MAX PLANCK INSTITUTE
FOR DEMOGRAPHIC RESEARCH

Konrad-Zuse-Strasse 1 · D-18057 Rostock · Germany · Tel +49 (0) 3 81 20 81 - 0 · Fax +49 (0) 3 81 20 81 - 202 · www.demogr.mpg.de

MPIDR Working Paper WP 2020-025 I May 2020
Revised August 2020
https://doi.org/10.4054/MPIDR-WP-2020-025

Brain Drain and Brain Gain in Russia: Analyzing International Migration of Researchers by Discipline using Scopus Bibliometric Data 1996-2020

Alexander Subbotin I subbotin@demogr.mpg.de
Samin Aref I aref@demogr.mpg.de

This working paper has been approved for release by: Emilio Zagheni (sekzagheni@demogr.mpg.de), Head of the Laboratory of Digital and Computational Demography.

Brain Drain and Brain Gain in Russia: Analyzing International Migration of Researchers by Discipline using Scopus Bibliometric Data 1996-2020

Alexander Subbotin[1,2][0000−0001−5016−0473] and
Samin Aref[2][0000−0002−5870−9253]

[1] Lomonosov Moscow State University, ul. Leninskiye Gory, 1
119991 Moscow, Russia
subbotin@demogr.mpg.de
[2] Laboratory of Digital and Computational Demography,
Max Planck Institute for Demographic Research,
Konrad-Zuse-Str. 1, Rostock 18057, Germany
aref@demogr.mpg.de

Abstract. We study international mobility in academia with a focus on migration of researchers to and from Russia. Using all Scopus publications from 1996 to 2020, we analyze bibliometric data from over half a million researchers who have published with a Russian affiliation address at some point in their careers. Migration of researchers is observed through the changes in their affiliation addresses. For the first time, we analyze origins and destinations of migrant researchers with respect to their fields and performance and compute net migration rates based on incoming and outgoing flows. Our results indicate that while Russia has been a donor country in the late 1990s and early 2000s, it has experienced a relatively symmetric circulation of researchers in more recent years. Using subject categories of publications, we quantify the impact of migration on each field of scholarship. Our analysis shows that Russia has suffered a net loss in almost all disciplines and more so in neuroscience, decision sciences, dentistry, biochemistry, and mathematics. For economics and environmental science, there is a relatively balanced circulation of researchers to and from Russia. Our substantive results reveal new aspects of international mobility in academia and its impact on a national science system which speak directly to policy development. Methodologically, our new approach of handling big data can be adopted as a framework of analysis for studying scholarly migration in other countries.

Keywords: High-skilled migration · Bibliometric data · Computational demography · Science of science · Scientometrics.

1 Introduction

In the interconnected world, most national science systems cannot be studied in a vacuum disregarding the impact of human mobility and migration. Countries are indeed affected by international migration of the highly skilled specialists including researchers. Our era has witnessed a large increase in high-skilled migration between countries, which poses new challenges both for researchers and policy makers. Russia is not an exception in the global international migration system: a large part of its population is actively on the move for various reasons [14]. Russia is also an attractive destination for some international migrants, especially migrants from former Soviet Union countries [6,30]. Moreover, some migrants may consider Russia as a transit stop for further migration to other countries [32]. Previous studies suggest that Russia is both a donor country and a recipient country [11,29] for migration in general. If the characteristics of migrants are taken into account, Russia is suggested to be more of a donor country [36,39,17], i.e. a country on the losing side of an international exchange of highly skilled individuals.

The number of researchers and their outputs in Russia are perhaps not as well-known as those of other developed countries. According to SciVal 2010-2019 data, Russia has over 440,000 researchers (comparable to Australia and Italy) who have produced nearly 700,000 pieces of scholarly publications (comparable to South Korea). Despite these features, Russia has been a relatively under-studied case in the scientometrics literature. Most studies on this topic are limited to qualitative explanations on the emigration of specialists which often do not go beyond suggesting the necessity of facilitating circular migration for Russia [38,15,24,17,31,37,36,27,34]. Therefore, a deeper analysis is needed to quantitatively study the international movements of researchers in Russia and its implications for different fields of science.

According to previous studies, a large number of scientists in mathematics [37,31], physics [5,31], and computer science [31,3] leave Russia. The major destination countries for the scholars from Russia are suggested to be the United States (US), Germany, France, the United Kingdom (UK), and Japan [18]. The movers are more often from major scientific centers in Moscow, St. Petersburg, Novosibirsk, and Yekaterinburg, and come from lower age groups [7,13], who otherwise have the potential of contributing to the Russian science system for a long time. Further research is needed to accurately quantify this phenomenon with respect to similarities and differences between all migrant researchers, their origin and destination countries and the interplay of their mobility patterns, level of experience, and research performance, and the impact on Russia.

Quantitative studies on international migration of researchers seem to be complicated by a lack of reliable, relevant, and comparable statistics. Recent studies on this topic use bibliometric data to detect migrant populations among researchers and obtain migration trajectories and flows for further analysis [23,4,19,22]. This method involves tracking the international movements of researchers through the changes in the affiliation addresses. The feasibility of this approach has been tested in previous studies that estimated migration flows among scholars

[23,4,19,22]. In this study, we adopt such a method for focusing on researchers who have published with a Russian affiliation address at some point in 1996-2020. We track the international movements of all such researchers to analyze the impact of migration on the Russian science system overall and in different fields of research.

2 Materials and Methods

2.1 Scopus publications of all authors with ties to Russia

The availability of millions of publications in the Scopus database (the largest database of peer-reviewed literature [12,25]), allows us to study scholarly migration in Russia by aggregating movements of each researcher who has affiliation ties to Russia at some point in 1996-2020 period (up to the end of April 2020). The unit of data is an *authorship record* which is the linkage between an author affiliation and a publication. The data linked to an authorship record provide proxies not only for the geographic locations of researchers, but also for their research areas. Scopus annotates subject codes to more than 25000 indexed publication venues based on the topics they cover. This allows us to analyze the disciplines of internationally mobile researchers based on the subjects of their publications.

There are more than 2 million publications in Scopus from over 659'000 individual authors who have published with a Russian address at some point over the 1996-2020 period. After retrieving this data, we focus on the scholars who also have countries of affiliation other than Russia in their publications. This step excludes those researchers who do not have any evidence of international mobility and authors who only have one publication. Given that migration is a rare event, the subset of the data we mostly focus on would be authorship records associated with 522'000 publications from nearly 30'000 internationally mobile researchers.

2.2 Data pre-processing

The Scopus author ID [16,2] allows us to identify authorship records of individual scholars and accordingly detect mobility events. However, there are data quality issues with Scopus author IDs and affiliations [22] which require some attention before movements can be detected. The affiliations are not standard, and they may have substantially differ formats. In a large majority of cases, an affiliation address has a country while there are 9'701 authors in our dataset who have records without a country. These come from 7'279 distinct publications. Inspired by [22], we use a neural network to predict the missing country information. The neural network takes affiliation address of an authorship record and predicts the country associated with it. For technical details of the development of such an algorithm, one may refer to [22]. We use 1 million records which have countries as training data (80%) and test data (20%). On 98.4% of the test data, the

neural network predicts the expected country. Ensuring the high accuracy of this method, we use the trained neural network for predicting missing countries.

Scopus author identification system is suggested to be reliable for analyzing migration of researchers [2] as most author IDs correctly identify one researchers. However, the Scopus author identification is not perfect: there may be several different individuals with the same name (or similar names) who are incorrectly assigned the same author ID. We approach this problem by applying an author disambiguation process [22] on the authorship records which are more likely to be impacted by the lack of accuracy in the Scopus author identification system. These records are selected from the extreme values in number of countries and number of publications. Authorship IDs which exceed either of two thresholds below are deemed suspicious and will be treated by an author disambiguation method. Threshold 1: being associated with more than 6 countries of affiliation. Threshold 2: being associated with more than 292 publications (an average of more than one publication per month across a period of 24 years and 4 months).

Among more than 659'000 distinct author IDs in our data, 3'563 author IDs are deemed suspicious. They are associated with 334'484 distinct publications (some publications are shared between them). We disambiguate these records using an unsupervised machine learning algorithm [22] inspired by the state-of-the-art methods in the literature [10] and assign revised author IDs using the method briefly described below. The idea behind the author disambiguation algorithm that we use is making pairwise comparisons between every two records with the same author ID and allocating scores which are higher if the two authorship records share similar traits and lower if they are dissimilar. Then, the scores are summed up and a distance matrix is calculated for all pairs of authorship records. Using agglomerative clustering from the scikit-learn package in Python [28], we obtain clusters of highly similar authorship records. Finally, a revised author ID is issued to each cluster [22]. Implementing this author disambiguation method to the subset of 3'563 suspicious author IDs, leads to 11'833 revised author IDs.

2.3 Four fields and 26 sub-fields of scholarship

According to All Science Classification Codes (ASJC), there are four major fields of science: *life sciences* (including five sub-fields[3]), *social sciences* (which includes six sub-fields[4]), *physical sciences* (including ten sub-fields[5]), and *health*

[3] (1) agricultural and biological sciences (2) biochemistry, genetics and molecular biology (3) immunology and microbiology (4) neuroscience and (5) pharmacology, toxicology and pharmaceutics

[4] (1) arts and humanities (2) business, management and accounting (3) decision sciences (4) economics, econometrics and finance (5) psychology and (6) other social sciences

[5] (1) chemical engineering (2) chemistry (3) computer science (4) earth and planetary sciences (5) energy (6) engineering (7) environmental science (8) materials science (9) mathematics and (10) physics and astronomy

sciences (which includes five sub-fields[6]). Each publication venue in Scopus is classified by possibly multiple ASJC codes which determine the fields and sub-fields of the topics they cover. At the level of ASJC four major fields of science, we consider that researchers can either belong to one of the four fields or they are multidisciplinary.

We initially compute the frequency (f) of each of the four major fields in the authorship records of each researcher. Then we calculate four Z-scores for each researcher based on the mean (μ) and standard deviation (σ) of frequencies of each major field using $Z = (f - \mu)/\sigma$. Based on the largest Z-score which exceeds $\alpha = 0.25$[7], we group the researchers to one of the four groups of health, life, physical, and social sciences. For 10% of researchers, neither of the Z-scores exceed the threshold of α, and we group them as *multidisciplinary*.

2.4 Detecting moves and types of migrant researchers

For analyzing scholarly migration, we borrow well known and fundamental concepts such as origin, destination and migrant from migration studies, and repurpose them for usage in an academic sense. Accordingly, a country of academic origin is the country appearing in the first publication of a researcher, while the destination country is determined by the most recent country affiliation. To refer to a researcher who have had an international move we use the term academic migrant (or migrant for brevity). We consider an international mobility event if the changes in affiliations across two different years are such that the mode of affiliation country changes for a researcher. We define four categories for academic migrants based on their countries of academic origin and destination. In our analysis, each migrant belongs to one of four categories as follows:
(1) *Immigrant* (origin: not Russia, destination: Russia),
(2) *Emigrant* (origin: Russia, destination: not Russia),
(3) *Return migrant* (origin: Russia, destination: Russia),
(4) *Transient* (origin: not Russia, destination: not Russia).

2.5 Quantifying contributions of researchers by sub-field

At the level of ASJC 26 sub-fields (disciplines), we consider that researchers are potentially active in and contributing to several of them. Therefore, we use *normalized contribution* to quantify the contribution of a given researcher to different fields in a normalized way. The normalized contribution $NC^{j}_{(d)}$ of researcher j in discipline d (among a total of n disciplines) is defined and formulated in Eq. (1) based on the relative frequency of discipline d in their authorship records. s^{j}_{d} is the frequency of discipline d in the authorship records of individual j. The denominator in Eq. (1) is the sum of frequencies of all disciplines in the authorship records of individual j.

[6] (1) medicine (2) nursing (3) veterinary (4) dentistry and (5) health professions

[7] Value of α is selected such that only 10% of researchers become multidisciplinary. Stricter limits based on a larger α lead to clearer boundaries between the four main fields and more individuals belonging to the multidisciplinary group.

$$NC^{j}_{(d)} = \frac{s^{j}_{d}}{\sum_{i=1}^{n} s^{j}_{i}} \quad i = 1, \ldots, n \quad \forall j \in \{1, \ldots, k\} \tag{1}$$

As an illustrative example, consider that the authorship records of an individual over their career are as provided in Table 1.

Table 1. Example authorship records with multiple subjects and countries

Author ID	DOI	ASJC Subject	Country	Year
X	22222	Mathematics	Russia	2012
X	33333	Chemistry, Energy	Russia, US	2013
X	44444	Mathematics, Chemistry	US	2015

Distinct digital object identifiers (DOIs) in Table 1 show that these authorship records are associated with three distinct publications. The normalized contributions of the researcher, who is identified by Author ID X, are

$$NC^{X}_{(\text{chemistry})} = 2/5, \quad NC^{X}_{(\text{energy})} = 1/5, \quad NC^{X}_{(\text{mathematics})} = 2/5.$$

To aggregate the normalized contributions for a discipline, the *normalized count* of migrants in discipline d can be used which is calculated by adding up all the normalized contributions of mobile researchers for discipline d as formulated in Eq. (2). The normalized count of migrants in discipline d, denoted as P_d, can be thought of as a weighted sum for the population of internationally mobile researchers in discipline d normalized based on giving fractional weights to individuals depending on how active they are in that discipline compared to their other disciplines. If the result for P_d is decimal we use arithmetic rounding.

$$P_d = \sum_{j=1}^{k} NC^{j}_{(d)} \tag{2}$$

Given that each mobile researcher belongs to one of the four categories of migrants, normalized counts can similarly be computed based on the normalized contributions associated with each type of migrant. Accordingly, we obtain $P^{\text{imm}}_{d}, P^{\text{emi}}_{d}, P^{\text{ret}}_{d}, P^{\text{tra}}_{d}$ respectively as normalized populations of immigrants, emigrants, return migrants, and transients in discipline d.

3 Results

We present the main results of our analysis in this Section which is structured as follows: Subsection 3.1 outlines the analysis of the geography of mobile researchers (common origin and destination countries). Subsection 3.2 concerns the origin and destination countries with respect to research performance normalized by age. Subsection 3.3 presents our estimates of net migration rates (to

evaluate brain circulation in Russia overall and by major fields). Subsection 3.4 explores disciplines of mobile researchers (to evaluate the impact of migration on each field of science in Russia).

3.1 Flows, origins, and destinations

Figure 1 illustrates the international paths for researchers to and from Russia over the 1996-2020 period. The five most common countries of academic origin for immigrants are US, Ukraine, Germany, France, and UK respectively. As destinations for emigrants US and Germany are again the most common countries respectively, followed by UK and France, while Ukraine is ranked the fifth among common destinations. Moreover, the scale of emigration to frequent destinations is more than twice the scale of immigration from frequent origins. We can see in Figure 1 that US and Russia are connected by two edges (whose directions are clockwise): blue (scholars moving from US to Russia) and pink (scholars moving from Russia to US). The pink edge is thicker than the blue one, which means that the researchers leaving Russia for US outnumber people coming from US to Russia. In this context, US, Germany, UK, and France are more likely to be destinations than being origins with respect to the imbalanced flows of immigrants and emigrants.

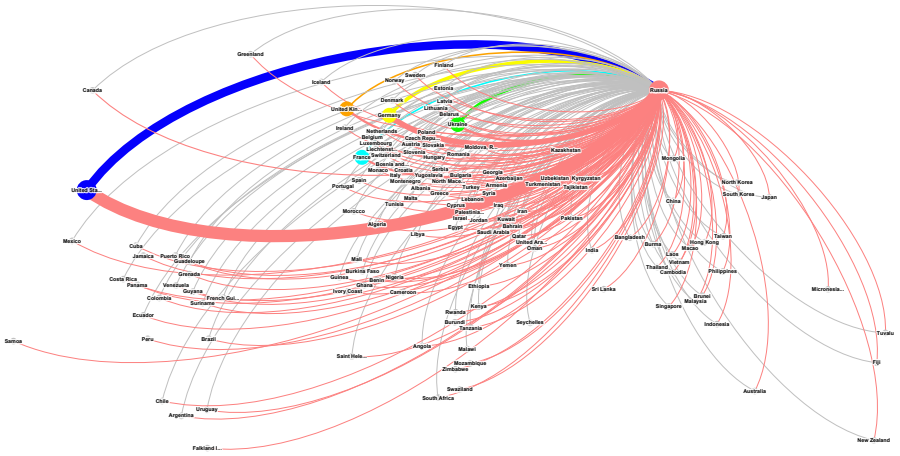


Fig. 1. Network of movements to and from Russia among researchers over 1996-2020. Directions of edges are clockwise. Common origins and destinations are shown with distinct colors. Colors of the flows are based on the origin country. Thickness of an edge is proportional to the flow it represents. See the figure on screen for high resolution.

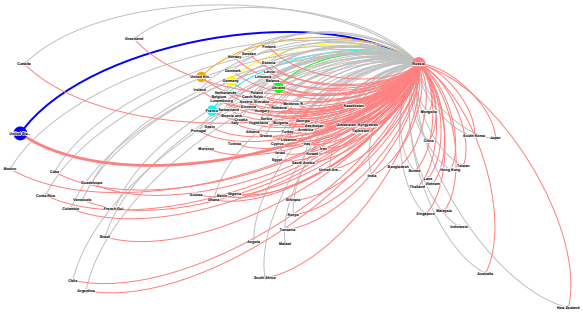
An exception among the top countries of origin is Ukraine with fewer emigrants from Russia than immigrants to Russia. The number of immigrants from

Ukraine to Russia is 2.5 times larger than the number of emigrants from Russia to Ukraine. So, Ukraine is more likely to be an origin than a destination from the perspective of the Russian science system. In the general population, the Russo-Ukrainian migration relationship has been substantial, and for a long time Russia has been the major destination for the Ukrainian migrants [8,26]. Also, from April 2014 to February 2016 more than one million people migrated from Ukraine to Russia following the 2014 events in Ukraine [24]. Canada, Finland, Sweden, the Netherlands, and China are more likely to be destinations than origins similar to Czech Republic, Austria, Norway, Spain, and South Korea. Canada and Finland are three times more likely to be a destination country. Also Switzerland and Australia are common academic destination countries without being among the top 15 origin countries.

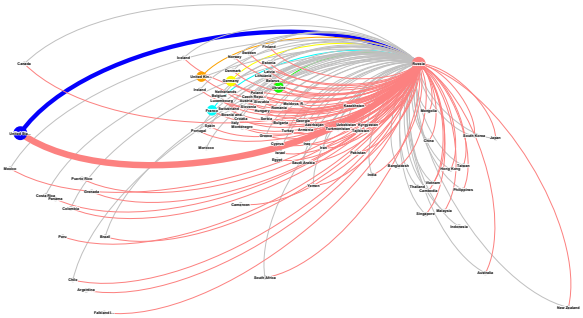
Contrary to this, Belarus and Uzbekistan are more likely to be a countries of origin rather than destination. The migration patterns of academics from Belarus and Uzbekistan might be explained by the historical trend in patterns of general migration which shows that Russia has been a primary destination for the migration flows from these countries [35,6]. Poland has almost an equal number of emigrants and immigrants which makes it equally likely to be an origin as well as a destination.

The rankings of most common origins and destinations for transient scholars, in general, match with minor differences. The number of outgoing transients from Russia to China, Switzerland, Poland, Sweden, the Netherlands, Kazakhstan, US, and Germany are higher than the respective number of incoming transients from these countries to Russia.

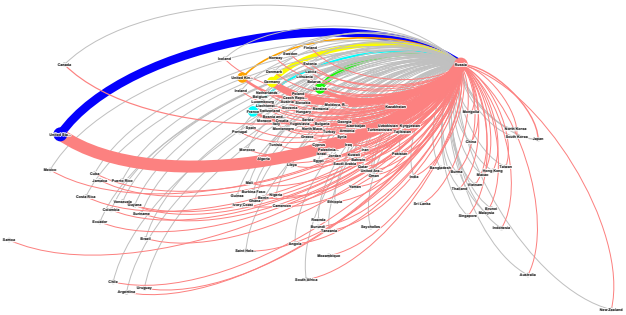
As for return migrants, we look at the intermediate country(ies) that a given return migrant has been affiliated with while being temporarily away from Russia. The five most common intermediate countries for return migrants are US, Germany, France, UK, and Ukraine which match the ranking of destinations for emigrants, with the exception of ranks of France and UK being swapped.Figure 2 shows the migration flows disaggregated by the four major fields of scholarship.



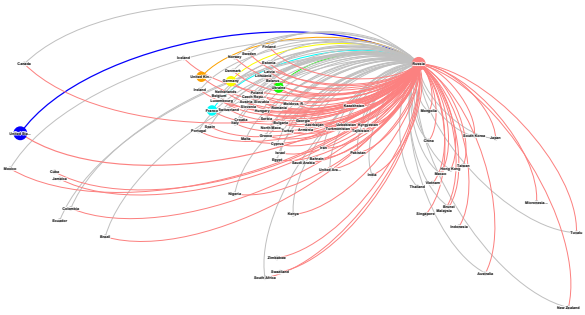
(a) Health sciences



(b) Life sciences



(c) Physical sciences



(d) Social sciences

Fig. 2. Migration flows among researchers in four major fields. Colors of the flows are based on the origin country. See the figure on screen for high resolution.

It can be seen in Figure 2 that Physical sciences has generally the largest flows followed by life sciences, health sciences, and social sciences in decreasing order. The top five destination countries in Figure 2 are among US, Germany, UK, Ukraine, France, and Kazakhstan while their order sometimes changes depending on the major field of science. US is always the most common destination. For life and health sciences, the second country is Germany followed by UK in the third place while for social sciences their order is reversed. For physical sciences, the second and third common destinations are Germany and Ukraine respectively. For US, Germany, UK, and France, the flow to Russia (disaggregated by the major field) is smaller than the respective flows in the opposite direction. However, in social sciences, the flows from US and UK to Russia are larger than the respective flows from Russia to these two countries. For Ukraine, in all four major fields, the flows to Russia is larger than the respective flows from Russia.

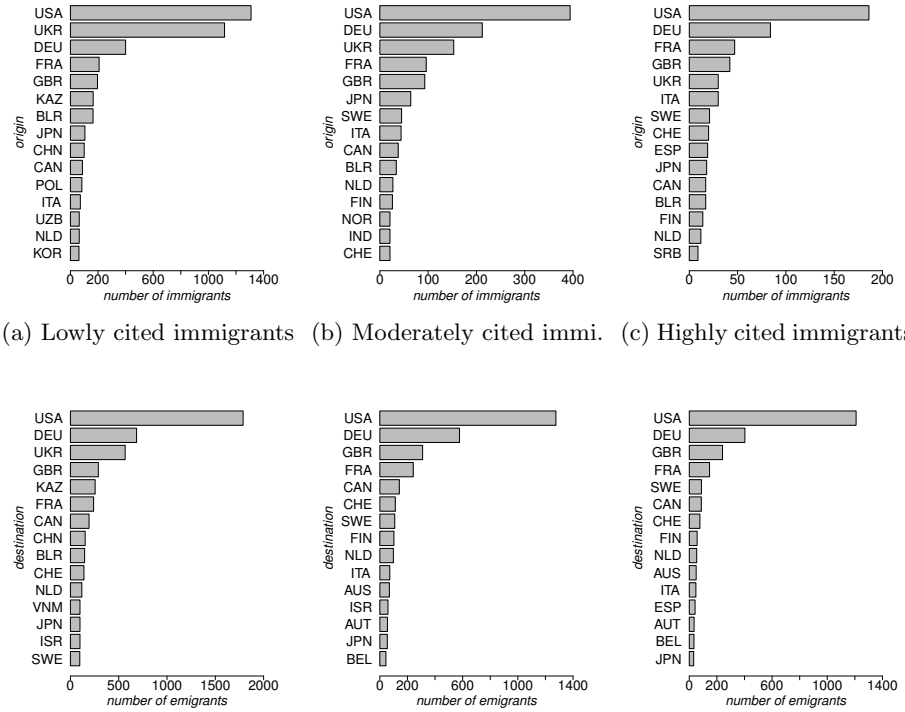
3.2 Geography by citation-based performance

We also look at common origins and destinations taking into account citations and academic age (the number of years since first publication [4]) of researchers. We calculate an annual citation rate by dividing the a researcher's total number of citations (as of 2020) by their academic age. It should be noted that there are disparities in citations between immigrants and emigrants and by fields of science. In table 3.2, the averages and standard deviations of annual citation rates are provided. We can see that the rates for emigrants are generally higher than those of the immigrants, with the exception of social sciences. This suggests that in most major fields, internationally mobile researchers who come to Russia perform lower than those leaving Russia in terms of total citations received controlling for the differences in years of academic experience (academic age).

Table 2. Average and standard deviation of annual citation by field and migrant type

	Life sci.	Social sci.	Physical sci.	Health sci.	Multidisciplinary
Immigrants	37.9 ± 83.3	10.3 ± 27.7	60.3 ± 280.5	26.6 ± 61.0	43.9 ± 139.0
Emigrants	66.6 ± 152.9	8.7 ± 15.2	103.8 ± 324.3	72.6 ± 168.9	75.7 ± 260.3

Among fields of science, there are substantial variations in citations patterns (the mean values are reported in the brackets): life sciences (58.5), social sciences (9.45), physical sciences (89.8), health sciences (55.9), and multidisciplinary (69.0). To obtain a citation-based measure of performance, annual citation of a researcher should be divided by the average of their field. After normalization, we identify three groups of migrants: *lowly cited migrants* (the field-normalized annual citation rate is less than 0.09); *moderately cited migrants* (the field-normalized annual citation rate is between 0.09 and 0.52); *highly cited migrant* (the field-normalized annual citation rate is above 0.52). Figure 3 shows the common origins of immigrants and the common destinations of emigrants categorized by citation-based performance.



(a) Lowly cited immigrants (b) Moderately cited immi. (c) Highly cited immigrants

(d) Lowly cited emigrants (e) Moderately cited emig. (f) Highly cited emigrants

Fig. 3. Top 15 origins for lowly (a), moderately (b), and highly cited (c) immigrants, and top 15 destinations for lowly (d), moderately (e), and highly cited (f) emigrants.

In Figure 3, it can be seen that US is on the lead as the most common origin of immigrants. Germany is more common among moderately and highly cited immigrants, while Ukraine is a more common origin among lowly cited immigrants (Subfigures 3a-3c). We can also see that some former Soviet Union countries (Ukraine, Kazakhstan, Belarus, Uzbekistan) appear in the top 15 origins for lowly cited immigrants, while only Ukraine and Belarus also appear among the most common countries of immigrants performing better with respect to citations.

US and Germany are the two most common destinations for all emigrants regardless of their citation-based performance (Subfigures 3d-3f). Ukraine, Kazakhstan, and Belarus are among the top destinations of lowly cited emigrants while none of these countries is common for moderately or highly cited emigrants. UK, France, and Canada also appear as common destinations for all categories of emigrants.

3.3 Net migration rates

Migration rates are commonly used measures of the difference between movements into and out of a certain area [21]. Net migration rate for a given area refers to the difference between in-migration and out-migration rates per 1000 people. A positive value means more people entering than leaving a given area. Using I_y and E_y to represent the number of scholars who have immigrated to Russia and emigrated from Russia respectively during year y, and M_y to represent Russia's population of scholars at the beginning of year y, the net migration rate NMR_y can be calculated according to Eq. 3.

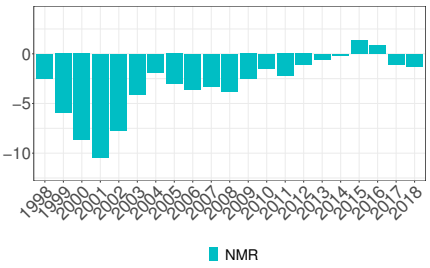
In-migration and out-migration rates can be computed based on I_y/M_y and E_y/M_y respectively which only take one direction of the flows into consideration. In Eq. 3, the denominator, M_y, is obtained from the original superset of our bibliometric data (which includes non-movers as well). It estimates the total number of researchers in Russia in year y based on the affiliation addresses associated with publication dates within a two year vicinity of year y. For this estimation, we assume that the researchers with Russian addresses have been in the country two years before and two years after the publication year unless there are evidence to the contrary (publications showing other countries for the researcher).

$$NMR_y = (I_y - E_y) \times 1000/(M_y) \tag{3}$$

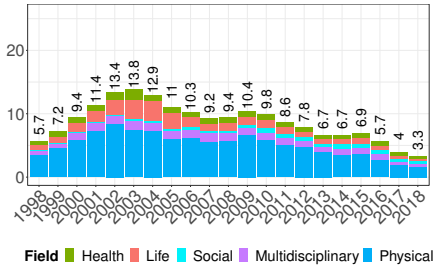
Subfigure 4a illustrates net migration rate in Russia over the 1998-2018 period. The lowest value of net migration rate is observed for 2001, and it was -10.5 per 1000 researchers. Looking at the in-migration rate in Subfigure 4b and the out-migration rate in Subfigure 4c for year 2001, one may see that in Russia 11.4 per 1000 researchers have migrated to Russia, and 21.9 per 1000 researchers have left Russia resulting in a negative net flow of 10.5 per 1000 researchers from Russia to the outside of Russia in year 2001. Subfigure 4a shows that over time, the net migration rate has generally increased. The highest rate is $+1.4$ which is for year 2015. Starting from 2015 we can see a slow downward trend in the net migration rates which ends with the value -1.3 for year 2018. Subfigures 4b- 4c also show the shares of a each major field of science in both in- and out-migration rates suggesting relatively similar compositions for emigrants and immigrants being mostly made up of researchers from physical sciences.

3.4 Impact of migration on each discipline

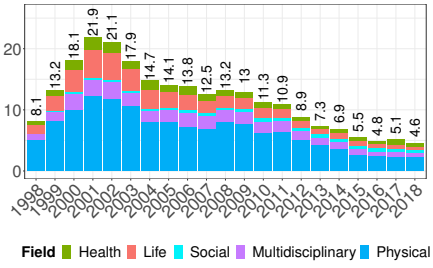
We further analyze movements of migrant scholars with respect to the 26 subfields of ASJC assocaited with their authorship records to find the the impact of migrations on different disciplines. We develop a measure inspired by net migration rate for each discipline to find the extent to which a given discipline in Russia is impacted by the imbalance of incoming and outgoing flows. To operationalize this idea, we start with the concepts of normalized contribution and normalized count formulated in Eqs. (1) and (2) discussed in Section 2. We evaluate the possible losses in each field by looking at the relative difference



(a) Net migration rate



(b) In-migration rate



(c) Out-migration rate

Fig. 4. Net migration (a), in-migration (b) and out-migration (c) rates per 1000 researchers in Russia over the 1998-2018 period

between the normalized counts of immigrants, return migrants, emigrants, and transients using a parsimonious measure to quantify *Field-based net brain drain* ($FNBD_d$) formulated in Eq. 4. Emigrants and transients increase the net drain of a national science system and therefore they have positive coefficients in Eq. 4. In contrast, we consider negative coefficients for immigrants and return migrants in Eq. 4 because these groups of migrants decrease the net drain of a national science system. A larger positive value of $FNBD_d$ means a larger loss due to the imbalance of migration flows in discipline d. The largest (smallest) value possible for $FNBD_d$ is 1 (-1) which is associated to a hypothetical situation where all migrants in discipline d are emigrants and transients (immigrants and return migrants) and the brain drain (brain gain) is therefore at its peak. Each term of Eq. 4 represents the impact of the respective group of migrants.

$$FNBD_d = (P_d^{\text{emi}}/P_d) + (P_d^{\text{tra}}/P_d) - (P_d^{\text{imm}}/P_d) - (P_d^{\text{ret}}/P_d) \tag{4}$$

To illustrate the application of $FNBD_d$, we use the discipline of computer science as an example. We obtain contributions of migrants to different fields using Eq. (1) and sum up the normalized contributions in the field of computer science for all four types of migrants (calculating $P_d^{\text{imm}}, P_d^{\text{emi}}, P_d^{\text{ret}}$ and P_d^{tra}) and all migrants together (calculating P_d) using Eq. 2. Accordingly, the normalized count of mobile researchers in computer science would be $P_d = 1329$ which includes all four types of migrants. Then, we use the formula in Eq. 4 to calculate

$FNBD$ for computer science which is equal to 0.135. The is interpreted as overall migration of researchers in Russia over the period 1996-2020 leading to to a 13.5% net drain in the field of computer science. Figure 5 shows the four terms of $FNBD$ and its total value for all disciplines. We can see in Figure 5 that Russia

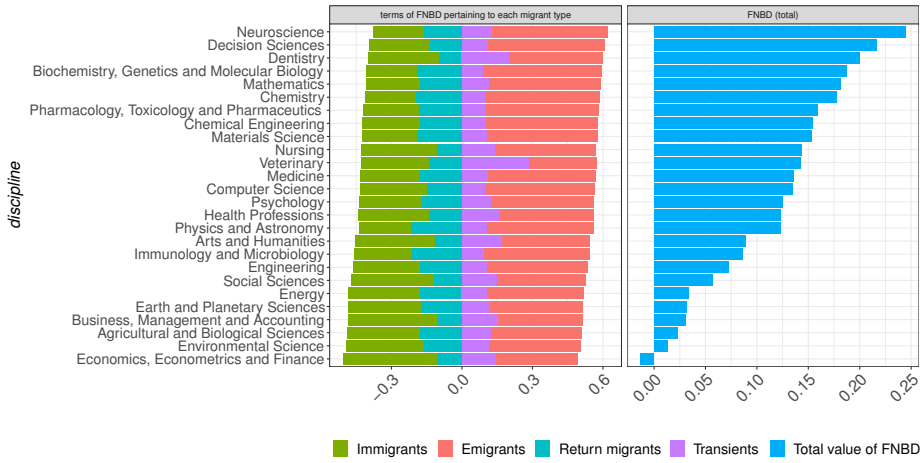


Fig. 5. Field-based net brain drain for different categories of migrants

has suffered a large loss in disciplines such as neuroscience (24.5%), decision sciences (21.6%), dentistry (20.0%), biochemistry (18.7%), mathematics (18.2%), chemistry (17.7%), pharmacology (15.9%), chemical engineering (15.4%), and materials science (15.3%). For almost all other disciplines, $FNBD$ values show a loss, but to a smaller degree. Interestingly, the values of $FNBD$ is close to zero suggesting a relatively balanced circulation of flows in economics (-1.3%) and environmental science (1.3%).

The results indicate that there is heterogeneity in the impact of international mobility of researchers on the fields of scholarship in Russia. This observation casts doubt on viewing a national science system as just one unit with a simple positive/negative response to international mobility, instead, the components of such a system could be differently impacted by the balance of migration flows or lack thereof. Note that, if we only consider terms of $FNBD$ pertaining to immigrants and emigrants, the alternative measurements also show Russia suffering a net loss in all disciplines because emigrants outnumber immigrants (an exception is economics that has the opposite pattern). However, return migrants and transients account for considerable proportions of migrations and therefore excluding them from the analysis could be contestable.

4 Limitations

A major limitation of this study, as well as a remarkable merit, is the use of bibliometric data, and the unique view they provide. The time required to conduct and publish research is an important factor [9] to keep in mind when interpreting the temporal component of the results on mobility patterns observed through bibliometric data. Also, it is noted that a one-time usage of an affiliation is not guaranteed to show a direct attachment to the country of affiliation [20]. Our conservative approach resolves this issue by detecting international moves only if the modal country of affiliation changes across different years. Moreover, we cannot observe and track any migration events that are not represented in publications indexed in Scopus. Bibliometric databases could be biased, and there could be an under-representation of some countries, scientific fields, and languages [12,25,33]. Also, given the fact that we are investigating a specific period of time, our data suffers from left-truncation.

Despite these limitations, bibliometric data facilitate a study on migration of researchers to cross the disciplinary boundaries and become more contemporary and extensive compared to what traditional data sources may allow [1]. This study makes several contributions to the literature: both methodological contributions in usage of bibliometric data and substantive contributions to the study of scholarly migration in Russia. A missing piece of the puzzle for understanding academic brain drain has been key migration statistics which our study provides for the first time for Russia; a commonly debated country of brain drain despite being under-explored by quantitative analysis.

5 Discussion and Summary

In this study, we used affiliation addresses from Scopus publications over 1996-2020 to present a comprehensive and detailed picture of academic migration in Russia. Our goal was to understand the patterns of scholarly migration by tracking the international movements of researchers and to identify the impact of such movements on the Russian science system overall and in each field of scholarship. The use of large-scale bibliometric data from Scopus allowed us to achieve this goal in a cross-disciplinary study of demography and scientometrics.In this study, we analyzed international mobility of scholars in Russia by different migration patterns with respect to their countries of origin and destination and their disciplines. We aim to extend this study to additional dimensions of analysis such as gender and level of experience as well as more detailed measures for performance and research quality and quantity for migrating researchers.

Our analysis of four categories of academic migrants revealed the similarities between their common countries of academic origin and destination while contrasting their differences in migration patterns and the impact on the Russian science system. US and Germany are the largest scientific hubs linked to Russia. Ukraine turned out to be one of the main donors of researchers to Russia, which could be partly explained by the patterns of Russo-Ukrainian migration

in the general population. Using data on major fields of science, we made comparisons between the international flows in health sciences, life sciences, physical sciences, and social sciences. Physical sciences has the largest flows followed by life sciences which seem to be the major fields where most technical knowledge and skills could be more easily transferable in different parts of the world compared to health sciences which may involve bodies of knowledge varying across countries (e.g. national medical protocols) and social sciences which may depend more on the language, culture, and context of societies.

We also analyzed citation data and observed large disparities between citation-based performance of migrants by their types and across different fields. Consistent with the generally observed pattern in scientometrics, the two fields of physical sciences have the most citations respectively followed by life sciences and health sciences (which are somewhat comparable) and social sciences (which has the lowest citations). Comparing by migrant types, emigrants from Russia have substantially higher citations compared to immigrants to Russia in all fields except for social sciences. This disparity could be the effect of a combination of reasons including the research performance of immigrants and emigrants and the difference in research opportunities in destination countries. Grouping the migrant researchers into three categories based on citation (normalized by age and field), we compared the origin countries of immigrants as well as destination countries of emigrants. As origins, Ukraine, Kazakhstan, Belarus, and Uzbekistan are more common among lowly cited immigrants. Similarly, Ukraine, Kazakhstan, and Belarus are ranked higher as destinations of lowly cited emigrants.

Using net migration rate, it was shown that while in the late 1990s and early 2000s Russia has been overall on the losing side of a brain circulation system, this has not been the case in recent years when net migration rate has been even positive for some years and generally oscillates around zero indicating a balance between the annual flows of incoming and outgoing researchers. Our results indicate that the overall lack of balance between incoming and outgoing flows of researchers has improved for Russia, but it could be still too early to call Russia a country of attraction for researchers.

As for analyzing the disciplines of migrating researchers, we introduced normalized measures for contributions of individuals in different fields. We also introduced a measure of net brain drain for quantifying the impact of international migration on each specific field of science. The analysis showed that over the time period 1996-2020, there has been a relatively large outflow of specialists in most fields of scholarship in Russia. Our results indicate that researchers leaving Russia in the fields of neuroscience, decision sciences, dentistry, biochemistry, mathematics, chemistry, pharmacology, chemical engineering, and materials science outnumber those who come to Russia supporting previous findings that specialists in these fields are actively leaving Russia [37,5,32,31,3] and going a few steps further by considering incoming flows and all fields of science.

The results that this research substantiated for the first time are generalizable within the limitations of bibliometric data which were discussed in Section 4. Keeping the possible caveats in mind, our substantive and methodological

contributions can be used in furthering our understanding of international migration in academia.For the specific case of this study, timely and detailed statistics were required. Our findings revealed new insights for Russia which is hoped to be used in policy development involving highly qualified professionals. The methodological contribution of this study can be applied to other countries as a framework of analysis to examine other national science systems and the impact of international mobility of researchers on them. This study has only scratched the surface on the study of migration among researchers and a new application of bibliometric data while many questions are yet unanswered for which we hope to have paved the way.

Acknowledgements

Authors highly appreciate the technical support from Tom Theile and discussions with Ebru Sanliturk and Emilio Zagheni which helped improving the article. This study has received access to the bibliometric data through the project "Kompetenzzentrum Bibliometrie" and the authors acknowledge their funder Bundesministerium für Bildung und Forschung (funding identification number 01PQ17001).

References

1. Alburez-Gutierrez, D., Aref, S., Gil-Clavel, S., Grow, A., Negraia, D., Zagheni, E.: Demography in the digital era: New data sources for population research. In: Arbia, G., Peluso, S., Pini, A., Rivellini, G. (eds.) Book of Short Papers SIS2019. pp. 23–30. Pearson (2019). https://doi.org/10.31235/osf.io/24jp7
2. Aman, V.: Does the Scopus author ID suffice to track scientific international mobility? A case study based on Leibniz laureates. Scientometrics **117**(2), 705–720 (2018), publisher: Springer
3. Antoshchuk, I.: Female computer scientists from Post-Soviet space: migration and academic career in the UK. In: 23rd International Conference on Science and Technology Indicators (STI 2018), September 12-14, 2018, Leiden, The Netherlands. Centre for Science and Technology Studies (CWTS) (2018)
4. Aref, S., Zagheni, E., West, J.: The demography of the peripatetic researcher: Evidence on highly mobile scholars from the Web of Science. In: International Conference on Social Informatics. pp. 50–65. Springer (2019)
5. Ball, D.Y., Gerber, T.P.: Russian Scientists and Rogue States: Does Western Assistance Reduce the Proliferation Threat? International Security **29**(4), 50–77 (2005)
6. Bedrina, E., Tukhtarova, Y., Neklyudova, N.: Migration from Uzbekistan to Russia: Push-Pull Factor Analysis. In: The International Science and Technology Conference "FarEastCon". pp. 283–296. Springer (2018)
7. Chepurenko, A.: The role of foreign scientific foundations' role in the cross-border mobility of Russian academics. International Journal of Manpower **36**(4), 562–584 (Jul 2015). https://doi.org/10.1108/IJM-11-2013-0256, `https://www.emerald.com/insight/content/doi/10.1108/IJM-11-2013-0256/full/html`
8. Cipko, S.: Contemporary migration from Ukraine. In: Migration Perspectives Eastern Europe and Central Asia, pp. 117 – 132. International Organization for Migration (2006)

9. Cohen, P.N.: Scholarly Communication in Sociology. Open Sociology (Mar 2019). https://doi.org/10.21428/4388219e, https://opensociology.pubpub.org/pub/scis, publisher: PubPub
10. D'Angelo, C.A., van Eck, N.J.: Collecting large-scale publication data at the level of individual researchers: a practical proposal for author name disambiguation. Scientometrics pp. 1–25 (2020), publisher: Springer
11. Di Bartolomeo, A., Makaryan, S., Weinar, A.: Regional Migration Report: Russia and Central Asia. Tech. rep., European University Institute (2014)
12. Falagas, M.E., Pitsouni, E.I., Malietzis, G.A., Pappas, G.: Comparison of pubmed, scopus, web of science, and google scholar: strengths and weaknesses. The FASEB journal **22**(2), 338–342 (2008)
13. Iontsev, V.A., Magomedova, A.G.: Demographic aspects of the development of human capital in Russia and its regions. R-Economy. 2015. Vol. 1. Iss. 3 **1**(3), 467–477 (2015)
14. Iontsev, V.A., Ryazantsev, S.V., Iontseva, S.V.: Emigration from Russia: new trends and forms. R-Economy. 2016. Vol. 2. Iss. 2 **2**(2), 216–224 (2016)
15. Iontsev, V.A., Zimova, N.S., Subbotin, A.A.: The Problems of "Brain Drain" in Russia and Member States of the Eurasian Economic Union (EAEU). RUDN Journal of Economics **25**(4), 510–517 (2017)
16. Kawashima, H., Tomizawa, H.: Accuracy evaluation of Scopus Author ID based on the largest funding database in Japan. Scientometrics **103**(3), 1061–1071 (2015), publisher: Springer
17. Kolesnikova, J., Camille, R., Kamasheva, A., Yue, Z.: Current Trends of Realization of the Intellectual Capital and Problems of Intellectual Migration. Procedia Economics and Finance **14**, 326–332 (Jan 2014). https://doi.org/10.1016/S2212-5671(14)00720-5, http://www.sciencedirect.com/science/article/pii/S2212567114007205
18. Korobkov, A., Zaionchkovskaia, Z.: Russian brain drain: Myths v. reality. Communist and Post-Communist Studies **45**, 327–341 (Sep 2012). https://doi.org/10.1016/j.postcomstud.2012.07.012
19. Kosyakov, D., Guskov, A.: Impact of national science policy on academic migration and research productivity in Russia. Procedia Computer Science **146**, 60–71 (2019), http://www.ipgg.sbras.ru/ru/science/publications/publ-high-pressure-evolution-of-encaged-hydrocarbons-in-046082
20. Kosyakov, D., Guskov, A.: Synchronous scientific mobility and international collaboration: case of Russia. In: Proceedings of the 17th International Conference on Scientometrics. pp. 1319–1328 (2019)
21. Lieberson, S.: The Interpretation of Net Migration Rates. Sociological Methodology **11**, 176–190 (1980). https://doi.org/10.2307/270863, https://www.jstor.org/stable/270863
22. Miranda-González, A., Aref, S., Theile, T., Zagheni, E.: Scholarly migration within Mexico: Analyzing internal migration among researchers using Scopus longitudinal bibliometric data. EPJ Data Science (2020). https://doi.org/10.1140/epjds/s13688-020-00242-x, (in press)
23. Moed, H.F., Halevi, G.: A bibliometric approach to tracking international scientific migration. Scientometrics **101**(3), 1987–2001 (Dec 2014). https://doi.org/10.1007/s11192-014-1307-6, http://link.springer.com/10.1007/s11192-014-1307-6
24. Molodikova, I.N., Yudina, T.N.: Migration strategies of Ukrainian migrants: EU or Russia. Contemporary Problems of Social Work **2**(3), 62–71 (2016)

25. Mongeon, P., Paul-Hus, A.: The journal coverage of Web of Science and Scopus: a comparative analysis. Scientometrics **106**(1), 213–228 (2016), publisher: Springer
26. Mukomel, V.: Migration of Ukrainians to Russia in 2014–2015. Discourses and Perceptions of the Local Population. Migration and the Ukraine Crisis p. 105 (2017), publisher: E-International Relations Bristol
27. Naumova, T.V.: Russia's "brain drain". Russian Social Science Review **39**(2), 49–56 (1998), publisher: Taylor & Francis
28. Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V.: Scikit-learn: Machine learning in Python. the Journal of machine Learning research **12**, 2825–2830 (2011), publisher: JMLR. org
29. Podolskaya, T., Medyakova, E., Kolesnikova, N.: Migration Policy and Problems of Ensuring Economic Security of Countries (2020). https://doi.org/10.4018/978-1-7998-0111-5.ch004, www.igi-global.com/chapter/migration-policy-and-problems-of-ensuring-economic-security-of-countries/237468, iSBN: 9781799801115 Library Catalog: www.igi-global.com Pages: 63-80 Publisher: IGI Global
30. Rasuly-Paleczek, G., Six-Hohenbalken, M.: Migration and its Impact on Armenia. A field practice. ASSA-Austrian Studies in Social Anthropology: Wien (2017)
31. Ryazantsev, S.: Russia needs a new migration policy. Russian Politics & Law **51**(3), 80–88 (2013), publisher: Taylor & Francis
32. Rybakovsky, L., Ryazantsev, S.: International migration in the Russian Federation. In: UN Expert Group Meeting on International Migration and Development, New York, July (2005)
33. Sugimoto, C.R., Larivière, V.: Measuring research: What everyone needs to know. Oxford University Press (2018)
34. Taylor, R.G., Mechitov, A.I., Schellenberger, R.E.: Transformation within the Russian academic community. International Education **26**(1), 29 (1996), publisher: College of Education, University of Tennessee
35. Titarenko, L.: The Republic of Belarus: Flows and Tendencies in Migration Processes. Borders, Migration and Regional Stability in the EU's Eastern Neighbourhood p. 172 (2016)
36. Ushkalov, I.G., Malakha, I.A.: The "Brain Drain" as a global phenomenon and its characteristics in Russia. Russian Social Science Review **42**(5), 79–95 (2001)
37. Volz, E.: Utechka Umov: The History, Implications, and Solutions concerning Russia's Post-Soviet Brain Drain. Journal of Undergraduate Research (Rochester) (2002)
38. Yurevich, M.A., Malakhov, V.A., Aushkap, D.S.: Global Experience in Interaction with Compatriot Scientists: Lessons for Russia. Herald of the Russian Academy of Sciences **89**(4), 342–350 (Jul 2019). https://doi.org/10.1134/S1019331619040129, https://doi.org/10.1134/S1019331619040129
39. Zubova, L.G.: The human potential of Russian science. Herald of the Russian Academy of Sciences **82**(4), 295–301 (2012)