



Mapping scientific mobility in leading Eurozone economies: insights from ORCID data analysis

Alicia Moreno-Delgado¹ · Marlon Cárdenas-Bonett² ·
Óscar de Gregorio-Vicente¹ · Julio Montero-Díaz¹

Received: 19 February 2024 / Accepted: 5 September 2024
© Akadémiai Kiadó, Budapest, Hungary 2024

Abstract

Research into the mobility of researchers has garnered increasing interest among institutions and governments. In this study, we use ORCID as a data source to analyse the mobility of researchers trained in Germany, France, Spain, the Netherlands, and Italy, the main economies of the Eurozone according to Gross Domestic Product (GDP). Our approach focuses on the connection between the place of education and employment, identifying graduates and their countries of employment through profiles on ORCID. We conduct a comparative analysis of preferred destinations, considering various levels of education, and develop a migration rate for researchers from these countries. The results reveal a clear preference for the United States and Great Britain among graduates, influenced by linguistic affinities and historical cultural relations. Regarding the migration rate, we observe that all countries retain more graduates than those who emigrate. France leads in emigration, followed by the Netherlands, Germany, Italy, and Spain. This analysis of researcher mobility in the Eurozone allows us to track migratory flows, identifying both sending and receiving countries. These findings are essential for the formulation of scientific and migration policies and contribute to understanding individual behaviour in building academic and professional careers.

Keywords ORCID · Scientific mobility · Internationalisation · Research policy · Europe

✉ Alicia Moreno-Delgado
aliciamorenodelgado@gmail.com

Marlon Cárdenas-Bonett
mf.cardenas@soprasterianext.com; marlon.cardenas@unir.net

Óscar de Gregorio-Vicente
oscar.degregorio@unir.net

Julio Montero-Díaz
julio.montero@unir.net

¹ Vicerrectorado de Transferencia, Universidad Internacional de La Rioja, Logroño, Spain

² Escuela Superior de Ingeniería y Tecnología, Universidad Internacional de La Rioja, Logroño, Spain

Introduction

Researcher mobility sparks the interest of institutions and governments to assess and develop scientific and migration policies. Researchers relocate from one country to another to continue their education, gain prestige, and enhance their career prospects. In this sense, we can distinguish between educating or training countries and employing countries. The former invest in the basic training of researchers and host them during this stage, while the latter receive them, facilitating their professional or research development.

Mobility is of considerable interest for various reasons, both to academics and from a scientific policy perspective (Gomez et al., 2020). It facilitates the exchange of knowledge, fosters international collaboration, and can enhance the overall quality of research. This topic has also been extensively reviewed in the literature (Gureyev et al., 2020; Netz et al., 2020; Siekierski et al., 2018) and addressed across various disciplines. Bibliometric analyses have examined publications and authors' affiliations to study mobility (Moed et al., 2013; Robinson-Garcia et al., 2019; Robinson-García et al., 2018; Sugimoto et al., 2017). However, this method does not capture all forms of mobility, as data is only obtained when it is associated with a publication (Robinson-Garcia et al., 2019). Additionally, information about higher education studies is unavailable unless the researcher publishes it while studying in college, which is uncommon. Therefore, it is necessary to resort to other sources that provide data with fewer biases. In this context, utilizing data directly provided by the researchers rather than relying on institutions or publications will help avoid the aforementioned distortions.

ORCID (Open Researcher and Contributor ID) is “a unique, persistent identifier free of charge to researchers” that aims to “enable transparent and trustworthy connections between researchers, their contributions, and their affiliations by providing a unique, persistent identifier for individuals to use as they engage in research, scholarship, and innovation activities” (About ORCID, n.d.). But it is not just an identifier; it is an opportunity as a source of scientometric data for meta-research. Researchers use this ID to be identified by their peers and by managers and institutions. It makes available their CV to the academic and scientific community at the level of information that the user decides. Journals also promote the use ORCID and frequently demand it as an indispensable requirement for authors. Additionally, managing entities also resort to this identifier for professional evaluation, hiring, or searching for specialists in a field and funding agencies and governments allow importing and synchronizing directly from ORCID for project calls.

However, there are a series of drawbacks that need to be considered when using ORCID as data source. Thus, consideration must be given to fake profiles, misuse (Baglioni et al., 2021; Teixeira da Silva, 2023), or ghost or disposable accounts (Teixeira da Silva, 2021a, 2021b). Additionally, there are researchers who, although they are prominent in academia and publish relevant works for a field, do not complete their profile as they do not need the promotion or visibility offered by ORCID; they only require the ID to publish in certain journals or complete profiles for competitive project applications. Among the identified biases, there is also the unequal representation of countries due to the use of ORCID as a central identifier, which could alter country-level comparison data (Costas et al., 2022). Furthermore, data standardisation needs improvement as it currently complicates analysis and classification. On the other hand, ORCID data may vary from those used to study researcher mobility with bibliometric data because highly skilled/educated individuals may not necessarily participate in academic publications (Orazbayev, 2017).

Although its use is not yet widespread and the existence of biases, ORCID provides self-declared biographical data that allows mapping academic trajectories from different perspectives (Costas et al., 2022). This information is valuable for meta-research as unlike other products, it places the researcher at the centre and provides various perspectives in an integrated manner: education, employment, and publications (Costas et al., 2022). Thus, some researchers have previously used ORCID to analyse the mobility of research professionals. Yan et al. (2020) explored the mobility of professors in the U.S. using ORCID data and found that they tend to move to institutions with a higher research intensity. This discovery suggests a preference for academic environments that promote research and innovation. In turn, the global study by Gomez et al. (2020) provided a broader view of scientific mobility patterns worldwide. In this paper, an increase in mobility within the same transnational region was observed, indicating a trend toward regional mobility rather than purely global migration. These results suggest that collaboration and knowledge exchange may be driven more by regional factors than by international movements. Similarly, the analysis by Raghupathi et al. (2023) examined migration records of PhDs from 194 countries, noting a recent decrease in migration trends. Despite the increase in the total number of PhDs, international mobility is stagnant, which could have significant implications for the global circulation of scientific knowledge. Lastly, the study by Kim (2022) focused on postdoctoral mobility in STEM fields, highlighting that the U.S. and European countries continue to be the primary destinations for the first postdoctoral position. However, gender differences in mobility trends were observed, with female postdoctoral researchers showing a greater propensity to remain in a single country or to avoid moving to the U.S. for their second postdoctoral position compared to their male colleagues.

Traditionally, Europe has exported highly qualified researchers to the United States. The United Kingdom has also been a country sought after to boost research careers. This trend is self-perpetuating as highly prolific scientists move and preferentially stay in central cities within the international scientist mobility network (Verginer & Riccaboni, 2021). The openness to migrate to these two countries persists, but new recipients have also emerged for distinct reasons such as linguistic, cultural, and historical proximity. In order to delve deeper into this phenomenon, this study analyses the mobility of researchers from the 5 main countries in the Eurozone in terms of Gross Domestic Product (GDP). This analysis will contribute to highlighting both the investment in education and the job opportunities in each country for these profiles. The data allows for traceability in the migratory flows of researchers, which countries train them, and which are the recipients. In other words: where the talent trained in each country ends up and will contribute to the evaluation of policies developed in this area allowing them to be compared with expectations and forecasts. Gureyev et al. (2020) identified the most prevalent topics in papers analysing scientific mobility. They observed that 38% of the analysed works focused on the factors inducing scientific mobility, another 38% on its impact, 15% on mobility flows, and only 5% on developing approaches to study it.

Considering that mobility flows are underrepresented in mobility studies, the main objective of this research is to investigate and analyse labour mobility among researchers from the European countries with the highest GDP, using ORCID data, to identify patterns and trends that better explain the factors influencing education and employment decisions in the scientific field. To achieve this, the following specific objectives are proposed:

- Identify researchers trained in the major Eurozone economies based on GDP: Germany, France, Spain, the Netherlands, and Italy.

- Analyse differences in researcher mobility based on the type of studies undertaken, distinguishing between master's and doctoral degrees.
- Identify the migration rate of users from the countries under study.

This study delves into the mobility flows and captures the situation at a specific moment, the year 2022. However, the explanation of the situation must take into account circumstances that occurred years before (for example, the economic crisis of 2008–2011; or the COVID-19 epidemic). Although it is currently not possible to quantify the specific effect of each phenomenon, the impact of official online university degrees should be considered in future studies, as they could hide graduates from one country (e.g., Spain) who actually reside in another. The data presented here does not discriminate this issue.

Methodology

The breadth and detail of the information analysed in this study directly depend on the data provided by ORCID users. This contribution occurs both at the time of their initial registration on the platform and through subsequent updates to their profiles. Additionally, ORCID implements mechanisms to track the academic activity of its users. These systems are designed to identify and suggest the inclusion of new relevant information in researchers' profiles. However, it is worth noting that the incorporation of such information into a user's ORCID profile is only carried out upon receiving their approval. For this study, data from the year 2022 were used, which were extracted in July 2023 for processing and analysis.

Data download and extraction

This study addressed the collection and analysis of data from researcher profiles registered on the ORCID platform. The implemented methodology involved several essential stages, including the definition of the target group, as well as the loading, transformation, validation, and parameterization of the data. An export of data from ORCID was used, containing information provided by researchers both at the time of registration and in subsequent updates. This dataset encompassed researcher profiles created between 2015 and 2023.

ORCID shares a public data file at least once a year, considering it as a vital part of the academic communication infrastructure, thus making it available to all stakeholders, including organizations that are not ORCID members.

There has been public ORCID data available since 2013, with the understanding that, from 2018 onwards, the file is provided only in XML (Extensible Markup Language) format, which can be transformed into JSON (JavaScript Object Notation) format files.

Each annual data file may include multiple files in tar.gz compression format. For instance, the 2013 file includes a single public tar file. However, the 2022 file includes a tar.gz file for all ORCID records (Summaries file) and a tar.gz file for all registration activities, from 1 to 10.

As indicated on the ORCID website (Working with Bulk Data, n.d.), it is important to note that: (a) the annual activities file contains eleven different compressed files, each of which contains a subset of the public activities available at the time of file generation, based on the last 3 digits of the ORCID identification; (b) the annual summaries file is a tar file containing summaries of public records of all ORCID available when the file

was generated, with the folder hierarchy defined by the last three digits of the ORCID identification:

The generated folder for each annual file varies depending on the year the file was generated and the corresponding XSD (XML Schema Definition) version, such that:

- From 2013 to 2017, within the generated folder, there are several folders, for example, JSON/ and XML/, in the specified format and XSD version.
- From 2018 onwards, within the generated folder, there is: (a) a summaries folder, which in turn contains various folders containing ORCID records in XML format, such that the records are aggregated into subfolders based on the shared last three digits of the ORCID identification; (b) multiple folders for each ORCID record with the complete activities in each ORCID record in XML format, separated by activity subsection.

XSD is a schema language used to describe the structure and constraints of the contents of XML documents in a very precise way, beyond the syntactic rules imposed by the XML language itself.

In this study, the decision was made to download the complete summary files (compressed) for all available years using the capability of a virtual machine. Thus, the research team conducted initial tests of downloading and analysing records with 2021 data to identify and address various issues: (a) the substantial increase in information volume when decompressing the initial annual tar.gz file into the resulting.tar file (also compressed), containing folders encoded according to the last three digits of the ORCID iD, in which the different.xml files with corresponding information records are located. For example, in the case of the 2021 Summary, this means going from around 20 GB initially to 230 GB after the first decompression, which results in more than 7.7 million XML files, each with its corresponding records; (b) managing the vast number of temporary files generated in the decompression processes; (c) addressing the loss of estimated free storage space due to the additive effect of (a) + (b); (d) optimising the antivirus system used for this volume of records; (e) the necessary backup system prior to data structure analysis and subsequent debugging and validation processes.

Once these phases were completed, the integration activity of the different.xml files into a single query database began. It should be noted that the original.xml files do not have the same vectorial structure (columns) among them. For this reason, the research team decided to use the MongoDB tool, which meets the aforementioned conditions and is oriented towards documents/files with non-vectorially structured data. This is because each researcher incorporates into their profile the information they consider relevant or have information to include in ORCID, so the field structure does not guarantee uniformity in the information of the records included in different years by different researchers.

During the structural integration, query indices (numeric or text) were defined to facilitate subsequent searches in the unified database, which could be simple or compound to expedite the process.

Once all tests were completed and both the equipment and starting processes were configured, the download of the 2022 data was initiated, which was collected in July 2023 for subsequent processing.

The total volume of information extracted, after the initial decompression, was 467 GB, corresponding to 15.4 million XML files. Extraction procedures were developed to transform these files into complete researcher profiles, represented as individual JSON documents. Given the inherent complexity of directly reading XML files, transformation protocols to JSON format were established.

The working environment included a global database hosting the original export, loading and transformation scripts, and a variety of reports, among other elements. Scripts were tested and executed by installing libraries in Jupyter, which is an open-source web application, to optimize real-time processes.

Following the conversion of data to JSON format, integration into MongoDB, a document database, was carried out. In this database, each researcher profile was stored as a unique JSON document, identified by their ORCID code. To ensure data relevance and accuracy, a strict loading, review, and validation process was established.

Several databases were set up to host both complete and partial information. The main database, named “researchers,” contained an exact replica of the original ORCID data, without processing or cleaning. This base, maintaining ORCID’s original format, presented a complex structure that did not facilitate direct queries. Other databases included “researchers_education,” which exclusively compiled researchers’ educational data, and “researchers_employment” for professional information. All these bases were linked by the ORCID code, identified in the “_id” field.

During extraction, various inconsistencies were identified, leading to the exclusion of certain records due to their limited contribution to the analysis. These included profiles with insufficient information on education, professional activity, and research, duplicate data, lack of data on country of origin or residence, and discrepancies between the information provided by researchers and that obtained by ORCID through its web extraction mechanisms. Exclusion criteria were also applied to records manifesting linguistic issues, especially due to the inclusion of character sets from Asian languages. These aspects presented significant challenges for ORCID’s automatic data extraction algorithms. It is noteworthy that such linguistic errors, as well as issues related to data insufficiency, were primarily detected in the database as a whole and were not part of the specific study objectives.

Data processing

The study focuses on the five leading Eurozone countries in terms of Gross Domestic Product (GDP): Germany, France, Spain, the Netherlands, and Italy, due to their economic significance in the region. While additional countries such as the United States and the United Kingdom are mentioned because of their prominence in scientific research and their status as preferred migration destinations, the selection was deliberately limited to maintain the coherence of the study.

ORCID includes a “country” field which can be understood as either the country of origin or the current country of residence. Here, the country corresponding to the education-country field has been chosen as the country of origin: the country where the ORCID user has been educated. This option has been used in other studies (Orazbayev, 2017) where the education-country field of the ORCID profile is used as the likely place of origin of the researcher.

Regarding the country of work, the employment-country field has been selected. For profiles where more than one workplace is listed, the latest, most current one has been chosen. ISO 3166-1 alpha-2 codes, which correspond to two-letter country codes, have been used to name the countries.

The data has been divided into two collections: *researchers_education* and *researchers_employment*. In MongoDB Shell, searches with JSON codes were performed by filtering for each study country in the *researchers_education* collection. Once filtered

by country, another collection was generated which merged with *researchers_employment*, and each document of the collection was merged for those sharing ORCID. Subsequently, a search was conducted in MongoDB Compass for the study country and the results were exported to create a new collection. In this way, ORCID profiles without education information were removed after unifying them with the *researchers_employment* collection. Next, the search {"documents.employment-country": {"\$exists": true}} was performed and the result was exported in JSON. With this step, records where the employment-country field was empty were eliminated.

Finally, with the necessary information, a CSV was generated with the required fields for this work (education-country for verification, employment-country, and ORCID), and a macro was applied to the worksheet to remove empty cells from the document. This first block thus includes any ORCID registrant who has obtained a university degree (bachelor's, master's, or doctoral) or any prior to the establishment of the European Higher Education Area in any of the five countries studied. It also includes non-European students who have completed studies for any of these degrees in these countries and are likely to have a higher tendency to return to their home countries. Also, European students outside the five countries of the study and in the same circumstances are included. In any case, their registration in ORCID indicates their predominantly research-oriented dedication, distinguishing them from those who completed their studies, especially master's degrees, with an exclusively professional purpose.

For the elimination of postgraduate studies, the same method was followed, but incorporating the extraction of documents that do not contain the characters of the *education-role-title* field. Not all concepts have been applied to all countries because the *education-role-title* field of each dataset per country was reviewed, and characters were added and applied to each country. These characters, whether for master's or doctoral degrees, were as follows:

MAESRIA/máster/Maestría/Magister/Marter's Degree/MastÃ~r/MastÃ~r/Mastar/Maste/MASTEAM/MASTEN/MSc/Master's/MBA/MBIM/MC Ecological Economics/Mcs Water Science and Technology/ME in Chemical Engineering/Meng/Mestra/Mestrado/Mester/Mestre/mFIA/MFA/MG/MGA/MGH/MGR/Mgter/MGS/Mres/MS/MA Architecture/Masters/Meisterklasse/MasterÂ's/M. A./M. P. A. Wissenschaftsmanagement/M. Sc./M. Eng/M. Ed./M. Eng./M. Sc/MA/MaestrÃa/MaÃ~trise/Mag./Magistra/Maitrise/"MaÃ~trise/"Maitrise"/Master/(Master)/(Msc)/Advanced Masters/Advanced Master/DPEA/DPLG/MPhil/M. S./M.A./M.Res./M2/M1/MÃ;ster/MÃ~itrise/MÃSTER/MaÃ~Ã~trise/MaÃ~trise/MaÃ~tre/MagÃster/MagistÃ~re/magistere/MastÃ~re/Mastaire/Masterâ€™s/Mastere/Masters/Master's/â€™Masterâ€™s/MPharm/MPhys/Mastersâ€™/LL. M/LL.M./LLM/M B A/M. Comn/M.Des./M.Div./M.Ed./masters/MHPE/MPH/Mr./Mr/Mrs/Mrs./MScEng/MScGH/MSci/MSEE/MSEN/MSPT/MSS/MSW/MT/MTD/MTh/Mtheol.

The following terms related to doctoral training have been removed:

Doutora/Doctora/Doctorado/Doctorando/Doctoranda/Doctoral/PhD/Ph.D/Doctorate/Doctorat/Doctor/Doktor/Doktora/Doktorand/Doktorandin/Doktorarbeit/Doktorat/Doktorate/Doktorate/PhD/Doktorgrad/PhD/Doktorin/Doktoringenieur/Doktor-Ingenieur/Doktors/Dokortitel/Dorctor/Dorktor rer. nat./Dotor of Engineering/Dotate/Dott/Doutorado/Doutorado-SanduÃche em Antropologia/Doutoramento/doutoranda/Doutoranda convidada/Doutorando/DOUTORANDO e pesquisador cientÃfico/DPhil/Dr/Dr./Docteur/"DOCTORAT/(PhD)/DOCTORATE/Doctrante/Dosteure/Doteure/Doteur/Dottoranda/Dottorato/Dottore/Doucteur/Doutor/Doutorad/Douturamento/DSc/Dottorando/Doctoral/Dottoressa/Dottorandi/Dott.ssa/Dott.ric/dotterssa/Dottor/Dott./

doctoratus/Doctorial/doctor's/Doctory/Docter/Docterandus/Doctoraal/Doctoraalexamen/Doctoraalstudie/Doctorandus/Doctorem/Doktoraal/doktorandus/Dotorandose/Drs.

In this filtering section, seven profiles with Italy (IT) as education-country were removed due to each profile containing more than 300 items and the difficulty in managing the information.

The last group, those with basic-level qualifications (bachelor's or equivalent title prior to the implementation of the EHEA, excluding those only listed as master's or doctoral students) and registered in ORCID, represent the most stable residency group. Firstly, due to the longer duration of their studies (compared to master's degrees); then due to the earlier age at which they are undertaken (typically starting between 17 and 18 years old). Like the previous case, their registration in ORCID indicates a predominantly research-oriented activity, and very frequently within university settings.

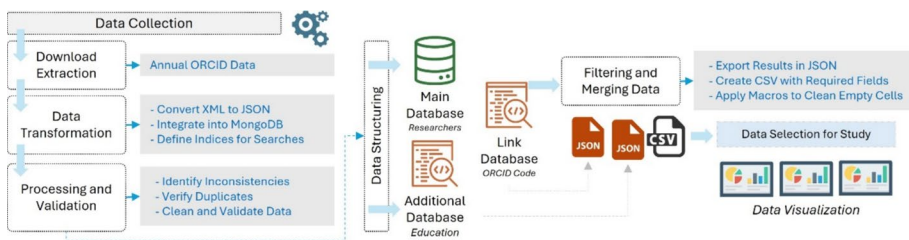
In summary, students educated in the countries under study are identified as those whose profiles mention any of these countries in the education-country section. The data are then refined by eliminating profiles referencing master's studies undertaken in the respective country, identified through specific keywords. These results form the second data block, which is further filtered to exclude profiles mentioning doctoral studies, also detected through country-specific keywords. The ultimate aim is to analyse graduates who have completed undergraduate studies in one of the countries under study and determine their current employment location, observing the process as a progressive pruning of a tree, and revealing the differences among various types of studies (Fig. 1).

Results

Orcid rate

Firstly, this work has identified ORCID users who have studied in each of the 5 countries analysed: Germany (DE), France (FR), Italy (IT), Spain (ES), and the Netherlands (NL). It has been determined in which country those who studied in each of them work. ORCID, with logical caution, allows the identification of graduates from institutions in these countries and identifies their workplaces. Therefore, an immigration rate can be generated.

The process was as follows: the number of ORCID identifiers from each country was first extracted and compared with the total population of each. It should be noted that this number of identifiers corresponds to profiles that have incorporated both the education-country and employment-country fields. The relationship between this total number



Data Processing and Integration Pipeline for ORCID Research Profiles

Fig. 1 Summary of the data extraction, structuring, and filtering process

Table 1 Population and ORCID rate

Country	Population	No. ORCID	ORCID rate
DE	83,196,076	61,350	0.074
FR	67,738,863	57,711	0.085
IT	59,133,173	64,987	0.110
ES	47,331,302	111,540	0.236
NL	17,533,048	28,625	0.163

Source: Compiled from data obtained from OECD data (<https://data.oecd.org/>) and ORCID

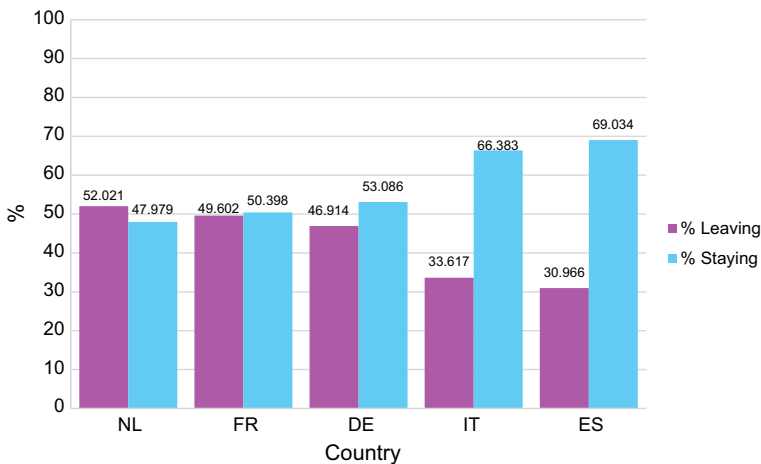


Fig. 2 Emigration rate. Percentage of graduates who emigrate and stay in Germany, Spain, France, Italy and the Netherlands

of identifiers and the population is the ORCID rate (Table 1). Spain has the highest rate: 0.236%. Germany is at the opposite end: 0.074%. Germany is precisely the country with the largest population in the group. Additionally, figures for Great Britain (GB: 0.218%) and the United States (US: 0.112%) are provided to facilitate comparison and contextualisation, as they are the main powers in the research field. The former is in second position after Spain, and the latter in fourth place. The data for Spain probably reflects a high number of students of Latin American nationalities who are undertaking or have already completed their doctoral studies. Comparison with the number of doctoral theses defended in each analysed country would help confirm this hypothesis.

The Atlantic bloc and linguistic affinities: educating countries and employing countries

The second step was to identify both the profiles of users who state that they have obtained a degree (usually a bachelor’s, master’s, or doctoral degree, or equivalent qualifications prior to the EHEA) in any of the analysed countries, as well as those who are working professionally in countries different from the first (Fig. 2). Additionally, it was noted which

countries were chosen by these latter individuals. The data suggests that there is a high percentage of graduates who decide to work outside the country where they completed their studies. Thus, these graduates exceed 50 per cent in the Netherlands (52.021%) and over 40 per cent in France (49.602%) and Germany (46.914%). In contrast, Spain has the lowest percentage of university graduates who migrate to other countries: only 30.966%.

If we look at the receiving countries (Fig. 3), graduates generally preferentially head to the United States and Great Britain: these are the places that receive the most university graduates from these five analysed European countries. More than 9% of the total opt for (and are accepted by) the United States and over 7.5% of the total go to Great Britain. Other countries with a percentage around 4 to 5% are Brazil, Switzerland, Colombia, and Germany.

If we look at each country (Table 2 highlighted in green), graduates from Germany who leave there preferentially opt for the United States, which hosts 11.9% of this group (German graduates working abroad). In second place, they choose Great Britain (7.3% of the total). Switzerland is the third preferred foreign choice for German graduates at 7.1%.

Great Britain is the preferred place to work for graduates from Italy and the Netherlands. Those graduating in Spain opt for Latin American countries: Colombia, Ecuador, and Mexico in that order. Those who have graduated in France and have decided to leave work in the USA (9.1%), Brazil (7.4%), and Great Britain (6%). Italian graduates, as observed in Table 2, move to the United States in second place, as do the Dutch, while Spain (for Italians) and Germany for the Dutch occupy the third position.

Therefore, the Atlantic science bloc centred around the United States and Great Britain is identified. All countries except Spain count these two research powerhouses among their top three preferred destinations. Within this bloc, but with some independence, are two others: one being European, marked by the cultural affinities of Romania and Germany; the other organized around Spain: The Ibero-American bloc.

To refine further and identify aspects that may determine employment, it was decided to exclude graduates with master's degrees from the analysis. This is intended to avoid including students who, having completed most of their studies in one country, went to

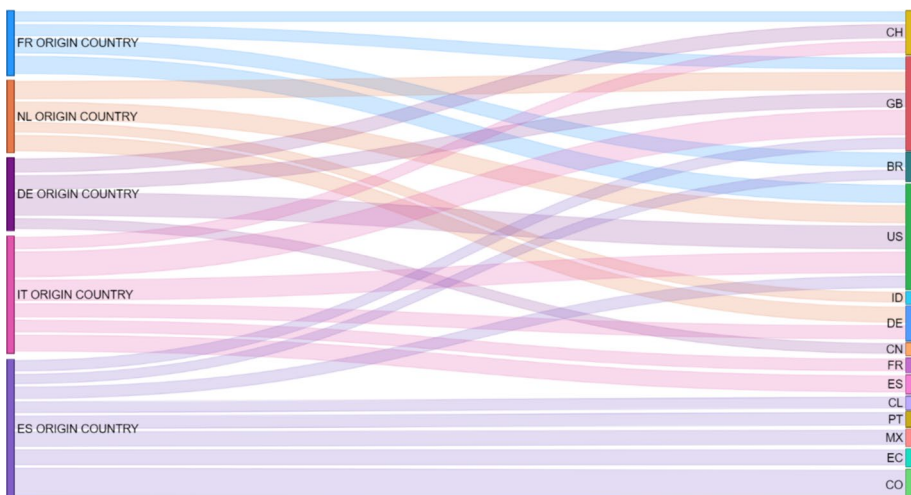


Fig. 3 Graduates' flow in percentages by countries above 5%

Table 2 Employment statistics of graduates with over 1% employment

Employer country	Graduates					No Master's in their home country					No Master's/Phd in their home country				
	DE	ES	FR	IT	NL	DE	ES	FR	IT	NL	DE	ES	FR	IT	NL
AR		2.3					2.0					1.7			
AT	4.8			2.1	1.3	4.8			2.0	1.4	5.7	1.2		2.2	1.7
AU	2.4		1.6	1.5	2.8	2.2		1.6	1.4	2.7	2.6		2.2	1.6	3.3
BE	1.5	1.4	2.6	2.8	4.7	1.3	1.5	2.4	2.5	4.4	1.6	2.3	3.2	2.7	6.1
BR	3.7	5.1	7.4	3.5	2.4	4.0	5.6	8.1	3.7	2.9	3.4	4.4	5.7	2.7	2.7
CA	1.8		3.5	1.4	1.8	1.8		3.7	1.5	1.8	2.0	1.2	4.0	1.3	1.4
CH	7.1	1.6	5.1	6.1	3.5	6.9	1.9	5.3	6.3	3.4	9.3	2.9	6.7	6.9	4.5
CL		5.8	1.3			1.0	6.1	1.3				3.4			
CN	5.4		3.3	1.4	4.1	6.2		3.9	1.6	5.7	4.1		3.0		3.6
CO	1.3	14.6	2.5	1.5	1.2	1.2	11.8	2.1	1.2	1.1		7.9	1.7	1.1	
DE		3.3	4.9	7.0	8.1		3.7	4.5	6.9	8.2		5.6	6.0	7.7	10.8
DK	2.3			2.1	2.2	2.2	1.0		2.0	2.2	2.5	1.6		2.1	2.7
DZ			1.2					1.3					1.0		
EC		8.0					5.6					3.6			
EG	1.4					1.6									
ES	3.3		4.6	8.5	3.7	3.0		4.1	8.7	3.1	3.5		5.6	10.1	4.2
ET					1.2										
FR	2.6	2.7		6.3	2.2	2.5	3.2		6.1	2.5	2.6	4.3		6.1	2.2
GB	7.3	5.6	6.0	12.9	9.2	7.1	6.6	5.9	13.3	9.5	8.6	9.7	7.2	15.0	10.7
ID	1.9				5.2	1.8				4.3					3.1
IE				1.1					1.0			1.1		1.0	
IN	1.9				1.4	2.0				1.3	1.1				
IR										1.1					
IT	2.4	2.2	3.1		2.8	2.5	2.6	3.2		2.9	2.6	3.0	3.5		3.3
LB			1.7					1.8							
LU			1.1					1.1					1.4		
MA			1.3					1.4					1.2		
MX	1.1	7.7	2.9			1.1	8.5	3.1				5.7	2.0		
NL	3.1	1.5	1.5	3.6		2.8	1.7	1.3	3.4		3.5	2.4	1.6	3.7	
NO	1.6				1.5	1.6				1.6	2.0				1.5
PE		4.7					3.3					3.3			
PL	1.4		1.1	1.4	1.1	1.4	1.1	1.5	1.2	1.9			1.4	1.6	1.7
PT	1.3	6.7	2.7	2.7	2.3	1.4	7.8	2.7	2.7	2.6	1.6	7.6	3.0	3.1	2.7
SA								1.0							
SE	2.7	1.2	1.4	2.3	2.3	2.6	1.4	1.3	2.2	2.3	3.3	2.0	1.5	2.3	2.7
TN			1.5					1.7					1.3		
TR	1.8				1.4	1.8				1.5	1.9		1.1		1.8
US	11.9	6.2	9.1	10.7	9.0	12.7	7.1	9.6	11.8	10.2	13.8	9.6	9.6	12.1	9.9
VE		1.1					1.1								
VN			1.6					1.6							
ZA					1.4					1.6					1.1

Graduation from darker to lighter indicates the percentage of employed graduates, with darker tones representing higher percentages and lighter tones representing lower percentages. Green: graduates employed above 1%. Blue: graduates employed above 1% who have not pursued a master's degree in their home coun-

Table 2 (continued)

try. Pink: graduates employed above 1% who have not completed master’s or doctorate studies in their home country. (Color Table online)

another country to pursue a postgraduate study, usually a master’s degree, which typically involves a relatively short stay (between one and two years in most cases). Therefore, the figures mainly refer to graduates or doctoral candidates in their respective countries who did not obtain their master’s degree in any of the five countries considered.

In Fig. 4, the United States and Great Britain stand out again as employers of this group, which includes both those who obtained their bachelor’s degree in one of the five countries analysed but did not pursue their master’s degree there, and those who came solely to obtain their doctorate. Additionally, there remains a potential group of graduates who completed their master’s degree outside the 5 countries in this study and then returned to obtain their doctorate, although they are expected to constitute a small number. In this case, there is even an increase in the emigration figures to these two core destinations, exceeding 10% in the case of the USA. Therefore, little variation is observed when analysing the data collectively.

When conducting a more detailed analysis, it is observed that in the case of Spanish graduates, Colombia maintains the top position followed by Mexico, which rises to second place, and Portugal, which occupies the third position (Table 2 highlighted in blue). Germany and France, although the percentages vary slightly, do not see their top 3 altered, while in the Netherlands only the first and second positions are exchanged, with the United States now being the preferred destination at 10.2%. Italy varies its percentages in the top three positions, increasing in Great Britain and the United States (13.3% and 11.8% respectively).

To further delve into the factors that may alter the results, it was decided to exclude doctoral studies. The aim was to focus solely on undergraduate or bachelor’s studies obtained in one of the five countries under study, as explained in previous sections in more detail. This limits the analysis to those university studies that are not postgraduate. By eliminating the concepts related to this type of education (master’s and doctoral degrees) in the

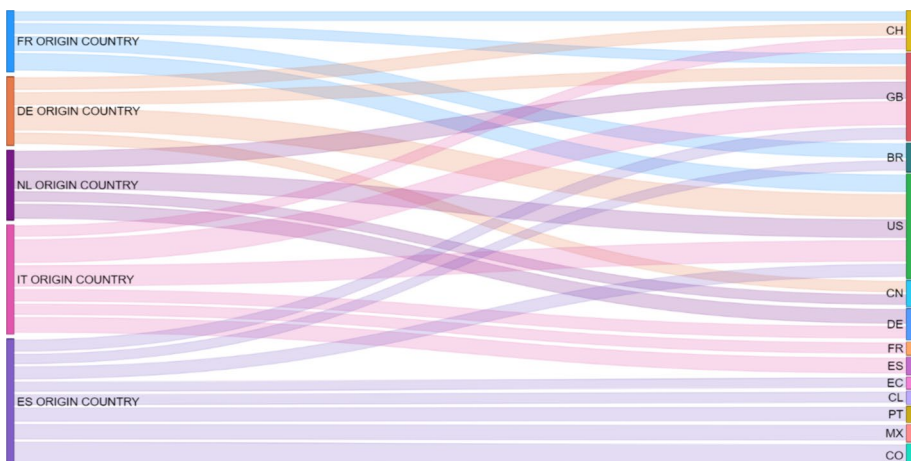


Fig. 4 Percentage distribution of graduates by countries above 5% without master’s degrees

5 countries, the results reflected in Fig. 5 were obtained. The differences in this case are more pronounced, and after the two leading powers in the percentages, Switzerland, Germany, and to a lesser extent, Spain stand out. Furthermore, the figures tend to accentuate the characteristics that define the Atlantic bloc of researchers and the linguistic, historical, and cultural affinities that initially appeared, albeit with some notable nuances.

In this case, more alterations in the country rankings are observed, as reflected in Fig. 6. German graduates favour the United States and Switzerland, followed by Great Britain: there is a shift in positions. In this scenario, research powerhouses also notably stand out in all examined countries. Switzerland appears in the top 3 in two of them: Germany and France. It is interesting to note that these two countries share borders with Switzerland and share official languages, which may explain their prominent position in preference for these. However, this pattern is not repeated in the case of Italy.

In Table 2 (highlighted in pink), the variation in positions can be better observed with data from a larger number of countries. As noted, Spain is the country that undergoes the most changes among the top positions. For example, Ecuador moves to the ninth place as an employer of graduates from Spain, and Mexico to the fifth. These changes could be explained by graduates who come to Spain to pursue postgraduate studies. This would also explain the high number of ORCID identifiers in this country. In the case of Germany, Switzerland moves to second place with the highest percentage of graduates at 9.3%, only behind the United States (13.8%), and followed by Great Britain (8.6%). The preferences of French graduates also change when excluding those with postgraduate studies in France. On this occasion, Brazil moves to fifth place (perhaps because Brazilian graduates obtained their postgraduate

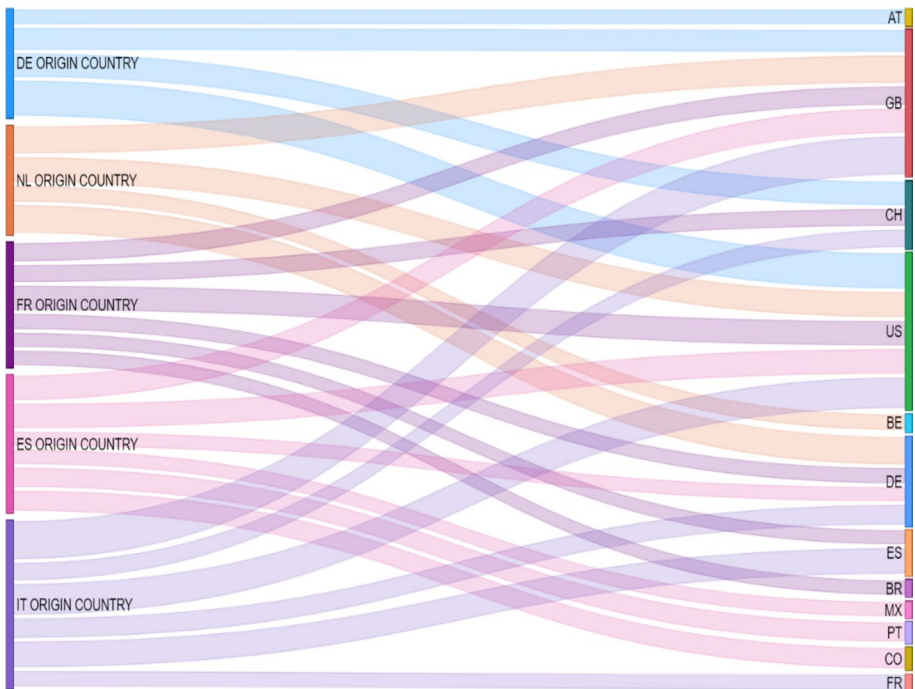


Fig. 5 Flow of graduates in percentages by countries above 5% without master's or doctoral studies

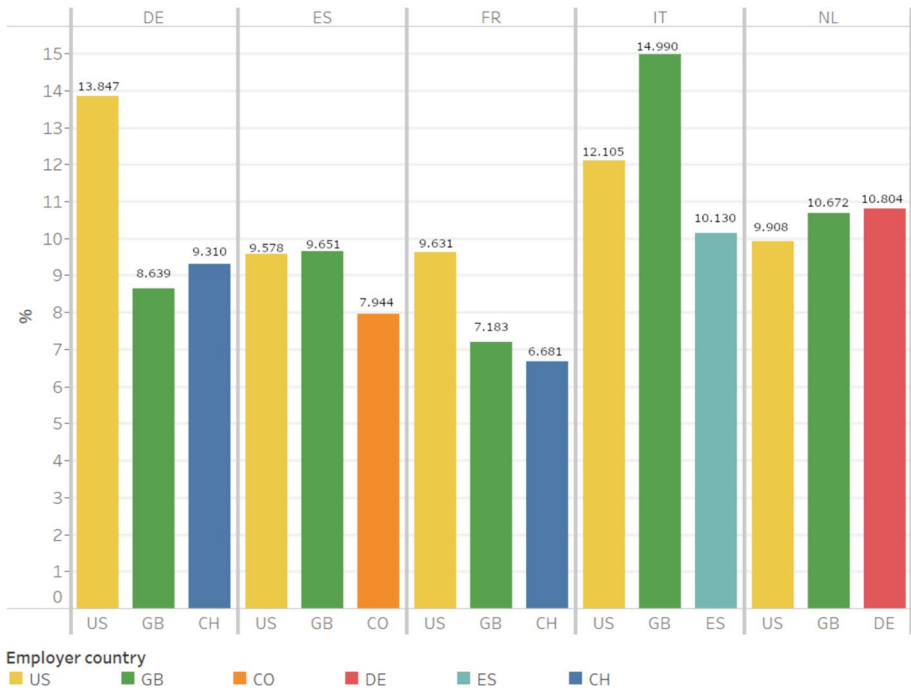


Fig. 6 The top 3 countries with the highest percentage of graduates employed excluding postgraduate studies (master’s and doctorate)

degrees in France and then disappeared back to their country)), while a significant number of French graduates preferred to specialize in the USA, Great Britain, and Switzerland (which occupy the first, second, and third positions) and then stayed there to work. The presence of Italian graduates in Spain could be explained by linguistic similarity added to the difficulties faced by that country in accessing higher education for its graduates, which later facilitates their integration into professional practice in universities where ORCID subjects are mostly concentrated.

Finally, once the destinations where individuals educated in the five analysed countries go to work have been identified, it becomes necessary to identify the emigration rate of these graduates, excluding postgraduate studies (master’s and doctorate). Figure 7 shows the France as the country with the highest emigration rate; 42.872% seek job or training opportunities outside the country. The Netherlands follows with 41.036% of graduates emigrating. Germany ranks third with a rate of 38.858%.

These data place Spain and Italy as the countries with the lowest rate, while the rest are positioned with percentages between 38 and 42%. The figures indicate a lower integration of Italy and Spain into the Atlantic research bloc within the group of the five analysed countries. The Spanish case is nuanced by its leadership within the Ibero-American context.

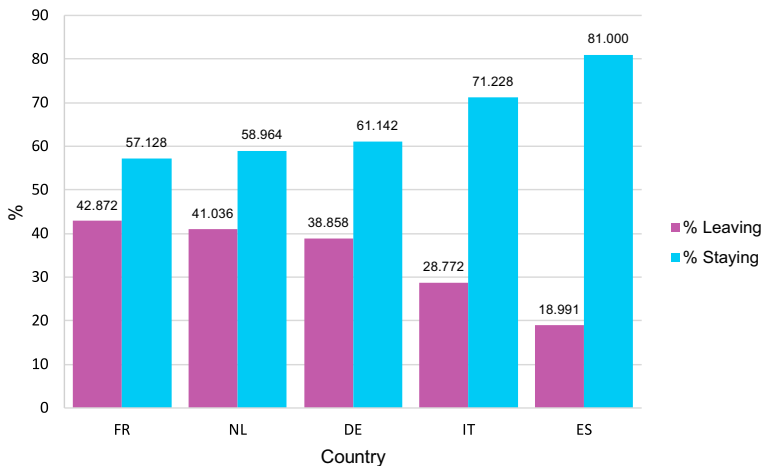


Fig. 7 Emigration rate. Percentage of graduates emigrating and remaining in Germany, Spain, France, Italy, and the Netherlands

Discussion and conclusions

This study aims to follow the path of previous research to validate ORCID as a data source and provide an analysis of researcher mobility. Although it was not created for this purpose, and biases described in the literature must be considered, the value of the information collected by the platform is undeniable, as is the inevitability of its use for such studies (Bohannon & Doran, 2017). This study has focused on understanding the relationship between the place of education and work of researchers from the main countries in the Eurozone. The responses provided are of interest for the analysis and development of scientific and migratory policies and for understanding how individuals behave and shape their academic/professional careers (El-Ouahi et al., 2021; Gomez et al., 2020). It must be understood that opting to research in another country immediately after obtaining a degree usually means embarking on a period of higher education prior to engaging in research tasks (and employment). In most cases, this involves pursuing a master’s degree and almost always a doctorate. In other words, the figures mostly refer to graduates who have chosen to further their higher education in another country, which typically implies opting for one with better academic standing in the discipline or field they intend to pursue in the future.

Here, it has been assumed that the country of education is the country of origin, as other authors have done (Orazbayev, 2017), and that the last employment location recorded in each individual’s profile is their current place of employment. With this, the results demonstrate the widespread preference for the United States and Great Britain among university graduates from the five European powers who end up working in research activities, consolidating a compact bloc in scientific geography: the Atlantic basin. Indeed, around this “inner sea” for science, the majority of academic and research migrations occur. This aspect is notable because the international economic centre is situated on the Pacific coast. It can be affirmed that there is a Western scientific world around the Atlantic that integrates Europe with the United States. Previous studies have also identified the USA as the dominant destination country for almost all the countries analysed (Franzoni et al., 2012; Veugelers & Van Bouwel, 2015). It is an

open world with barely any difficulties in moving around, where English is spoken. As this language is today's academic and, indeed, research "lingua franca", the establishment and development of large-scale projects are more widespread, and personal integration possibilities are not complicated by the social interaction spaces that campuses usually offer initially. That's why a significant number of graduates at various levels (undergraduate, master's, or doctoral) are encouraged to attempt to enter a higher level of research activity. Another issue is what that attempt leads to.

The second major talent movement stream is organized around what could be called linguistic affinities, in line with Moed et al. (2013) description. In this trend, two levels must be distinguished. The first is marked by broad adherence to linguistic Romania or Germania. The second is more specific: the same language. Regarding the former, the movements of Germans to Switzerland as the third preferred place is an example. This is confirmed by the trend of the Netherlands towards Germany (also the third option after the United States and Great Britain) as reflected in Table 2. The French and Italian preference for Switzerland would follow the same line due to the multilingualism of the Swiss Confederation and its transitional status between the Latin and Germanic linguistic worlds. With lower figures, Belgium presents a similar situation due to its bilingualism.

The second internal migration stream of graduates from the five countries considered is articulated around Romania. The first thing that is observed is a hierarchy: there are graduates in Italy who go to France to develop their research project professionally, but there is barely any movement in the opposite direction. The same occurs with Spain, which receives Italian graduates, in line with the results published by Veugelers and Van Bouwel (2015), but hardly sends any of its own to Italy. Portugal receives graduates from Spain as the largest bloc: in fact, it is the non-Spanish-speaking country that receives the most graduates from Spain.

Within the same language context, two blocks of very different volume can be distinguished, but which probably respond to similar causes. Language identity is the key external manifestation, but behind it, it is necessary to align the corresponding traditional historical-cultural relationships and international trajectory: the old European colonial empires. For example, Canada, with significant restrictions on migration, barely receives graduates from Europe who do not come from France.

But the most notable case is Spain's. First, graduates in Spain are the only ones from the five European countries considered in this study who do not choose the United States and Great Britain as their first option for professional development. The figures are clear: the presence of graduate researchers trained in Spain in Colombia (14.5%), Mexico (7.7%), Ecuador (8%), Chile (5.8%), Brazil (5%), and Portugal (6.7%) in proportions higher or very similar to those corresponding to those who go to the United States (6.1%) or Great Britain (5.6%) (Fig. 3 and Table 2). Language identity and historical and cultural relations, in addition to the strength and consistency of reciprocal migration flows, even allow us to speak of a Spanish-speaking university space that maintains its headquarters in the former metropolis. In this movement of graduates in Spain, it must be considered that a high percentage of graduates who migrate to Ibero-American republics are natives of those countries who come to Europe (Spain) to pursue their undergraduate, master's, or doctoral studies and then return to their places of origin and nearby Portugal to practice their profession. Other authors have delved into the relevance of linguistic proximity in student mobility as it simplifies their academic, cultural, and socioeconomic integration (Ovchinnikova et al., 2023). It can therefore be said that there is a particular universe within the Atlantic bloc of researchers that speaks Spanish. It is precisely this feature that tilts their preferences,

probably for obtaining their higher degrees (master's and doctoral) but also for top-tier ones: graduates or, previously, undergraduates.

This study presents some limitations that should be considered when interpreting its findings. Firstly, inaccuracies in categorising graduates by their educational level may arise due to limitations in the available data. Secondly, variations in ORCID registration among countries could introduce biases, especially towards certain regions, and researchers in countries with low ORCID adoption rates are more likely to be disconnected from their profiles (Porter, 2022). Additionally, focusing solely on countries with a high GDP could overlook important trends in researchers' mobility. Furthermore, it's important to note that the data collected from ORCID relies on users' self-reporting, which could introduce biases due to the lack of independent verification of the information.

The conclusions of this study are presented in line with the objectives set. The data provided by ORCID allow us to identify the destinations chosen by graduates from the five countries under study: Germany, France, Spain, Italy, and the Netherlands. In this regard, it can be concluded that, regardless of the type of study and based on the information provided by ORCID users, the USA is the preferred country for Germans and French; Great Britain is the top choice for Italians and Dutch; Colombia is the preferred destination for those educated in Spain. It should be noted that, although the United States continues to be the most influential country in science, as European and Asian scientists become more prolific and influential, the trend is beginning to shift, and scientists are moving to closer countries (Gomez et al., 2020). It should be noted that this proximity is not only physical but also cultural, what has been referred to here as linguistic similarity and cultural affinities; despite English being the lingua franca of researchers. Perhaps because the lives of researchers are not limited to their professional activities.

When filtering out master's studies and eliminating all profiles with these postgraduates in the countries under study, the USA is the choice of Germans, French, and Dutch; Great Britain is the choice for Italians, and Colombia is the country chosen by Spaniards. By further eliminating doctoral studies to obtain a result as close as possible to the origin or nationality of the graduates, it can be concluded that the USA is the preferred country for Germans and French, Great Britain is the main choice for Spaniards and Italians, and Germany is the preferred destination for the Dutch. It is worth noting that all five countries under study have the USA and Great Britain in their top 3 destinations, with each country introducing another different one: Germany to Switzerland, Spain to Colombia, Italy to Spain, the Netherlands to Germany, and France to Switzerland. Other previous studies based on ORCID identified more frequent mobility between nearby countries (Gomez et al., 2020). The most significant changes are observed in the case of Spain, where its graduates migrate to Great Britain and the United States as preferred countries, followed by Colombia. On the one hand, Spanish graduates who aspire to further postgraduate education and then professionally integrate into research teams there migrate to the recognized powers for this reason: the United States. The figures reflect a phenomenon explained precisely by the exclusion of master's and doctoral graduates, who, in respectable figures—those shown by the exclusion of postgraduates—come to Spain to obtain these degrees and then return to their countries of origin and nearby Portugal to practice their profession. There remains a residue of graduates in Spain who, whether native or not of those countries, migrate to work as researchers there and completed their specialized training and research in various places among the five in this study.

These results can be explained by the historical relationship between countries or by their proximity. Linguistic and cultural ties between Spain and Latin American countries, or the Netherlands with countries like Indonesia (former colony). The destination countries

are more diverse among the 5 countries analysed if we consider the full range of studies (including postgraduates). However, as we eliminate postgraduate studies, the destinations become more similar, and the USA and Great Britain stand out in all of them. This may be due to the number of foreign students who decide to attend the 5 countries analysed, to which they feel connected for various reasons, to pursue master's or doctoral degrees. This may also explain the high number of ORCID profiles in Spain compared to the other nations analysed. In this case, it could be a country where a large number of Latin American students come to pursue postgraduate studies.

Regarding the migration rate, it is concluded that all countries retain more graduates than they lose. France has the highest emigration rate (42.87%), followed by the Netherlands (41.04%). On the other hand, Spain has the lowest migration rate, with only 18.99%, followed by Italy with 28.77%. Meanwhile, in Germany, 38.86% emigrate. If we establish a ranking considering this rate, France would top the list, followed by the Netherlands, Germany, Italy, and Spain.

Future work will focus on studying research trajectories over time, rather than at a specific moment, and delving deeper into profiles considering other characteristics such as publications and positions held in organizations.

Author contributions Alicia Moreno-Delgado: Conceptualization; Investigation; Formal Analysis; Visualization; Writing—original draft, review & editing. Marlon Cárdenas-Bonet: Material preparation; Data collection; Formal Analysis; Data structure; Writing—review & editing. Óscar de Gregorio-Vicente: Material preparation; Data collection; Formal Analysis; Data structure; Writing—review & editing. Julio Montero-Díaz: Conceptualization; Supervision; Writing—review & editing. All authors read and approved the final manuscript.

Data availability The repository of the load scripts is available at <https://github.com/mfcardenas/StudyResearcherMobility> ORCID.

Declarations

Conflict of interest The authors have no conflicts of interest to declare that are relevant to the content of this article.

References

- About ORCID. (n.d.). Retrieved 8 June 2024, from <https://info.orcid.org/what-is-orcid/>
- Baglioni, M., Manghi, P., Mannocci, A., & Bardi, A. (2021). We can make a better use of ORCID: Five observed misapplications. *Data Science Journal*, 20, 38–38. <https://doi.org/10.5334/dsj-2021-038>
- Bohannon, J., & Doran, K. (2017). Introducing ORCID. *Science*, 356(6339), 691–692. <https://doi.org/10.1126/science.356.6339.691>
- Costas, R., Corona, C., & Robinson-García, N. (2022). *Could ORCID play a key role in meta-research? Discussing new analytical possibilities to study the dynamics of science and scientists*. SocArXiv. <https://doi.org/10.31235/osf.io/sjck6>
- El-Ouahi, J., Robinson-García, N., & Costas, R. (2021). Analyzing scientific mobility and collaboration in the Middle East and North Africa. *Quantitative Science Studies*, 2(3), 1023–1047. https://doi.org/10.1162/qss_a_00149
- Franzoni, C., Scellato, G., & Stephan, P. (2012). Foreign-born scientists: Mobility patterns for 16 countries. *Nature Biotechnology*, 30(12), 1250–1253. <https://doi.org/10.1038/nbt.2449>
- Gomez, C. J., Herman, A. C., & Parigi, P. (2020). Moving more, but closer: Mapping the growing regionalization of global scientific mobility using ORCID. *Journal of Informetrics*, 14(3), 101044. <https://doi.org/10.1016/j.joi.2020.101044>

- Gureyev, V. N., Mazov, N. A., Kosyakov, D. V., & Guskov, A. E. (2020). Review and analysis of publications on scientific mobility: Assessment of influence, motivation, and trends. *Scientometrics*, 124(2), 1599–1630. <https://doi.org/10.1007/s11192-020-03515-4>
- Kim, H. (2022). *Global mobility of the recent STEM postdoctoral workforce registered in ORCID*. OSF. <https://doi.org/10.31235/osf.io/cugv6>
- Moed, H. F., Aisati, M., & Plume, A. (2013). Studying scientific migration in Scopus. *Scientometrics*, 94(3), 929–942. <https://doi.org/10.1007/s11192-012-0783-9>
- Netz, N., Hampel, S., & Aman, V. (2020). What effects does international mobility have on scientists' careers? a systematic review. *Research Evaluation*, 29(3), 327–351. <https://doi.org/10.1093/reseval/rvaa007>
- Orazbayev, S. (2017, April 6). *International stocks and flows of students and researchers reconstructed from ORCID biographies* [MPRA Paper]. <https://mpra.ub.uni-muenchen.de/79242/>
- Ovchinnikova, E., Mol, C. V., & Jones, E. (2023). The role of language proximity in shaping international student mobility flows. *Globalisation, Societies and Education*, 21(4), 563–574. <https://doi.org/10.1080/14767724.2022.2070132>
- Porter, S. J. (2022). Measuring research information citizenship across ORCID Practice. *Frontiers in Research Metrics and Analytics*. <https://doi.org/10.3389/frma.2022.779097>
- Raghupathi, V., Ren, J., & Raghupathi, W. (2023). Exploring the nature and dimensions of scientific mobility: insights from ORCID database—a visualization approach. *International Journal of Technology Diffusion (IJTD)*, 14(1), 1–31. <https://doi.org/10.4018/IJTD.331090>
- Robinson-García, N., Sugimoto, C. R., Murray, D., Yegros-Yegros, A., Larivière, V., & Costas, R. (2018). Scientific mobility indicators in practice: International mobility profiles at the country level. *Profesional De La Información*. <https://doi.org/10.3145/epi.2018.may.05>
- Robinson-García, N., Sugimoto, C. R., Murray, D., Yegros-Yegros, A., Larivière, V., & Costas, R. (2019). The many faces of mobility: Using bibliometric data to measure the movement of scientists. *Journal of Informetrics*, 13(1), 50–63. <https://doi.org/10.1016/j.joi.2018.11.002>
- Siekierski, P., Lima, M. C., Borini, F. M., & Pereira, R. M. (2018). International academic mobility and innovation: A literature review. *Journal of Global Mobility: The Home of Expatriate Management Research*, 6(3/4), 285–298. <https://doi.org/10.1108/JGM-04-2018-0019>
- Sugimoto, C. R., Robinson-García, N., Murray, D. S., Yegros-Yegros, A., Costas, R., & Larivière, V. (2017). Scientists have most impact when they're free to move. *Nature*. <https://doi.org/10.1038/550029a>
- Teixeira da Silva, J. A. (2021a). Abuse of ORCID's weaknesses by authors who use paper mills. *Scientometrics*, 126(7), 6119–6125. <https://doi.org/10.1007/s11192-021-03996-x>
- Teixeira da Silva, J. A. (2021b). ORCID: Issues and concerns about its use for academic purposes and research integrity. *Annals of Library and Information Studies (ALIS)*. <https://doi.org/10.56042/alis.v67i4.36267>
- Teixeira da Silva, J. A. (2023). ORCIDs of “researchers” with the names of vegetables. *International Journal of Vegetable Science*, 29(4), 291–293. <https://doi.org/10.1080/19315260.2023.2211576>
- Verginer, L., & Riccaboni, M. (2021). Talent goes to global cities: The world network of scientists' mobility. *Research Policy*, 50(1), 104127. <https://doi.org/10.1016/j.respol.2020.104127>
- Veugelers, R., & Van Bouwel, L. (2015). Chapter 8—Destinations of Mobile European Researchers: Europe versus the United States. In A. Geuna (Ed.), *Global Mobility of Research Scientists* (pp. 215–237). Academic Press.
- Working with bulk data*. (n.d.). Retrieved 8 June 2024, from <https://info.orcid.org/documentation/integration-guide/working-with-bulk-data/>
- Yan, E., Zhu, Y., & He, J. (2020). Analyzing academic mobility of U.S. professors based on ORCID data and the carnegie classification. *Quantitative Science Studies*, 1(4), 1451–1467. https://doi.org/10.1162/qss_a_00088

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.