MAX PLANCK INSTITUTE
FOR DEMOGRAPHIC RESEARCH

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

MPIDR Working Paper WP 2022-019 l August 2022
https://doi.org/10.4054/MPIDR-WP-2022-019

Homecoming after Brexit: evidence on academic migration from bibliometric data

Ebru Şanlıtürk
Samin Aref
Emilio Zagheni l zagheni@demogr.mpg.de
Francesco C. Billari

Homecoming After Brexit: Evidence on Academic Migration from Bibliometric Data

Ebru Sanliturk,[1,2] Samin Aref,[3,2] Emilio Zagheni,[2] Francesco C. Billari[1,4]

[1] Laboratory of Digital and Computational Demography, Max Planck Institute for Demographic Research, 18057 Rostock, Germany
[2] Department of Social and Political Sciences, Bocconi University, 20136 Milan, Italy
[3] Department of Mechanical and Industrial Engineering, University of Toronto, ON M5S 3G8 Toronto, Canada
[4] Carlo F. Dondena Centre for Research on Social Dynamics and Public Policy, Bocconi University, 20136 Milan, Italy

*To whom correspondence should be addressed; E-mail: zagheni@demogr.mpg.de.

Abstract

This study assesses the initial effects of the 2016 Brexit referendum on the mobility of academic scholars to and from the United Kingdom (UK). We leverage bibliometric data from millions of Scopus publications to infer changes in the countries of residence of published researchers by the changes in their institutional affiliations over time. We focus on a selected sample of active researchers whose movements are traceable for every year between 2013 and 2019, and measure the changes in their international migration patterns. While we do not observe a brain drain following Brexit, we find evidence that the mobility patterns of scholars began to change following the referendum. Among the active researchers in our sample, we find that their probability of leaving the UK increased by approximately 86% if their academic origin (country of first publication) was an EU country. For scholars with a UK academic origin, we observe that after Brexit, their probability of leaving the UK decreased by approximately 14%, and their probability of moving (back) to the UK increased by around 65%. Our analysis points to a compositional change in the academic origins of the researchers entering and leaving the UK as one of the first impacts of Brexit on the UK and EU academic workforce.

Keywords: High-skilled migration | Brexit | Bibliometric data | Migration of scholars

Acknowledgements

We would like to thank Tom Theile for the technical support in different parts of this project. This study has received access to the bibliometric data through the project "Kompetenzzentrum Bibliometrie" and the authors acknowledge their funder Bundesministerium für Bildung und Forschung (funding identification number 01PQ17001).

On 1 January 2021, the free movement of people between the United Kingdom (UK) and the European Union (EU) ended. Regulations and requirements for professionals moving between the UK and the EU switched to a point-based visa system that was intended to favour the immigration of migrants deemed crucial for the UK economy. The UK's decision to leave the EU (the so-called *Brexit*) is likely to have profound consequences for migration to and from the UK, including for researchers who still benefit from a special visa as part of the Global Talent program. Supporters of the so-called "Brexit camp" argue that the standing of the UK in the global competition for talent will improve, as it will be able to increase its attractiveness to scholars from outside the EU. Critics point to the lower level of attractiveness of the UK for top researchers, especially for EU nationals, who may face additional obstacles to working in the UK, including legal barriers for themselves, their families, and their collaborators, as well as the prospect of diminished access to EU resources.

The Brexit process dates back to 23 June 2016, when the referendum on whether the UK should remain in the EU was held. The electorate's choice to leave, which was fuelled by the idea that the UK should "take back control" of immigration (Gietel-Basten, 2016), created an unprecedented situation of political discontinuity that led to widespread uncertainty about the status of immigrants in the UK. It has been shown that changes in migration policy affect the decisions of researchers to migrate internationally (Arrieta et al., 2017; Scellato et al., 2015), which, in turn, have an impact on the scientific and technological development of the countries involved (Mahroum, 2005; Moser et al., 2014). Brexit can be seen as a clear example of a shift in migration policy that could impede the international circulation of scholars, which is known to enhance research performance by facilitating knowledge recombination (Scellato et al., 2017; Wible, 2017; Sugimoto et al., 2017), and to be fundamental to scientific discovery, especially in its most innovative forms (Fernández-Zubieta et al., 2016). Researchers and academic institutions were rattled by the outcome of the vote: in the weeks leading up to the referendum, leaders from 103 universities, including from all of the top UK institutions, openly expressed their opposition to Brexit, stating that: "Cutting ourselves out of the world's largest economic bloc would undermine our position as a global leader in science and innovation (Independent, 2016)".

While it is too early to assess the long-term consequences of Brexit on the migration of researchers, here we analyse large-scale bibliometric data in order to offer insights into the recent trends and compositional changes in the population of researchers moving to and from the UK. We use data from Scopus, a comprehensive bibliometric database that includes detailed metadata on over 80 million scientific publications, and is considered a source of highly precise individual-level data on published researchers and their affiliations (Kawashima and Tomizawa, 2015; Aman, 2018). Using these data, we can infer international migration patterns by examining changes in the institutional affiliation of authors. In 2015, the precision of Scopus individual-level data on researchers (Scopus author ID) was estimated to be 99% (Kawashima and Tomizawa, 2015), whereby a precise author ID is a unique number that is associated only with the publications of that same author. Previous studies on the migration of researchers have used either highly accurate data with low coverage (Bohannon, 2017), bibliometric databases with high coverage and a focus on specific types of researchers (Chinchilla-Rodríguez et al., 2018a), or ad-hoc surveys that may include biases due to non-response (Scellato et al., 2015). As there are trade-offs in using each of these data sources, we invested in further refining the quality of the Scopus data for use in migration research by enhancing the disambiguation of authors and tackling other data quality challenges (see Data section), and thus further improving the accuracy of inferences of migration events from bibliometric data. This approach enables us to strike a suitable balance between coverage levels, data quality, and timeliness in studying scholarly migration before and after Brexit.

This paper is structured in five parts. We begin by establishing the conceptual framework, both by providing a literature review on high-skilled migration and the use of bibliometric data in migration studies, and brief background information on Brexit as a policy change. Next, in two separate sections on data and methods, we explain how we process and use bibliometric data in this research, and the methods we employ to both infer migration events and conduct our statistical analysis, respectively. We then present the outcomes of both our descriptive and statistical analyses, and close the paper with a discussion of the implications of our results.

Background and Conceptual Framework

High-Skilled Migration and Policy Change

The international circulation of scholars is essential to fostering scientific knowledge, especially in its most innovative forms (Agrawal et al., 2017; Fernández-Zubieta et al., 2016). For instance, nearly half of the world's most-cited physicists reside outside their country of birth (Hunter et al., 2009). The international migration and mobility of academics and researchers is a subfield of high-skilled migration that rightly commands attention from researchers and policymakers alike (Sugimoto et al., 2016; Czaika and Parsons, 2017; Czaika, 2018; Chinchilla-Rodríguez et al., 2018a). For these reasons, it is of paramount importance that we understand the dynamics of the in-flows and out-flows of scholars across countries, and the underlying determinants of the international mobility of researchers.

In the international migration literature, academic migration that is studied within the framework of the brain drain and brain gain relationships can be aptly framed using the concept of *brain circulation* (Saxenian, 2005). The brain circulation concept assumes that high-skilled migration should be considered as a means of knowledge transfer through reciprocal migration flows, and is therefore a circular exchange rather than a one-way loss. While many factors influence the decisions of scholars to move (Azoulay et al., 2017), a key determinant is the policy environment in their country of residence and in the destination country. More specifically, policy changes may substantially affect the decisions of researchers to migrate internationally (Franzoni et al., 2014; Scellato et al., 2015; Franzoni et al., 2015; Arrieta et al., 2017), which can, in turn, affect the scientific and the technological development of the countries involved (Mahroum, 2005; Moser et al., 2014).

Big Bibliometric Data and Academic Migration

The early studies that used of bibliometric data were based on only a limited volume of data, and focused more on citation counts as the unit of measure to assess scientific impact, scientific progress (Martin and Irvine, 1983), and institutional research performance (Moed et al., 1985). The assessment of scientific performance by using bibliometric data influenced not just scholars, but also policymakers during the 1990s, especially under the New Public Management framework (Mingers and Leydesdorff, 2015). In recent decades, the volume of data used for bibliometric analyses has expanded over time, and the data now extend beyond the country and the institutional levels, thus creating what could be called *big bibliometric data*. As the literature on measuring scientific performance using bibliometric data has continued to grow (Sugimoto and Larivière, 2018), big bibliometric data have paved a new way of study for migration research (Alburez-Gutierrez et al., 2019).

Migration studies using bibliometric data rely on information on the movements of researchers. Following the network-based approach to investigating high-skilled migration (Meyer, 2001) and scientific migration (Ackers, 2005), the use of bibliometric data to study the migration and mobility of researchers started to receive some attention (Laudel, 2003). The feasibility of employing this method to examine the migration and mobility patterns of scholars was demonstrated first for a select group of countries (Moed et al., 2013; Halevi and Moed, 2013). More recently, the literature on scientific migration using bibliometric data has expanded with the publication of studies addressing co-affiliation and collaboration networks (Sugimoto et al., 2016; Chinchilla-Rodríguez et al., 2018b; Aref et al., 2018), the identification of migration and mobility events (Robinson-García et al., 2019), and the mobility patterns of highly mobile researchers (Aref et al., 2019).

In addition, bibliometric data have also been used to investigate certain demographic characteristics of researchers. Bibliometric data have, for example, been employed in prominent studies to examine gender disparities and their influence on scientific performance (Larivière et al., 2013), the academic ages of researchers

(Nane et al., 2017), and the impact of academic age on international mobility (Sugimoto et al., 2017).

The Case of the United Kingdom

The notion of brain circulation has long been a subject of scientific debates in the UK. Indeed, the term *brain drain* was coined in this very context. During the early-1960s, the Royal Society published a report on the increase in the number of scientists and engineers emigrating from the UK to the USA and Canada that referred to this situation as "a drain of scientists and drain of talent" (Oldfield et al., 1963). The drain of scientists and talent out of the UK was later defined as a "brain drain" (Johnson, 1965).

Concerns about the brain drain lessened during the 1970s, as British policymakers started to view brain drain as an inevitable part of globalisation, and as the US became less appealing for scientists due to its role in the Vietnam War (Godwin et al., 2009). However, in the 1980s, fears that British science was declining reappeared. In the STEM fields, the UK's share of global publications and citations decreased by 10% and 15%, respectively, between 1973 and 1982; with the sharpest declines occurring in the physics, engineering, and technology fields, at over 20% (Irvine et al., 1985). In reaction to these concerns, the initiative *Save British Science* was launched in 1986. The initiative called upon the government to take action and to support research, as "opportunities are missed, scientists emigrate and whole areas of research are in jeopardy (Noble, 2016)". Research from the early 1990s reported that the scientific performance of Britain was growing in some areas, but the overall relative decline was continuing (Martin, 1994).

Although the general impression of the performance of British science has been rather pessimistic since the early 1960s, the lack of scientific investment and the emigration of scientists should not be seen as the only underlying reasons for this trend. The gradual decrease in British scientific publications, should also be considered in the light of the global increase in English-language publications by non-native authors, especially since the 1990s. Bibliometric data indicate that by 2018, the United Kingdom accounted for 3.82% of global publication output, and was in sixth place in the world rankings for publication output (White, 2019). The negative evaluation of the UK's scientific performance based on bibliometric data analyses and the impression that British science has been declining may be due to the increased ability of scientists worldwide to publish in English, which mitigated the native English speaker bias. Furthermore, from the late 1960s onwards, the emigration of scientists from the UK to the US and Canada has been offset by the immigration of scientists from developing countries (and/or Commonwealth countries) to the UK (Nature, 1967; Watanabe, 1969; Godwin et al., 2009).

These migration patterns were again disrupted when the UK decided to withdraw from the EU as a result of the referendum held on 23 June 2016, and when Brexit became official on 31 January 2020. Despite longstanding fears that Britain has been losing researchers to other countries (Irvine et al., 1985; Martin et al., 1987; Martin, 1994), the UK remains one of the leading nations in the world in terms of scientific research. In 2019, the UK was the G20 country with the largest share of the top 10% of high-quality scientific publications globally (Adams et al., 2019). Moreover, the UK was the highest ranking EU member state in terms of top 1% highly cited scientific publications in 2016 with 1.63% [1], whereby its rate

[1] The figure refers to the percentage of the scientific publications produced in a country, that

ranked in the third place globally after Switzerland and the US, and exceeded the EU average (0.95%) by a considerable margin (Pereira et al., 2020, Figure 6.1-8). The UK received 7.86 billion euros of net research funding from the EU within the framework of the Horizon 2020 program, which made it the member state with the second-largest share of funding received from the budget, after Germany (EU Commission, 2021). The strong ties that British science and technology have established with the EU are among the reasons why some researchers have raised concerns about a potential loss of these relationships due to Brexit (Golding, 2017).

Data

Source of Raw Bibliometric Data

The main data we use in this study have been obtained from Scopus, a database containing detailed meta-data on over 80 million scientific publications. For each publication, the database includes the individual author IDs, the publication year, the affiliation country(ies) linked to publications, and the All Science Journal Classification (ASJC) code for fields of each publication venue (journal, conference proceedings, etc.). To obtain the raw bibliometric data, we queried all Scopus data from a relational database using SQL. The query involved two steps: (1) obtaining the author IDs of all authors who have published at least once with a UK affiliation, and (2) obtaining data on all publications over the 1996-2019 period from the list of author IDs produced in the previous step. Through this process, we obtained exhaustive data on 26,748,770 author-publication linkages (*authorship record*) involving more than 1,619,000 published researchers with ties to the UK and their 12,365,837 Scopus publications over the period 1996-2019 period. The

are among the top 1% of worldwide most-cited publications.

raw data were then pre-processed for use in our analyses. The pre-processing steps mainly addressed the challenges posed by missing values for the country variable and author name ambiguity.

Data Pre-Processing

There were two technical challenges associated with the raw bibliometric data that we had to address before they could be used to analyse scholarly migration: (1) missing countries and (2) author name ambiguity.

In the extract of the raw bibliometric data that were obtained through queries based on affiliation ties to the UK, the country variable for a small number of records was missing. We modified the neural network algorithm developed in (Miranda-González et al., 2020) to use it to predict the missing values. This neural network algorithm was trained and tested on a large sample of authorship records for which the country variable was available. The trained neural network algorithm took the affiliation address as an input and predicted the country associated with the affiliation address with a high degree of accuracy. Below, we provide some statistics on our implementation of this method for handling missing values. For more technical details on the development of the neural network, refer to (Miranda-González et al., 2020).

Of the 26,748,770 authorship records we obtained in total, we identified 208,762 authorship records in our raw data that did not have a country variable. These records with missing values involved 111,899 author profiles and 147,579 distinct publications. We used one million authorship records that all had the country variable as training data (80%) and as testing data (20%). From each record, we used the variables on the institution's name, address, city, and postal code as the predictors of the target variable, which was the country of affiliation. The neural network predicted the correct country for 96.2% of the test data. As this level of accuracy was deemed acceptable, we used the trained neural network to predict the missing countries. Each prediction came with a reliability score. To preserve the more reliable predictions, we discarded predictions with scores below 0.8 (55,516 authorship records (0.2%) out of the 26,748,770 records in total).

Scopus provides author IDs to identify the publications of each individual researcher. These author IDs appear to be sufficiently reliable for analysing migration of researchers (Aman, 2018), as previous research has shown that 98.3% of author IDs precisely identify one researcher[2] (Paturi and Loktev, 2020). Despite the high degree of precision of the Scopus author IDs, we consider Scopus as an imperfect source of digital trace data for studying migration of researchers. The lack of precision in the Scopus author IDs implies that, on average, 1.7% author IDs might involve publications from multiple individuals who may share the same name. To handle this problem systematically, we applied a conservative author disambiguation process (Miranda-González et al., 2020; D'Angelo and van Eck, 2020) to the author profiles that were more likely to be affected by the precision flaws of the Scopus author IDs. The author disambiguation algorithm we implemented was based on recent developments in the use of unsupervised learning for disambiguating bibliometric data (D'Angelo and van Eck, 2020). This algorithm

[2]According to the latest accuracy evaluation in August 2020 the precision of the Scopus author profiles is 98.3%, and the completeness is 90.6% (Paturi and Loktev, 2020). In this context, precision is the percentage of author profiles that only contain publications of one individual. Completeness is the ratio of individual researchers whose publications are all in one author profile.

was designed based on a conservative approach: it assumes that every two authorship records are from distinct individuals unless sufficient evidence is found to demonstrate the similarity of the two records. We considered the author profiles that exceeded either of these thresholds suspicious and they were treated by the disambiguation algorithm. These author IDs were associated with a suspiciously high number of countries or a suspiciously high number of publications. The two thresholds were being associated with (1) more than six countries of affiliation or (2) more than 292 publications. The number 292 was chosen to imply that a given author ID had an average of more than one publication per month across a period of 24 years and four months. These thresholds were chosen by trial and error. The aim of this screening of outliers was to reduce the risk that the lack of precision in 1.7% of author profiles , which may have represented more than one individual researcher, would lead to the overestimation of migration.

Of the 1,619,000 author IDs in our data, there were 14,441 author IDs (less than 0.9%) that exceeded at least one of the above thresholds, and were therefore considered suspicious. These author profiles were associated with $2,783,657$ publications[3]. These suspicious authorship records were put through our disambiguation algorithm (Miranda-González et al., 2020), which processed them first by comparing every pair of records with the same author ID and then by assigning revised author IDs using the method briefly described below. For each group of records with the same author ID, a similarity matrix was created based on a comparison of the author names, co-author names, subjects, funding information, and grant numbers for every pair of records. The entries of the matrix were higher if the

[3]Note that some publications might be shared between different individual authors.

two authorship records had similar traits, and were lower if their traits were dissimilar. We then implemented an agglomerative clustering from the scikit-learn Python library (Pedregosa et al., 2011) to obtain clusters of highly similar authorship records. The agglomerative clustering algorithm started by placing each record in its own cluster, and then merging pairs of clusters in a successive manner if doing so minimally increased a given linkage distance. Therefore, the clustering stopped when the linkage distance could not be increased further (Pedregosa et al., 2011). In the last step, each cluster was issued a revised author ID. The suspicious records with the revised author IDs were then merged with the rest of the data, which made up the pre-processed data. While this pre-processing step could not increase the precision of the Scopus author IDs to 100%, it reduced the possible impact of Scopus author ID precision flaws on migration estimates.

A Focus on Active Researchers

Migration is well-known to be a selective process. However, in part because of a lack of data, the measurement of high-skilled migration has typically been based on broad categories, like levels of educational attainment or sectors of the economy. Unobservable characteristics that may be related to the potential for breakthroughs are more difficult to measure. The results of our analyses using the disambiguated Scopus data show that while migrant researchers were outnumbered by those who remained affiliated with UK institutions only, the scientific impact of migrants was substantially larger. For example, our data indicate that migrant scholars received, on average, 90% more citations per year. In this study, we focus in particular on the migration of upper-tier researchers who were consistently active in producing scientific publications over the period under study (hereafter *active researchers*).

By concentrating on the top end of the distribution, we aim to identify those groups who are typically the targets of immigration policies intended to attract top talent.

Methods

Detecting Migration Events

We build on previous research on bibliometric data to define academic migration. Throughout this paper, the country of academic origin is used as a shorthand for the country of first publication. The academic origin is not considered as a proxy for the nationality of a scholar, but as the country that is most likely to have invested in the pre- or post-doctoral period of academic development of the individual that led him/her to become a published researcher, regardless of his/her nationality (Robinson-García et al., 2016; Robinson-García et al., 2019; Aref et al., 2019; Subbotin and Aref, 2021; Zhao et al., 2021, 2022). For each year and for each scholar, we assessed the mode country of affiliation, given that some researchers were affiliated with multiple countries in a given year. We used a calendar year as the time unit, per the definition of long-term international migrant as a change of the country of usual residence for a period of at least one year (IOM, 2019, p. 125), which is also the definition used by the Office for National Statistics (ONS) in the UK (ONS, 2020). Migration across countries was defined as a change in this mode country. For example, a scientist who published with an affiliation(s) mostly from Germany for the year 2016, and then published with an affiliation(s) mostly from the UK for the year 2017, was considered by our algorithm to have moved from Germany to the UK in the year 2016. To be precise, the year of the move was calculated based on the rounded mid-point between the last year when the researcher had Germany as a mode country of affiliation and the first year when the researcher had the UK as the mode country of affiliation. Because of the time it takes to conduct and publish research, the publication years did not necessarily match the years of move. We should, however, point out that according to our method, when a continuously active researcher has at least one publication every year, the move year becomes the last year of the common usage of the old affiliation. For a researcher, with less frequent publications, the potential gap between the actual move year and the move year that our algorithm estimated could be larger.

Inferring the migration events retrospectively from publications, posed a challenge of the right-censoring of the data. As not every researcher necessarily publishes in every year, the number of movers at the end of our time period was inevitably underestimated, and this underestimation cannot be corrected until more recent data becomes available. For the last few years of our dataset, we were only able to identify the most immediate migration events, and thus assume that the number of migration events that we detected is an underestimate. Therefore, we used the partial information we have for 2020 to detect migration events, but did not include 2020 in the analysis, as the estimates would be unreliable. Furthermore, to prevent the right-censoring from biasing our results, we restricted our sample to the researchers for whom locational signals from affiliation countries are available for every year of the analysis; a group to whom we refer to as *active researchers*. While we identified 946,991 published researchers with ties to the UK for the 1996-2019 period, only around 11% of them (102,058) were classified as active, irrespective of whether they were internationally mobile. We used the dataset that included all researchers in the descriptive and visual analyses only, while applying additional caution in our statistical analyses and interpretations, bearing the right-censoring issue in mind.

Therefore, our sample for the statistical analysis consisted of researchers who were either continuously active (had at least one publication for each year of the analysis period), or published with such a frequency that with the abovementioned inference of migration events, their location information could be identified for the seven consecutive years between 2013 and 2019. Restricting our sample to active researchers enabled us to observe the migration patterns of researchers (with respect to the UK) who would be considered the potential target of policies to attract talent due to their productivity, as measured by their publications.

It should be noted that focusing on a selected group of people (i.e., active researchers) also enabled us to create a panel dataset and to observe how the migration patterns of a large group of researchers with relatively high levels of scientific productivity and ties to the UK changed in the years before and after the Brexit referendum. As this was a strongly balanced panel data, we also avoided the problem of attrition.

Inferring Gender

The most likely gender of each active researcher included in the dataset was inferred from the first names of the researcher using the *genderizeR* package in R (Wais, 2006). Studies of big bibliometric data analysis typically rely on various gender estimation algorithms (Krapf et al., 2016). However, since these algorithms were initially developed for marketing rather than for research purposes, they are more accurate when applied to certain populations than to others. Generally, the gender inference algorithms work better for Anglo-Saxon and European names, for which the training sample is large. In contrast, as Asian and African names are underrepresented in the training data, the predicted gender is less accurate for these names. Moreover, for unisex names, the probability of the inferred gender being reported is low, which indicates that the result is unreliable. Therefore, for our analysis on gender, we used three categories: female, male, and unknown. The last category contained all of the authorship records for which gender could not be estimated or the gender estimation lacked accuracy. The accuracy of the gender estimation was based on the probability reported by the *genderize* function from the *genderizeR* package. We used two different thresholds of probability to infer the most likely gender of researchers: 75% and 90%. We considered the gender inference accurate enough if the reported probability of being male or female for a given name was at least 75%. If the gender estimation failed to meet this criterion, the predicted gender was tagged as unknown. For robustness checks, we created a separate gender variable that used the same logic, but had a minimum threshold of 90%. The distribution of estimated gender is presented in *Table 4* in the Online Appendix.

Statistical Modelling

To quantify the changes in the brain circulation patterns in the UK after the Brexit referendum, we narrowed down our focus to the sample of active researchers in our statistical analysis. The sample of active researchers formed a strongly balanced panel data for the seven years between 2013 and 2019. We applied a random-effects logistic regression model to the panel data of the 45,316 internationally mobile researchers who were classified as active between 2013 and 2019. The dependent variable was binary, taking the value of 1 to represent a year with an out- or an in-migration event, and the value of 0 otherwise for each active researcher and for each year between 2013 and 2019. Our main explanatory variables were Brexit, represented by a binary variable taking the value of 1 after 2016, and the value of 0 before 2016, and academic origin. Academic origin was defined as the country of the author's primary institutional affiliation when the author published his/her first article, going as far back as 1996 (see Detecting Migration Events).

The panel data we compiled consist of 45,316 internationally mobile, active researchers with at least one UK-affiliated publication throughout their career, for whom location (of residence) information is available via Scopus-indexed publication references or inferences of migration events over the years 2013 to 2019. Our strongly balanced panel data with annual observations for each active researcher in the sample, which were derived from the information from multiple publications, provided us with a robust resource for our statistical analyses. In addition, the random-effects model allows us to explore the potential effects of the time-invariant variables, such as academic origin, and of the control variables such as scientific field and gender. Therefore, we have selected the individual-specific random-effects model as the main model for our analysis. However, for robustness, we apply and present also the results of its replication using simple logistic regression.

We consider the following two models for the emigration and immigration of active researchers respectively:

$$MovesOut_{i,t} = ln(P/(1-P))_{i,t} = \alpha + \beta_1 Brexit_t + \beta_2 Origin_{i,t}$$
$$+\beta_3(Brexit \times Origin)_{i,t} + \sum_{(k=9)}^{K} \beta_k X_{i,t} + w_i + \tau_t$$

$$MovesIn_{i,t} = ln(P/(1-P))_{i,t} = \alpha + \beta_1 Brexit_t + \beta_2 Origin_{i,t}$$
$$+\beta_3(Brexit \times Origin)_{i,t} + \sum_{(k=9)}^{K} \beta_k X_{i,t} + w_i + \tau_t$$

In the random effects logistic regression equations above, dependent variables $MovesOut$ ($MovesIn$) represent the binary variable that takes the value of 1, when in a given year t the researcher i leaves the UK or moves to the UK respectively. We consider scientific immigration and scientific emigration as two different models, acknowledging that moving into and out of a country may follow a different pattern, as is also observed in the descriptive graphs. The main explanatory variables are denoted by the interaction term $BrexitxOrigin$, while the control variables are represented by X. The variable $Brexit$ is a binary variable that is equal to 1 for the years 2016 to 2019, and is equal to 0 otherwise. The control variables include academic age and dummy variables for having higher than average publication and citation counts, scientific field, and gender.

Similar to the approach used to define academic origin, academic age is measured based on the first publication. The year of first publication is considered to be the academic birth year of a researcher, and the researcher's academic age is calculated in a dynamic way for the following years. The scientific field dummy variable is based on the All Science Journal Classification (ASJC) field codes tagged by Scopus, and consists of four general categories: life sciences[4], social sciences[5], physical sciences[6], and health sciences[7]. The publication and citation count variables are calculated over the entire dataset (starting with the year 1996) that are available to us for each researcher in the active researchers sample. The gender variable is created using the method explained in the Gender Inference subsection above.

Results

Descriptive Analysis

In order to understand the changing characteristics of brain circulation in the UK, we first consider the descriptive statistics. We visually explore the dynamic flows of researchers both moving to and from the UK, by academic origin, before (*Figure 1*) and after Brexit (*Figure 2*). In the online appendix, we also illustrate the trends of the outgoing and incoming researchers by academic origin in the UK (*Figure 3*).

The results of our descriptive analysis of longitudinal Scopus bibliometric data suggest that if the post-Brexit trends we have observed continue, Brexit may trigger a change in the composition of the British scientific workforce. Although we

[4]Life sciences include Agricultural and Biological Sciences, Biochemistry, Genetics and Molecular Biology, Immunology and Microbiology, Neuroscience, and Pharmacology, Toxicology and Pharmaceutics

[5]Social sciences include Arts and Humanities, Business, Management and Accounting, Decision Sciences, Economics, Econometrics and Finance, Psychology, and other Social Sciences

[6]Physical sciences include Chemical Engineering, Chemistry, Computer Science, Earth and Planetary Sciences, Energy, Engineering, Environmental Science, Materials Science, Mathematics, and Physics and Astronomy

[7]Health sciences include Medicine, Nursing, Veterinary, Dentistry, and Health Professions

used a comprehensive source of data on published researchers, conducting an empirical analysis with these data was a challenge due to the lack of observations in the years the authors did not publish. For the visualisations using the dataset with no restrictions (to active researchers), we consider a sharp decline only as a potential decline in the pattern that should be re-assessed in future work. We expect the slope of the trend to change upwards in the coming years, when more recent data become available that enable us fill the data gaps for the most recent years. Therefore, under these circumstances, observing an increasing trend in such visualisations with the minimum estimates for the most recent years, instead of a sharp decline, would be striking.

Despite this challenge, the results of the descriptive analysis in *Figure 3* point to a potential change in researchers' patterns of movement out of and to the UK by academic origin. Indeed, the *Figure 3* shows a slight but steady increasing trend in leaving the UK for researchers with an EU country academic origin up to 2018. The decreasing trend between 2018 and 2019 is probably due to the right-censoring in the data. To avoid overestimating immobility during the years without any publications, we focused on a subset of the active migrant researchers: i.e., the same subset we used in the statistical analysis ($N = 45,316$). We categorised the academic origins into four groups: EU countries, the USA, the UK, and other. The migration trends leaving and entering the UK of the active researchers in each academic origin category are shown in *Figure 4*. According to *Figure 4*, following the Brexit referendum, the number of active researchers leaving the UK with an EU country academic origin increased, while the number of active researchers leaving the UK with a UK academic origin decreased. In the online appendix, *Figure 6* also displays a similar picture of the shares of active researchers leaving and entering the UK by academic origin. The compositional changes after the Brexit referendum are clear among the active researchers moving to (leaving) the UK, as the share of those with an EU academic origin decreased (increased), while the share of those with a UK academic origin increased (decreased).

Statistical Analysis

The results of the random-effects logistic regressions to assess the out-migration and in-migration patterns of active researchers between 2013 and 2019 are presented in *Table 1.* For comparison, *Table 2* shows the results when estimating the parameters of the logistic model without random effects. These results are shown as robustness checks: i.e., when considering moving out of the UK (leaving) and moving to the UK (entering), respectively. The results of the empirical analysis corroborate the implications of the initial descriptive analysis, and confirm the statistical significance of the changes in migration patterns. *Table 1* shows that the odds of moving to the UK after Brexit were 44% higher for active researchers with a UK academic origin than for the baseline group of active researchers with an academic origin other than the UK, an EU country, or the US. Without the interaction with the Brexit variable, the odds of moving to the UK were 64% lower for the active researchers with a UK academic origin than for the baseline group. Furthermore, after Brexit, the odds of leaving the UK were 36% higher for an active researcher with an EU academic origin than for a researcher with an academic origin other than the EU, the UK, or the US. Without the condition of Brexit, this trend would be reversed, and the odds of leaving the UK would be 21% lower, for an active researcher with an EU academic origin than for an active researcher from the baseline group. In the online appendix, *Figure 7* shows the changing patterns of the odds of moving out of and moving to the UK in an odds ratio plot.

Table 3 and *Figure 5* show the probabilities of leaving and entering the UK, calculated based on the results of the random-effects logistic regression, for the active researchers by academic origin, before and after Brexit, respectively. *Figure 5* shows that among the active researchers, the probability of leaving the UK after Brexit declined only for those with a UK academic origin. For active researchers, the probability of leaving the UK fell from 5.25% to 4.54%, which represents a decrease of 14%. All of the active researchers except for those with a UK academic origin became increasingly likely to leave the UK after Brexit. The change in the probability of leaving was largest for the active researchers with an EU academic origin, rising from nearly 2.96% to 5.51%, which represents an increase of approximately 86%. Thus, our results support the argument that active researchers with an EU academic origin, who constituted an important share of the academic population in the UK, became significantly more likely to leave the UK after the Brexit referendum than they were before the vote. *Figure 5* also presents the predicted probability of entering the UK for active researchers by academic origin. While the trends in the probability of entering the UK before and after Brexit did not change significantly for active researchers with a US or an EU academic origin, the probability increased significantly for active researchers with a UK and another academic origin. For active researchers with a UK academic origin, the probability of moving to the UK increased from 1.97% to 3.24% after Brexit, representing a change of approximately 65%. For active researchers with another academic origin (a non-EU country other than the UK and the US), there was a statistically significant increase (by 20%) in the probability of moving to the UK after Brexit, from 5.15% to 6.20%. It is, however, important to note that for active researchers with a US or an EU academic origin, the odds of entering the UK decreased, as shown in *Table 1*; while the marginal probability displayed stable patterns with no statistically significant changes. This is likely because the baseline group used for the odds ratio calculations was other academic origin.

Discussion

Our aim in this study was to estimate the immediate effects of the Brexit referendum in 2016 on the mobility of top talent in British academia, by following migration patterns of scholars using bibliometric data. Our analysis did not find a pattern of brain drain for the period after the referendum and before Britain's withdrawal from the EU became official. This finding suggests that the migration policies that the UK implements after the Brexit in 2020 likely bear more importance for the migration decisions of internationally mobile researchers than the uncertainty of the intermediary period between 2016 and 2019. This hypothesis for researchers, based on the results of bibliometric data analysis, seems to be in line with the long-term international migration estimates for the general population in the UK. According to the estimates published by the ONS (2021), the EU migration patterns with respect to the UK changed drastically in 2020. From 2018 to 2020, these estimates showed an increasing trend of emigration and a decreasing trend of immigration for the EU nationals. In 2020, due to the global pandemic, immigration and emigration estimates fell by almost 50%, for every group other than the emigration estimates for EU nationals. Coupled with the fall in immigration estimates for all groups, this created an estimate of negative net migration for EU nationals in the UK, by a great margin. We should underline that while we did not observe a pattern of brain drain for the time period of the study, our results uncovered a significant pattern of compositional change in the academic origins of researchers entering and leaving the UK as an early consequence of Brexit. This trend has the potential to make British academia more insular unless it is addressed by future academic migration policies.

The descriptive analyses showed the changes in the migration behaviour of internationally mobile researchers by academic origin, following the Brexit referendum. Without restricting our dataset to active researchers, we observed a slight increase in the trend towards leaving the UK for researchers whose academic origin was an EU country, despite the bias in the data. When we narrowed our focus to active researchers to obtain a more accurate picture, we observed that, after the Brexit referendum, the share of active researchers leaving the UK with an EU academic origin increased continuously, surpassing the share of active researchers leaving the UK with a UK academic origin. The reverse trend was also observed for active researchers entering the UK, whereby the share of incoming active researchers with a UK origin surpassed the share of incoming active researchers with an EU origin, and increased continuously after the Brexit referendum. The visualisations suggest that the trend of compositional change by academic origin, for leaving rather than entering the UK first started in 2014, and continued after the Brexit referendum. While the reasons for this shift require further exploration, it should also be noted that anti-immigrant and anti-foreigner sentiments did not start with the Brexit referendum in 2016, as they first became the focus of public discussions following the success of the Brexit supporters in the 2014 EU Parliament elections.

Statistical analysis confirmed the significance of the changing migration patterns that emerged from this simple visualisation. The marginal probabilities for leaving and entering the UK before and after the Brexit referendum by academic origin, shown in *Figure 5*, which were calculated based on the results of the random-effects logistic regressions, support the implications of the compositional changes outlined in the descriptive analyses. We found that for active researchers with a UK academic origin, the probability of moving (back) to the UK increased by approximately 65% following Brexit, rising from nearly 2.0% before Brexit to 3.2% after Brexit. In contrast, the probability of leaving the UK among this group declined by roughly 14% following Brexit, falling from nearly 5.3% before Brexit to 4.5% after Brexit. Active researchers with a UK academic origin constituted the only group of active researchers in this categorisation by academic origin for whom the probability of leaving the UK decreased after Brexit. For all of the other three groups in our analysis, the probability of leaving the UK increased after Brexit. In terms of the probability of leaving the UK after Brexit, the most striking result was observed for active researchers with an EU academic origin, who represented a large fraction of the foreign-trained scholars in the UK. For an active researcher with an EU academic origin, the probability of leaving the UK rose by approximately 86%, from nearly 3.0% before Brexit to 5.5% after Brexit.

For active researchers with an EU or a US academic origin, we did not observe a statistically significant change in the probability of moving to the UK when the periods before and after Brexit were compared. However, for active researchers with an academic origin other than the UK, the EU, or the US, our analysis showed that the probability of moving to the UK increased after Brexit, and although this change was relatively small, it was statistically significant.

In our study, we refrained from making causal assumptions. Instead,we provided evidence for scholarly migration patterns associated with the Brexit referendum. As bibliometric data are not real-time data, we were only able to observe scholarly migration patterns retrospectively. Bearing in mind that Brexit became official in January 2021, and that the available bibliometric data for 2020 were incomplete at the time of our statistical analysis (which therefore ended in 2019), we considered the observed patterns as a reaction to the Brexit referendum. Further research is needed when the relevant data become available to observe the *Brexit effect* after 2021, and to enable us to make causal claims.

While it is too early to assess the full scope of the impact of Brexit on the migration of active scholars, our evidence on active researchers indicates that the anticipation of the erection of legal barriers between the UK and the EU has already had an influence on migration flows. Top researchers cannot be attracted only by offering them visas and funding for themselves. Many of them have families that need long-term prospects along many dimensions, and the scientists themselves are key drivers in attracting other successful researchers (Waldinger, 2016). This suggests that, over the longer term, the disruption of Brexit may lead to a reduction in the circulation of scholars between the EU and the UK, that could have a negative impact not only on the UK, but also on the EU and the international science system. To counteract this trend, explicit changes in science collaboration policies between the UK and the EU are needed. The early signs are not completely encouraging. On the positive side, the UK has signed a cooperation agreement for Horizon Europe, including for the European Research Council, in which the UK has been highly successful. On the negative side, the UK will not participate in the Erasmus+ collaboration, the flagship program of scientific exchange between university students in Europe. Instead, the UK has been focused on developing its own Turing scheme, which will not offer placements for teaching and college staff. Developing and funding new programs that favour visiting periods abroad for productive international scholars, including for their families, should become a priority to help compensate for the barriers to the circulation of researchers between the EU and the UK that Brexit has erected.

References

Ackers, L. (2005). Moving people and knowledge: Scientific mobility in the European Union. *International Migration*, 43(5):99–131.

Adams, J., Rogers, G., and Szomszor, M. (2019). *The Annual G20 Scorecard: Research Performance 2019.* Institute for Scientific Information.

Agrawal, A., McHale, J., and Oettl, A. (2017). How stars matter: Recruiting and peer effects in evolutionary biology. *Research Policy*, 46(4):853–867.

Alburez-Gutierrez, D., Aref, S., Gil-Clavel, B. S., Grow, A., Negraia, D. V., and

Zagheni, E. (2019). Demography in the digital era: new data sources for population research. In *Smart statistics for smart applications: book of short papers SIS2019*, pages 23–30. Pearson.

Aman, V. (2018). Does the Scopus author id suffice to track scientific international mobility? A case study based on Leibniz laureates. *Scientometrics*, 117(2):705–720.

Aref, S., Friggens, D., and Hendy, S. (2018). Analysing scientific collaborations of New Zealand institutions using Scopus bibliometric data. In *Proceedings of the Australasian Computer Science Week Multiconference*, page 49, New York, NY, USA. ACM.

Aref, S., Zagheni, E., and West, J. (2019). The demography of the peripatetic researcher: Evidence on highly mobile scholars from the web of science. In Weber, I., Darwish, K. M., Wagner, C., Zagheni, E., Nelson, L., Aref, S., and Flöck, F., editors, *Social Informatics*, pages 50–65, Cham. Springer International Publishing.

Arrieta, O. A. D., Pammolli, F., and Petersen, A. M. (2017). Quantifying the negative impact of brain drain on the integration of European science. *Science Advances*, 3(4):e1602232.

Azoulay, P., Ganguli, I., and Zivin, J. G. (2017). The mobility of elite life scientists: Professional and personal determinants. *Research Policy*, 46(3):573 – 590.

Bohannon, J. (2017). Restless minds. *Science*, 356(6339):690 – 692.

Chinchilla-Rodríguez, Z., Bu, Y., Robinson-García, N., Costas, R., and Sugimoto,

C. R. (2018a). Travel bans and scientific mobility: Utility of asymmetry and affinity indexes to inform science policy. *Scientometrics*, 116(1):569–590.

Chinchilla-Rodríguez, Z., Miao, L., Murray, D., Robinson-García, N., Costas, R., and Sugimoto, C. R. (2018b). A global comparison of scientific mobility and collaboration according to national scientific capacities. *Frontiers in Research Metrics and Analytics*, 3:17.

Czaika, M. (2018). *High-skilled migration: Drivers and policies*. Oxford University Press, New York, NY, USA.

Czaika, M. and Parsons, C. R. (2017). The gravity of high-skilled migration policies. *Demography*, 54(2):603–630.

D'Angelo, C. A. and van Eck, N. J. (2020). Collecting large-scale publication data at the level of individual researchers: a practical proposal for author name disambiguation. *Scientometrics*, pages 1–25. Publisher: Springer.

EU Commission (2021). United Kingdom Horizon 2020 country profile. url: `https://webgate.ec.europa.eu/dashboard/extensions/CountryProfile/CountryProfile.html?Country=United%20Kingdom` accessed on 26 July 2022.

Fernández-Zubieta, A., Geuna, A., and Lawson, C. (2016). Productivity pay-offs from academic mobility: Should I stay or should I go? *Industrial and Corporate Change*, 25(1):91–114.

Franzoni, C., Scellato, G., and Stephan, P. (2014). The mover's advantage: The superior performance of migrant scientists. *Economics Letters*, 122(1):89–93.

Franzoni, C., Scellato, G., and Stephan, P. (2015). International mobility of research scientists: Lessons from GlobSci. In *Global Mobility of Research Scientists*, pages 35–65. Elsevier.

Gietel-Basten, S. (2016). Why Brexit? The toxic mix of immigration and austerity. *Population and Development Review*, 42(4):673–680.

Godwin, M., Gregory, J., and Balmer, B. (2009). The anatomy of the brain drain debate, 1950–1970s: Witness seminar. *Contemporary British History*, 23(1):35–60.

Golding, B. T. (2017). The impact of Brexit on British science: Also sprach Zunderland. *Future Medicinal Chemistry*, 9(3):271–273.

Halevi, G. and Moed, H. (2013). A bibliometric approach to tracking international scientific migration. *Scientometrics*, 1.

Hunter, R. S., Oswald, A. J., and Charlton, B. G. (2009). The elite brain drain. *The Economic Journal*, 119(538):F231–F251.

Independent (2016). EU referendum: An open letter to UK voters from leaders of 103 British universities. The Independent, url: accessed on 26 May 2022.

IOM (2019). Glossary on migration, iml series no. 34.

Irvine, J., Martin, B. R., Peacock, T., and Turner, R. (1985). Charting the decline in British science. *Nature*, 316(6029):587–590.

Johnson, H. G. (1965). The economics of the "brain drain": The Canadian case. *Minerva*, 3(3):299–311.

Kawashima, H. and Tomizawa, H. (2015). Accuracy evaluation of Scopus author

ID based on the largest funding database in Japan. *Scientometrics*, 103(3):1061–1071.

Krapf, S., Kreyenfeld, M., and Wolf, K. (2016). Gendered authorship and demographic research: An analysis of 50 years of demography. *Demography*, 53(4):1169–1184.

Larivière, V., Ni, C., Gingras, Y., Cronin, B., and Sugimoto, C. R. (2013). Bibliometrics: Global gender disparities in science. *Nature News*, 504(7479):211.

Laudel, G. (2003). Studying the brain drain: Can bibliometric methods help? *Scientometrics*, 57(2):215–237.

Mahroum, S. (2005). The international policies of brain gain: a review. *Technology Analysis & Strategic Management*, 17(2):219–230.

Martin, B. R. (1994). British science in the 1980s — Has the relative decline continued? *Scientometrics*, 29(1):27–56.

Martin, B. R. and Irvine, J. (1983). Assessing basic research: Some partial indicators of scientific progress in radio astronomy. *Research Policy*, 12(2):61–90.

Martin, B. R., Irvine, J., Narin, F., and Sterritt, C. (1987). The continuing decline of British science. *Nature*, 330(6144):123–126.

Meyer, J.-B. (2001). Network approach versus brain drain: Lessons from the diaspora. *International Migration*, 39(5):91–110.

Mingers, J. and Leydesdorff, L. (2015). A review of theory and practice in scientometrics. *European Journal of Operational Research*, 246(1):1–19.

Miranda-González, A., Aref, S., Theile, T., and Zagheni, E. (2020). Scholarly migration within Mexico: Analyzing internal migration among researchers using Scopus longitudinal bibliometric data. *EPJ Data Science*, 9:1–26.

Moed, H. F., Aisati, M., and Plume, A. (2013). Studying scientific migration in Scopus. *Scientometrics*, 94(3):929–942.

Moed, H. F., Burger, W. J. M., Frankfort, J. G., and Van Raan, A. F. J. (1985). The use of bibliometric data for the measurement of university research performance. *Research Policy*, 14(3):131–149.

Moser, P., Voena, A., and Waldinger, F. (2014). German Jewish émigrés and US invention. *American Economic Review*, 104(10):3222–55.

Nane, G. F., Larivière, V., and Costas, R. (2017). Predicting the age of researchers using bibliometric data. *Journal of Informetrics*, 11(3):713–729.

Nature (1967). Gaps and Drains. *Nature*, 214(5085):223–223.

Noble, D. (2016). From 'Save British Science', to 'CaSE', to the future. http://www.sciencecampaign.org.uk/news-media/guest-blog/from-save-british-science-to-case-to-the-future.html. Accessed on 16 May 2022.

Oldfield, R. C., Simmons, J. A., Jeffery, J. W., Cooper, W. M., Eden, R. J., Jones, G. O., Kondic, V., Mcmichael, J., Pike, E. R., Andrews, K. W., Bragg, W. L., Bagguley, D. M. S., Baker, J. M., Cooke, A. H., Elliott, R. J., Griffiths, J. H. E., ter Háar, D., Hatton, J., Hill, R. W., Houghton, J. T., Kuhn, H. G., Kurti, N., Robinson, F. N. H., Rollin, B. V., Rosenberg, H. M., Sanders, J. H., Series, G. W., Wilks, J., Woodgate, G. K., Hutchinson, G. W., Runcorn, S. K., GOLD, T., Albon, N., Haddow, A., Hartwell, R. M., Holmes, J., Garrett, P. J., Garrett, S., Knowles, P. F., Moss, G. P., Prince, A., Sachs, G., Weitzman, D., Williamson, A. R., and Holliday, P. (1963). Emigration of scientists from the United Kingdom: Report of a committee appointed by the council of the Royal Society. *Minerva*, 1(3):358–380.

ONS (2020). Defining and measuring international migration. url: https://www.ons.gov.uk/peoplepopulationandcommunity/populationandmigration/internationalmigration/articles/definingandmeasuringinternationalmigration/2020-02-14 accessed on 26 July 2022.

ONS (2021). Long-term international migration, provisional: year ending december 2020. url: https://www.ons.gov.uk/peoplepopulationandcommunity/populationandmigration/internationalmigration/bulletins/longterminternationalmigrationprovisional/yearendingdecember2020 accessed on 26 July 2022.

Paturi, M. and Loktev, A. (2020). The best gets better: Scopus data quality, measured. In *Pure International Conference.* url: brighttalk.com/webcast/13819/456949/prcn2020-the-best-gets-better-scopus-data-quality-measured accessed on 16 May 2022.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., and Dubourg, V. (2011). Scikit-learn: Machine learning in Python. *the Journal of Machine Learning Research*, 12:2825–2830. Publisher: JMLR. org.

Pereira, T., Prange-Gstöhl, H., Szkuta, K., and Truco, M. (2020). Scientific, technological and innovation performance. In European Commission and

Directorate-General for Innovation and Research, editor, *Science, research and innovation performance of the EU, 2020 : a fair, green and digital Europe*, chapter 6, pages 358–450. Publications Office.

Robinson-García, N., Cañibano, C., Woolley, R., and Costas, R. (2016). Scientific mobility of early career researchers in Spain and the Netherlands through their publications. In *21st International Conference on Science and Technology Indicators-STI 2016. Book of Proceedings*, Valencia, Spain.

Robinson-García, N., Sugimoto, C. R., Murray, D., Yegros-Yegros, A., Larivière, V., and Costas, R. (2019). The many faces of mobility: Using bibliometric data to measure the movement of scientists. *Journal of Informetrics*, 13(1):50–63.

Sanliturk, E., Aref, S., Theile, T., Zagheni, E., and C. Billari, F. (2021). Dataset of international migration among published researchers from and to the United Kingdom over 1996-2020. FigShare data repository, url: https://doi.org/10.6084/m9.figshare.14207369.

Saxenian, A. (2005). From brain drain to brain circulation: Transnational communities and regional upgrading in India and China. *Studies in Comparative International Development*, 40(2):35–61.

Scellato, G., Franzoni, C., and Stephan, P. (2015). Migrant scientists and international networks. *Research Policy*, 44(1):108–120.

Scellato, G., Franzoni, C., and Stephan, P. (2017). A mobility boost for research. *Science*, 356(6339):694–694.

Subbotin, A. and Aref, S. (2021). Brain drain and brain gain in Russia: Analyzing international mobility of researchers by discipline using Scopus bibliometric data 1996-2020. *Scientometrics*.

Sugimoto, C. R. and Larivière, V. (2018). *Measuring research: What everyone needs to know*. Oxford University Press, Oxford, New York.

Sugimoto, C. R., Robinson-García, N., and Costas, R. (2016). Towards a global scientific brain: Indicators of researcher mobility using co-affiliation data. *ArXiv*. url: https://arxiv.org/pdf/1609.06499.

Sugimoto, C. R., Robinson-García, N., Murray, D. S., Yegros-Yegros, A., Costas, R., and Larivière, V. (2017). Scientists have most impact when they're free to move. *Nature News*, 550(7674):29.

Wais, K. (2006). Gender prediction methods based on first names with genderizer. *The R Journal*, 8(1):17–37.

Waldinger, F. (2016). Bombs, brains, and science: The role of human and physical capital for the creation of scientific knowledge. *The Review of Economics and Statistics*, 98(5):811–831.

Watanabe, S. (1969). The brain drain from developing to developed countries. *Int'l Lab. Rev.*, 99:401.

White, K. (2019). Publications output: US trends and international comparisons. science & engineering indicators 2020. nsb-2020-6. *National Science Foundation*.

Wible, B. (2017). Crossing borders along an endless frontier. *Science*, 356(6339):694–694.

Zhao, X., Aref, S., Zagheni, E., and Stecklov, G. (2021). International migration in academia and citation performance: An analysis of German-affiliated researchers by gender and discipline using Scopus publications 1996-2020. In *Proceedings of the 18th International Conference on Scientometrics and Informetrics*. ISSI.

Zhao, X., Aref, S., Zagheni, E., and Stecklov, G. (2022). Return migration of German-affiliated researchers: analyzing departure and return by gender, cohort, and discipline using Scopus bibliometric data 1996–2020. *Scientometrics*.

Tables

	Leaving	**the UK**	**Entering**	**the UK**
	Logit Coef.	Odds Ratio	Logit Coef.	Odds Ratio
Post-Brexit	0.356***	1.428***	0.152***	1.164***
	(0.0399)	(0.0570)	(0.0361)	(0.0420)
EU origin	-0.231***	0.794***	0.168***	1.183***
	(0.0387)	(0.0307)	(0.0279)	(0.0331)
UK origin	0.378***	1.459***	-1.013***	0.363***
	(0.0330)	(0.0482)	(0.0358)	(0.0130)
US origin	0.0122	1.012	0.173***	1.189***
	(0.0534)	(0.0540)	(0.0401)	(0.0477)
Post-Brexit # EU origin	0.304***	1.356***	-0.0984**	0.906**
	(0.0517)	(0.0701)	(0.0450)	(0.0408)
Post-Brexit # UK origin	-0.513***	0.599***	0.367***	1.443***
	(0.0474)	(0.0284)	(0.0508)	(0.0732)
Post-Brexit # US origin	0.112	1.118	-0.122*	0.885*
	(0.0716)	(0.0800)	(0.0638)	(0.0564)
Academic age	-0.108***	0.898***	-0.119***	0.888***
	(0.00141)	(0.00127)	(0.00146)	(0.00130)
Above average publications	-0.245***	0.783***	-0.185***	0.831***
	(0.0186)	(0.0145)	(0.0177)	(0.0147)
Above average citations	0.0425**	1.043**	0.00654	1.007
	(0.0198)	(0.0206)	(0.0193)	(0.0194)
Social sciences	0.154***	1.167***	0.247***	1.281***
	(0.0330)	(0.0385)	(0.0292)	(0.0374)
Health sciences	-0.0298	0.971	-0.0184	0.982
	(0.0285)	(0.0277)	(0.0279)	(0.0274)
Physical sciences	-0.118***	0.888***	-0.0866***	0.917***
	(0.0218)	(0.0194)	(0.0206)	(0.0189)
Life sciences	-0.0584**	0.943**	-0.0911***	0.913***
	(0.0247)	(0.0233)	(0.0238)	(0.0218)
Male (75% probability)	0.00787	1.008	0.144***	1.155***
	(0.0277)	(0.0279)	(0.0277)	(0.0320)
Female (75% probability)	-0.0735**	0.929**	0.117***	1.124***
	(0.0306)	(0.0285)	(0.0300)	(0.0337)
Constant	-1.897***	0.150***	-1.597***	0.202***
	(0.0672)	(0.0101)	(0.0383)	(0.00776)
Observations	317,212	317,212	317,212	317,212
Number of researchers	45,316	45,316	45,316	45,316

Table 1: Results of the random-effects logistic regression models for leaving and entering the UK. For both models, the first column shows the logit coefficients and the second column shows the odds ratios. *** p<0.01, ** p<0.05, * p<0.1

Leaving the UK							**Entering the UK**					
	RE Model Odds R.	Logit Odds R.	RE Model Odds R.	Logit Odds R.	RE Model Odds R.	Logit Odds R.	RE Model Odds R.	Logit Odds R.	RE Model Odds R.	Logit Odds R.	RE Model Odds R.	Logit Odds R.
Brexit	1.428***	1.428***	1.720***	1.720***	1.759***	1.759***	1.164***	1.164***	1.341***	1.341***	1.279***	1.279***
	(0.0570)	(0.0571)	(0.0854)	(0.0856)	(0.0881)	(0.0881)	(0.0420)	(0.0420)	(0.0652)	(0.0652)	(0.0628)	(0.0628)
EU Origin	0.794***	0.794***	0.792***	0.792***	0.780***	0.780***	1.183***	1.183***	1.183***	1.183***	1.082***	1.082***
	(0.0307)	(0.0307)	(0.0307)	(0.0307)	(0.0306)	(0.0306)	(0.0331)	(0.0331)	(0.0331)	(0.0331)	(0.0322)	(0.0322)
UK Origin	1.459***	1.459***	1.460***	1.460***	1.274***	1.274***	0.363***	0.363***	0.363***	0.363***	0.426***	0.426***
	(0.0482)	(0.0476)	(0.0483)	(0.0477)	(0.0422)	(0.0422)	(0.0130)	(0.0130)	(0.0130)	(0.0130)	(0.0157)	(0.0157)
US Origin	1.012	1.012	1.015	1.015	1.047	1.047	1.189***	1.189***	1.192***	1.192***	1.069	1.069
	(0.0540)	(0.0540)	(0.0543)	(0.0542)	(0.0558)	(0.0558)	(0.0477)	(0.0477)	(0.0478)	(0.0478)	(0.0449)	(0.0449)
Brexit #	1.356***	1.356***	1.356***	1.356***	1.358***	1.358***	0.906**	0.906**	0.906**	0.906**	0.907**	0.907**
EU Origin	(0.0701)	(0.0701)	(0.0701)	(0.0701)	(0.0707)	(0.0707)	(0.0408)	(0.0408)	(0.0409)	(0.0409)	(0.0416)	(0.0416)
Brexit #	0.599***	0.599***	0.599***	0.599***	0.596***	0.596***	1.443***	1.443***	1.442***	1.442***	1.449***	1.449***
UK Origin	(0.0284)	(0.0283)	(0.0284)	(0.0283)	(0.0284)	(0.0284)	(0.0732)	(0.0732)	(0.0733)	(0.0733)	(0.0746)	(0.0746)
Brexit #	1.118	1.118	1.118	1.118	1.119	1.119	0.885*	0.885*	0.885*	0.885*	0.886*	0.886*
US Origin	(0.0800)	(0.0800)	(0.0801)	(0.0801)	(0.0808)	(0.0808)	(0.0564)	(0.0564)	(0.0565)	(0.0565)	(0.0573)	(0.0573)
Constant	0.150***	0.150***	0.140***	0.140***	0.169*	0.169*	0.202***	0.202***	0.156***	0.156***	0.401**	0.401**
	(0.0101)	(0.00632)	(0.00999)	(0.00641)	(0.154)	(0.154)	(0.00776)	(0.00776)	(0.00673)	(0.00673)	(0.165)	(0.165)
Obs.	317,212	317,212	317,212	317,212	316,967	316,967	317,212	317,212	317,212	317,212	316,687	316,687
Num. Res.	45,316		45,316		45,281		45,316		45,316		45,241	
ICC	1.39e-06		1.39e-06		2.47e-07		2.00e-07		1.99e-07		6.56e-07	
Sigma	0.00214		0.00214		0.000901		0.000811		0.000809		0.00147	
Chi2	6612		6502		6577		9787		9787		9888	
Pseudo R2		0.0485		0.0492		0.0677		0.0763		0.0785		
Controls	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Year FE	NO	NO	YES	YES	YES	YES	NO	NO	YES	YES	YES	YES
Dest. FE	NO	NO	NO	NO	YES	YES	NO	NO	NO	NO	YES	YES

Table 2: Comparison of the odds ratio values in random effects logit model (RE Model Odds Ratios) and simple logit model (Logit Odds Ratios) for robustness, including the addition of year and destination country (*Dest. FE*) controls in second and third columns, for active researchers leaving and entering the UK. *Controls* refer to the control variables used in the random effects model shown in *Table 1*, which are included in all regressions shown above. These variables are; academic age, above average publications, above average citations, dummy variables for field of work (social, health, life and physical sciences) and inferred gender with 75% probability. Abbreviations: Obs. refers to *observations*, Num. Res. refers to *number of researchers* and ICC refers to the *Intraclass Correlation*. Robust standard errors in parentheses *** $p<0.01$, ** $p<0.05$, * $p<0.1$

Leaving the UK				
	Marginal Probability	Std. Error	95%	CI
Pre-Brexit # Other Origin	0.0368	0.0010	0.0349	0.0388
Pre-Brexit # EU Origin	0.0296	0.0008	0.0281	0.0310
Pre-Brexit # UK Origin	0.0525	0.0008	0.0509	0.0541
Pre-Brexit # US Origin	0.0373	0.0016	0.0341	0.0404
Post-Brexit # Other Origin	0.0515	0.0011	0.0493	0.0536
Post-Brexit # EU Origin	0.0551	0.0009	0.0533	0.0569
Post-Brexit # UK Origin	0.0454	0.0008	0.0437	0.0470
Post-Brexit # US Origin	0.0577	0.0019	0.0539	0.0615
Entering the UK				
	Marginal Probability	Std. Error	95%	CI
Pre-Brexit # Other Origin	0.0515	0.0011	0.0494	0.0536
Pre-Brexit # EU Origin	0.0601	0.0010	0.0583	0.0620
Pre-Brexit # UK Origin	0.0197	0.0005	0.0186	0.0207
Pre-Brexit # US Origin	0.0604	0.0018	0.0568	0.0640
Post-Brexit # Other Origin	0.0592	0.0012	0.0569	0.0615
Post-Brexit # EU Origin	0.0631	0.0009	0.0613	0.0650
Post-Brexit # UK Origin	0.0324	0.0007	0.0310	0.0338
Post-Brexit # US Origin	0.0620	0.0020	0.0582	0.0659

Table 3: Marginal probabilities of leaving and entering the UK before (pre) and after (post) Brexit, with standard errors and 95% confidence intervals.

Figures

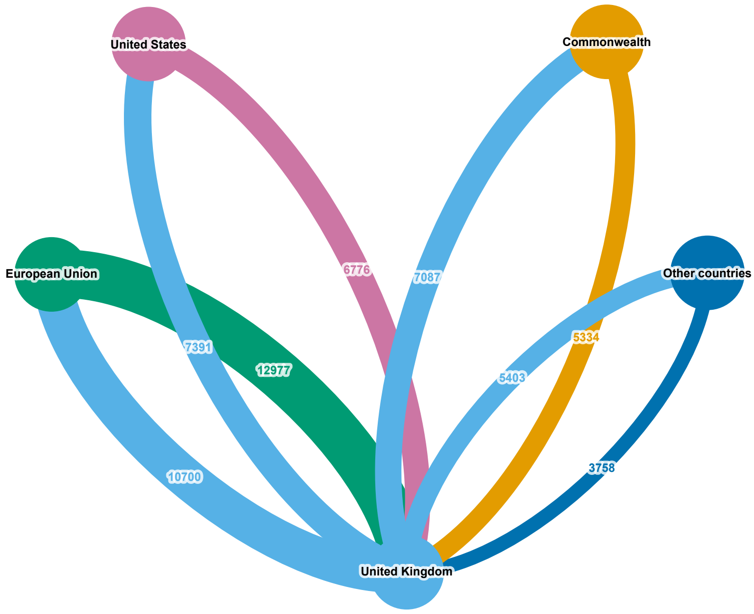

Figure 1: Migration flows and the overall patterns of scholarly migration in the three years prior to the Brexit referendum: the EU had the largest flows to and from the UK followed by the US, the Commonwealth countries, and all *other* countries in a decreasing order. The edges represent the migration flows in 2013-2015. The direction of the edges is clockwise. The colours of the edges are based on the origin node.

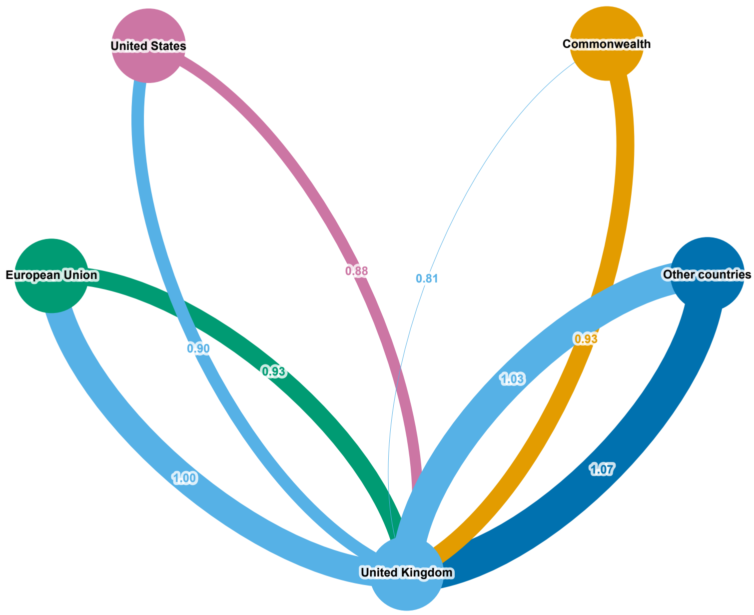

Figure 2: In the three years after the Brexit referendum, the flows have mostly decreased, except for the flows from the UK to the EU (which has remained the same), and the flows between the UK and *other* countries (which have increased in both directions). The edges represent the changes in migration flows in 2016-2018 compared to in 2013-2015. For example, the weight of the (green) directed edge from the EU to the UK is 0.93. This indicates that the total flow from the EU to the UK in the three years after the Brexit referendum was equal to 0.93 of the corresponding flow during the three years before the Brexit referendum. The direction of the edges is clockwise. The colours of the edges are based on the origin node.

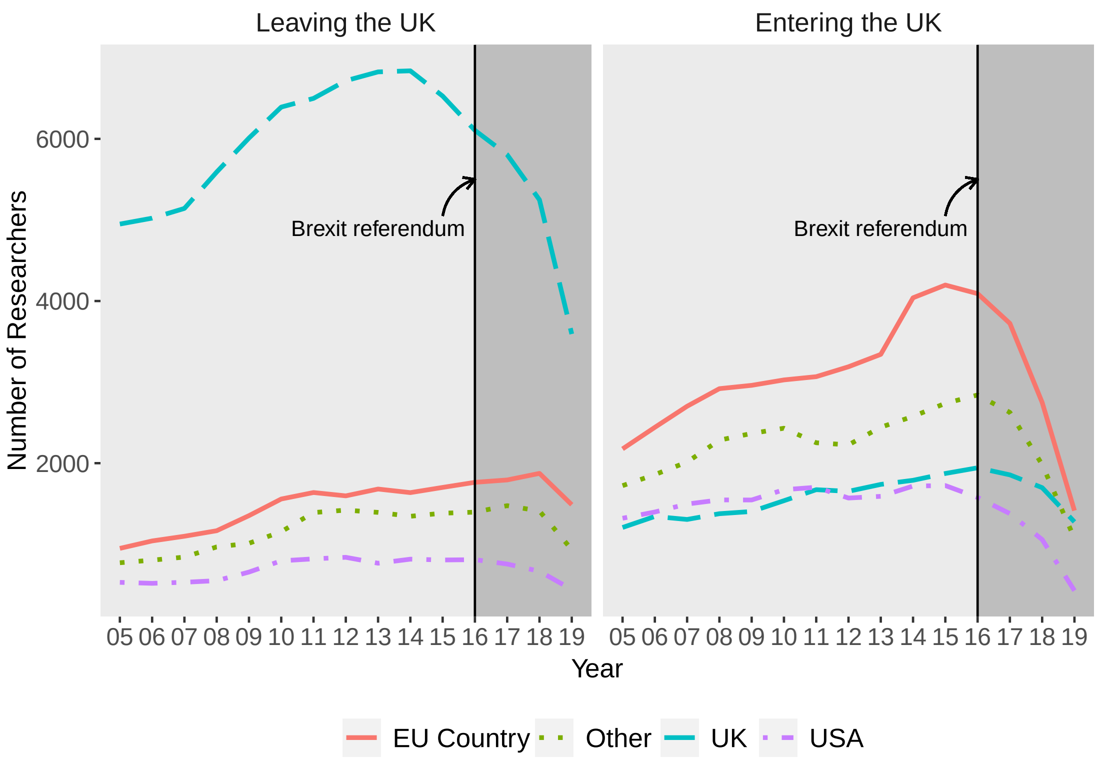

Figure 3: The figure is a descriptive representation of the number of all researchers in our data set leaving and entering the UK by country of academic origin. Instead of starting in the year 2013, we report the numbers for a the longer period of 2005-2019. On the left side, the patterns of moving out of the UK are shown based on the annual total number of researchers by academic origin. In contrast, on the right side, the researchers' patterns of moving to the UK are presented by the annual total numbers and by academic origin. The year of the Brexit referendum, 2016, is marked with a black vertical line. The sharp decline observed in the last years should be interpreted as a result of right-censoring. The slope is expected to partially flatten with the introduction of more recent publication data and related improvements for the inference of migration events. However, it is important to note that we observe a slightly increasing trend for all researchers leaving the UK whose academic origin was an EU country after the year of Brexit and despite the right-censoring, except for the year 2019, for which we see the impact of right-censoring in the data for all groups.

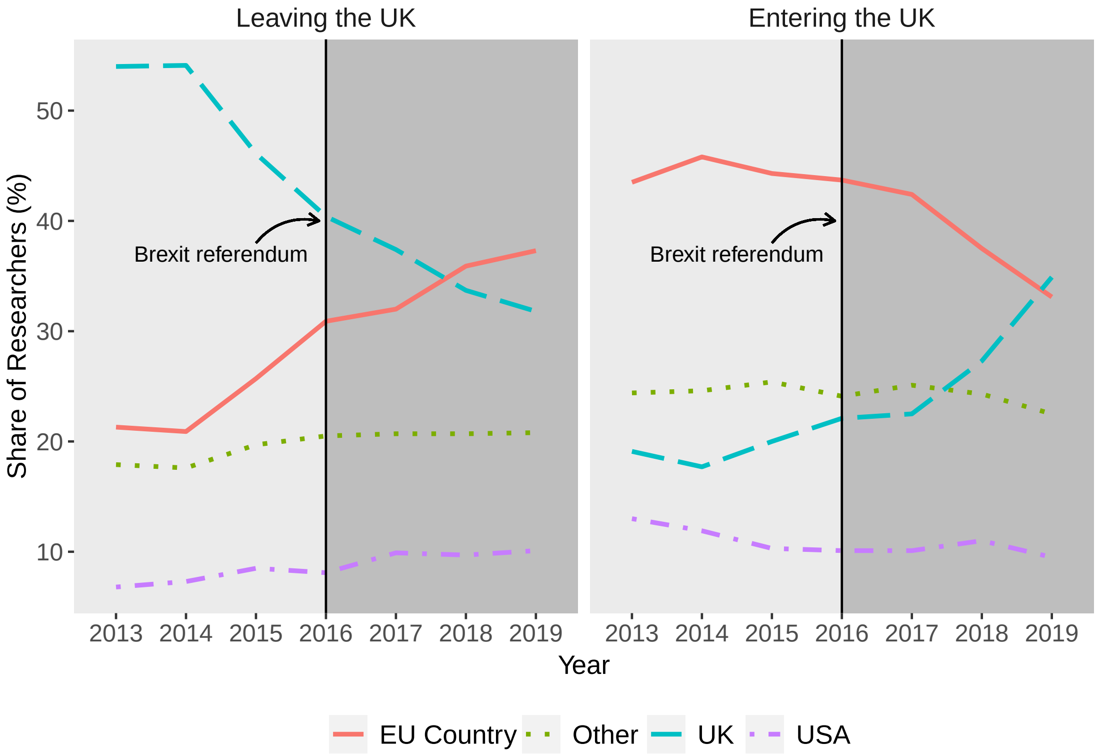

Figure 4: Shares of active researchers leaving (emigration) or entering (immigration) the UK by country group of academic origin from 2013 to 2019 ($N = 45,316$). The shares reflect the percentages of the four academic origin groups among all active researchers leaving and entering the UK in a given year, respectively. The year of the Brexit referendum, 2016, is marked with a black vertical line. Building on the descriptive analysis in *Figure 4*, we observe that the changing patterns of active researchers by academic origin are more prominent when we focus on their shares among all active researchers leaving or entering the UK, instead of on the sheer numbers. The changes after the Brexit referendum were more remarkable for active researchers entering the UK. In 2015, among all active researchers entering the UK, the share of researchers with an EU country academic origin was above 40%, while the share of researchers with a UK academic origin was around 20%. By 2019, we see that the share of researchers with an EU country academic origin decreased by approximately 10 percentage points. The share of researchers with a UK academic origin increased instead by more than 10 percentage points, thus accelerating an increasing trend right before the Brexit, that brought both categories to around the same level of above 30%.

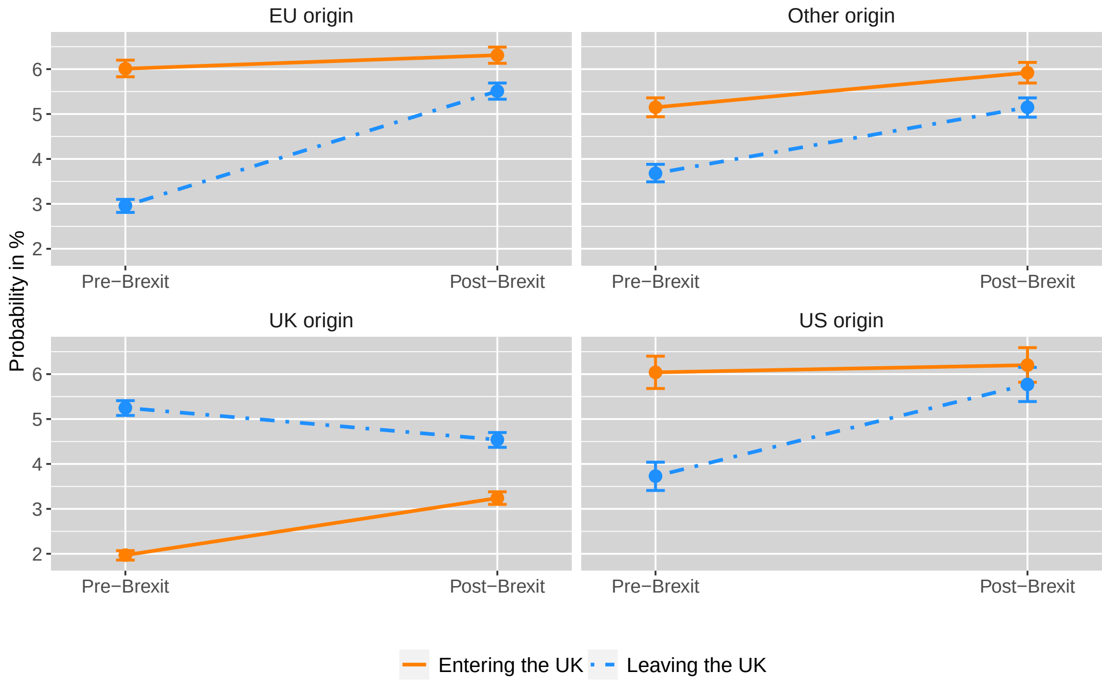

Figure 5: The figure shows the marginal probabilities of entering (immigration) and leaving (emigration) the UK before and after Brexit with 95% confidence intervals, for active researchers ($N = 45,316$) by academic origin. The marginal probabilities of entering the UK are represented by solid orange lines, whereas the marginal probabilities of leaving the UK are represented by dot-dash blue lines. The category of other academic origin includes any country that is not the UK, the US, or part of the EU. A look at the probabilities of entering the UK before and after Brexit shows that the probability of leaving the UK increased after Brexit for all academic origin groups except for the UK academic origin group. The probability of leaving the UK decreases by around 14% after Brexit for active researchers with a UK academic origin, falling from 5.3% to 4.5%. The biggest change was observed for active researchers with an EU country academic origin, as their probability of leaving the UK increased by approximately 86%, from almost 3.0% in the pre-Brexit period to 5.5% in the post-Brexit period. Regarding the probability of entering the UK, we can see that for active researchers with an EU or a US academic origin, there was no statistically significant change. For the other academic origin group, there was a small but statistically significant increase in the probability of entering the UK after Brexit. The most striking change was in the probability of active researchers with a UK academic origin of moving (back) to the UK, as it increased by approximately 65%, from nearly 2.0% before Brexit to 3.2% after Brexit.

Appendix

Captions for Movies S1 to S3

The animated maps in movies S1-S3 demonstrate the international movements of researchers between the UK and EU countries. The country names are labelled if the number of researchers moving from a given EU country is in the fourth quartile (the highest 25%) of the same type of movements (to or from the UK) in all the seven years combined, to allow for comparison between years.

Movie S1 is accessible at saref.github.io/SI/SAZB2021/All-moves-to-UK.mp4. It is an animated map that demonstrates the annual number of researchers who moved from EU countries to the UK in the 2013-2019 period. The researchers included in the data used to create this animated map are internationally mobile and have at least one publication with a UK institution affiliation in their academic career. The active researcher status used in the statistical analysis and the academic origin of the researchers did not play a role in the drawing of this animated map. The relative decrease in the number of researchers moving from an EU country to the UK after 2016 found in this descriptive analysis is in line with the results of the statistical analysis of this study. However, in contrast to the statistical analysis, in which we focused only on active researchers, this analysis includes all researchers who have moved during the given time period, regardless of their frequency of publishing. It is possible that the decrease in the last two years (2018-2019) may be due to right-censoring of the data, and thus to our inability to detect migration events towards the end of the study period. Therefore the decrease in the movements in the last two frames of the animation should be interpreted cautiously.

Movie S2 is accessible at saref.github.io/SI/SAZB2021/All-moves-from-UK.mp4. It is an animated map that demonstrates the annual number of researchers who moved from the UK to EU countries in 2013-2019. The researchers included in the data used to create this animated map are internationally mobile and have at least one publication with a UK institution affiliation in their academic career. The active researcher status used in the statistical analysis and academic origin of the researchers do not play a role in this animated map. In this animated map, the changes that can be observed after the year of the Brexit referendum (2016) are rather small. However, we are able to observe an increase in the number of movements after 2016 for two years, despite the right-censoring problem in the data, as explained for Movie S1. We observe that Italy, Spain, and France appear to become more popular destinations for researchers leaving the UK after 2016. Caution is needed in interpreting this descriptive analysis, as the numbers of researchers and the numbers of moves may be underestimated, especially during the last two years.

Movie S3 is accessible at saref.github.io/SI/SAZB2021/Active-moves-to-UK.mp4. It is an animated map that demonstrates the annual number of active researchers who moved from EU countries to the UK in 2013-2019. The researchers included in the data used to create this animated map are active researchers who are internationally mobile and have at least one publication with a UK institution affiliation in their academic career. The academic origin of the researchers does not play a role in this animated map. Note that in contrast to the statistical analysis, in which the country from which an active researcher moves out (emigration) to enter the UK could be any country, in this animated map, we restrict the country of emigration to EU countries. As this descriptive analysis focuses on active researchers who have a location reference for each year, the right-censoring issue encountered in movies S1 and S2 does not apply in this case. The animated map shows that there was a declining trend in the number of researchers moving to the UK from EU countries, especially after the Brexit referendum in 2016.

Dataset S1

Dataset S1 contains one of the migration flows of researchers between the UK and other countries disaggregated by years and countries. The dataset will become publicly available in a *FigShare* repository (Sanliturk et al., 2021) upon publication of the article.

Dataset S1 is shared under a CC BY-NC-SA Creative Commons v 4.0 license (Attribution-NonCommercial-ShareAlike). This means that other individuals may remix, tweak, and build upon these data non-commercially, as long as they provide citations to the data repository (Sanliturk et al., 2021) and the reference article, and license the new creations under the identical terms.

The dataset is provided in a comma-separated values file (.csv file), and each row represents the migration flow of research-active scholars from one country to another country in a specific year. Either the origin country or the destination country is the United Kingdom (coded as GBR).

The data may be used to produce migration models or possibly other measures and estimates. They can also be used as an edge list for creating a network model of migration flows (directed weighted edges) between the UK and other countries (nodes).

Additional Figures and Tables for Online Appendix

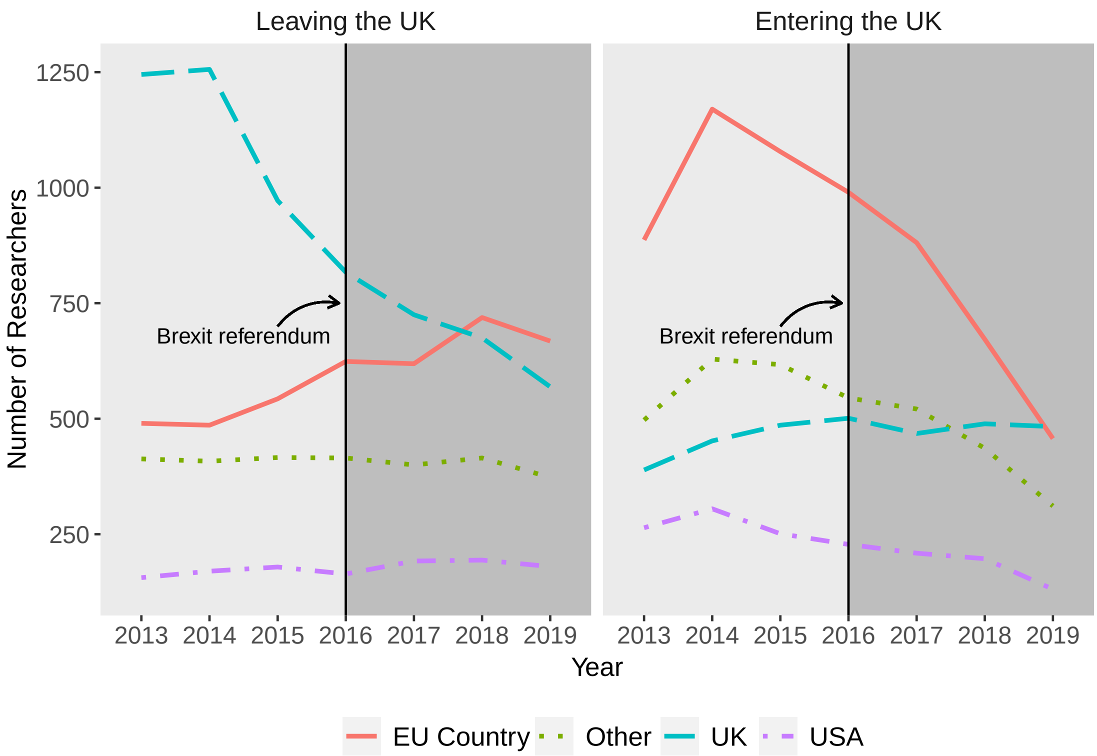

Figure 6: The figure depicts the number of active researchers leaving (emigration) and entering (immigration) the UK by country of academic origin, from 2013 to 2019 ($N = 45,316$). The panel on the left shows the patterns of active researchers leaving the UK, while the panel on the right shows the patterns of active researchers entering the UK. The year of the Brexit referendum, 2016, is marked with a black vertical line. The figure shows that, for active researchers whose academic origin was an EU country, the declining trend of entering the UK sharpened after the Brexit referendum, while the number of researchers in this group who were leaving the UK was increasing. In contrast, for the active researchers whose academic origin was the UK, the declining trend of leaving the UK continued, and this group was the only academic origin group who did not to have a decreasing trend of entering the UK after 2016.

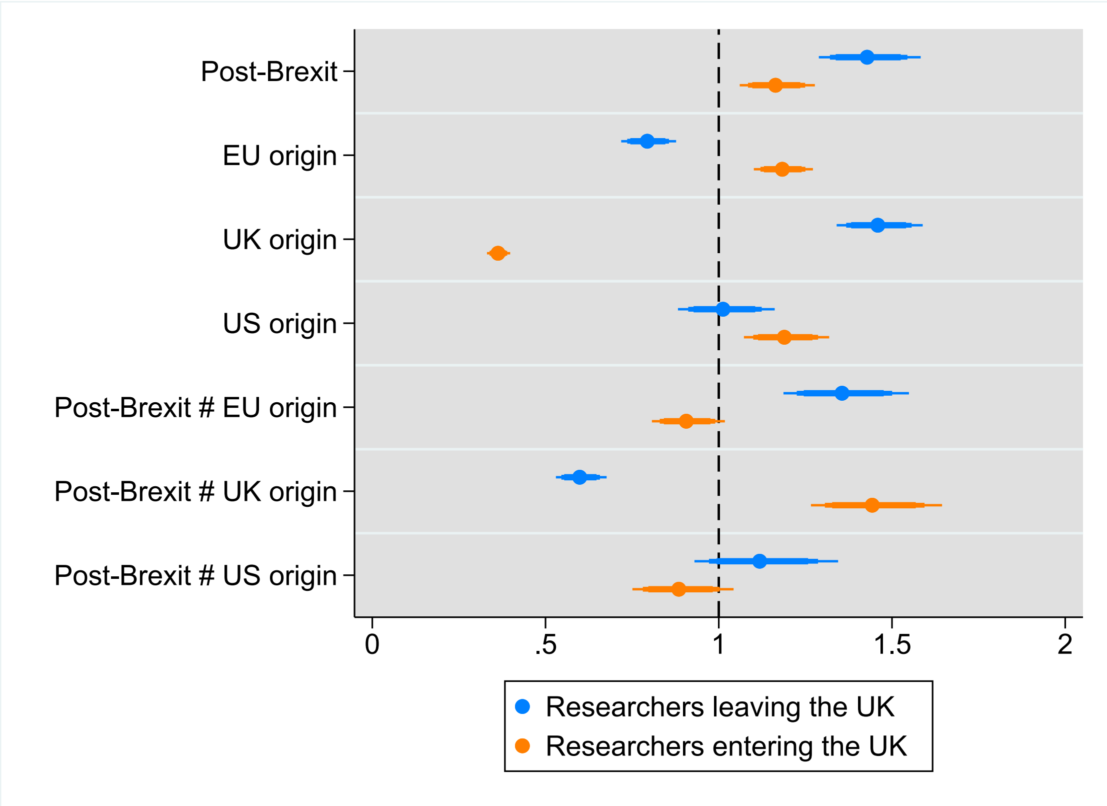

Figure 7: The plot of odds ratios with 99%, 95%, and 90% confidence intervals for entering and leaving the UK for active researchers ($N = 45,316$), with interactions by Brexit years and academic origin, between 2013 and 2019. Results are obtained by fitting a random-effects logistic regression model. Researchers with an EU academic origin were, on average, more likely to enter the UK and less likely to leave the UK, but the Brexit referendum counteracted this trend. Conversely, researchers with a UK academic origin became, on average, less likely to leave the UK and more likely to move back to the UK, in the post-Brexit years.

	Male	Female	Unknown	Total
Gender by 75% threshold	32019	10427	2870	45316
Gender by 90% threshold	30401	9813	5102	45316

Table 4: The table shows the number of active researchers and gender estimation by 75% and 90% probability thresholds.