Globalization of Science: Evidence from authors in academic journals by country of origin[[1]](#footnote-1)\*

Vít Macháček1,2

1 vit.machacek@cerge-ei.cz

CERGE-EI, Politických vězňů 7, 110 00 Prague (Czechia)

2 Institute of Economic Studies, Faculty of Social Sciences at the Charles University, Opletalova 26, 110 00 Prague (Czechia)

# Abstract

# The scientific community shares a common sense that by default, science should be globally oriented. This study measures the tendency to publish in globalized journals on a large dataset of journals indexed in the Scopus database. Based on data on 34 964 journals indexed in the Scopus Source List (Scopus 2018), we derived eight globalization indicators. These were subsequently scaled-up to the level of 174 countries and 27 disciplines between 2005 and 2017. The methodology draws from the pioneering work of Zitt and Bassecoulard (1998; 1999). The paper is accompanied by the interactive publication available at <http://www.globalizationofscience.com>

# Advanced countries tend to have high globalization that is not varying across disciplines. Social sciences are similarly globalized as life sciences. The globalization in the former Soviet bloc is lower, especially in social sciences or health sciences. China has profoundly globalized its science system; gradually moving from the lowest globalization rates to the world average. Contrary Russia was constantly among the least globalized during the whole period, with no upward trend.

# Introduction

The scientific community shares a common sense that by default science should be globally oriented. Academic researchers should present their results to their peers across the world and publish in journals contributed by researchers from the whole world (*globalized journals)*.

On the other hand, there are concerns about the side-effects of ever-growing globalization of scientific communication. Globalized journals that are predominantly based in industrial countries might underestimate the needs of other countries (Chavarro 2016, Rafols et al 2014). Local journals then offer a platform for knowledge that is relevant locally, but is not of concern at the global level.

Journal globalization (from now on just *globalization*) can naturally differ across disciplines. In social sciences, the local research can be more important than in physics. However the cross-country heterogeneity within a single discipline points towards the research evaluation in the country and the research culture in a broader context. Large differences between globalization in Economics in Netherlands and in the Czech Republic cannot be easily explained by the research content. They point at the incentives provided by the research system of given countries.

The main goal of the paper is to measure globalizationacross countries and disciplines as well as its development in time and to identify the sources of its variance. The more researchers publish in the same journals as their peers abroad, the more globalized their research is. The global dimension of the audience is emphasized (hence globalization), but also alternative specifications based on language and institutional concentration are added to increase robustness of findings.

The intuition suggests that the *globalized* journals will be disseminating science better than journals operating in only a handful of countries. Publishing in globalized journals improves the researchers’ visibility on the international scene (XXX). The more authors contribute to the journal, the higher competition may theoretically enforce higher quality. International publishing leads to higher competition faced by local researchers and to a larger emphasize on novelty in the international context (XXX). The systemic tendency to publish in non-globalized journals indicates the local researchers’ lack of motivation to open up to the global stream of knowledge.

In performance-based research systems where journals are used as major evaluation tool journals can be prone to gaming practices (Rijcke et al 2016, Hicks 2012, Good et al. 2015). Arguably, local journals are more vulnerable as they are easier to “control”. Good et al. (2015) describe practices of “establishing working paper series and promoting them as if they were refereed journals” (p. 97) in the Czech formula-based Evaluation Methodology.

Researchers shape the content of published articles already in the early stage of the research according their journal submission plans (Gläser and Laudel 2016). This also applies to globalization. Editors of globalized journal would naturally emphasize topics, which they consider relevant for the broadest possible audience. But that does not necessarily fits into the local research needs. Local journals may be sometimes better at targeting the local issues, but they may be more prone to nepotism and gaming (XXX).

|  |
| --- |
| **\*** |
|  |
| *\* are* |

In Western Europe and North America, research has undergone a transition from the national to the transnational model of publications already in the 1980s and the beginning of the 1990s (Zitt et al. 1998). Thirty years later, the non-globalized journals still play a substantial role in the countries from the former Soviet bloc (see Moed 2018, Kirchik and Gingras 2012). For example in Russia, Romania or Croatia more than 25 % of the national reseach output is published in what we refer to as domestic journals (Figure 1).

Since the transition the global research landscape changed dramatically. It grows both in terms of size and interconnectedness (Wilsdon 2011). New countries incorporate their research into the global knowledge flows (Gazni, Sugimoto, Didegah 2012; Wagner et al. 2015) and collaboration distances decrease (Waltman 2011). The international collaboration drives the growth of the research output (Adams 2013). Developing countries invest heavily to improve its research infrastructure (Wilsdon 2011) and international visibility (Zhou and Glanzel 2010).

Zitt and Bassecoulard (1998 and 1999) suggested a methodology for determining journal internationality. It also allow to scale up from journal level to the national level. However, since then any systematic evidence on the journals’ internationality is missing. This paper applies similar methodology to study the changes in the globalization landscape in the new millenia. How is China globalizing its fast growing research output (Leyersdorf and Zhou 2006)? How about Russia and Eastern Europe?

Based on data on 34 964 journals indexed in the Scopus Source List (Scopus 2018), we derived 8 indicators of journal globalization. These were subsequently scaled-up to the level countries, disciplines and time. The final dataset consists of globalization scores for 174 countries across 27 narrow disciplines and 4 broad disciplines (plus 1 ubiquitous *All disciplines*) between 2005 and 2017.

The following section describes used methodology and its main limitations; in the second section we describe collection of data and its characteristics and the third describes results. The last section concludes. The paper is accompanied by an already released interactive study available at <http://www.globalizationofscience.com> (Macháček and Srholec 2019). Readers can spend their time with the interactive application, as it offers an intuitive way of exploring the results of this paper.

# Globalization Indicators

The academic journals are an essential platform for scientific knowledge dissemination. They allow scientists across the world to keep up-to-date with the latest discoveries and to present their results to the global audience. Simultaneously they serve as a major research evaluation tool and researchers often face a significant pressure to publish “internationally”.

There is no consensus on the definition of “internationality”. Journal can be considered international based on country of origin of authors, country of origin of editors, language of publication or even having “international” in its name. Each definition of interationality can lead to different ranking in the degree of the internationality (Buela-Casal et al. 2006). One must be careful when measuring internationality.

To ensure robustness the paper uses eight *globalization indicators* assessing each country and discipline’s globalization for each year. The two-step methodology builds on the work of Zitt and Bassecoulard (1998) and (1999). First, we calculate the globalization indicators for each journal in our dataset in each year. Subsequently, the journal-level indicators are aggregated up to the level of countries and disciplines.

*Journal-level Indicators*

The indicators are intentionally constructed diverse. They vary in terms of input data as well as the approach to globalization. The indicators are not perfect, but each is imperfect in a different way. When combined, they can yield a robust picture of development of globalization.

Four indicators – *euclid*, *cosine*, *wGini* and *top3* (see Table 1 for details) are designed to account for a strong concentration of research in a few countries[[2]](#footnote-2). They are based on the idea that globalized journals have a structure that closely resembles the global structure of the whole discipline. Researchers from the whole world have equal probability to be published regardless of their affiliation country. These indicators measure how the distribution of authors corresponds to the distribution of authors in the entire discipline (column *Bench.* in Table 1).

Two indicators are simple shares of documents in the journal fulfilling a simple condition. The *englishShare* is a share of documents written in English and the indicator *localShare* is the share of authors originating in the same country as the publisher of the journal. Last two indicators – *giniSimpson, instTOP3* – are diversity measures. The first is Gini-Simpson Index applied on country data. The second is a simple ratio of three largest affiliations on the total number of documents.

The indicators are based on three different data sources (column *Data* in Table 1). Six indicators employ affiliation countries of authors. These are complemented with one indicator based on language and one based on affiliation names. Half of the indicators is constructed using the whole distribution of the underlying data, i.e. each document enters the calculation (column *Dist.*). Two indicators analyze only documents by most important contributors (countries and affiliation).

, and are the number of documents with authors affiliated to the country *c* or institution *i,* in journal *j* or discipline *d*, in year *y*.

is the number of documents with authors from the same country as the publisher of journal *j* in the year *y*. is the number of English-written documents in the journal *j* in year *y*.

denotes the total number of documents in the journal *j* in year *y*. Note that documents by authors from multiple countries are fully attributed to each country, i.e. .

The vectors and represent the country distribution of authors of the journal and the discipline , in which and . While is calculated separately in each year *y*, relates to the whole period. The benchmark distribution is always calculated from all available periods so that the development in time takes into account the world trend.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Table 1: Globalization Indicators** | | |  |  |  |
| *Indicator*  *(shortcut)* | *Formula* | *Data* | *Bench.* | *Dist.* | *Source\** |
| *Description* | | | | |
| Euclidian distance (*euclid)* |  | Country | Yes | Full | ZB (1998) |
| Euclidian distance of journal and discipline country distribution | | | | |
| Cosine distance  (*cosine*) |  | Country | Yes | Full | ZB (1998) |
| Cosine distance of journal and discipline country distribution | | | | |
| GiniSimpson Index (*GiniSimpson*) |  | Country | No | Full | Aman (2016) |
| Gini-Simpson diversity of journal country distribution | | | | |
| Weighted Gini  (*wGini*) |  | Country | Yes | Full | Own |
| Gini Index of authors' countries weighted by discipline country distribution | | | | |
| Largest Contributors Surplus (*top3*) |  | Country | Yes | Partial | Own |
| Surplus of three largest contributing countries over its share in discipline | | | | |
| Institutional Diversity  (*instTOP3*) |  | Institutional | No | Partial | Own |
| Share of three largest institutions on all documents | | | | |
| English Documents  (*englishShare*) |  | Language | No | Share | BC et al (2006) |
| Share of English-written documents | | | | |
| Local Authors  (*localShare)* |  | Country | No | Share | ZB (1998) |
| Share of documents from a journal's domicile | | | | |
| \* ZB is Zitt and Bassecoulard; BC is Buela-Casal | | | | | |

## Aggregation

In the second stage, the journal-level indicators were aggregated to the level of countries and disciplines. The resulting globalization score is a weighted average of individual journals scaled between 0 and 1, where 0 is the lowest globalization across all years, countries and disciplines within the particular indicator and 1 is the highest.

To increase robustness and decrease volatility, the aggregation was only performed when the authors from the country published in at least 30 journals. This leads to gaps in results, particularly in the small disciplines and small countries.

The globalization of science in country , discipline and year expressed by an indicator is calculated from the set of journals assigned to discipline as follows:

is the share of documents with authors from country *c* in journal on all documents from the country , discipline in year , is the globalization indicator *i* of journal in the discipline *d* and year *y*.

Subsequently, the aggregated globalization index was standardized between 0 and 1 and converted to an ascending scale to simplify the interpretation of the results:

in which and is minimum and maximum value of the indicator *i* across all years, countries and disciplines and equals -1 for the minimizing indicators (low values for high globalization) and 1 otherwise.

It is possible to compare globalizations between countries, discipline, and in time. However, the meaningful comparison between indicators is not possible due to the large heterogeneity of underlying distributions. The same value from two indicators cannot be interpreted as corresponding levels of globalization.

## Limitations

The major drawback of the methodology is the representativeness of the underlying data. We refer to the *Globalization of Science*, but it might be more convenient to refer to the *Globalization of Science in Journals then Indexed in the Database*. Citation databases may index the research output across units unevenly. The representativeness issue is present in all major dimensions – countries, disciplines, and in time.

Bibliometric databases probably represent larger portion of the research output in the countries of scientific core than in those at the periphery (Rafols et al 2014 shows an example of rice research). With a reasonable assumption that the under-representation is affecting more the non-globalized part of journals the results show the upper bound of globalization of the periphery countries.

The results are sensitive to journal-indexation decisions of the database. For example, in 2009 Scopus reacted to criticism by increasing its coverage of SSH journals by 39 % (Hicks and Wang 2010). Short-term changes must be interpreted with caution. Large year-by-year jumps are not necessarily caused by fundamental changes of the researchers’ behavior but are often driven by adding (or removing) journals in the database. Also long-term gradual changes may be partly driven by indexation decisions.

The bibliometric databases can cover disciplines unevenly as well. These concerns not only contain the coverage of journals within the particular database, but also use of journals as a publication platform. For disciplines that rely on other publication venues such as books the results may be distorted.

# Data

The analysis is based on Scopus data. Scopus indexes approximately twice more journals than its main competitor, Web of Science (Mongeon and Paul-Hus 2016). It is more likely to contain the more localized part of the scientific output in the country.

The data for all 34 964 journals indexed in the Scopus Source List (Scopus 2018) were downloaded using Scopus API in August 2018[[3]](#footnote-3). For each journal in each year between 2005 – 2017, we downloaded the country and institutional distribution of authors and the distribution of languages. Data were limited to articles, reviews, and conference papers.

The Scopus Search API was requested with the following query:

*ISSN(AAAA-BBBB) AND DOCTYPE(AR OR RE OR CP) AND PUBYEAR = YYYY*

in which AAAA-BBBB is the journal's ISSN and YYYY is the year. Rather than publication-level data, the aggregate distribution is collected. For each journal in each year, we collect the number of articles affiliated to each country, language, and institution.

Scopus Journal Classification (see Scopus 2019) is used to assign journals to disciplines. Both more narrow classification (*Major Subject Classification;* referred to as *narrow disciplines*) and broad classification on 4 disciplines (*Broad Subject Clusters –* Life Sciences, Physical Sciences, Health Sciences and Social Sciences) is used, supplemented by all-encompassing discipline *All* (referred to as *broad disciplines*). In the rest of the paper, the broad classification will be stressed, but the narrow results are also available in both interactive application and the downloadable data.

Journal-based discipline classification is a rough brush as it is not possible to assign documents directly to disciplines, which is relatively common. In our dataset large part of journals (20 % according to *broad* classification and 50 % of narrow) are assigned to more than one discipline. The used methodology fully attributes all articles in the journals to all assigned disciplines. This may cause distortion, especially with respect to large interdisciplinary journals such as PLOS One as these tend to be highly globalized.

Only minor data cleaning was required after downloading the data. Approximately 5 % of publications are attributed to the *undefined* country. These were excluded from the analysis. Undefined publications were also subtracted from the total number of publications in the journal. The data for Russia and the Russian Federation and Yugoslavia and Serbia were merged.

The resulting database contains information on more than 22 million documents. The Scopus indexation grows relatively fast (by an average pace of 4 %). We track 1.29M documents published in 2005 up to 2.09M in 2017. The growth momentum was generally faster in the first half of analyzed period.

Whereas research output grows in almost all countries, it is much faster in *Developing countries*. The share of documents from Advanced *countries* decreased from 75 % in 2005 to less than 60 % in 2017, while it increased in from 19 % to more than 33 % in *Developing countries*. Interestingly, the share of *Transition countries* stays stable throughout the period – cca 6.5 %.

# Results

The reported results are normalized relative to all observations within a single indicator. and 1 always refer to the least and most globalized country, discipline and the period within all results available within a single indicator. Neglecting representiveness issue, the methodology allows for any comparison within an indicator – between countries, disciplines and development. However, the comparison between indicators is not possible as the journal-level indicators distributions is too heterogenous.

After excluding all dependent territories except Hong Kong, the computation algorithm yielded results for 174 countries. The minimum requirement on the number of journals to calculate results lead to gaps in the data. In 2017, the data were available for 171 countries in the discipline *All,* approximately XXX countries in broad disciplines and for less than 100 countries in most of narrow disciplines. Naturally, the larger researcher the research production in the country, the more globalization scores is computed.

The results are robust to varying indicators. The correlation coefficients reported in appendix X are generally high. 25 out of 28 coefficients exceed 0.5, and a half coefficients are higher than 0.7. Also, visual check in the interactive application reveals that most globalization paths in time and relative rankings are similar across indicators. The most in-between indicator is Euclidian distance with a correlation coefficient higher than 0.75 in 8 out of 9 indicators. Therefore Euclidian distance is used as a main indicator of globalization. By default, we refer to it when not stated otherwise.

The distortion resulting from journals being associated

In what follows, only the major trends of globalization are summarized. The data are available in much greater detail in the interactive application, where the readers are invited to explore the situation in country, discipline, or period of their interest. The data are also available as an Appendix 1.

Of the three dimensions the globalization is computed for, countries are the most important for explaining its variance.

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Table 2: Variance explained () by countries C, disciplines D and time T by ANOVA** | | | | | | | | | |
|  |  | Gini-Simpson | cosine | euclid | instTOP3 | localShare | shareEnglish | top3 | weightGini |
| Broad disc. | C | 37% | 48% | 37% | 44% | 19% | 29% | 32% | 42% |
| D | 20% | 16% | 22% | 17% | 34% | 14% | 25% | 18% |
| T | 1,2% | 0,2% | 0,1% | 0,4% | 2,5% | 0,7% | 0,2% | 0,8% |
| Narrow disc. | C | 55% | 67% | 58% | 64% | 37% | 47% | 53% | 61% |
| D | 14% | 4% | 11% | 4% | 23% | 10% | 9% | 5% |
| T | 2,6% | 0,3% | 0,2% | 0,9% | 4,2% | 1,7% | 0,3% | 1,5% |
| Note: ANOVA based on regression for all observations, | | | | | | | | | |

The results cover the period between 2005 and 2017. When analyzing cross-country differences, the period 2015 – 2017 is considered. The three-year window allows to partially balance the effect of jumps caused by changes in the Scopus indexation.

The narrow disciplines are neglected. The data availability in smaller countries is lower than in broader disciplines, and the identification of general trends makes more sense in the broader context. Generally, the narrow disciplines usually follow similar paths as the broad disciplines. However, the volatility naturally grows with the detail.

The research is most globalized in *advanced countries*, followed by *developing countries* and *transition countries*. Table 3 presents the means and standard deviations of all globalizations within country groups between 2015 and 2017. In all disciplines, the mean is highest in *advanced countries* and the lowest in *transition countries*. The *life sciences* and *physical sciences* are more globalized than *social sciences* and *health sciences*.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Table 3: Means and standard deviations of globalizations within country groups in 2015-2017 measured by Euclidian distance by disciplines (wider definition)** | | | | | |
| Economic status | All | Life | Physical | Social | Health |
| Advanced countries | .75 | .80 | .79 | .75 | .72 |
| (.03) | (.03) | (.03) | (.05) | (.04) |
| Developing countries | .70 | .74 | .75 | .68 | .68 |
| (.06) | (.07) | (.05) | (.10) | (.08) |
| Transition countries | .61 | .68 | .66 | .53 | .58 |
| (.09) | (.10) | (.11) | (.13) | (.13) |
| *Source: own calculation; IMF (2003); Scopus; Standard deviations in brackets.* | | | | | |

The science in *advanced countries* is not only highly globalized but also relatively invariant. The figures are similar in all disciplines, and their standard deviations are low. In all disciplines, research in advanced countries tends to be globalized. Only research in Japan and South Korea is somewhat lower (globalization 0.66 in discipline *All*). In Germany, France, and Spain, the *life sciences* and *physical sciences* are highly globalized (approx. 0.78), but *social sciences* and *health sciences* tend to be below standards (<0.65).

In *transition countries,* the globalization is generally lower in all disciplines. The means vary between 0.53 in Social Sciences, to 0.68 in Life Sciences. The variance is also larger within disciplines. The standard deviations are approximately three times larger than in advanced countries. It is hard to characterize this diverse group. However, there is only one country whose characteristics fits well with the patterns common in advanced countries – Estonia.

In some countries, the globalization is relatively high in *life sciences* and *physical sciences* (>0.7), but significantly lower in *social sciences* and *health sciences* (<0.55). This applies mainly countries from the former Soviet bloc in Europe, such as Czechia, Poland, Slovakia, Slovenia, or Croatia. However, there are also countries who perform badly in most disciplines – Belarus, Ukraine, Kazakhstan, or Russia.

Russia is a prime example of the strongly isolated research system (see Figure 2). In 2017, Russia ranked as the first or second least globalized country in all broad disciplines. Even when extended to the narrow definition of disciplines, Russia is among the last 3 countries in 23 out of 25 disciplines where the data are available.

Where there is a will, there is a way. At the beginning of the analyzed period, in 2005, China was the least globalized country in the world in *All disciplines,* and among the 5 least globalized in all broad disciplines (see Moed 2002). China invests heavily in the modernization of its science infrastructure (Wilsdon 2011). Figure 2 shows a rapid transformation of the Chinese system that resulted in the relatively fast growth of globalization. During the analyzed period, the globalization grew fast across all disciplines. In 2017 China was still less globalized than *advanced countries*, but its progress is visible.

|  |
| --- |
| **Figure 2: Globalization in broad disciplines in China and Russia in time (Euclidian distance)** |
|  |
| *Source: own calculation, Scopus* |

|  |
| --- |
| **Figure 3: Globalization of science on the world map (2017; Euclidian distance; All disciplines)** |
|  |
| *Source: own calculation; Scopus; the darker the color, the higher globalization* |

The map on Figure 3 shows that in all BRIICS[[4]](#footnote-4) countries, the globalization is below (or very close to at maximum) average in all broad disciplines in 2017. The long-term growth trend in the past 12 years is only seen in China. In Indonesia, the globalization even slightly decreases in time in all broad disciplines. The path to globalization is far from being guaranteed for the developing countries.

COUNTRY IS MUCH STRONGER PREDICTOR OF GLOBALIZATION THAN DISCIPLINE AND YEAR. PERHAPS TRY <https://scikit-learn.org/stable/modules/generated/sklearn.tree.DecisionTreeClassifier.html?fbclid=IwAR2HrJuop9xe9G3REaTJjvlpYTVxXUkQu9byFHk3w8J960XswQZdO4u-7KU#sklearn.tree.DecisionTreeClassifier.feature_im +portances_>

TRY SIMPLE CLUSTERING TO DELINEATE COUNTRY GROUPS (K-MEANS ? )

STABILITY OF RESULTS WITH REGARD TO DISCIPLINE MULTIPLE DISCIPLINE ATTRIBUTION

THINK AGAIN OF DECOMPOSING CHINESE GROWTH BETWEEN GROWTH OF CHINESE IMPORTANCE AND GROWTH OF GLOBALIZATION IN CHINA

# Conclusions

# Appendix

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Appendix X: Full-Sample Correlation Matrix of Indicators** | | | | | | | |
| *Indicator* | *euclid* | *cosine* | *GiniSimpson* | *wGini* | *top3* | *instTOP3* | *sEnglish* | | *localShare* |
| ***euclid*** | **1.00** | **.83** | **.87** | **.90** | **.93** | **.81** | **.61** | | **.75** |
| *cosine* | .83 | 1.00 | .64 | .93 | .75 | .69 | .47 | | .41 |
| *GiniSimpson* | .87 | .64 | 1.00 | .79 | .72 | .67 | .64 | | .78 |
| *wGini* | .90 | .93 | .79 | 1.00 | .86 | .73 | .55 | | .56 |
| *top3* | .93 | .75 | .72 | .86 | 1.00 | .79 | .51 | | .67 |
| *instTOP3* | .81 | .69 | .67 | .73 | .79 | 1.00 | .43 | | .57 |
| *sEnglish* | .61 | .47 | .64 | .55 | .51 | .43 | 1.00 | | .61 |
| *localShare* | .75 | .41 | .78 | .56 | .67 | .57 | .61 | | 1.00 |

*Pearson correlation coefficients of all available data for each indicator; Source: Scopus; own calculation*

# Journal Globalization

International focus is an important pillar of research policies across the world. FIRST INTRO PARAGRAPH

Originally academic journals developed in national contexts. They were written in national languages and often focused exclusively on national researchers. In 1970s and 1980s the journals used by Western researchers transformed from “national to transnational mode of publication” (Zitt 1998). Journals started to be published in English and published researchers from other countries on regular basis. Traditinal national boundaries in large extent disappeared.

While some acknowledge the internationalization of research leads to more intensive exchange of ideas,(XXX), others also point to its potential drawbacks (YYY).

There are signs that similar transition was not finished in Eastern Europe.

Researchers have incentives to shape the journals according to their benefits. These benefits do not have to be in line with other dissemination goals.

Moreover, scientific environment is conservative and sometimes evolves only slowly.

This combination can lead to situation

Zitt (1998) describes the transition from national to transnational model of communication that happened during 1980s. He even describes it as “almost complete” (p. 30). However, the analysis has been done on only small subset of mostly developed western countries (USA, UK, Germany, Netherlands, Japan, France, Italy, Spain, and exception of Russia).

The globalization of science (from now on just globalization) contains a piece of useful information about this transformation process. The researchers in the country and discipline with high globalization publish their work in journals operating worldwide. On the contrary, when globalization is low, local researchers rely on journals that significantly deviate from the world distribution of journals. Typically, they are local journals contributed by researchers from just one or a narrow group of countries. These journals are relevant for only a small group of local researchers.

|  |
| --- |
|  |
|  |
|  |

Domestic journals were originally published in national languages. However, they often adjust to the transforming environment by translating their content into English while keeping their author-base local (Kirchik and Gingras 2012, Moed 2018). Language of publication is not necessarily a good indication of the journal’s globalization. Nevertheless, one indicator based on language is included in the paper.

Many journals operate worldwide and forcing a single-country domicile can be too strict. The concept of domestic journals outlined above also does not take into account the size of the research sector in the country. The composition of the editorial board can be a useful measure of journal internationality, but its data availability is limited. Moreover, it is easier to fake than the authors contributing to the journal. The journal's authors provide a direct measure of the journal globalization that is both available and relatively reliable.

The globalization is informative about the incentives provided by the research evaluation system in the country and the power relations in the local research system. Incentives influence researchers’ decision on where to publish (Franzoni et al. 2011). The decision where to submit a paper is a key stage of the research-production process, which involves a highly strategic behavior of the researcher (see Heintzelman and Nocetti, 2009) His decision is linked to the strength of the “adverse mechanisms” - persistent proximity networks hindering the internationalization process (Zitt and Bassecoulard 2004). The local ties between researchers, institutions, research managers, and policy-makers can diverge the knowledge flows from international, to more local paths.

To our knowledge, the only paper providing analyzing globalization in countries is Zitt and Bassecoulard (1999). Based on journal-level indicators suggested in their previous paper (1998), they analyzed countries and discipline according to their internationalization. They acknowledge a “general trend towards internationalization,” with some exceptions such as Russia. Since then, any systematic evidence is missing.

**Data**

The analysis is based on Scopus data. Scopus indexes approximately twice more journals than its main competitor, Web of Science (Mongeon and Paul-Hus 2016). It is more likely to contain the more localized part of the scientific output in the country.

The data for all 34 964 journals indexed in the Scopus Source List (Scopus 2018) were downloaded using Scopus API in August 2018. For each journal in each year between 2005 – 2017, we downloaded the country and institutional distribution of authors and the distribution of languages. Data were limited to articles, reviews, and conference papers.

The Scopus Search API was requested with the following query:

*ISSN(AAAA-BBBB) AND DOCTYPE(AR OR RE OR CP) AND PUBYEAR = YYYY*

in which AAAA-BBBB is the journal's ISSN and YYYY is the year. Rather than publication-level data, the aggregate distribution is collected. For each journal in each year, we collect the number of articles affiliated to each country, language, and institution.

Scopus Source List also contains Scopus Journal Classification (see Scopus 2019) used to assign journals to disciplines. Both more narrow classification (*Major Subject Classification;* referred to as *narrow disciplines*) and broad classification on 4 disciplines (*Broad Subject Clusters*) is used, supplemented by all-encompassing discipline *All* (*broad disciplines*).

Approximately 5 % of publications affiliated to the *undefined* country were excluded from the analysis. Undefined publications were also subtracted from the total number of publications in the journal. The data for Russia and the Russian Federation and Yugoslavia and Serbia were merged.



, Scopus indexed 184 documents from Romania in Social Sciences. One year later, this figure almost tripled to 471, while the globalization decreased from 0.5 to 0.2. This development is driven by a sudden indexation of several Romanian journals[[5]](#footnote-5) that publish mainly Romanian researchers.

**Results**

The reported results are normalized relative to all observations within a single indicator. and 1 always refer to the least and most globalized country, discipline and the period within all results available within a single indicator. The methodology allows for any comparison within an indicator. However, the comparison between indicators is not possible as the journal-level indicators distributions vary.

After excluding all dependent territories except Hong Kong, the computation algorithm yielded results for 174 countries. The minimum requirement on the number of journals to calculate results lead to gaps in the data. Naturally, the broader the discipline definition and the larger the research production, the more countries are available. In 2017, the data were available for 171 countries in the discipline *All,* but for most of the narrow disciplines, the results are calculated for less than 100 countries. The results availability also grows in time together with output indexed in Scopus.

The results are robust to varying indicators. The correlation coefficients reported in table 2 are generally high. 25 out of 28 coefficients exceed 0.5, and a half coefficients are higher than 0.7. Also, visual check in the interactive application shows that most globalization paths in time and relative rankings are similar across indicators.

The most in-between indicator is Euclidian distance with a correlation coefficient higher than 0.75 in 8 out of 9 indicators. Therefore Euclidian distance is used as a main indicator of globalization. By default, we refer to it when not stated otherwise.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Table 2: Full-Sample Correlation Matrix of Indicators** | | | | | | | |
| *Indicator* | *euclid* | *cosine* | *GiniSimpson* | *wGini* | *top3* | *instTOP3* | *sEnglish* | | *localShare* |
| *euclid* | 1.00 | .83 | .87 | .90 | .93 | .81 | .61 | | .75 |
| *cosine* | .83 | 1.00 | .64 | .93 | .75 | .69 | .47 | | .41 |
| *GiniSimpson* | .87 | .64 | 1.00 | .79 | .72 | .67 | .64 | | .78 |
| *wGini* | .90 | .93 | .79 | 1.00 | .86 | .73 | .55 | | .56 |
| *top3* | .93 | .75 | .72 | .86 | 1.00 | .79 | .51 | | .67 |
| *instTOP3* | .81 | .69 | .67 | .73 | .79 | 1.00 | .43 | | .57 |
| *sEnglish* | .61 | .47 | .64 | .55 | .51 | .43 | 1.00 | | .61 |
| *localShare* | .75 | .41 | .78 | .56 | .67 | .57 | .61 | | 1.00 |

*Pearson correlation coefficients of all available data for each indicator; Source: Scopus; own calculation*

In what follows, only the major trends of globalization are summarized. The data are available in much greater detail in the interactive application, where the readers are invited to explore the situation in country, discipline, or period of their interest. The data are also available in Appendix 1.

The results cover the period between 2005 and 2017. When analyzing cross-country differences, the period 2015 – 2017 is considered. The three-year window allows to partially balance the effect of jumps caused by changes in the Scopus indexation.

The narrow disciplines are neglected. The data availability in smaller countries is lower than in broader disciplines, and the identification of general trends makes more sense in the broader context. Generally, the narrow disciplines usually follow similar paths as the broad disciplines. However, the volatility naturally grows with the detail.

The research is most globalized in *advanced countries*, followed by *developing countries* and *transition countries*. Table 3 presents the means and standard deviations of all globalizations within country groups between 2015 and 2017. In all disciplines, the mean is highest in *advanced countries* and the lowest in *transition countries*. The *life sciences* and *physical sciences* are more globalized than *social sciences* and *health sciences*.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Table 3: Means and standard deviations of globalizations within country groups in 2015-2017 measured by Euclidian distance by disciplines (wider definition)** | | | | | |
| Economic status | All | Life | Physical | Social | Health |
| Advanced countries | .75 | .80 | .79 | .75 | .72 |
| (.03) | (.03) | (.03) | (.05) | (.04) |
| Developing countries | .70 | .74 | .75 | .68 | .68 |
| (.06) | (.07) | (.05) | (.10) | (.08) |
| Transition countries | .61 | .68 | .66 | .53 | .58 |
| (.09) | (.10) | (.11) | (.13) | (.13) |
| *Source: own calculation; IMF (2003); Scopus; Standard deviations in brackets.* | | | | | |

The science in *advanced countries* is not only highly globalized but also relatively invariant. The figures are similar in all disciplines, and their standard deviations are low. In all disciplines, research in advanced countries tends to be globalized. Only research in Japan and South Korea is somewhat lower (globalization 0.66 in discipline *All*). In Germany, France, and Spain, the *life sciences* and *physical sciences* are highly globalized (approx. 0.78), but *social sciences* and *health sciences* tend to be below standards (<0.65).

In *transition countries,* the globalization is generally lower in all disciplines. The means vary between 0.53 in Social Sciences, to 0.68 in Life Sciences. The variance is also larger within disciplines. The standard deviations are approximately three times larger than in advanced countries. It is hard to characterize this diverse group. However, there is only one country whose characteristics fits well with the patterns common in advanced countries – Estonia.

In some countries, the globalization is relatively high in *life sciences* and *physical sciences* (>0.7), but significantly lower in *social sciences* and *health sciences* (<0.55). This applies mainly countries from the former Soviet bloc in Europe, such as Czechia, Poland, Slovakia, Slovenia, or Croatia. However, there are also countries who perform badly in most disciplines – Belarus, Ukraine, Kazakhstan, or Russia.

Russia is a prime example of the strongly isolated research system (see Figure 2). In 2017, Russia ranked as the first or second least globalized country in all broad disciplines. Even when extended to the narrow definition of disciplines, Russia is among the last 3 countries in 23 out of 25 disciplines where the data are available.

Where there is a will, there is a way. At the beginning of the analyzed period, in 2005, China was the least globalized country in the world in *All disciplines,* and among the 5 least globalized in all broad disciplines (see Moed 2002). China invests heavily in the modernization of its science infrastructure (Wilsdon 2011). Figure 2 shows a rapid transformation of the Chinese system that resulted in the relatively fast growth of globalization. During the analyzed period, the globalization grew fast across all disciplines. In 2017 China was still less globalized than *advanced countries*, but its progress is visible.

|  |
| --- |
| **Figure 2: Globalization in broad disciplines in China and Russia in time (Euclidian distance)** |
|  |
| *Source: own calculation, Scopus* |

The map on Figure 3 shows that in all BRIICS[[6]](#footnote-7) countries, the globalization is below (or very close to at maximum) average in all broad disciplines in 2017. The long-term growth trend in the past 12 years is only seen in China. In Indonesia, the globalization even slightly decreases in time in all broad disciplines. The path to globalization is far from being guaranteed for the developing countries.

|  |
| --- |
| **Figure 3: Globalization of science on the world map (2017; Euclidian distance; All disciplines)** |
|  |
| *Source: own calculation; Scopus; the darker the color, the higher globalization* |

**Conclusions**

The paper measured the globalization of science in 174 countries and 31 disciplines between 2005 and 2017. Using the data on articles, reviews and conference papers from journals indexed in the Scopus Source List we developed a two-step methodology that scales the globalization of individual journals up to the level of countries, disciplines and years. Multiple indicators of globalization (based on country data, language data, and institutional data) were constructed to verify the robustness of our findings. The paper is also accompanied by an interactive application available at <http://www.globalizationofscience.com/>.

The transition from national to the transnational mode of communication took place in the 1980s and 1990s in Western countries. A similar transition is still not finished in the countries of the former Soviet bloc. Especially in the Social Sciences and the Health Sciences, the role of non-globalized journals is high. An example of Russia shows that the lack of globalization can be very persistent. An example of China shows that under certain conditions, the widespread integration into a global research publication flows is possible, and it can even be relatively fast.

The granular data on the globalization of science can serve as a good starting point for further exploration for researchers, research managers, policy-makers, and other research evaluation practitioners. These professionals should be asking whether the level of globalization in their context is adequate. However, further research necessary to understand the role of globalized journals in modern science. The relationship between quality and globalization is far from straightforward.

# References

Adams, J. (2013). Collaborations: The fourth age of research. Nature, 497(7451), 557.

Aman, V. (2016). Measuring internationality without bias against the periphery. In 21st International Conference on Science and Technology Indicators-STI 2016. Book of Proceedings.

Buela-Casal, G., Perakakis, P., Taylor, M., & Checa, P. (2006). Measuring internationality: Reflections and perspectives on academic journals. Scientometrics, 67(1), 45-65.

Franzoni, C., Scellato, G., & Stephan, P. (2011). Changing incentives to publish. Science, 333(6043), 702-703.

Gazni, A., Sugimoto, C. R., & Didegah, F. (2012). Mapping world scientific collaboration: Authors, institutions, and countries. Journal of the American Society for Information Science and Technology, 63(2), 323-335.

Heintzelman, M., & Nocetti, D. (2009). Where should we submit our manuscript? An analysis of journal submission strategies. The BE Journal of Economic Analysis & Policy, 9(1).

IMF (2003) World Economic Outlook (Statistical Appendix; pp. 163-169). Available at: <https://www.imf.org/external/pubs/ft/weo/2003/02/>

Kirchik, O., Gingras, Y., & Larivière, V. (2012). Changes in publication languages and citation practices and their effect on the scientific impact of Russian science (1993–2010). Journal of the American Society for Information Science and Technology, 63(7), 1411-1419.

Macháček, V. and Srholec, M. (2019) Globalization of Science: Evidence from Authors in Academic Journals by Country of Origin. 6/2019. Institute for Democracy and Economic Analysis (IDEA).

Moed, H. (2002). Measuring China" s research performance using the Science Citation Index. Scientometrics, 53(3), 281-296.

Moed, H. F., Markusova, V., & Akoev, M. (2018). Trends in Russian research output indexed in Scopus and Web of Science. Scientometrics, 116(2), 1153-1180.

Mongeon, P., & Paul-Hus, A. (2016). The journal coverage of Web of Science and Scopus: a comparative analysis. Scientometrics, 106(1), 213-228.

Scopus (2019) What is the complete list of Scopus Subject Areas and All Science Journal Classification Codes (ASJC)? Available at: <https://service.elsevier.com/app/answers/detail/a_id/15181/supporthub/scopus/related/1/> [accessed May 28, 2019]

Scopus (2018) Scopus Source List (May 2018 version). Available at: <https://www.scopus.com/sources>

Wagner, C. S., Park, H. W., & Leydesdorff, L. (2015). The continuing growth of global cooperation networks in research: A conundrum for national governments. PLoS One, 10(7), e0131816.

Wilsdon, J. (2011). Knowledge, networks and nations: Global scientific collaboration in the 21st century.

Zitt, M., & Bassecoulard, E. (1998). Internationalization of scientific journals: a measurement based on publication and citation scope. Scientometrics, 41(1-2), 255-271.

Zitt, M., & Bassecoulard, E. (1999). Internationalization of communication a view on the evolution of scientific journals. Scientometrics, 46(3), 669-685.

Zitt, M., & Bassecoulard, E. (2004). Internationalisation in science in the prism of bibliometric indicators. In Handbook of quantitative science and technology research (pp. 407-436). Springer, Dordrecht.

Zitt, M., Perrot, F., & Barré, R. (1998). The transition from “national” to “transnational” model and related measures of countries' performance. Journal of the American Society for Information Science, 49(1), 30-42.

Zhou, P., & Glänzel, W. (2010). In-depth analysis on China’s international cooperation in science. Scientometrics, 82(3), 597-612.

# Appendix 1 and 2 – Complete results available in CSV files data.csv and countries.csv. The files are available at: [https://github.com/vitekzkytek/GlobalizationPaper/blob/master/appendix/](https://github.com/vitekzkytek/GlobalizationPaper/blob/master/appendix/data.csv)

1. \* Financial support from the research programme Strategy AV21 of the Czech Academy of Sciences and from the Charles University Grant Agency (GA UK; project no 1062119) is gratefully acknowledged. All usual caveats apply. [↑](#footnote-ref-1)
2. US alone accounts for 21 %, of the research output in 2018, China for additional 19 %. Scopus search: PUBYEAR = 2017 AND DOCTYPE (ar OR re OR cp) on January 24th 2020 . [↑](#footnote-ref-2)
3. We gratefully acknowledge Elsevier for providing us an extended access to the Scopus API. [↑](#footnote-ref-3)
4. Brazil, Russia, India, Indonesia, China and South Africa [↑](#footnote-ref-4)
5. Namely Amfiteatru Economic Journal, Transylvanian Review of Administrative Sciences, Journal of Economic Computation and Economic Cybernetics Studies and Research etc. [↑](#footnote-ref-5)
6. Brazil, Russia, India, Indonesia, China and South Africa [↑](#footnote-ref-7)