***	0.007***
	(0.013)	(0.011)	(0.010)	(0.010)	(0.002)	(0.001)	(0.001)	(0.001)
Ind. Collab	0.317***	0.186**	0.268***	0.188*	0.053**	0.012**	0.031***	0.010
	(0.106)	(0.090)	(0.095)	(0.107)	(0.023)	(0.006)	(0.011)	(0.006)
Total Pubs (log)	0.144	1.211***	0.557*	0.253	0.047	0.209***	0.104***	0.131
	(0.414)	(0.423)	(0.315)	(0.758)	(0.054)	(0.049)	(0.036)	(0.08)
English Official	1.267***	0.651**	0.929***	1.001***	0.059	0.043**	0.050*	0.066***
	(0.445)	(0.271)	(0.324)	(0.338)	(0.051)	(0.020)	(0.029)	(0.021)
GDPpc (log)	−0.171	0.634	−0.067	0.146	−0.097	−0.002	−0.075**	−0.034
	(0.592)	(0.435)	(0.352)	(0.650)	(0.063)	(0.036)	(0.033)	(0.053)
Popul. (log)	−1.053**	−1.723***	−1.292***	−0.820	−0.099**	−0.176***	−0.122***	−0.099**
	(0.410)	(0.345)	(0.336)	(0.537)	(0.041)	(0.030)	(0.031)	(0.048)
Constant	−0.277		−0.165		0.342		0.302	
	(2.652)		(2.853)		(0.220)		(0.239)	
Observs.	1686		1686		1686		1686	
R-squared	0.730		0.729		0.649		0.644	

Note 1: Robust standard errors in parentheses: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.
Note 2: Linear regression, absorbing indicators. Std. error adjusted for 126 clusters (countries).

research of better quality and also perpetuation of higher levels of citation impact. Our models do not suggest big differences between South and North in this regard, however they do show that previous performance is strongly associated with future citation impact; for example, for the PPtop10% indicator, those countries that have 1% more papers in the top 10% more-cited publications than others, have around 0.6% more papers in that same "excellent" tier during the next period.

As for the scientific output variable, those countries that produce more publications in specific subject areas also have higher citation rates per paper. This is intuitive, as, in theory, these are the subject areas in which countries have a higher scientific capacity. At the same time, given the effect of scale, researchers in the same subject areas probably cite their own compatriots more frequently, which thus increases the number of citations received by their country. This covariate represents not only the scientific output, but also the intensity of involvement in scientific activities of a country in a specific area (as gross expenditure on R&D, and number of researchers are usually highly-correlated with number of publications). As this effect is significantly higher in the Global South, one implication of this result is that the importance of generating a higher critical mass in a specific field, in order to produce research with more influence in the world, seems to be larger in the South. For instance, those countries in the South that have 50% more publications in a subject area than others, have, on average, 1% more papers in the top 10% most-cited publications in the world. In the North this relation is significantly smaller.

With regards to international collaboration, it is well known in the literature that citation impact is typically greater when research groups collaborate amongst each other, and the benefit strengthens when co-authorship is international (van Raan, 1998). The rationale behind this is that scientists are likely to develop new and alternative ways of thinking when they interact with other scientists with diverse areas of expertise and backgrounds (Hollingsworth, 2006). Co-publication allows access to a larger social network, which consequently leads to increased visibility, which in turn is reflected in higher citation rates (Goldfinch et al., 2003). This cross-fertilization is amplified by international collaboration, as scientists who produce co-authored papers with foreign scientists are more likely to belong to elite research groups within their own countries (Adams, 2013). As countries in the Global South depend a lot on international scientific networks to produce research that has visibility and impact (see Fig. A1 in Appendix A), we would expect that this positive relation to be higher in the countries from the Global South. Yet our results seem to show that the importance of international collaboration is not significantly different in both groups of countries. Specifically, those countries that have 10% more internationally co-authored publications in a subject area have, on average, 0.85% more publications in the top 10% in that subject area.

An interesting finding is that industry collaboration seems to matter for citation impact, especially in the Global South. This is because most co-publications with industry are co-authored by staff from the large R&D-intensive technology companies in science-based industrial sectors, such as biotechnology, pharmaceuticals, electronics, chemicals, and computers (Godin, 1996). This indicator can be seen as being a "knowledge linkage indicator" (Tijssen, 2012) between multinational R&D-intensive technology companies and public research organizations. This type of collaboration with industry is very likely to be driven by the need for access to international R&D networks, advanced research facilities, and contributions by scientists and research teams of international repute. Whereas from the industry side, researchers may be tempted to publish because they aspire to be active members of a research community and they want to be regarded as such by their peers, together with the other objectives of making corporate research findings public (Godin, 1996; Tijssen et al., 2009). The industry side may feel a particular appeal to collaborate with scientists from the South as a way of reaching specific resources, or of testing new medicines.

We found that this type of collaboration may be relevant for countries in the Global South, not only for updating their technological capabilities, but also for increasing their visibility and impact in the scientific community of their field. However, it is relevant to acknowledge that co-authorships with the industry are far from common in science, which thus represents a case of corner outcomes with an edge at zero, and a continuous distribution for strictly positive values (our sample as mean value of 1.73%). Our results show that, in line with Tijssen (2012), the intensity of science-industry co-authorship is lower in African and Latin American countries, than in countries in the North. Therefore, we should be cautious when interpreting this result, as the incidence of few publications in collaboration with industry can substantially change this indicator (high sensibility). Further to this, if we add a variable to our model that interacts industry collaboration intensity with international collaboration intensity, then the covariate industry collaboration changes signal and the significance disappear. At the same time, the international collaboration parameter remains positive and significant. Therefore, it is not clear whether the positive and significant effect of industry collaboration intensity in citation impact in our general model is due to the industry "effect", or whether it just occurs because most industry collaborations are also international collaborations.

For those countries that have English as an official language, our results show that the relation is positive and significant in almost every model specification. As the majority of scientific journals are written in English, and as articles published in a non-English language have less potential readers, this positive relation was an expected result. In the Global South, an Anglophone colonial history and concomitant opportunities for partnerships with English speaking countries (e.g., by hosting international research institutes) may have a significant effect on their citation impact.

These results also indicate that, contrary to what has been revealed by King (2004), the relation between GDPpc and citation impact is not strictly positive. One would expect wealthier countries to have more resources to apply to science, and that therefore they would perform better in terms of citation impact. However, for countries in the Global South, the coefficients are negative, and in the North they are positive and negative (non-significant), depending on the model specification. We also tried to understand whether the U relationship show in Figs. 2 and 3 holds in our model with all countries. However, when we include the variable GDPpc squared in the regression (see Table A3 in the Appendix A), the coefficient is positive, but non-significant. These results indicate that there are other elements beyond wealth intensity that matter for research quality in the South, namely: previous performance, a higher level of international collaboration, and more publications in the specific subject area.

Finally, for country size, in terms of total scientific output, there is no clear pattern, as our results differ depending on the model specification. However, countries with higher population seem to have, on average, less citation impact than smaller countries. A possible interpretation for this is that smaller countries are more involved in international collaborations to produce their scientific articles. This may be so, as when we interact country size with level of international collaboration, the negative effect of population size is no longer significant, and the interaction variable seems to capture this effect. For example, Frame and Carpenter (1979) also argued that the scientific size of a nation determines the need for international collaboration. Small countries have fewer opportunities to find collaborators inside their own country when compared to larger ones, and thus have a greater need for research partners from other countries (Narin et al., 1991). Our results do not show significant difference between the South and North.

To complement this analysis, in the Appendix A we carry out three different robustness checks. In Table A4, in order to explore the performance of countries with different levels of international collaboration, we create two sub-groups (i.e. low international collaboration intensity and high international collaboration intensity) in both Global South and Global North. In Table A5, we carry out the same analysis as in Table 2, but instead of separating the world into South and North, we use four broad world regions to see whether there are significant differences between them. In Table A6, we computed our model for all countries in our sample, using subject area groups, in order to check whether the results are consistent in most areas of knowledge. In general, these results are consistent with the previous models. In Table A4, we show that previous citation impact contributes more in the lower international collaboration group, whereas number of publications and level of collaboration with industry, all have a higher effect in the higher international collaboration group. In Table A5, the main findings are that previous citation impact is more relevant in Africa, and that Latin America & the Caribbean is the region where international collaboration has a higher effect on citation impact. Finally in Table A6, we found that both level of international collaboration and previous citation impact are positively and significantly associated with higher levels of citation impact in almost all areas. The same also occurs in eight areas for scientific output in the specific area, and in eleven areas for English as an official language.

5. Discussion & conclusion

In this article, bibliometric and econometric analysis were used to identify which countries are producing research with higher scientific influence, and also to understand which factors lead to these higher results. We focused particularly on the Global South, as the scientific output of some of these countries has been converging recently with that of the Global North.

We found some evidence suggesting that the determinants of citation impact may not coincide across countries in different wealth intensity levels. While previous citation impact, level of international collaboration, and publication output in a specific scientific field are all important determinants of citation impact among all nations, we observed that the variable number of publications in a specific disciplinary area appears to be substantially more important in the South than in the North. This covariate represents not only the scientific output, but also the intensity of involvement in the scientific activities of a country in a specific area. The agglomeration effects that may arise in some disciplines in scientific communities that are generally much smaller than their counterparts in the North seem therefore to be relevant. This implies that the importance of concentrating resources and of generating higher critical masses in specific fields, in order to produce research with more influence in the world, is seemingly greater in the South.

As for our lag dependent variable, we confirmed that it has an important effect on citation impact in both groups of countries. Societies vary in their capacity to produce major scientific discoveries over time. This happens because they are influenced in various ways by previous historical processes and institutional settings. This type of path dependency in scientific knowledge arises as a consequence of researchers in different types of organizations in the same country engaging in a great deal of common learning and socialization, which is transmitted across time and across organizations.

Eventually, these path-dependencies can be transformed into virtuous or vicious circles of development. Due to the fact that countries in the Global South have, on average, few "excellent" researchers with the know-how and tacit knowledge needed to engage in virtuous circles, one potential implication is that a "brain-drain" may have a severe negative effect on their scientific performance. If their few best "minds" leave to carry out research abroad and do not come back, or do not interact with their national colleagues, then the tacit knowledge will decline and the potential spillovers that they generate will stop being used for their countries' benefit.

With regards to the level of international collaboration, as has been widely shown by past research, a positive and significant relationship exists with citation impact. With the advances in information and communication technology and institutional changes, scientists can more easily obtain relevant knowledge by collaborating with other peers with diverse areas of expertise and backgrounds. Accessing external complementary knowledge and skills through networking, namely with scientists working in more developed environments, seems to be extremely relevant for performing research with high impact. However, interestingly, our results suggest that, contrary to what could be expected, this covariate does not seem more relevant in the South than in the North. This therefore indicates that the interest in pursuing international collaborations seems to be equally relevant in both environments.

Our analysis also suggests that industry collaboration seems to be positively associated with citation impact, especially in the Global South. However, it is not clear whether the positive and significant effect of industry collaboration intensity on citation impact in our general model is due to the "industry effect", or just occurs because the relatively few industry collaborations performed by the South also happen to be international collaborations.

In our regressions we also used country controls. We found that smaller countries (population wise) and countries with English as an official language perform on average better than others in some model specifications. A possible interpretation of this finding is that smaller countries rely more on international collaborations to produce their scientific articles. This was confirmed by assessing the interaction of country size with level of international collaboration. When such a possibility was tested, the negative effect of population size was no longer significant, with the interaction variable seemingly capturing this effect. For those countries that have English as an official language, as the majority of scientific journals are written in English, and as articles published in other languages have less potential readers, this positive relation was an expected result. Besides this, those countries that have English as an official language usually have a colonial legacy with Anglo-Saxon countries (US, UK, Canada, Australia), and consequently have more collaboration with them (Mêgnigbêto, 2013; Pouris, 2010). As these are the leading countries in many scientific fields, this positive relation is therefore reinforced. Finally, there is no clear relation between wealth intensity as measured by GDPpc and citation impact. It would be expected that wealthier countries would have more resources to apply to science, and therefore would perform better in terms of citation impact. However, we found that other elements beyond wealth intensity are much more relevant for the research quality of nations.

It is worth noting in relation to the groups of countries that we have assumed in this paper, that there could be other alternatives for their classification. The division suggested by us allocates countries to one of two groups, respectively Global South and Global North, or alternatively OECD and Non-OECD Countries. In doing this, we mainly took into consideration differences in economic wealth. However, as it has been shown by research on the same topic, countries across the globe may cluster into different groups depending also on geographical, political, ideological, cultural, or demographic lines. For example, Moya-Anegón and Herrero-Solana (2013) established a typology of three main groups of countries worldwide, according to the thematic characterization of scientific output in journals of international visibility. Their results show that each of these groups accounted for specific behavioural models, reflecting the distinctive characteristics of knowledge production in each country.

Furthermore, scientific performance worldwide may be influenced by the reputation of the affiliation, this being determined by the institution that the authors belong to (Peters and Ceci, 1982), or the country of the address of the submitted publications. Smith et al. (2014) investigated specifically whether the country where an author is based influences the notoriety of manuscripts. Their study found that, generally, international co-authorship enhanced scientific performance, but more specifically, they found that specific combinations of countries for the authorship of papers influence differently the performance of published papers, with the effect of these specific combinations also varying across disciplinary areas. This result suggests that the complexity of the factors determining scientific performance across countries may go well beyond the relationships stipulated by our econometric model.

Another open question has to do with the adequacy of bibliometric indicators in different socio-economic contexts. In this article we were aware that potential biases could arise from applying bibliometric indicators to countries belonging to the Global South. It is widely accepted that these types of indicators capture poorly certain types of research and that they encourage certain scientific activities and behaviours, including a shift towards English publications (Hicks et al., 2015), diversion of research away from local or national issues (Hicks et al., 2015), scientific supply poorly aligned with societal needs (Sarewitz and Pielke, 2007) and bias toward positive reporting (Fanelli, 2011), etc. As the Global South has a "lower" status in the scientific enterprise, these effects may be aggravated within this group. Another important issue when measuring citation impact is to be aware that it is a relative indicator. For example, if a country has the same citation impact (measured in intensity) as the US, but it has 1000 times less publications than the US, then evidently the actual absolute impact (scientific, societal and economical) of its research in the world is completely different. Therefore, even when using thresholds, as we did, indicators measuring citation impact should always be interpreted within their context, as we have done in the examples of Mozambique and Panama.

Another possible limitation stems from this study being mainly carried out in a macro perspective, based on bibliometric indicators.

This has certainly impaired our understanding of the specificities of the national scientific systems. The level of knowledge about science in lower income contexts would certainly improve by complementing this analysis with a more qualitative approach, such as researching why specific institutions in the Global South have such high performance levels, and by understanding their interactions. Furthermore, as the relational dimension (who do you collaborate with? What is the strength of the relationships?) seems to matter for citation impact (Gonzalez-Brambila et al., 2013), improving this model by using measures of network centrality, instead of level of international collaboration could give us a better understanding of the role of scientific network co-authorships for citation impact. Lastly, as each subject area has singular characteristics, we believe that our model can be expanded and adapted to each area, by including other independent variables that are specific to each field.

With regards to the normative implications, our findings allow us to draw some potentially relevant indications. Lower and middle-income countries with globally small scientific communities would better concentrate their resources in generating higher critical masses in specific fields, in order to produce research with a higher impact. Furthermore, the interest in pursuing international collaborations seems more than justified. International scientific collaborations have been pursued more intensely by smaller countries, which is comprehensible, given the fact that larger countries may have larger numbers of researchers in every single major discipline, and thus the need to collaborate abroad does not arise as much as in the former case. However, even for the larger countries, there may be good reasons for scientists to seek collaboration abroad, at least in some fields, thus balancing this orientation without jeopardizing the cohesion of their research systems.

These recommendations assume that increasing the impact of scientific publication in the South is an important objective, and that such impact is directly related to the quality of the research produced. However, it may be relevant to bear in mind the distinction between academic and practical impact. Although one may assume that in the long-term, both these impacts may coincide, wise policy-makers in the Global South may recognize these may well not coincide for shorter time-spans and in specific geographic or institutional conditions.

Finally, the science policy-making process needs to keep in mind the strong path-dependencies that dominate scientific activities globally. Despite the success stories of a few lower and middle-income countries that have forged ahead in scientific matters during the most recent decades, most countries in the Global South remain held back by the chains of path-dependency. Overcoming such path-dependencies implies persistence, continuous investment, and far-reaching institutional change, as these successful cases have confirmed.

Acknowledgements

The authors would like to thank Robin Cowan, Bart Verspagen, Johann Mouton and Pierre Mohnen for their advice and insights. This article also benefited from comments made by the participants of the workshop entitled “Science and development: Growth, expansion, and the role of the university system in South Africa” (Stellenbosch, 2016), the DRUID16 conference (Copenhagen, 2016), the 2016 Schumpeter Conference (Montreal, 2016) and the STI2016 conference (Valencia, 2016). This research did not receive any specific grant from funding agencies in the public, commercial or not-for-profit sectors.

Appendix A.

Table A1
Significant determinants of citation impact, based on previous studies (not exhaustive).

Level and Factors	What associates with higher citation	Prior literature
Article		
Number of authors	Four or more authors	Leimu and Koricheva (2005)
Length of the abstract	Longer abstract	Leimu and Koricheva (2005)
Journal impact factor (JIF)	Articles in journals with higher JIF	Peters and van Raan (1994); Didegah and Thelwall (2013)
Number of references	More references	Peters and van Raan (1994)
Impact of references	Higher no. of citations	Bornmann et al. (2012); Didegah and Thelwall (2013)
Length of the paper	Longer paper	Peters and van Raan (1994)
Type of document	Reviews	Peters and van Raan (1994)
Language	English journal and paper	Peters and van Raan (1994)
Author		
Country of origin	Native English-speaking authors	Leimu and Koricheva (2005)
Previous performance	More citations in the past	Merton (1968)
Institution		
Size	Universities with a large publication output	Moed et al. (2011)
Number of institutions	More institutions	Narin et al. (1991)
Specialization intensity	Weak negative effect	Moed et al. (2011)
Country		
Economic development	Higher GDP per capita	King (2004)
Number of countries of affiliation	More countries	Glänzel et al. (1995); Katz and Hicks (1997); Narin et al., 1991; Puuska et al. (2013); van Raan (1998)
Country of affiliation	English speaking country	Leimu and Koricheva (2005)

Source: Own elaboration.

Table A2
Descriptive statistics in the Global South and the Global North.

Variables	South (GDPpc)		North (GDPpc)	non-OECD	OECD
	Countries	54	72	89	37
	Obs.	490	1196	928	758
PPtop10%_0812	Mean	7.64	10.11	7.54	11.66
	Std. Dev.	5.22	5.13	5.02	4.67
PPtop10%_0307	Mean	7.02	9.03	6.78	10.49
	Std. Dev.	4.88	4.54	4.62	4.01
FNCS_0812	Mean	0.91	1.06	0.89	1.18
	Std. Dev.	0.52	0.39	0.46	0.35
FNCS_0307	Mean	0.81	0.76	0.79	1.07
	Std. Dev.	0.33	0.96	0.32	0.3
Pubs Area	Mean	2481	5580	2139	7790
	Std. Dev.	10419	16226	8036	19835
Int. Collab (%)	Mean	63.33	53.09	59.01	52.46
	Std. Dev.	22.95	17.11	21.7	15.82
Ind. Collab (%)	Mean	0.96	2.07	1.14	2.51
	Std. Dev.	1.52	2.74	1.81	2.97
Total Pubs	Mean	40023	116912	43185	163870
	Std. Dev.	143143	241811	109194	291501
English Official	Mean	0.27	0.24	0.25	0.25
	Std. Dev.	0.45	0.43	0.43	0.43
GDPpc	Mean	1814	24950	6949	32033
	Std. Dev.	1166	17143	8945	16305
Population	Mean	1.63E+07	3.26E+07	1.00E+08	3.44E+07
	Std. Dev.	3.48E+08	5.38E+07	2.64E+08	5.65E+07

Source: Own elaboration.

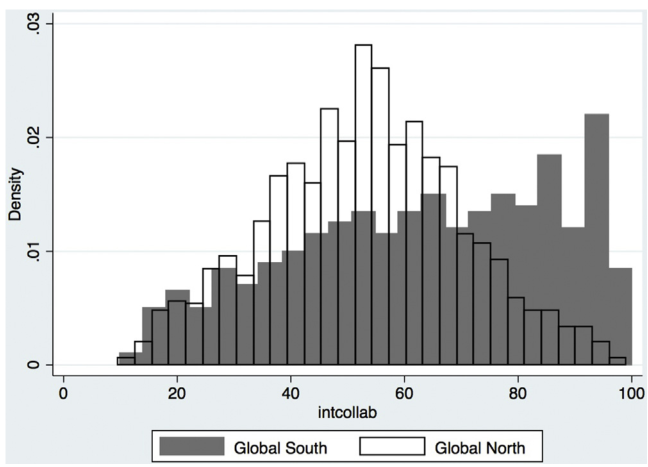

Fig. A1. Distribution of international collaboration levels. South vs North (2008-2012).
Source: Own calculations based on InCites[TM]. Note: Vertical axis shows the density of observations in a specific level of international collaboration in 2008-2012; Horizontal axis shows level of international collaboration.

Table A3
Determinants of citation impact in all countries.

Variables	PPtop10%_0812	FNCS_0812
PPtop10%_0307	0.590***	
	(0.040)	
FNCS_0307		0.606***
		(0.060)
Pubs area (log)	1.507***	0.066*
	(0.324)	(0.033)
Int. Collab	0.085***	0.007***
	(0.008)	(0.001)
Ind. Collab	0.214**	0.017***
	(0.084)	(0.006)
GDPpc (log)	−2.867	−0.350*
	(2.140)	(0.205)
GDPpc2̂ (log)	0.378	0.038
	(0.291)	(0.027)
English Official	0.910***	0.053**
	(0.254)	(0.021)
Total Pubs (log)	0.803**	0.134***
	(0.335)	(0.037)
Popul. (log)	−1.451***	−0.137***
	(0.330)	(0.029)

Table A3 (*Continued*)

Variables	PPtop10%_0812	FNCS_0812
Constant	6.764	0.992**
	(4.486)	(0.436)
Observations	1686	1686
R-squared	0.726	0.638

Note 1: Robust standard errors in parentheses: *** p < 0.01, ** p < 0.05, * p < 0.1.
Note 2: Linear regression, absorbing indicators. Std. error adjusted for 126 clusters (countries).

Table A4
Determinants of citation impact in the Global South and the Global North by international collaboration groups.

	PPtop10%_0812			
	South (GDPpc)		North (GDPpc)	
Variables	Int. Collab > = average (56%)	Int. Collab< average (56%)	Int. Collab > = average (56%)	Int. Collab< average (56%)
PPtop10%_0307				
	0.482***	0.526***	0.663***	0.688***
	(0.074)	(0.066)	(0.094)	(0.049)
Pubs Area (log)	2.461***	1.188*	0.322	0.495
	(0.924)	(0.622)	(0.432)	(0.391)
Int. Collab	0.125***	0.153***	0.056***	0.047***
	(0.023)	(0.024)	(0.017)	(0.012)
Ind. Collab	0.325**	0.249**	0.098	0.052
	(0.129)	(0.103)	(0.204)	(0.081)
Total Pubs (log)	0.031	2.339***	1.042*	0.641
	(0.667)	(0.730)	(0.568)	(0.528)
English Official	1.510**	0.758*	−0.327	0.674**
	−0.58	−0.387	−0.46	−0.307
GDPpc (log)	0.208	−0.307	−0.278	2.064***
	(0.744)	(0.696)	(0.854)	(0.658)
Popul. (log)	−1.504**	−2.292***	0.096	−0.956**
	(0.632)	(0.567)	(0.428)	(0.426)
Constant	−1.069		−5.708	
	(3.627)		(3.564)	
Observations	815		871	
R-squared	0.682		0.805	

Note 1: Robust standard errors in parentheses: *** p < 0.01, ** p < 0.05, * p < 0.1
Note 2: Linear regression, absorbing indicators. Std. error adjusted for 126 clusters (countries).

Table A5
Determinants of citation impact in four World regions (Africa; Asia; LA&C – Latin America & Caribbean; E&NA&P – Europe & North America & Pacific)

	Dependent variables							
	PPtop10%_0812				FNCS_0812			
Variables	Africa	LA&C	Asia	E&NA&P	Africa	LA&C	Asia	E&NA&P
PPtop10%_0307	0.690***	0.477***	0.532***	0.598***				
	(0.062)	(0.104)	(0.077)	(0.052)				
FNCS_0307					0.811***	0.354**	0.558***	0.602***
					(0.171)	(0.144)	(0.068)	(0.084)
Pubs Area (log)	1.139**	2.349***	1.888***	1.494***	0.098*	0.197***	0.074*	0.052
	(0.523)	(0.490)	(0.524)	(0.425)	(0.056)	(0.064)	(0.044)	(0.048)
Int. Collab	0.069***	0.122***	0.069***	0.106***	0.004**	0.013***	0.006***	0.009***
	(0.012)	(0.025)	(0.014)	(0.012)	(0.002)	(0.003)	(0.001)	(0.001)
Ind. Collab	0.353***	−0.097	0.109	0.252**	0.057	0.005	0.009	0.016**
	(0.111)	(0.081)	(0.078)	(0.100)	(0.036)	(0.006)	(0.006)	(0.007)
Total Pubs (log)	−0.610	−0.040	0.341	0.110	−0.062	0.034	0.122***	0.102
	(0.551)	(0.669)	(0.461)	(0.697)	(0.106)	(0.055)	(0.043)	(0.077)
English Official	0.478	0.340	1.193**	1.120***	0.033	0.085	0.049	0.080***
	(0.401)	(0.871)	(0.509)	(0.374)	(0.066)	(0.082)	(0.035)	(0.026)
GDPpc (log)	−0.351	−0.892	0.571	0.374	−0.092**	−0.263**	−0.023	−0.037
	(0.650)	(0.785)	(0.406)	(0.577)	(0.043)	(0.132)	(0.035)	(0.054)
Popul. (log)	0.028	−0.566	−0.996**	−0.925*	−0.007	−0.013	−0.104***	−0.107**
	(0.401)	(0.423)	(0.411)	(0.511)	(0.068)	(0.071)	(0.034)	(0.051)
Constant	−1.814				0.233			
	(3.051)				(0.274)			
Observations	1686				1686			
R-squared	0.740				0.661			

Note 1: Robust standard errors in parentheses: *** p < 0.01, ** p < 0.05, * p < 0.1
Note 2: Linear regression, absorbing indicators. Std. Err. adjusted for 126 clusters (countries).
Note 3: Africa = 221 observations (28 countries); LA&C = 195 observations (16 countries); Asia = 329 observations (34 countries); E&NA&P = 832 observations (48 countries).

Table A6
Determinants of citation impact in the World (2008-2012) by subject area using the percentile ranking (PPtop10%)

Independ. variables Subject Areas	Obs.	PPtop10%_0307	Pubs area (log)	Int. Collab	Ind. Collab	English Official	GDPpc (log)	Total Pubs (log)	Popul. (log)
Agricultural Sciences	91	0.548***	2.451**	0.112***	−0.098	1.273**	0.188	1.997*	−2.247***
		(0.093)	(1.029)	(0.018)	(0.178)	(0.613)	(1.138)	(1.086)	(0.831)
Biology & Biochemistry	79	0.755***	0.998	0.052***	0.043	0.218	1.503	−0.546	−0.365
		(0.109)	(2.040)	(0.016)	(0.200)	(0.489)	(1.038)	(2.255)	(0.747)
Chemistry	95	0.597***	0.888	0.014	0.258*	1.856***	2.114***	−0.515	0.133
		(0.116)	(1.252)	(0.016)	(0.150)	(0.658)	(0.635)	(1.418)	(0.557)
Clinical Medicine	115	0.512***	1.758	0.153***	0.598***	1.240**	−0.931	1.871*	−2.166***
		(0.069)	(1.116)	(0.023)	(0.184)	(0.544)	(0.788)	(1.071)	(0.686)
Computer Science	70	0.328***	3.083**	0.135***	−0.156	0.458	−1.678	1.728	−2.227
		(0.103)	(1.374)	(0.036)	(0.111)	(0.611)	(1.812)	(1.633)	(1.409)
Economics & Business	53	0.405***	6.723***	0.054**	0.650	−0.276	−0.267	−3.571**	−0.842
		(0.121)	(2.446)	(0.025)	(0.438)	(0.745)	(2.499)	(1.409)	(1.348)
Engineering	87	0.680***	1.968	0.039	−0.045	1.617**	3.255***	−3.574***	2.044**
		(0.171)	(1.593)	(0.028)	(0.112)	(0.684)	(1.130)	(1.172)	(1.027)
Environment/Ecology	92	0.548***	0.138	0.127***	0.445	2.095***	0.299	2.500	−1.529*
		(0.105)	(2.307)	(0.023)	(0.436)	(0.723)	(1.048)	(1.873)	(0.921)
Geosciences	87	0.429***	2.476**	0.165***	−0.155	1.222**	1.908	1.165	−0.866
		(0.081)	(1.242)	(0.024)	(0.112)	(0.577)	(1.293)	(1.273)	(0.948)
Immunology	85	0.021	2.715**	0.145***	0.052	1.994***	2.000**	0.474	−1.381
		(0.107)	(1.131)	(0.026)	(0.084)	(0.495)	(0.907)	(1.620)	(0.836)
Materials Science	76	0.769***	5.158***	0.072***	0.041	2.520***	0.341	−2.997	−1.104
		(0.087)	(1.772)	(0.027)	(0.094)	(0.758)	(1.317)	(1.959)	(0.897)
Mathematics	76	0.747***	0.349	0.027	−0.460	0.093	−1.892	1.664	−1.169
		(0.153)	(2.146)	(0.045)	(0.619)	(0.756)	(1.895)	(2.627)	(1.576)
Microbiology	71	0.545***	2.116	0.109***	0.276	2.578***	−1.552	3.024	−4.137***
		(0.121)	(1.699)	(0.027)	(0.241)	(0.778)	(1.649)	(2.094)	(1.322)
Molecular Biology & Genetics	66	0.521***	−2.434	0.081**	1.333***	0.844	0.461	4.430	−1.489
		(0.171)	(3.959)	(0.033)	(0.305)	(0.608)	(1.400)	(4.721)	(1.144)
Neuroscience & Behavior	66	0.625***	−0.796	0.098***	0.136	0.381	1.924	1.262	0.635
		(0.096)	(1.591)	(0.037)	(0.198)	(0.742)	(1.771)	(1.631)	(1.243)
Pharmacology & Toxicology	77	0.358***	−0.906	0.012	0.331***	1.826**	1.198	2.233**	−2.364***
		(0.093)	(1.095)	(0.020)	(0.074)	(0.711)	(1.093)	(1.075)	(0.835)
Physics	92	0.676***	3.228	0.156***	−0.030	0.242	−2.246	−0.145	−1.474
		(0.227)	(2.836)	(0.046)	(0.289)	(1.140)	(2.193)	(3.486)	(2.119)
Plant & Animal Science	103	0.599***	−0.009	0.078***	0.776**	0.759*	0.941	1.903*	−0.391
		(0.100)	(1.042)	(0.026)	(0.363)	(0.452)	(0.906)	(1.061)	(0.824)
Psychiatry/Psychology	53	0.554***	2.959*	0.064**	1.119***	1.120	−2.062	−1.071	−1.328
		(0.163)	(1.619)	(0.029)	(0.331)	(0.737)	(1.769)	(1.680)	(1.313)
Social Sciences, general	98	0.406***	2.558	0.157***	0.664	0.498	−0.765	0.976	−1.289
		(0.149)	(1.580)	(0.041)	(1.372)	(0.563)	(1.319)	(1.511)	(1.310)
Space Science	54	0.574***	3.767**	0.117**	0.568***	1.364	−7.324***	7.294***	−8.620***
		(0.114)	(1.505)	(0.053)	(0.212)	(0.994)	(2.423)	(2.261)	(1.907)
Constant		−0.830							
		(2.798)							
Observations		1686							
R-squared		0.800							

Note 1: Robust standard errors in parentheses: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.
Note 2: Linear regression. Std. error adjusted for 126 clusters (countries).

References

Adams, J., 2013. Collaborations: the fourth age of research. Nature 497 (7451), 557–560 http://doi.org/10.1038/497557a.

Akhondzadeh, S., 2013. Iranian science shows world's fastest growth: ranks 17th in science production in 2012. Avicenna J. Med. Biotechnol. 5 (3), 139.

Albuquerque, E. da M.e., 2004. In: Moed, H.F., Glänzel, W., Schmoch, U. (Eds.), Science and Technology Systems in Less Developed Countries. Springer, Netherlands, pp. 759–778.

Barnard, H., Cowan, R., Müller, M., 2012. Global excellence at the expense of local diffusion, or a bridge between two worlds? Research in science and technology in the developing world. Res. Policy 41 (4), 756–769 http://doi.org/10.1016/j.respol.2011.12.002.

Bernardes, A.T., Albuquerque, E. da M.e., 2003. Cross-over, thresholds, and interactions between science and technology: lessons for less-developed countries. Res. Policy 32 (5), 865–885 http://doi.org/10.1016/S0048-7333(02)00089-6.

Bornmann, L., Leydesdorff, L., 2013. Macro-Indicators of citation impacts of six prolific countries: inCites data and the statistical significance of trends. PLoS One 8 (2) http://doi.org/10.1371/journal.pone.0056768.

Bornmann, L., Schier, H., Marx, W., Daniel, H.-D., 2012. What factors determine citation counts of publications in chemistry besides their quality? J. Informetr. 6 (1), 11–18 http://doi.org/10.1016/j.joi.2011.08.004.

Cohen, W.M., Levinthal, D.A., 1990. Absorptive capacity: a new perspective on learning and innovation. Adm. Sci. Q. 35 (1), 128–152 http://doi.org/10.2307/2393553.

Collazo-Reyes, F., 2014. Growth of the number of indexed journals of Latin America and the Caribbean: the effect on the impact of each country. Scientometrics 98 (1), 197–209 http://doi.org/10.1007/s11192-013-1036-2.

Confraria, H., Godinho, M.M., 2014. The impact of African science: a bibliometric analysis. Scientometrics 102 (2), 1241–1268 http://doi.org/10.1007/s11192-014-1463-8.

DFID (2014). What is the Evidence on the Impact of Research on International Development? London.

Didegah, F., Thelwall, M., 2013. Determinants of research citation impact in nanoscience and nanotechnology. J. Am. Soc. Inf. Sci. Technol. 64 (5), 1055–1064 http://doi.org/10.1002/asi.22806.

Fagerberg, J., Godinho, M.M., 2004. Innovation and Catching-Up. Oxford University Press, Oxford, pp. 514–543 (Fagerberg).

Fairclough, R., Thelwall, M., 2015. More precise methods for national research citation impact comparisons. J. Informetr. 9 (4), 895–906 http://doi.org/10.1016/j.joi.2015.09.005.

Fanelli, D., 2011. Negative results are disappearing from most disciplines and countries. Scientometrics 90 (3), 891–904 http://doi.org/10.1007/s11192-011-0494-7.

Frame, J.D., Carpenter, M.P., 1979. International research collaboration. Soc. Stud. Sci. 9 (4), 481–497 http://doi.org/10.1177/030631277900900405.

Garfield, E., 1979. Citation Indexing: Its Theory and Application in Science, Technology, and Humanities. Isi Press.

Glänzel, W., Rinia, E.J., Brocken, M.G.M., 1995. A bibliometric study on highly-cited European physics papers in the 80 s based on physics briefs source publications and SCI citation data. Res. Eval. 5 (2), 113–122.

Godin, B., Gingras, Y., 2000. What is scientific and technological culture and how is it measured? A multidimensional model. Public Underst. Sci. 9 (1), 43–58 http://doi.org/10.1088/0963-6625/9/1/303.

Godin, B., 1996. Research and the practice of publication in industries. Res. Policy 25 (4), 587–606 http://doi.org/10.1016/0048-7333(95)00859-4.

Goldemberg, J., 1998. What is the role of science in developing countries? Science 279 (5354), 1140–1141 http://doi.org/10.1126/science.279.5354.1140.

Goldfinch, S., Dale, T., DeRouen Jr., K., 2003. Science from the periphery: collaboration, networks and periphery effects in the citation of New Zealand crown research institutes articles, 1995–2000. Scientometrics 57 (3), 321–337 http://doi.org/10.1023/A:1025048516769.

Gonzalez-Brambila, C.N., Veloso, F.M., Krackhardt, D., 2013. The impact of network embeddedness on research output. Res. Policy 42 (9), 1555–1567 http://doi.org/10.1016/j.respol.2013.07.008.

Hicks, D., Wouters, P., Waltman, L., de Rijcke, S., Rafols, I., 2015. Bibliometrics: the leiden manifesto for research metrics. Nature 520 (7548), 429–431 http://doi.org/10.1038/520429a.

Hollingsworth, J.R., 2006. The Dynamics of American Science: An Institutional and Organizational Perspective on Major Discoveries. Campus Verlag, Frankfurt, New York, pp. 361–380.

InCites^{TM} (2014). Thomson Reuters. This data is reproduced under a license from Thomson Reuters to the University of Lisbon. Retrieved from https://incites.thomsonreuters.com.

Jin, B., Rousseau, R., 2004. Evaluation of research performance and scientometric indicators in China. In: Handbook of Quantitative Science and Technology Research. Kluwer Academic Publishers, Dordrecht, pp. 497–514 http://doi.org/10.1007/1-4020-2755-9_23.

Katz, J.S., Hicks, D., 1997. How much is a collaboration worth? A calibrated bibliometric model. Scientometrics 40 (3), 541–554 http://doi.org/10.1007/BF02459299.

King, D.A., 2004. The scientific impact of nations. Nature 430 (6997), 311–316 http://doi.org/10.1038/430311a.

Lall, S., 2000. In: Kim, L., Nelson, R.R. (Eds.), Technologizing Change and Industrialization in the Asian Newly Industrializing Economies: Achievements and Challenges. Cambridge University Press, Cambridge, pp. 13–68.

Lawrence, P.A., 2003. The politics of publication. Nature 422 (6929), 259–261 http://doi.org/10.1038/422259a.

Leeuwen, T., Van Moed, H., Tijssen, R., 2001. Language biases in the coverage of the Science Citation Index and its consequencesfor international comparisons of national research performance. Scientometrics 51 (1), 335–346 http://doi.org/10.1023/A:1010549719484.

Leimu, R., Koricheva, J., 2005. What determines the citation frequency of ecological papers? Trends Ecol. Evol. 20 (1), 28–32 http://doi.org/10.1016/j.tree.2004.10.010.

Mêgnigbêto, E., 2013. International collaboration in scientific publishing: the case of West Africa (2001–2010). Scientometrics 96 (3), 761–783 http://doi.org/10.1007/s11192-013-0963-2.

Marx, W., Bornmann, L., 2014. On the causes of subject-specific citation rates in Web of Science. Scientometrics 102 (2), 1823–1827 http://doi.org/10.1007/s11192-014-1499-9.

May, R.M., 1997. The scientific wealth of nations. Science 275 (5301), 793–796.

Mazzoleni, R., Nelson, R.R., 2007. Public research institutions and economic catch-up. Res. Policy 36 (10), 1512–1528 http://doi.org/10.1016/j.respol.2007.06.007.

McCaffrey, D.F., Mihaly, K., Lockwood, J.R., Sass, T.R., 2012. A review of Stata commands for fixed-effects estimation in normal linear models. Stata J. 12, 406–432.

Merton, R.K., 1968. The Matthew effect in science. The reward and communication systems of science are considered. Science 159 (3810), 56–63 http://doi.org/10.1126/science.159.3810.56.

Merton, R.K., 1973. The Sociology of Science: Theoretical and Empirical Investigations. University of Chicago Press.

Moed, H.F., de Moya-Anegón, F., López-Illescas, C., Visser, M., 2011. Is concentration of university research associated with better research performance? J. Informetr. 5 (4), 649–658 http://doi.org/10.1016/j.joi.2011.06.003.

Moed, H.F., 2005. Citation Analysis in Research Evaluation, vol. 9. Springer-Verlag, Berlin/Heidelberg.

Moulton, B.R., 1990. An illustration of a pitfall in estimating the effects of aggregate variables on micro unit. Rev. Econ. Stat. 72 (2), 334–338.

Moya-Anegón, F., Herrero-Solana, V., 2013. Worldwide topology of the scientific subject profile: a macro approach in the country level. PLoS One 8 (12), e83222 http://doi.org/10.1371/journal.pone.0083222.

Narin, F., Stevens, K., Whitlow, E.S., 1991. Scientific co-operation in Europe and the citation of multinationally authored papers. Scientometrics 21 (3), 313–323 http://doi.org/10.1007/BF02093973.

Nelson, R.R., 2005. The roles of research in universities and public labs in economic catch-up. In: Santangelo, G. (Ed.), Technological Change and Economic Catch-Up. Edward Elgar Publishing.

Perkmann, M., King, Z., Pavelin, S., 2011. Engaging excellence? Effects of faculty quality on university engagement with industry. Res. Policy 40 (4), 539–552 http://doi.org/10.1016/j.respol.2011.01.007.

Peters, D.P., Ceci, S.J., 1982. Peer-review practices of psychological journals: the fate of published articles, submitted again. Behav. Brain Sci. 5 (2), 187 http://doi.org/10.1017/S0140525X00011183.

Peters, H.P.F., van Raan, A.F.J., 1994. On determinants of citation scores: a case study in chemical engineering. J. Am. Soc. Inf. Sci. 45 (1), 39–49 http://doi.org/10.1002/(SICI)1097-4571(199401)45:1<39:AID-ASI5>3.0.CO;2-Q.

Pouris, A., 2010. A scientometric assessment of the Southern Africa Development Community: science in the tip of Africa. Scientometrics 85 (1), 145–154 http://doi.org/10.1007/s11192-010-0260-2.

Pudovkin, A.I., Garfield, E., 2009. Percentile rank and author superiority indexes for evaluating individual journal articles and the author's overall citation performance. CollNet J. Scientometr. Inf. Manage. 3 (2), 3–10.

Puuska, H.-M., Muhonen, R., Leino, Y., 2013. International and domestic co-publishing and their citation impact in different disciplines. Scientometrics 98 (2), 823–839 http://doi.org/10.1007/s11192-013-1181-7.

Radosevic, S., Yoruk, E., 2014. Are there global shifts in the world science base? Analysing the catching up and falling behind of world regions. Scientometrics 101 (3), 1897–1924 http://doi.org/10.1007/s11192-014-1344-1.

Rafols, I., Ciarli, T., Chavarro, D., 2015. Under-reporting research relevant to local needs in the global south. Database biases in the representation of knowledge on rice. In: 13th Globelics International Conference, Havana.

Salter, A.J., Martin, B., 2001. The economic benefits of publicly funded basic research: a critical review. Res. Policy 30 (3), 509–532.

Sarewitz, D., Pielke Jr., R.A., 2007. The neglected heart of science policy: reconciling supply of and demand for science. Environ. Sci. Policy 10 (1), 5–16 http://doi.org/10.1016/j.envsci.2006.10.001.

Shibayama, S., Baba, Y., 2015. Impact-oriented science policies and scientific publication practices: the case of life sciences in Japan. Res. Policy 44 (4), 936–950 http://doi.org/10.1016/j.respol.2015.01.012.

Smith, M.J., Weinberger, C., Bruna, E.M., Allesina, S., 2014. The scientific impact of nations: journal placement and citation performance. PLoS One 9 (10), e109195 http://doi.org/10.1371/journal.pone.0109195.

StataCorp, 2013. Stata Statistical Software. StataCorp., College Station, TX.

Testa, J. (2011). The Globalization of Web of Science: 2005–2010. Philadelphia.

Thelwall, M., 2016. Are the discretised lognormal and hooked power law distributions plausible for citation data? J. Informetr. 10 (2), 454–470 http://doi.org/10.1016/j.joi.2016.03.001.

Tijssen, R.J.W., Leeuwen van, T.N., Wijk van, E., 2009. Benchmarking university-industry research cooperation worldwide: performance measurements and indicators based on co-authorship data for the world's largest universities. Res. Eval. 18 (1), 13–24 http://doi.org/10.3152/095820209X393145.

Tijssen, R.J.W., 2012. Co-authored research publications and strategic analysis of public-private collaboration. Res. Eval. 21 (3), 204–215 http://doi.org/10.1093/reseval/rvs013.

UNESCO, 2010. UNESCO Science Report 2010. United Nations Educational, Scientific and Cultural Organization, Paris.

UNESCO, 2015. UNESCO Science Report: Towards 2030. UNESCO Publishing, Paris.

Vargas, R.A., de Souza Vanz, S.A., Stumpf, I.R., 2014. The role of National journals on the rise in Brazilian Agricultural Science Publications in Web of Science. J. Scientometr. Res. 3 (1), 28.

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

van Raan, A.F.J., 2003. The use of bibliometric analysis in research performance assessment and monitoring of interdisciplinary scientific developments. Technikfolgenabschätzung-Theorie Und Praxis/Technol. Assess. Theory Pract. 1 (12), 20–29.

van Raan, A.F.J., 2004. In: Moed, H.F., Glänzel, W., Schmoch, U. (Eds.), Measuring Science. Springer, Netherlands, pp. 19–50.

Waltman, L., van Eck, N.J., 2012. Source normalized indicators of citation impact: an overview of different approaches and an empirical comparison. Scientometrics 96 (3), 699–716 http://doi.org/10.1007/s11192-012-0913-4.

Waltman, L., van Eck, N.J., van Leeuwen, T.N., Visser, M.S., van Raan, A.F.J., 2011. Towards a new crown indicator: some theoretical considerations. J. Informetr. 5 (1), 37–47 http://doi.org/10.1016/j.joi.2010.08.001.

Waltman, L., 2016. A review of the literature on citation impact indicators. J. Informetr. 10 (2), 365–391 http://doi.org/10.1016/j.joi.2016.02.007.

Wang, L., 2016. The structure and comparative advantages of China's scientific research: quantitative and qualitative perspectives. Scientometrics 106 (1), 435–452 http://doi.org/10.1007/s11192-015-1650-2.

Zhou, P., Leydesdorff, L., 2006. The emergence of China as a leading nation in science. Res. Policy 35 (1), 83–104 http://doi.org/10.1016/j.respol.2005.08.006.