Greek chemical engineers. Are they static or mobile? Evidence from the national archive of PhD theses

Evi Sachini, Konstantinos Sioumalas-Christodoulou, Charalampos Chrysomallidis, Galatios Siganos and Nikolaos Karampekios
National Documentation Centre, Athens, Greece

Abstract

Purpose – The purpose of this paper is to identify the geographical location of researchers.

Design/methodology/approach – Combine standard bibliometric databases with social media data.

Findings – The majority of the population of the sample (71.8%) – Greek chemical engineers – are static. A significant portion of the mobile researchers (28.2%) returned to their country of origin (25.6%). Performing network analysis, the cluster of countries corresponding to the mobile category of researchers is identified and depicted.

Originality/value – Herein, this study introduce a new, national data set on doctorate holders that will allow multiple bibliometric analyses in the future. Also, this study is among the few (Gendronneau *et al.*, 2019) that combines standard bibliometric databases with social media data. In cases where multiple affiliations per year pose a difficulty in understanding the geographical location of each individual, LinkedIn data were used. The analysis sheds light on a field of science that is not extensively examined in terms of brain circulation. While similar publications focus on physicians (i.e. cardiologists – Dyachenko and Mironenko, 2018), this paper focus on a subset of doctorate holders in engineering.

Keywords Greece, Mobility, Bibliometrics, Chemical engineers, PhD holders, Social data

Paper type Research paper

1. Introduction

The geographical mobility of the highly educated is a theme of increased interest, as it gives rise to a number of considerations. For example, in relation to the international flow of human capital (Hunter, 2013) or to the realisation that open countries (as opposed to closed ones) have a strong science system (Wagner and Jonkers, 2017). This is associated with the finding that scientists have the most impact when they are free to move, according to Sugimoto *et al.* (2017).

The theme is of interest not only for academic reasons but also for policy ones, as domestic education considerations consider this outward mobility a loss of national resources, as migrated scientists eventually work in another, hosting country that has probably not bear the cost of their education and training (Cervantes and Guellec, 2002). Others view the phenomenon as making countries lose a dynamic part of their human capital such as young and talented scientists, that may be regarded to be a prerequisite for society's and economy's further development and growth (Martineau *et al.*, 2004). This is an argument worth considering especially in the case of Greece, which only very recently started bouncing back from an extreme decline of macroeconomic indicators during the 2009–2016 crisis that forced its educated workforce to emigrate abroad (Labrianidis and Pratsinakis, 2014; Labrianidis, 2017). To restore economic growth – a potential that has been just scaled back due to the 2020 Covid-19 pandemic – a country needs all the help it can get from its educated science diaspora. While luring them back is unfounded, as top-notch positions are highly paid globally and go against the intrinsic nature of the scientific endeavour that gives a premium to mobility and network creation, establishing a digitally-enabled, knowledge- and/or entrepreneurial-based relations is a potential middle ground (Labrianidis *et al.*, 2019). Especially, as the majority of scientists do not sever ties with their homeland but may build a chain of affiliations, benefitting it, in an indirect way (Sugimoto *et al.*, 2017).

However, no analysis nor policy implementation is possible without the mapping and identification of the highly educated that would shed light on their geographical, knowledge, etc., mobility. Short of nation-wide, periodical censuses that probe into the mobility of researchers and scientists, existing large data sets on the outputs produced by those highly educated stands as the main sources of understanding this mobility. Analysis of the bibliometric output of these populations is central to this identification process (Laudel, 2003; Moed and Halevi, 2014; Robinson-Garcia, 2019). Tech-based data sets such as patents and trademarks, as well as the exploitation of the trove of data found in social media, stand as parallel avenues.

The current issue and full text archive of this journal is available on Emerald Insight at: **https://www.emerald.com/insight/2514-9326.htm**

Collection and Curation
© Emerald Publishing Limited [ISSN 2514-9326]
[DOI 10.1108/CC-06-2020-0017]

Received 12 June 2020
Revised 17 July 2020
Accepted 26 July 2020

2. Scope and objective

The paper's scope is to examine the mobility of a specific subset of the Greek highly educated, thus contributing to the wider scientific discussion on scientists' mobility, the way to identify it, as well as its policy implication in a global and national context.

The objective is to identify the geographical mobility of PhD holders in chemical engineering. Specifically, the publication activity of PhD holders by Greek universities during the 2008–2019 period were analysed. This was conducted by making use of relevant bibliometric data for the purposes of tracing their affiliation. In doing so, and in those cases that multiple affiliations per year pose a difficulty in understanding their mobility, we make additional use of social media data.

3. Contribution

Herein, we introduce a new, national data set on doctorate holders that will allow multiple bibliometric analyses in the future. Also, this study is among the few (Gendronneau *et al.*, 2019) that combines standard bibliometric databases with social media data. In cases where multiple affiliations per year pose a difficulty in understanding the geographical location of each individual, LinkedIn data were used.

Our analysis sheds light on a field of science that is not extensively examined in terms of brain circulation. While similar publications focus on physicians (i.e. cardiologists – Dyachenko and Mironenko, 2018), we focus on a subset of doctorate holders in engineering.

On a topical note, despite the high profile of Greek scientists (Yuret, 2017), surprisingly little has been done to identify the mobility models of Greek scholars. This is an attempt to do so.

4. Structure

The next section outlines the data, the collection process and the methodology. Section "data and methods" describes the process of collection, retrieval and validation of the data. Section "definitions, model and analysis" includes the methodological overview, the taxonomy model upon which the international geographical mobility analysis is based and unfolds the conducted bibliometric analysis' approaches. Moreover, it covers the taxonomy model. It is followed by the "findings" section where descriptive statistics on the researcher's mobility and affiliations, as well as graphs concerning the depiction of scientific connections between the affiliated countries are laid out. Additionally, insight into the main outcomes of the object of study, including the researcher mobility and the scientific connections between the affiliated countries' status over time, is provided. This is succeeded by the discussion where the findings are contextualised and set within specific considerations. Last is the formulation of points for further research.

5. Data and methods

The data collection and the overall analysis was conducted in March 2020. PhD holders were identified within the National Archive of PhD Theses (EADD). EADD is the national registry collecting PhD theses from all Higher Education Institutions in Greece, as well as those PhD theses awarded to Greeks by foreign universities and certified by the Hellenic NARIC (DOATAP). It contains more than 45.000 PhDs and spans a period of more than 30 years (1985–2020).

In this study, evidence from the EADD database was used. Emphasis was put on the period pertaining to the recent Greek economic crisis that has (also) resulted in the relative brain-drain debate (Labrianidis and Pratsinakis, 2014). Therefore, although the availability of data concerning previous years was possible, it was decided to limit the time-span for the 2008–2018 period. Within this time-frame, 443 individuals were identified within EADD as having their PhD thesis on "chemical engineering sciences" (Table 1) according to the Frascati classification system (Frascati42). This specific scientific subfield belongs to the field of Engineering and Technology (Frascati6). In our analysis, "chemical engineering sciences" include all the available subfields such as "chemical engineering (plants, products)" and "chemical process engineering".

Importantly, the classification "chemical engineering sciences" was attributed by holders themselves during their registration (as newly PhD holders) in EADD following the classification protocol as implemented by standardised librarian and information services' schemas. The variables of those 443 individuals extracted from the registry were as follows: "name", "surname", "father's name", "sex", "PhD ID number", "PhD year", "PhD institution" and "PhD department".

Table 1 presents the annual distribution of the 443 Chemical Engineers for the 2008–2018 period. It is observed that, in 2018, only three individuals were registered as PhD holders. As EADD's database is regularly updated, the low number of registered individuals can be explained by the time gap between the (actual) PhD acquisition and PhD thesis registration in EADD. Despite the low number of PhD holders in 2018 with respect to the other years, with the aim of exploring the most recent mobility attitudes of the scientific manpower, the authors decided to include such a group into the analysis.

All 443 had obtained their PhD from a Greek university. During registration, only 29.7% (132 individuals) had their names written in both Greek and Latin alphabet. For this research, it was decided to proceed to a transliteration process for the remaining 311 individuals. This was conducted on a manual basis and the methods used followed by Karakos (2003) and Chalamandaris *et al.* (2006).

Next, the PhD holders were linked to their Scopus author-name using the EADD PhD holders' full name as the key merging variable. Scopus is a source-neutral abstract and citation database, curated by independent subject-matter

Table 1 Distribution of PhD holders in "chemical engineering sciences"

2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018
44	88	45	50	44	30	33	38	39	29	3

experts indexing 75+ million records places and 5.000+ publishers. Towards this and given the small amount of our population group (443 individuals) the use of manual author name record linkage was preferred over the appliance of automatic name disambiguation techniques to allow for thorough supervision of each process step and consider critical aspects (transliteration issues). In cases where linkage with Scopus author-name fell upon setbacks due to names transliterated in English bearing resemblance to foreign names (<1%), the use of secondary internet sources (such as Linkedin, etc.,) were used to make sure that the PhD holder and the Scopus author was the same individual, for example, "Μιλένκοβιτς Γιοβάνα" ("Milenkovic Jovana").

5.1 Data scraping

Of the 443 individuals, 424 were identified inside the Scopus database and were associated with 476 Scopus author IDs. The remaining 19 (=443 – 424) individuals were not identified inside Scopus [1].

Of the rest 424 individuals, 51 had multiple Scopus IDs, whereas the remaining 373 were associated with unique Scopus IDs. This paper targeted all 424 individuals.

Retrieval of the Scopus IDs enabled the locating and downloading of information relevant to researchers' bibliographic profile. Specifically, the following variables related to the bibliometric performance (Waltman and Noyons, 2018) of each recipient were the subject of the retrieval process, namely, year of publications and the affiliation of a research unit (hereafter abbreviated to "country"). This information was deemed relevant for the research study. To retrieve the country in which each paper was published (i.e. affiliation county) the author's affiliation needed to be extracted. This is based on the assumption that the geographical location of the affiliated institution named by the author as her/his postal address stands as the country to which the specific article should be "measured" for – an important assumption that cuts through the entire paper. An algorithm (via forming XPath queries) using as input the Scopus IDs and output the aforementioned bibliometric variables was implemented within the Python (3.7.3) environment.

Within the Python (3.7.3) environment, two nested dictionaries for each Scopus ID were created. The main one used the Scopus author IDs as "keys". The second used the publication year as "keys". The affiliation country(ies) was used as "value". Below, an example describing the structure of the finalised author profile is presented:

- {"Scopus author ID": {Year: (Affiliation "Country"}}

  For example, Author with Scopus ID "12345" has a publication affiliated with the National Technical University of Athens, Greece, in 2020 and another one in 2019 with Sapienza University of Rome, Italy.
- {"12345": {2020: ("Greece"), 2019: ("Italy")}}

As said, the target population consisted of 424 individuals. Yet, the total number of individuals researchers with publications indexed in Scopus and within the scope of this study was 379. The remaining 45 researchers (=424 – 379) had solely registered publications prior to their PhD acquisition. This was decided so because, for each individual, the year after their PhD acquisition was considered as the starting point of their publication record. Given that the paper focusses on the interval 2008–2019, these 45 authors were excluded from the collected data set, as publications for years prior to their PhD were censored [2].

6. Definitions, model and analysis

6.1 Definitions

The following distinct affiliation instances were observed:

- *Single-affiliation*: Researcher affiliation with only one country in all publications in a given year (either Greece or any other single country).
- *Multiple-affiliation*: Researcher has many publications in different countries within the same year.
- *Co-affiliation*: Researcher affiliated with two or more different countries within a single publication.

The following three patterns were detected:

1. Pattern 1: Authors with single-affiliation instances – a pattern observed in 294 authors.
2. Pattern 2: Authors with multiple-affiliation and co-affiliation instances – a pattern observed in 85 authors.
3. The final data set included 379 authors. In total, 77.6% of the total were identified as fitting Pattern 1, while 22.4% as fitting Pattern 2.

6.2 Model

Herein, "mobility" is defined as the (unique) researcher's geographical location captured within a period of a year. Concerning the individuals falling under Pattern 1, all publications had single affiliations. For individuals falling in Pattern 2, both multiple and co-affiliations were identified, indicating that within a period of a year an author may be affiliated with more than one country. For the purposes of this paper, the understanding was that the representation of a researcher's geographical location during that specific year (in which, a co-affiliation and/or a multiple affiliations was identified) had to be singled out.

We turned to the bibliography. Attempting to identify a standard manner in which the bibliometrics' bibliography treats the specific issue of geographical mobility, our search results were limited. Most studies only briefly address the issue (Moed and Halevi, 2014) or focus solely on taxonomizing mobility according to its (multiple or co) affiliation patterns as derived from scientific publications (Robinson-Garcia, 2019). One case that did address the issue, it was conducted by using the per-capita gross domestic product (GDP) as the sub-principle for singling out the geographical location (Krause *et al.*, 2007). Using a macro-economic indicator to canonize bibliometric performance does not make a lot of sense.

Given the inability to locate a past bibliography upon which to base our analysis, we turned to self-made rules. As such, the selection process that was followed in cases of multiple or co-affiliation publications stood on the following axiom: select that country, which appeared in the succeeding or previous publications. To do so, a comparison with the countries appearing in the nearest succeeding publication was conducted. In cases where identical instances of a specific country were found, then this specific country was singled out as the attributed country for the year in which multiple or co-affiliation publications were identified. If not, a comparison with the countries of the nearest previous publication was conducted.

This led to as far as five publications ahead or behind being taken as a rule of thumb until a common country was identified. This approach presented a number of problems. Namely, to automatically single out a country appearing in the majority of newer or older publications is a major assumption that entailed the risk of confounding the actual location.

Given this reality, an alternative path was explored. A path that made use of social media and its abundant availability. As such, we opted for using LinkedIn as a validation method. LinkedIn is the world's leading professional cloud wherein each individual provides a voluntary and detailed description of their own (professional) life trajectory. This selection was made easier because of the fact that the number of individuals falling under Pattern 2 (n = 85) is rather small and allowed for a manual process to identify the geographical movement of these individuals on an annual basis. Specifically, in those cases of multiple or co-affiliation publications, the LinkedIn account of the specific researcher was probed. What was looked for was posted information (by the researcher) detailing the specific year in which the bibliometric analysis indicated multiple or co-affiliation publications. Assuming that the researcher himself/herself knew best the location he/she was during the specific period, that information was harvested for all 85 individuals [3]. Also, in cases of potential author name disambiguation issues, data existing in the original EADD database such as "PhD Institution", "PhD graduation year", etc., allowed cross-checking of the LinkedIn posted details.

In LinkedIn, data with reference to the career (or job) timeline of each individual are presented in a "month-year" format. Thus, there exist cases in which an individual can be affiliated with at least two different countries in the same year. In our case, throughout the whole sample, nine individuals (10.6%) were identified as having two different affiliations within the same year. Given that mobility was defined as the researcher's *unique* geographical location captured within a one-year period additional examination took place.

To single out such a country, the following assumption was made. Taking into consideration that January is the initial month of every year, each year is divided into two semesters as follows: 1st semester from January to June and 2nd semester from July to December. For those cases that had two different affiliations within the same year, if the time interval (1) was greater than 6 months, the corresponding affiliation country is considered as the valid one [4]. Contrarily, if the time interval was less than 6 months, the upcoming affiliation country is considered the valid one. Below, an example describing the selection process is presented.

6.2.1 Example

"Mr. X" was affiliated with National Technical University of Athens, Greece, from May 2017 – April 2018. From April 2018 and on, the affiliation changed to University of Oxford, Oxford, UK. Thus, for the year 2018, as his affiliation in Greece was from January 2018 to April 2018 (4 months $<$ 6 months), it is assumed that for the year 2018, "Mr.X" was affiliated with the UK.

The mobility pattern of a researcher can be determined by different mobility events. Herein, a taxonomic model similar to Robinson-Garcia (2019) is used. Proposing a general classification for analysing scientific mobility using institutional affiliation changes, it provides an analysis in delineating the various mobility patterns (static and mobile researchers).

A mobility event refers to each of the different possible combinations of affiliation instances a researcher can have in a specific time interval. The following basic notation aims to identify different mobility events.

1 Mobility events
2 Notation

- A: Author's first affiliation country.
- B: Any other affiliation country (or set of countries) different from A.
- t_p and t_{p+1}: Each refers to a particular point in time (i.e. the year of publication).

Specifically, the snapshot of time at t_p the affiliated country is considered the initial country where the author began publishing, t_{p+1} refers to the next tracked point in time (year) in the publication record of the author.

Two additional elements are related to such mobility events over time, then:

1 *Directionality*: Indicates whether it is possible to reliably establish if the author has chronologically published first to A and then to B.
2 *Country rupture*: Indicates when an author's affiliation country(ies) at t_p is not found among the affiliation country (ies) of the author at t_p. In other words, there is a rupture in the countries between t_p and t_{p+1}.

Adding the variable of time to each researcher's scientific profile, two distinct mobility events emerge. Table 2 shows the classified Events E1, E2 delineating the overall scientific mobility of the researchers. All researchers started their scientific career with affiliated publications that concerned a single country (t_p). Event E1 concerns publications affiliated only with a single country at a specific year or interval, respectively. Regarding Event E2, researcher publications are affiliated with more than one country associated with a specific affiliation type.

The proposed classification system below defines mobility at the level of the individual researcher as measured across their overall scientific output.

1 Mobility classification schema

- Static class: Researchers lacking any mobility event while showing affiliation instances only one specific country i.e. researchers exclusively attributed to Event E1.
- Mobile class: Researchers associated with a directional mobility event and at least a point or

Table 2 Taxonomy of mobility events tracked through affiliation countries given each consecutive publication time points t_p and t_{p+1}

Mobility event	t_p	t_{p+1}	Directionality	Country rupture
E1	A	A	No	No
E2	A	B	Yes	Yes

period of rupture with their country of origin, i.e. researchers attributed to Event E2.

6.3 Analysis

Initially, the researchers' affiliation status was examined. This status is characterised by researchers who have:

- *only foreign* affiliations [researchers with authored/co-authored publications only under non-Greek institutional affiliation(s)].
- *only Greek* affiliations [researchers with authored/co-authored publications only under Greek institutional affiliation(s)].
- *Greek and foreign* affiliations [researchers with authored/co-authored publications under both Greek and foreign institutional affiliation(s)].

Researchers indicating *only Greek* affiliations were identified as fitting into the static class. Researchers indicating *only foreign* affiliations and associated with mobility events E2, as well as researchers with *Greek and foreign* affiliations were included in the mobile class.

Concerning the mobile class, to schematically point out the scientific career paths between the target country (Greece) and different groups of countries a graph was constructed.

An undirected graph was constructed in which each node represents the affiliation country and each edge represents at least one publication from the current country to another. As a result, two nodes (two different countries) are linked if a researcher has shared at least one publication. The graph was constructed within the Python (3.7.3) environment, saved in an output file (.gml) and used as input for the software visualisation tool VOSviewer (hereafter abbreviated to "VOS"). As a first step of the conducted graph analysis, the associated strength was selected as the most appropriate similarity measure. Such a measure was used for normalising the co-occurrence frequencies between every pair of nodes, thus meaning, the frequency of any *linked* pair of nodes. Formally, given two nodes i and j, their association strength is given by (Van Eck, 2014):

$$s_{ij} = \frac{2mc_{ij}}{c_i c_j} \tag{1}$$

Here s_{ij} denotes the association strength of nodes i and j, c_{ij} denotes the number of links (co-occurrence links between countries in this case) between nodes i and j ($c_{ij} = c_{ji} \geq 0$), where c_i denotes the total number of links of node i (similarly for node j) and m denotes the total number of links in the network, that is:

$$c_i = \sum_{i \neq j} c_{ij} \text{ and } m = \frac{1}{2}\sum_i c_i \tag{2}$$

VOS aims to locate items in a low-dimensional space in such a way that the distance between any two items reflects the similarity. For each pair of nodes i and j, VOS requires a similarity s_{ij} ($s_{ij} \geq 0$) as input. VOS treats the similarities of s_{ij} as measurements on a ratio scale. The similarities s_{ij} were calculated using the association strength defined in equation (1). After the construction of the normalised network, the next step was to define the position of the nodes in the network in a two-dimensional space in such a way that strongly-related nodes (countries) are located close to each other while weakly-related nodes (countries) are located far away from each other. In the case of node mapping, using VOS we find for each node i a vector $x_i \in \mathbf{R}_p$ that indicates the location of node i in a p-dimensional map (here $p = 2$). In VOS, the locations of the nodes on a map were determined by minimizing:

$$V(x_1, \ldots x_n) = \sum_{i<j} s_{ij} d_{ij}^2 - \sum_{i<j} d_{ij} \tag{3}$$

with respect to $x_1, \ldots, x_n$. Here d_{ij} represents the distance between nodes i and j and is given by:

$$d_{ij} = \|x_i - x_j\| = \sqrt{\sum_{k=1}^{2} (x_{ik} - x_{jk})^2} \tag{4}$$

Equation (3) can be interpreted in terms of attractive and repulsive forces between nodes. The first term in equation (3) represents an attractive force, and the second term represents a repulsive force. The higher the association strength of two nodes, the stronger the attractive force between the nodes. As the strength of the repulsive force between two nodes does not depend on the association strength of the nodes, the overall effect of the two forces is that nodes with a high association strength are pulled towards each other while nodes with a low association strength are pushed away from each other.

In the case of clustering, for each node i it was necessary to find a positive integer x_i that indicates the cluster to which node i belongs to. Here the distance d_{ij} in equation (3) is defined as follows:

$$d_{ij} = \{0, \quad x_i = x_j \ \frac{1}{\gamma}, \quad x_i \neq x_j \tag{5}$$

The parameter γ in equation (5) is referred to as the resolution parameter ($\gamma \geq 0$). The larger the value of this parameter, the larger the number of clusters that are obtained. In our case, this parameter was set equal to 1. Additionally, the minimum cluster size (the minimum number of items inside each cluster) was also set equal to 1. In our case, the closer the two countries are located to each other, the stronger their scientific connection.

7. Findings

The findings section is divided into two, closely associated, sub-sections. The first one identifies and distributes the authors according to their mobility status in relation to their respective affiliations over the 2008–2019 period. Also, it posits the mobile class of authors to the countries in which their institutional affiliations are located, identifying specific "neighbourhood" of countries. Section 2 seeks to identify the whereabouts of the mobile class in relation to the "when" factor. That is, it is attempted to geolocate that group of individuals taking into consideration their most recent publication trace.

7.1 Distribution of authors

According to the analysis, the static class amounts to 272 individuals. With 71.8%, they stand to be the majority of the researchers (Table 3). It consists of researchers with only Greek (265) and only foreign (7) affiliations. That is, 265 individuals have published only in journals that are published under Greek institutional affiliations and 7 individuals in journals under foreign ones. Importantly, concerning these 7 individuals, their entire publication career paths are located within the "premises" of a single country and do not alternate from one "foreign" to another "foreign" country.

The mobile class amounts to 107 individuals, thus constituting 28.2% of the total researcher population (Table 4). This class includes 103 researchers *with Greek and foreign* affiliations and 4 researchers associated with *only foreign* affiliations. That is, 103 individuals have published in journals that are published under Greek institutional affiliations, as well as foreign ones and 4 researchers have published in journals that are institutionally associated with a number of "foreign" countries, over their publication career path.

Given that 265 authors, according to their bibliometric profile, appear to never have left Greece, the mobile class of researchers and the countries in which their institutional affiliations are located were focussed on. As presented in Table 5, their publications are located in specific "neighbourhoods" of countries. That is, specific groups of researchers have been identified to publish in specific groups of countries.

For example, the publication career paths of none of the authors do not contain publications in journals with institutional affiliations in any combination containing the following five countries: The Netherlands, Luxembourg, Denmark, Cyprus and Canada. On the other hand, some countries such as the USA and the UK, although clustered in different groups, they are connected with each other. This should be attributed to the fact that the link strength (total number of co-occurrences) between those countries is lower with respect to other countries inside their cluster, the resolution parameter γ was set equal to 1 and the minimum cluster size (no. of items) was set equal to 1 (Section 6.3). Such countries (the USA, The Netherlands, France, the UK, Italy, Belgium and Sweden) can be interpreted as "bridge nodes", as they link together diverse groups of countries.

Visualising Table 5 in Figure 1, the chain of affiliations between country-level institutions are presented. The majority of the mobile class of researchers are positioned in a specific group of countries. According to the node size in Figure 1, as well as Figure 2, that presents the frequency distribution of those countries, the UK (17.6%), the USA (16.2%), Germany (12.8%), France (9.6%), Belgium (8.8%), Cyprus (6.1%), Sweden (5.6%), The Netherlands (4.8%), Italy (4.5%) and Canada (3.5%) are the hosting countries of institutions in which the majority of the mobile class of researchers tend to publish throughout their scientific career.

These countries are the top 10 with the highest co-occurrence frequency with the target country (Greece). At the same time, they attract the highest number (89.5%) of the mobile class of researchers. Contrarily, countries such as Switzerland (2.5%), Denmark (1.6%), Luxembourg (1.6%), United Arab Emirates (1.1%), Qatar (0.8%), China (0.8%), Portugal (0.8%), Saudi Arabia (0.5%), Spain (0.5%) and Australia (0.3%) are the top 10 in terms of low co-occurrence frequency with the target country (Greece) and attract the lowest number (10.5%) of researchers.

Table 3 Distribution of authors with respect to their mobility status over the time interval 2008–2019

Authors	Frequency	(%)
Static	272	71.8
Mobile	107	28.2
Total	379	100

Table 4 Distribution of authors with respect to their affiliation status over the time interval 2007–2019

Authors	Frequency	(%)
only foreign	11	2.9
only Greek	265	69.9
Greek and foreign	103	27.2
Total	379	100

Table 5 Distribution of cluster's items of the mobile class

Clusters	Countries	Items
Red	The US, The Netherlands, France, Germany, Spain and China	6
Green	UK, Italy, Luxembourg, Portugal and Qatar	5
Blue	Greece, Denmark, Saudi Arabia and the United Arab Emirates	4
Yellow	Belgium, Cyprus, Switzerland and Sweden	4
Purple	Canada and Australia	2

7.2 The mobile ones. Where are they?

To assess the most recent geographical location of the mobile class of researchers, it was deemed sufficient to check the country of the last institutional affiliation recorded within the bibliometric database. This would provide evidence as to "where" these individuals are in relation to "when". Concerning the "when", obviously, the more recent this time element was found to be in the bibliometric database the easier it would be to make specific claims concerning their "present" location.

Given that the database contained bibliometric information over time, it was assumed that if the most recent affiliation country coincided with the oldest/first recorded affiliation country, then it would be assumed that the mobile class of researchers had returned to their country of origin. Contrarily, if a discrepancy between the oldest/first and most recent affiliation countries was identified, it would be presumed that the researcher currently has not returned to the country of origin and is located in a different institutional affiliation.

According to the data (Table 3), 107 individuals were classified as Mobile. Of those, 101 individuals (94.4%) had

Figure 1 Distribution of researcher scientific career paths corresponding to the mobile researchers

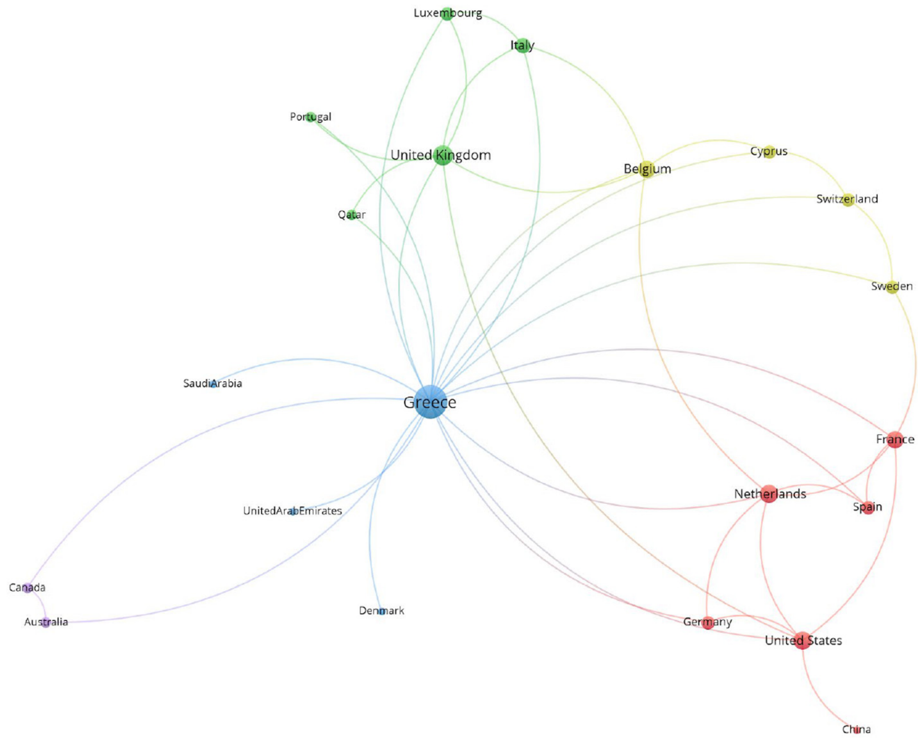
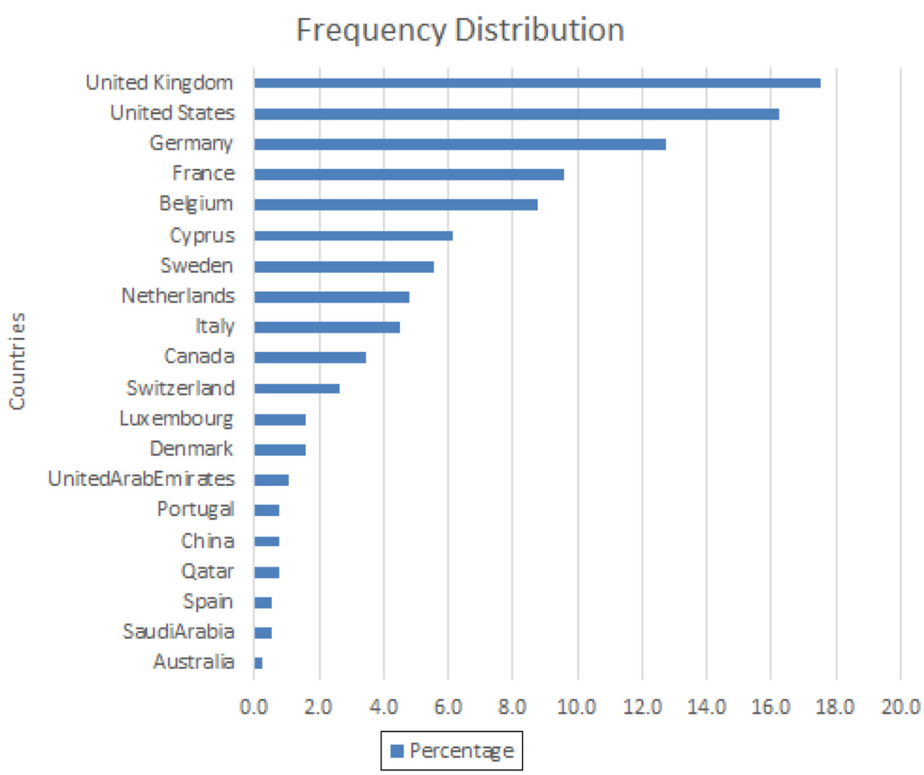


Figure 2 Frequency distribution of the 20 affiliation countries corresponding to the mobile class of researchers



initiated the scientific career in Greece, which is their first publication after their PhD acquisition was registered under a Greek institutional affiliation. The remaining 6 (5.6%) had initiated their careers in countries such as the USA, France, Germany, The Netherlands, Austria and Belgium.

Table 6 presents the annual distribution of the mobile class of researchers. This is done in a manner that spares between those that had Greece as the country from which their first (F) institutional affiliation is located and those that had a foreign one. The same holds concerning their last (L) institutional affiliation – it is divided between a Greek and a foreign one. This allows us to identify the exact number of researchers for which the most recent affiliation country coincides with the oldest/first one.

As the vast majority of researchers (24 + 77 = 101 individuals – 94.4%) had initiated their scientific career in Greece, w.l.o.g, it was decided to perform the analysis with respect to the two classes shown in Table 6 ("Greece" and "Other").

Identifying their last affiliation status, being in Greece or abroad, is of little value if the parameter of time is not taken into consideration. As such, Table 6 presents the distribution of the mobile class of individuals in relation to the country the last affiliation was recorded.

For example, using as a reference point the year (2019), the following observations can be made:

- the large majority of the mobile class of researchers [72.9% (78 out of 107)] have been identified as having published at least one publication on that year.
- In terms of "when", their publication has been geographically positioned/located in the countries corresponding to their last/most recent affiliation.
- In terms of "where", out of the 78, 18 have been recorded as having (co)authored a publication under a Greek institutional affiliation.
- Taking into consideration equations (2) and (3), one can argue with an increased level of confidence that this population sub-group has been identified as being "present" to their country of origin (as identified by their first publication).
- Using a wider net that considers their first publication to be their PhD thesis, the same can be argued in favour of the two individuals that have their first publication recorded in "Other" country and they are last in "Greece". As a result, it can be argued that 25.6% (20 out of 78) of the mobile class of researchers have returned to their country of origin, Greece.
- The remaining 58 individuals [74.4% (58 out of 78)] are identified as having (co)authored a publication under institutional affiliations in "Other" countries and as such have not returned to their country of origin.

  The last observation can be made in reference to the years prior to 2019.
- The further back in time one goes (the "when" factor), decreases the level of confidence one can attribute to their "present" location.

8. Discussion and concluding remarks

Following past studies on the mobility of scientists (Dyachenko and Mironenko, 2018), this study examined Greek doctorate holders in chemical engineering. Through novel combinatorial approaches (triangulation of doctorate-level data with Scopus and LinkedIn), the authors identified the mobility patterns of a specific population of Greek PhDs.

Results indicate that 71.8% of researchers are static. Most of them are static in their country of origin, Greece, indicating a single country publication trajectory. The mobile ones constitute 28.2% of the total population. Most mobiles did not sever ties with their country of origin, Greece, but instead built a chain of affiliations that linked nations together. Such chains are represented as groups of countries (clusters), in which the scientific connections between different countries can be visualised.

Of those mobile individuals, their most recent publications indicate that a substantial portion of individuals (25.6%) have returned to their country of origin in terms of where their publication career path initiated from. The remaining 74.4% of individuals have been identified to be in other countries, part of the global brain circulation.

Alike perspectives, do not consider a temporary stay abroad as a migration process with clear winners and losers – brain gain and brain drain; rather a reciprocal process allowing individuals and countries or regions to benefit from current collaborations and future returns – brain circulation. In every case, taking into consideration that a large proportion of researchers tends to publish to countries with typically significant scientific impact, international mobility should be encouraged.

On a policy level, validating Sugimoto *et al.* (2017), it was shown that the highly educated Greeks have not severed ties with their mother country. Instead, they have preserved a vibrant bibliographic relationship indicating their ability to benefit it with multiple manners.

9. Next steps

One avenue is to extend our analysis to complementary bibliometric data sources (Altmetrics.com, conference

Table 6 Annual distribution of a mobile class of researchers according to the first (F) and last (L) country of affiliation their publication was recorded

Countries \ Years	2012	2013	2014	2015	2016	2017	2018	2019	Total
("Greece" as F) \ ("Greece" as L)	0	0	0	1	2	0	3	18	24
("Greece" as F) \ ("Other" as L)	1	1	0	3	3	6	7	56	77
("Other" as F) \ ("Greece" as L)	0	0	0	0	0	0	0	2	2
("Other" as F) \ ("Other" as L)	0	0	2	0	0	0	0	2	4
Total	1	1	2	4	5	6	10	78	107

proceedings, etc.) to broaden the scope of the identification of geographical mobility. Also, identifying potential publication variations relative to sex, open versus closed access publications, as well as potential changes in relation to a researcher's original field of science (as derived from tertiary education degrees) and the field of science of the journal(s) wherein he/she has published stand as promising avenues.

Given that EADD is a registry containing thousands of Greek scholars, similar studies in other scientific groups are an extra option. Also, author-level combinations of science outputs (publications) with technology-based outputs (e.g. patents, industrial designs) stand as parallel avenues.

Finally, the issue of developing an algorithm-based software that could harvest in an automated manner the posted data on LinkedIn walls would greatly enable the tackling of the annual discerning the actual geographical location in cases of multiple or co-affiliation publications per year. On this, the authors are aware of the potential issues that pertain to such an automated approach, i.e. error-prone due to name disambiguation.

Notes

1 For those 19 individuals, further examination was conducted. Within the context of our object of study – exploration of their mobility attitudes, the methodology described in Section 6.2 (*Model*) was used also for this population. Specifically, 7 out of those 19 (36.8%) individuals were identified by using social media data, mainly LinkedIn and, secondary, Facebook, as well as internet sources (homepages, available CV's). According to the results, all individuals followed domestic scientific career paths, that is all authors were associated with *only Greek affiliations* (See Section 6.3, *Analysis*) and, thus, fall under the static class. The remaining 12 (=19 – 7) researchers were not found in none of the means explored and as such were assumed to be as non-existent.

2 As the analysis focusses on the time interval 2008-2018, the mobility attitudes of these 45 authors were examined in a separate manner. On this, we followed the same methodological framework (Section 6.2, *Model*). According to the results, all individuals followed domestic scientific career paths, that is all authors were associated with *only Greek affiliations* (Section 6.3, *Analysis*) and, thus, fall under the static class.

3 In this paper, the authors adopted an exploratory approach linking bibliometric data with social media. For the purposes of this analysis, data from LinkedIn were used. These data proved to be sufficient in complementing the bibliometric data in identifying the mobility of the researchers. Thus, no other internet (homepages/available CV's of researchers) or social media (Facebook/Twitter accounts) sources were probed. On the latter, such social media sources are not centred around the provision of detailed professional-oriented content, as is the case with LinkedIn, and thus entail a high degree of subjectivity and incomparability. Concerning CV's, it is to be noted that web availability should not be taken for granted.

4 Using January as a reference point, the time passed (measured in months) until the change of the affiliation country.

References

Cervantes, M. and Guellec, D. (2002), "The brain drain: old myths, new realities", *OECD Observer*, available at: www.oecdobserver.org/news/archivestory.php/aid/673/Thebraindrain:Oldmyths,newrealities.html#sthash.Dbm1chlV.dpu (accessed 21 January 2019).

Chalamandaris, A. Protopapas, A. Tsiakoulis, P. and Raptis, S. (2006), "All Greek to me! An automatic Greeklish to Greek transliteration system", available at: www.cs.brandeis.edu/~marc/misc/proceedings/lrec-2006/pdf/390_pdf.pdf (accessed 21 March 2019).

Dyachenko, E. and Mironenko, A. (2018), "Publication trajectories of Russian cardiologists: ways to attain international visibility", in: *23rd International Conference on Science and Technology Indicators, Leiden, NL, 12-14 September 2018*.

Gendronneau, C., Wiśniowski, A., Yildiz, D., Zagheni, E., Fiorio, L., Hsiao, Y., Stepanek, M., Weber, I., Abel, G. and Hoorens, S. (2019), *Measuring Labour Mobility and Migration Using Big Data*, Publications Office of the European Union, Luxembourg.

Hunter, P. (2013), "Brain drain, brain gain or brain sharing?", *EMBO Reports*, Vol. 14 No. 4, pp. 315-318.

Karakos, A. (2003), "Greeklish: an experimental interface for automatic transliteration", *Journal of the American Society for Information Science and Technology*, Vol. 54 No. 11, pp. 1069-1074.

Krause, J. Mele, S. and Lindqvist, C.M. (2007), "Quantitative study of the geographical distribution of the authorship of high-energy physics journals. No. CERN-OPEN-2007-014", available at: https://cds.cern.ch/record/1033099 (accessed 21 March 2019).

Labrianidis, L. (2017), "Why high school students aspire to emigrate: evidence from Greece", *Journal of International Migration and Integration*, Vol. 18 No. 1, pp. 107-130.

Labrianidis, L. and Pratsinakis, M. (2014), "Outward migration from Greece during the crisis", *Project funded by the National Bank of Greece through the London School of Economic's Hellenic Observatory 2014 call*, available at: www.lse.ac.uk/europeanInstitute/research/hellenicObservatory/CMS%20pdf/Research/NBG_2014_-Research_Call/Final-Report-Outward-migration-from-Greece-during-the-crisis-revised-on-1-6-2016.pdf (accessed 4 March 2019).

Labrianidis, L., Sachini, E. and Karampekios, N. (2019), "Establishing a Greek diaspora knowledge network through 'knowledge and partnership bridges", *Science and Diplomacy*, Vol. 8 No. 1.

Laudel, G. (2003), "Studying the brain drain: can bibliometric methods help?", *Scientometrics*, Vol. 57 No. 2, pp. 215-237.

Martineau, T., Decker, K. and Bundred, P. (2004), "Brain drain of health professionals: from rhetoric to responsible action", *Health Policy*, Vol. 70 No. 1, pp. 1-10.

Moed, H.F. and Halevi, G. (2014), "A bibliometric approach to tracking international scientific migration", *Scientometrics*, Vol. 101 No. 3, pp. 1987-2001.

Robinson-Garcia, N.S.Y. (2019), "The many faces of mobility: using bibliometric data to measure the movement of scientists", *Journal of Informetrics*, Vol. 13 No. 1, pp. 50-63.

Sugimoto, C., Robinson-Garcia, N., Murray, D., Yegros-Yegros, A., Costas, R. and Larivière, V. (2017), "Scientists have most impact when they're free to move", *Nature*, Vol. 550 No. 7674, available at: www.nature.com/polopoly_fs/1.22730!/menu/main/topColumns/topLeftColumn/pdf/550029a.pdf (accessed 21 January 2019).

Van Eck, N.J. (2014), "Visualizing bibliometric networks", in *Measuring Scholarly Impact*, Springer, New York, NY.

Wagner, C. and Jonkers, K. (2017), "Open countries have strong science", *Nature*, Vol. 550 No. 7674, available at: www.nature.com/news/open-countries-have-strong-science-1.22754 (accessed 21 January 2019).

Waltman, L. and Noyons, L. (2018), *Bibliometrics for Research Management and Research Evaluation*, Centre for Science and Technology Studies, Leiden.

Yuret, T. (2017), "An analysis of the foreign-educated elite academics in the United States", *Journal of Informetrics*, Vol. 11 No. 2, pp. 358-370.

Further reading

Jacob, J.H., Lehrl, S. and Henkel, A.W. (2007), "Early recognition of high-quality researchers of the german psychiatry by worldwide accessible bibliometric indicators", *Scientometrics*, Vol. 73 No. 2, pp. 117-130.

Moed, H.F., Aisati, M. and Plume, A. (2013), "Studying scientific migration in Scopus", *Scientometrics*, Vol. 94 No. 3, pp. 929-942.

Sugimoto, C. Robinson-Garcia, N. and Costas, R. (2016), "Towards a global scientific brain: indicators of researcher mobility using co-affiliation data", available at: https://arxiv.org/ftp/arxiv/papers/1609/1609.06499.pdf (accessed 25 January 2019).

Corresponding author

Nikolaos Karampekios can be contacted at: nkarampekios@ekt.gr