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 faces an everlasting pressure to publish internationally. 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 seven 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). Advanced countries tend to have high globalization that is not varying across disciplines. Social sciences and health sciences are less globalized than physical and 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 remains among the least globalized during the whole period, with no upward trend.

# Introduction

The scientific community faces an everlasting pressure to present research results internationally (Buela-Casal et al. 2006, Chavarro et al. 2017). Academic researchers are expected to publish in globalized journals, to which contribute authors regardless of their place of origin. Globalized journals promote competition, facilitate cross-fertilization of ideas and boost research collaboration across national borders. In contrast, a tendency to publish in non-globalized (or local) journals indicates the lack of researchers’ motivation to open to the global stream of knowledge.

Since the 1980s 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 et al. 2012; Wagner et al. 2015) and collaboration distances increase (Waltman et al. 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).

In Western Europe and North America, research has undergone a transition from the national to the transnational model of journal publications already in the 1980s and the beginning of the 1990s (Zitt et al. 1998). Thirty years later, however, the non-globalized journals still play a substantial role in the countries from the former Soviet bloc (see Moed et al. 2018, Kirchik et al. 2012). For example, in Russia, Romania or Croatia more than 25 % of the national research output is published in journals with at least a third of authors from the same country, where the journal publisher is based (see Figure 1). Clearly, local publishing is still common in Eastern Europe.

Zitt and Bassecoulard (1998 and 1999) suggested a methodology for determining journal globalization. It also allows to scale up from journal level to the national level. However, since then any systematic evidence is missing. This paper applies similar methodology to study the changes in the globalization landscape in the new millennium and apply it to much larger dataset. Based on data on 34 964 journals indexed in the Scopus Source List (Scopus 2018), we derived 7 indicators of journal globalization. These were subsequently scaled-up to the level of countries, disciplines and in time. The final dataset consists of average globalization scores and distribution of research output into journals by globalization quartiles.

This paper further contributes to the discussion on journal internationalization with data on differences in the propensity of researchers to publish in more or less globalized journals across countries, disciplines and over time. The major output of this research – The Globalization of Science dataset – provides this evidence for 174 countries, 27 narrow and 4 broad disciplines between 2005 – 2017. It is possible to compare globalization in a single discipline across countries, as well as to compare disciplines within a single country. Readers can take advantage of an interactive app [globalizationofscience.com](http://www.globalizationofscience.com) that is specifically designed for these kinds of comparisons.

|  |
| --- |
| **Figure 1: Share of research output in domestic\* journals in Europe in 2015-2017** |
|  |
| *\*Domestic journals are defined as journals with at least 33% authors from the same country as the publisher of the journal. Publisher country from Scopus Source List (Scopus 2018) is used to identify the domicile of the journal. Articles, reviews, and conference papers are included.* *Source: Own calculation based on Scopus (2018).* |

# Globalization Indicators

This paper perceives the globalization from the perspective of academic journals. The more researchers publish in the same journals as their peers abroad, the more globalized their research is. The global dimension is emphasized (hence globalization), but also alternative specifications based on language and institutional concentration are added to increase robustness of findings.

There is no consensus on the definition of “internationality”. Journal can be considered international based on the country of origin of authors, country of origin of editors, language of publication or even having “international” in its name. The journals publishing only Czech and Slovak authors, for instance, are de facto international. Naturally, each definition of internationality can lead to different results (Buela-Casal et al. 2006).

To ensure robustness the paper uses 7 *globalization indicators* assessing each journals globalization for each year. Most indicators are based on the geographical diversity of authors, but we include also one indicator based on language and one on institutional concentration. 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

Table 1 provides detailed overview of the indicators. For each journal *j* in the dataset, a set of globalization indicators *i* was calculated for each year *y*. Calculation is derived separately for each discipline *d*. The journal globalization is denoted .

, and are the number of documents with authors affiliated to the country *c* or organization *o,* 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 considers the world trend.

Three indicators – *euclid*, *cosine* and *top3* 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 benchmark 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). Five indicators employ affiliation countries of authors. These are complemented with one indicator based on language and one based on affiliation names. Three indicators are 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).

The indicators are constructed to be 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.

|  |  |  |  |
| --- | --- | --- | --- |
| **Table 1: Globalization Indicators** |  |  |  |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| *Indicator*  *(shortcut)* | *calculation* | *Data* | *Bench-mark* | *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 | | | | |
| 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  \*\* the underlying data for these indicators are sorted by descending order. The computation algorithm only considers the three most important | | | | | |

## 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 indicator and 1 is the highest.

The globalization of science in country , discipline and year expressed by an indicator is calculated from the set of journals assigned to discipline as an average of individual journals globalization weighted by the share of documents flowing into the journal:

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.

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

The methodology allows to compare globalizations between countries, discipline, and in time. Most indicators are independent on the size of research both in the country and in the discipline. The limitation is representativeness of the underlying data (see *Limitations* in the Data section). However, the meaningful comparison between indicators is not possible due to large heterogeneity of the underlying distributions[[2]](#footnote-2). The same value of two indicators cannot be interpreted as indicating the comparable level of globalization.

# Data

The analysis is based on data from the Scopus citation database. Scopus indexes more journals than Web of Science (SCI-Expanded, SSCI and A&HCI combined; see 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 Journal Classification (see Scopus 2019) is used to assign journals to disciplines. Both narrower classification to 27 disciplines (*Major Subject Classification;* referred to as *narrow disciplines*) and broad classification to 4 disciplines (*Broad Subject Clusters –* life sciences, physical sciences, health sciences and social sciences; referred to as *broad disciplines*) is used, supplemented by all-encompassing total for all disciplines combined. The most granular level of Scopus classification – *Scopus Subject Areas* – was neglected due to concerns to representativeness and threat of journals’ false identification (see Wang and Waltman 2016). In the rest of the paper, the broad classification will be stressed, but the narrow results are also available in both the 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. In our dataset large part of journals (20 % according to *broad* and 50 % to narrow classification) are assigned to more than one discipline. The used methodology fully attributes all journals documents to all assigned disciplines. This may cause distortion, especially due to large interdisciplinary journals that index research from various unrelated disciplines.

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

## Limitations

The major drawback of globalization approach 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 by Scopus*. Citation databases may index publications, journals or authors unevenly across 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, Chavarro 2017, Mongeon and Paul-Hus 2016). With a reasonable assumption that the globalized journals are more likely to be indexed than non-globalized the results can be interpreted as the upper bound of globalization of the periphery countries.

The results are sensitive to Scopus journal-indexation decisions. For example, in 2009 Scopus reacted to criticism by increasing its coverage of social sciences and humanities journals by 39 % (Hicks and Wang 2010). Short-term changes must be interpreted with caution. Sudden 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 changes may be partly driven by indexing larger portion of existing journals.

The bibliometric databases can cover disciplines unevenly as well. Mongeon and Paul-Hus (2016) report significant under-representation of social sciences and humanities in both Scopus and WoS. These concerns not only contain the coverage of journals within the database, but also use of journals as a publication platform. For disciplines relying on other publication venues such as books the results may be distorted.

# Results[[3]](#footnote-3)

After excluding all dependent territories except Hong Kong, the computation algorithm yielded globalization scores in three major dimensions - 174 countries (), narrow 32 disciplines () and 13 years (). The panel is unbalanced as the rule of publishing in at least 30 journals in given country, discipline and year is applied. In 2017, for example, the data were available for 171 countries in the total figures*,* 125 – 155 in broad disciplines and for less than 100 countries in 21 out of 27 narrow disciplines. Naturally, the larger the research production in the country and the discipline, the more globalization scores are available. The data coverage also grows in time together with growth of research output and indexation of journals by Scopus.

## Results robustness

The different indicators yield similar estimates of globalization. The various indicators are generally correlated (see Table A1 in the Appendix). 25 out of 28 coefficients exceed 0.5, and a half of coefficients is higher than 0.7. The most in-between indicator is Euclidian distance with a correlation coefficient higher than 0.75 with all other indicators. That’s why by default, we refer to it when not stated otherwise.

Results are not excessively distorted by relying on journal-based discipline classification, although the documents in journals publishing in multiple disciplines are fully attributed to each of the discipline. To test it, globalization was recalculated using only journals with 1 broad discipline (80 % of journals) and with 1 narrow disciplines (50 % of journals). The distortion is acceptable. The correlation between original all globalization scores computed from all journals and one calculated from journals with only 1 broad discipline is 0.83. When counting only journals with 1 narrow discipline assigned the correlation is still 0.78.

## Countries, disciplines or time?

At the beginning of the analysis we decompose the globalization variance into dimensions – countries *c*, disciplines *d* and years *y*. Table 2 shows explained variance of ANOVA regression, on all globalization scores of given indicator *i*, where , , are treated as categorical variables. allows to decompose across the factors used in the regression. The regression is run separately for broad and narrow disciplines. This simple regression already explains approximately 70 % of the variance in the data in case of broad disciplines and 60 % of variance in narrow disciplines.

Countries are by far the most important factor as they systematically explain more than 50 % of (in some cases even 90 %). The effect is stronger for broad disciplines. Disciplines are weaker predictor of globalization, but it is still relatively important factor (especially in case of narrow disciplines).

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Table 2: Variance decomposition ( and ) by countries , disciplines *d* and years *y* by ANOVA** | | | | | | | | |
| Disc. |  | Gini-Simpson | cosine | euclid | instTOP3 | localShare | shareEnglish | top3 |
| Broad | *c* | 55 % | 67 % | 58 % | 64 % | 37 % | 47 % | 53 % |
|  | *d* | 14 % | 4 % | 11 % | 4 % | 23 % | 10 % | 9 % |
|  | *y* | 2.6 % | 0.3 % | 0.2 % | 0.9 % | 4.2 % | 1.7 % | 0.3 % |
|  |  | 72 % | 72 % | 69 % | 69 % | 66 % | 60 % | 63 % |
| Narrow | *c* | 36 % | 46 % | 36 % | 42 % | 18 % | 28 % | 31 % |
|  | *d* | 20 % | 16 % | 23 % | 18 % | 34 % | 14 % | 24 % |
|  | *y* | 1.1 % | 0.2 % | 0.1 % | 0.4 % | 2.3 % | 0.6 % | 0.3 % |
|  |  | 57 % | 62 % | 58 % | 60 % | 56 % | 43 % | 56 % |
| Note: ANOVA based on regression for all observations. Type II sum of squares are used. | | | | | | | | |

Perhaps surprisingly, the role of time is negligible. This suggests that although Scopus decisions on inclusion of new journals into the database sometimes may have impact on situation in individual countries, it does not drive the global trends of globalization.

## Globalization by country groups and broad disciplines

The country groups by IMF (2003) serve as a good starting point for describing results. The countries are divided into three categories: (a) *Advanced countries* cover the richest countries in the world mainly in Western Europe, North America, and Eastern Asia. This group should capture countries of the western core (31 countries); (b) *Developing countries* – contain the rest of the World, including China (115 countries); and (b) *Transition countries* consist mainly of the post-soviet countries in Central and Eastern Europe, Central Asia. and Cuba, including new EU member states (28 countries). Exact classification of countries is available in the Appendix A5.

|  |
| --- |
| **Figure 2: Density of globalization by IMF (2003) economic status and by broad disciplines** |
|  |
| *Note: Each box represents distribution of for all countries and years in given country groups. Source: own calculation, Scopus.* |

Figure 2 depicts a distribution of globalization across all countries and years calculated by Euclidian distance grouped by economic status and broad disciplines. Each “*box*” represents a distribution of all globalization scores within given country group (x axis) and discipline (color) across all years. All the distributions are negatively skewed, meaning that globalization is more concentrated in the upper part of the distributions. Although Euclidian distance is reported here, the same patterns as described below are present in each of the indicators.

The cross-country data reveal that globalization in *Advanced countries* is high and relatively invariant. On the contrary, in *Transition countries*, the globalization is lower and much more heterogenous. The distribution is more spread across the whole spectrum. *Developing countries* are in between the groups.

Putting it all together, in the *Advanced countries*, the globalization is relatively high and differences between disciplines are minor. In *Transition countries* the mean globalization is lower and the gap between disciplines is larger. The *life sciences* and *physical sciences* are less globalized especially in the former USSR countries. The *social sciences* and *health sciences* are less globalized across *Transition countries*, but outside of the former USSR, the globalization gradually increases in time.

The differences between the *Advanced* and *Transition countries* can be depicted in the European Union, which has members from both country groups. The Advanced countries in the EU are referred to as the Western countries and the Transition countries as the Eastern. The figure 3 shows the estimated density of globalization measured by Euclidian distance in all member states of the EU in broad disciplines. Each density is calculated based on globalization scores between 2005 and 2017.

## Globalization in the European Union

The *life sciences* and *physical sciences* are highly globalized in almost all EU members, although the mean globalization is higher in the Western countries. The globalization also tends to be relatively stable in time. Of course, there are exceptions – see Romania or Poland – but these are exclusively in the East.

The differences between the East and the West seems to prevail in s*ocial sciences* and *health sciences*. They also tend to be globalized in Western countries. The gap between these disciplines and the *life* and *physical* sciences are usually small. The only exception are countries speaking with important scientific languages, but not English - France, Germany and Spain. In these countries, *social* and *health sciences,* are less globalized.

In the East, *social sciences* and *health sciences* are significantly less globalized. The globalization tends to grow, especially after 2010, but the gap between these disciplines in the West and the East is still significant in 2017.

This gap between East and West is depicted on figure 4 that breaks down the research output of the Austria, Bulgaria, Czech Republic, Denmark, France, Germany, Netherlands, Poland, Portugal, Slovakia and Spain by globalization of journals. Bulgaria, Czech Republic, Poland and Slovakia represent the *East* and the rest belongs to the *West*. Note that Germany, France, Portugal and Spain speak with important scientific languages.

|  |
| --- |
| **Figure 3: Density of globalization in the European Union** |
|  |
| *Note: Each box represents distribution of for all countries and years in given country. Source: own calculation, Scopus. Countries are sorted by .* |

First, all journals were divided into quartiles by their Euclidian distance globalization in 2017. In the second step all documents from the given country and discipline published in journals with computed globalization were assigned to the quartiles. Quartiles are marked Q1 – Q4, where Q1 are 25% of journals with highest globalization and Q4 with smallest. The darker the color on the figure, the higher the globalization.

The cross-country differences within *physical* and *life sciences* are relatively small. Approximately 40 % of all documents is published in journals with highest globalization (*Q1*) and additional 25-30% in the second quartile. There is slightly higher share of Q4 the East, especially in *life sciences*.

|  |
| --- |
| **Figure 4: Research output of selected EU countries divided into journals by globalization (2017, Euclidian distance)** |
|  |
| *Source: own calculation. Scopus* |

In Czech Republic, Poland and Spain almost 40 % of all documents is published in the Q4 journals in *social sciences*. The same figure is approx. 30 % in France and less than 20 % in Germany. In the Netherlands and Denmark only 3 % of all documents is published in Q4 journals. Not surprisingly, the larger portion of research is published in Q4, the less is published in the most globalized journals. Similar patterns can also be found in the *health sciences*.

## Globalization in Russia and China

Russia is a prime example of a strongly isolated research system (see figure 5). 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. Moreover, Russian science does not indicate any significant changes. More than 50 % of their research output is published in the Q4 journals in all broad disciplines. In *social sciences*, it is 74 %.

This is in sharp contrast with the case of China. Although at the beginning of the analyzed period, in 2005, China was the least globalized country in the world in *All disciplines* (see Moed 2002)*,* and among the 5 least globalized in all broad disciplines. Rapid transformation of the Chinese system resulted in the relatively fast growth of globalization. During the analyzed period, the globalization grew fast across all broad disciplines. In 2017, almost 30 % of their research output in All disciplines was published in Q1 journals.

|  |
| --- |
| **Figure 5: Globalization in broad disciplines in China and Russia in time (Euclidian distance)** |
|  |
| *Source: own calculation. Scopus* |

# Discussion and conclusion

The most important outcome of this paper is a rich dataset measuring the globalization of journal publications of researchers across countries, disciplines and in time. Data calculated from more than 34 964 Scopus-indexed journals are available on 174 countries, 27 narrow and 4 broad disciplines and 13 years between 2005 and 2017. The globalization is measured using 7 different indicators of globalization. For each country, discipline, year and indicator we report average globalization score and the distribution of the research output across quartiles. Globalization indicators of individual journals are also provided. See the appendix A2, A3 and A4 for details. Globalization scores can also be inspected in the [interactive app](http://www.globalizationofscience.com/).

The globalization is generally high (and the differences between disciplines small) in the countries of the “core” – in the Western Europe and North America. This finding is in line with Zitt et al. (1998) who announced “*almost complete transition from national to the transnational mode of communication*” already at the end of the last century.

The low globalization is symptomatic for the countries of the former soviet bloc, and especially for the former Soviet Union. The journals publishing authors from only a handful of countries (or even just one) are an important publishing platform, especially in the social sciences and health sciences. In Russia, Kazakhstan or Ukraine, the lack of globalization is present in almost all disciplines. In Russia, more than half of all publications in the dataset were published in 25 % least globalized journals (by Euclidian distance). The attempts to game the research evaluation system are commonly described (Moed et al. 2018, Rafols et al. 2016, Good et al. 2015, Kirchik et al. 2012). Much more research is needed to understand the purpose, quality, publishing standards and perhaps even ethics standards in local journals ecosystems in Eastern Europe.

The globalization of science gives information about a transformation of research outside the developed world. There is a striking difference between China, which gradually globalized all its broad disciplines and symptomatically low globalization in Russia. In Brazil, the globalization grows from 2010 and in India, it is not. In Indonesia, the globalization even seems to decline throughout the analyzed period. However, especially in developing countries, the data can be sensitive to the Scopus indexation decision.

The social sciences and health sciences are less globalized than life sciences and physical sciences. In combination with under-representation of socially scientific research in Scopus (Mongeon and Paul-Hus 2016) and an assumption that Scopus is more likely to index globalized than non-globalized journal, one might conclude that at least in case of social sciences, the true gap to life sciences can in many cases be even wider. However, the local journals often have strong incentives to get indexed in Scopus and sometimes they succeeded[[4]](#footnote-4).

It’s (almost) all about incentives (Franzoni et al. 2011). Countries are the most important factor explaining the variance of globalization. Disciplines are much weaker predictor. The incentives provided by national research evaluation schemes and the country research culture in general have much higher impact on resulting globalization than the discipline in which the research is done.

Especially in the environment where journals serve not only as a platform of scientific communication, but also as a major evaluation tool, journals are prone to gaming practices (Rijcke et al. 2016, Good et al. 2015, Rafols et al. 2016, Macháček and Srholec, 2017a). Arguably, local journals can be “*kidnapped*” to serve as a platform for “*cheap publications*” on expense of scientific quality[[5]](#footnote-5). Such practices could become widespread, especially in performance-based research systems with weak research evaluation culture extensively relying on simplistic metrics such as Impact Factor (see [DORA](https://sfdora.org/) or [Leiden Manifesto](http://www.leidenmanifesto.org/)).

Research internationalization is beneficial. Both the international competition and cooperation can spur innovation and lead to scientific breakthroughs (Adams 2013). However, science policy must consider the potential benefits and disadvantages linked to the international publishing requirements. More effort is needed to understand the impact of globalization on scientific visibility, collaboration and impact as well as the link of globalized research to the societal needs (Glaser and Laudel 2016, Chavarro et al. 2016). Hopefully, the data on globalization can be of use for research on the role of globalized and national journals in the 21st century.

It should be acknowledged, however, that there are concerns about the side-effects of ever-growing globalization of scientific communication (Chavarro et al. 2014, Chavarro et al. 2017). Local journals may offer a platform for research that would be not be published in “mainstream” journals - publications of junior researchers, publications bridging the knowledge between international science and local communities or topics that are overlooked by the mainstream research (Evans et al. 2014).

More research is needed to understand the results in more detail, focusing on some specific area or a discipline. Any detailed research must deal with the representativeness issue. The distortion is acceptable on a global scale that is described in this paper. However, Scopus might contain country-specific or discipline-specific niches that can drive part of the results. Scopus might not cover substantial part of targeted research at all (Rafols et al. 2014) or besides purely academic journals it might cover also professional journals (Macháček and Srholec 2017b). The more detail, the higher caution is needed (Waltman 2018).

# 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.

Chavarro (2017). Universalism and Particularism: Explaining the Emergence and Growth of Regional Journal Indexing Systems. PhD Thesis. SPRU – Science and Technology Policy Research Unit, University of Sussex

Chavarro, D., Tang, P., & Rafols, I. (2014). Interdisciplinarity and research on local issues: evidence from a developing country. Research Evaluation, 23(3), 195-209.

Chavarro, D., Tang, P., & Ràfols, I. (2017). Why researchers publish in non-mainstream journals: Training, knowledge bridging, and gap filling. Research policy, 46(9), 1666-1680.

Evans, J. A., Shim, J. M., & Ioannidis, J. P. (2014). Attention to local health burden and the global disparity of health research. PloS one, 9(4).

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

Garfield, E., & Merton, R. K. (1979). Citation indexing: Its theory and application in science, technology, and humanities (Vol. 8). New York: Wiley.

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.

Gläser, J., & Laudel, G. (2016). Governing science: How science policy shapes research content. European Journal of sociology/Archives Européennes de sociologie, 57(1), 117-168.

Good, B., Vermeulen, N., Tiefenthaler, B., & Arnold, E. (2015). Counting quality? The Czech performance-based research funding system. Research Evaluation, 24(2), 91-105.

Hicks, D., & Wang, J. (2011). Coverage and overlap of the new social sciences and humanities journal lists. Journal of the American Society for Information Science and Technology, 62(2), 284-294.

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. (2017a) Predatory Journals in Scopus. 2/2017. Institute for Democracy and Economic Analysis (IDEA).

Macháček. V. and Srholec. M. (2017b) Místní časopisy ve Scopusu (in Czech). 17/2017. Institute for Democracy and Economic Analysis (IDEA).

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. (2010). Measuring contextual citation impact of scientific journals. Journal of informetrics, 4(3), 265-277.

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.

Ràfols, I., Molas-Gallart, J., Chavarro, D. A., & Robinson-Garcia, N. (2016). On the dominance of quantitative evaluation in ‘peripheral’countries: Auditing research with technologies of distance. Available at SSRN 2818335.

Rijcke, S. D., Wouters, P. F., Rushforth, A. D., Franssen, T. P., & Hammarfelt, B. (2016). Evaluation practices and effects of indicator use—a literature review. Research Evaluation, 25(2), 161-169.

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>

Wang, Q., & Waltman, L. (2016). Large-scale analysis of the accuracy of the journal classification systems of Web of Science and Scopus. Journal of Informetrics, 10(2), 347-364.

Waltman, L., Tijssen, R. J., & van Eck, N. J. (2011). Globalisation of science in kilometres. Journal of informetrics, 5(4), 574-582.

Waltman, L. (2018). Responsible metrics: One size doesn't fit all. CWTS Blog, March 29th 2018. Available at: <https://www.cwts.nl/blog?article=n-r2s294>. Downloaded on February 20th, 2020.

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.

Whitley, R., & Gläser, J. (2007). The changing governance of the sciences. Sociology of the sciences yearbook, 26.

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.

Zhou, P., & Leydesdorff, L. (2006). The emergence of China as a leading nation in science. Research policy, 35(1), 83-104.

# Appendix

## Table A1: Correlation matrix of Globalization measured by various indicators

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| *Indicator* | *euclid* | *cosine* | *GiniSimpson* | *top3* | *instTOP3* | *sEnglish* | *localShare* |
| ***euclid*** | **1.00** | **.83** | **.87** | **.93** | **.81** | **.61** | **.75** |
| *cosine* | .83 | 1.00 | .64 | .75 | .69 | .47 | .41 |
| *GiniSimpson* | .87 | .64 | 1.00 | .72 | .67 | .64 | .78 |
| *top3* | .93 | .75 | .72 | 1.00 | .79 | .51 | .67 |
| *instTOP3* | .81 | .69 | .67 | .79 | 1.00 | .43 | .57 |
| *sEnglish* | .61 | .47 | .64 | .51 | .43 | 1.00 | .61 |
| *localShare* | .75 | .41 | .78 | .67 | .57 | .61 | 1.00 |

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

## Supplementary file Appendix A2: Mean globalization scores for all countries, disciplines, years and indicators

The CSV file is available at:

<https://github.com/vitekzkytek/GlobalizationPaper/blob/master/public_data/globalization_scores.csv>

## Supplementary file Appendix A3: Distribution of research output into journals by globalization (Euclidian distance) in broad disciplines

The CSV file is available at:

<https://github.com/vitekzkytek/GlobalizationPaper/blob/master/public_data/globalization_quartiles.csv>

## Supplementary file Appendix A4: Individual journal globalizations (Euclidian distance, broad set of disciplines, 2017)

Note that journals with multiple disciplines are computed for each discipline. The results will differ for journals with multiple disciplines for benchmark-based indicators.

The CSV file is available at:

<https://github.com/vitekzkytek/GlobalizationPaper/blob/master/public_data/globalization_journals.csv>

## Appendix A5: IMF (2003) country classification

Advanced countries: Andorra, Australia, Austria, Belgium, Canada, Cyprus, Denmark, Finland, France, Germany, Greece, Hong Kong, Iceland, Ireland, Israel, Italy, Japan, Liechtenstein, Luxembourg, Monaco, Netherlands, New Zealand, Norway, Portugal, San Marino, Singapore, South Korea, Spain, Sweden, Switzerland, Taiwan, United Kingdom, United States

Developing countries: Afghanistan, Algeria, Angola, Antigua and Barbuda, Argentina, Bahamas, Bahrain, Bangladesh, Barbados, Belize, Benin, Bhutan, Bolivia, Botswana, Brazil, Brunei, Burkina Faso, Burundi, Cambodia, Cameroon, Cape Verde, Central African Republic, Chad, Chile, China, Colombia, Comoros, Congo, Costa Rica, Cote d'Ivoire, Cuba, Democratic Republic Congo, Djibouti, Dominica, Dominican Republic, Ecuador, Egypt, El Salvador, Equatorial Guinea, Eritrea, Ethiopia, Fiji, Gabon, Gambia, Ghana, Grenada, Guatemala, Guinea, Guinea-Bissau, Guyana, Haiti, Honduras, India, Indonesia, Iran, Iraq, Jamaica, Jordan, Kenya, Kiribati, Kuwait, Laos, Lebanon, Lesotho, Liberia, Libya, Madagascar, Malawi, Malaysia, Maldives, Mali, Malta, Marshall Islands, Mauritania, Mauritius, Mexico, Micronesia, Morocco, Mozambique, Myanmar, Namibia, Nauru, Nepal, Nicaragua, Niger, Nigeria, North Korea, Oman, Pakistan, Palau, Palestine, Panama, Papua New Guinea, Paraguay, Peru, Philippines, Qatar, Rwanda, Saint Kitts and Nevis, Saint Lucia, Saint Vincent and the Grenadines, Samoa, Sao Tome and Principe, Saudi Arabia, Senegal, Seychelles, Sierra Leone, Solomon Islands, Somalia, South Africa, Sri Lanka, Sudan, Suriname, Swaziland, Syria, Tanzania, Thailand, Timor-Leste, Togo, Tonga, Trinidad and Tobago, Tunisia, Turkey, Tuvalu, Uganda, United Arab Emirates, Uruguay, Vanuatu, Venezuela, Vietnam, Yemen, Zambia, Zimbabwe

Transition countries: Albania, Armenia, Azerbaijan, Belarus, Bosnia and Herzegovina, Bulgaria, Croatia, Czechia, Estonia, Georgia, Hungary, Kazakhstan, Kyrgyzstan, Latvia, Lithuania, Macedonia, Moldova, Mongolia, Montenegro, Poland, Romania, Russia, Serbia, Slovakia, Slovenia, Tajikistan, Turkmenistan, Ukraine, Uzbekistan

1. \* Financial support from the Czech Academy of Sciences for the R&D&I Analytical Centre (RaDIAC) and from the Charles University Grant Agency (GA UK; project no 1062119) is gratefully acknowledged, as well as Elsevier for providing extended access to Scopus API. Earlier versions of the paper were presented at CWTS QSS Internal Meeting and on ISSI 2019 conference in Rome. I thank the participants at these events for their useful comments and suggestions. I also thank Martin Srholec, Ludo Waltman and Nees Jan van Eck for their guidance in the process of writing. All usual caveats apply. [↑](#footnote-ref-1)
2. For example, the mean value of is 0.94, because the distribution of the score is skewed towards 100 %. [↑](#footnote-ref-2)
3. This section only captures the most fundamental world trends of globalization. The reader is encouraged to take advantage of the interactive application and supplementary data to understand the globalization in the full detail. [↑](#footnote-ref-3)
4. Macháček and Srholec (2017b) describe on the Czech and Slovak journals cases of notoriously local journals, even though Scopus includes *International diversity of authors* within their indexation criteria*.* [↑](#footnote-ref-4)
5. Good et al. (2015) for example 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. [↑](#footnote-ref-5)