The Extent and Effect of the Brain Drain in Germany (Work in Progress)

Sonia Conchi[1] and Carolin Michels[2]

[1] *Sonia.Conchi@isi.fraunhofer.de*
Fraunhofer Institute for Systems and Innovation Research ISI, Breslauer Strasse 48, 76139 Karlsruhe, Germany

[2] *Carolin.Michels@isi.fraunhofer.de*
Fraunhofer Institute for Systems and Innovation Research ISI, Breslauer Strasse 48, 76139 Karlsruhe, Germany

Introduction

A global international exchange between scientists is essential to drive progress. Thus, scientists form international networks to promote the exchange of information in their research field (cf. Hunger 2003). Exchanges can take place via any means of communication or through personal contact, e.g. research visits or even working abroad.

Sharing knowledge and information can be an asset for a country, but in the case of researcher mobility, the ultimate gain depends on whether researchers return to their homeland. The concept of "Brain drain" is used to refer to scientists who migrate from one country to another with no intention of returning (Grubel, 1994) – an action that has international, economic and political impacts, especially in developing countries (see, e.g. Lowell 2002).

“Brain circulation”, on the other hand, which is defined as a research visit abroad with a fixed duration and the intention to return, is a politically and economically endorsed activity. Such exchanges early on in a researcher’s career can help him/her to learn more, build up international networks, and acquire inter-cultural competences. However, if a researcher decides not to return from abroad, these skills are lost for the original country. Moreover, besides not adding to the labor market (causing opportunity costs), all national investments in the researchers’ previous education and qualification are lost to the national economy. This so called brain drain hurts the national economy in contrast to a beneficial brain circulation.

A pervasive problem in Germany (GER) is the lack of skilled labor which is forecasted to escalate in the future due to demographic change.

The actual extent of brain drain is difficult to measure, because there are no official records capturing exactly how many skilled workers migrate, which countries they prefer, for how long they stay or when and if they return (OECD 2001).

Bibliometric data makes it possible to track scientists by the affiliations on their publications. Furthermore, bibliometric databases like, e.g. Elsevier’s Scopus, capture information about the author in addition to the publication form and source, i.e. his/her name, address and organization. Thus, it is possible to track authors over time, from country to country and from organization to organization. The core question of our paper is whether brain drain is a serious problem for the German research system.

In this paper, brain drain is studied based on data from Scopus - one of the biggest bibliometric databases worldwide. We created a data set of all German scientists who published in 2000. The international movements of these scientists are then tracked for a period of 10 years.

The first objective was to find out how many German scientists migrated in this period, how long they stayed abroad and how many of them returned. The second objective was to analyze where they moved to and returned from.

Literature overview

Various studies have been made on the extent and influence of the brain drain.

Allmendinger (2003) performed a qualitative study of the migration incentives of German scientists and identified the US, Great Britain and Switzerland as the most popular destinations for German scientists. Janson (2006) compared the career opportunities for scientists between the United States and Germany and showed that the US has a perceived better work situation than GER.

In a study by Büchtemann (2001), the migration of German scientists to the US was analyzed using different databases such as "Current Population Survey" (CPS), "National Survey of College Graduates" (NSCG), "Science & Engineering Statistic Database" (SESTAT), "Foreign Scholars Survey" and "Survey of Earned Doctorates" (SED). They found that Germans living in America are often employed in research, development and teaching, particularly in the technical and natural sciences and in medicine. Furthermore, the number of German doctoral recipients increased over time as did the number of those intending to stay in the US.

Using publication data to examine the mobility of scientists is a relatively new approach. Roberge (2012) used Scopus data for an analysis of Canadian researcher migration. Their findings for Canada reveal a net migration flow on a very low level, providing evidence that Canada is not facing a brain drain.

In a very recent study, Moed et al. (2013) also showed that it is possible to trace scientists and their mobility using Scopus data. They compared two different data sets (diachroneous approach and synchronous approach) for five countries (Germany, Italy, the Netherlands, the UK and the USA). According to their study, countries to which German scientists migrated most frequently were the USA, UK, Switzerland, France, Austria and the Netherlands – a finding we were able to corroborate to some extent.

Our research addresses the following research questions:

1. How many German scientists migrated in the period between 2000 and 2010?
2. How long did they stay abroad and how many of them returned?
3. Which are the most popular destinations for German scientists?

The hypotheses we test here are:

H1: Only a small percentage of German scientists leave GER for a long time or permanently.

H2: Migrating scientists publish more than scientists who remain in GER.

H3: The share of scientists publishing abroad is increasing over time.

H4: If migrating scientists have not returned after two years, they will stay away for longer period or permanently.

H5: The most popular countries are the US, the UK and Switzerland.

Other hypotheses could be tested with this database and would be our next steps of research:

H5: The US is particularly attractive to natural scientists.

H6: Migrating researchers publish more in international collaborations than non-migrating researchers.

H7: Migrants are more frequently cited (have a higher quality of research).

H8: US immigrants perform better than other migrants, etc.

The document is structured as follows: First, we present our method for tracking the movements of German authors and validate the Scopus data for a subset of authors with a CV-based approach. In the results section, we begin with a descriptive analysis of the author set. Then we explain the movements of German authors in detail. In particular, we describe the share of migrating researchers, how long they stay abroad and the countries of migration and countries from which they return. We conclude with a discussion of brain drain and an outlook to further research.

Methodology

In this study, the migrations of German scientists are depicted in a descriptive analysis. The data preparation is impeded by the necessity of reflecting the actual brain drain. Scopus has one advantage over other databases as it includes e-mail addresses, which enables a qualitative survey to validate the results.

The basic data set contains all the scientists who published in 2000 and the previous two years with a German affiliation. This minimizes the proportion of immigrants included in the basic set.

With this selection, a data set of 64,925 German scientists was generated, whose publication activity was traced for 10 years. Authors were tracked using the author ID in Scopus.

The data set includes the following information for each publication: author ID, last name, first name, country of publication from 2000 until 2010, e-mail address, classification code of the main research field and the number of publications per year.

Publications could not be identified for every author in the starting set in all subsequent years. It was not possible to deduce whether these authors did not publish in these years, or if their publications were not included in Scopus. However, for these cases, we have no information on the authors' location in the respective years and assumed that the scientists stayed in the country of their last known affiliation.

We identified three groups in the data set with different publication addresses related to countries:

1. Scientists who published in one year only with German addresses.
2. Scientists who published in one year with at least one German address and also with at least one foreign address.
3. Scientists who published in one year with only a foreign address.

The scientists of the first two groups were combined for the analysis. We concentrated on the migratory scientists who temporarily renounced the German address in their publications.

To track the movements of authors (migratory scientists) for each year between 2000 and 2010, we divided their actions into four groups:

1. Scientists who remained in GER are those who publish now (observation year) and also in the previous year with a German address (Stay GER).
2. Scientists who only swapped their German address for a foreign address between years (Move abroad).
3. Scientists who keep publishing with a foreign address (Stay abroad).
4. Scientists who publish in the observation year with a German address, but were associated in the previous year with foreign addresses only (Move to GER).

Figure 1 illustrates the four possible movements and their labeling in the remainder of this paper.

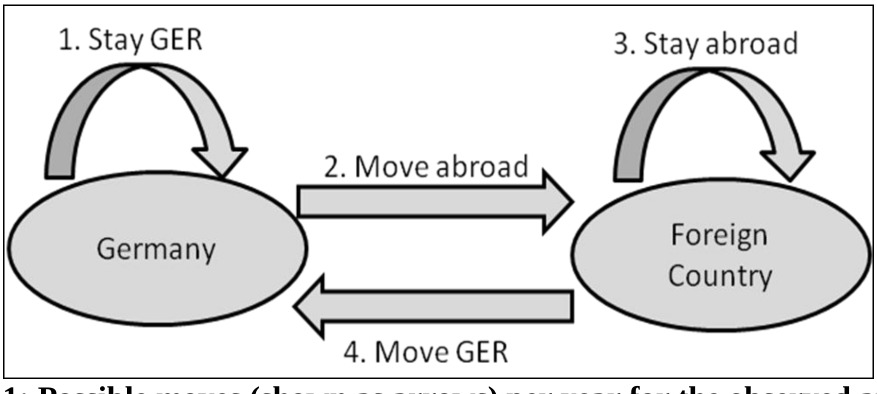

Figure 1: Possible moves (shown as arrows) per year for the observed authors.

The aim of this study was to track real migration and not short-term stays abroad. Therefore, authors with multiple addresses in one year who had at least one address in GER were assigned to GER. This was assumed to be a sign that they had not completely left their home country.
There were many cases in which authors published in one year from different non-GER countries. Countries were weighted in this case. For example, if an author published in two countries, each country was weighted with ½.
To validate the Scopus data, we randomly selected 50 scientists in our data set who published continually from 2000 until 2010. A manual internet search was performed to find out where the scientists had stayed in the period 2000 to 2010 according to their own CVs. The probability of consistency was calculated by comparing these results with our data set.

Data validation based on CVs
Moed et al. (2013) have pointed out two important errors in Scopus: "merge authors" and "split authors". The former occur when multiple authors are lumped together under the same author ID, for example, when there are two authors with the same name. To reduce this error, one could identify author IDs in the data set that have several e-mail addresses, an immense number of publications and have published in one year from more than three different countries or in various scientific fields. In a second step, their CVs could be compared with the bibliometric data to find out which author IDs are affected.
A "split authors" error occurs if one author has multiple author IDs. According to Moed et al (2013), it is difficult to identify the affected author IDs, in particular when the name of an author has changed.
However, other errors like incomplete or incorrect data also occur that are independent of the author ID. To validate our data set from Scopus, we compared the data with the CVs of the authors. The aim of this validation step is to find out whether and how precisely Scopus data can track the movement of scientists from one country to another.
We randomly selected 50 authors from our main data set of 64,925 authors, who published at least one paper each year between 2000 and 2010.
A high proportion of these - namely 41 of 50 authors - could be identified. To compare the Scopus data and the CV data in detail, six possible codes were assigned that depicted the level of agreement between the Scopus and the CV data for each author in each year (see description below Table 2). The level of agreement for one author could change over time if a CV was found. This is especially the case if the above mentioned merge or split errors occur and information is lost (codes 3 and 4) or wrongly attributed to an author (codes 2 and 4). In fact, there were only 11 authors for which the CV and the Scopus data were in accordance over the whole observation period (code 1). Table 2 shows the results of the distribution of the codes for each year and on average. The latter is calculated to show the average agreement per year between both data sources.

Table 2: Results of the data validation (n=50)

Code	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	Average
0	4%	4%	4%	2%	4%	4%	0%	0%	0%	0%	2%
1	46%	42%	60%	68%	62%	62%	66%	58%	70%	64%	60%
2	28%	26%	8%	4%	10%	10%	10%	20%	8%	12%	14%
3	2%	2%	6%	4%	4%	4%	4%	2%	2%	4%	3%
4	0%	4%	2%	2%	0%	0%	0%	2%	2%	0%	1%
5	20%	22%	20%	20%	20%	20%	20%	18%	18%	20%	20%

Codes: 0 = no match of Scopus data and CV; 1 = Full match of Scopus data and CV; 2 = match with additional countries in Scopus; 3 = match with additional countries in CV; 4 = match with additional countries in Scopus and in CV; 5 = CV could not be found on the internet or country was not specified in the CV.

On average, the CV data and the Scopus data were completely different in 2% of the cases. However, this occurred only for single years of authors and never over the whole observation period, so that merge author or split author errors can be eliminated; if a merge error had occurred, the respective code for the affected years would have been 2 and 4 and codes 3 or 4 would have been assigned for a split error.
The confirmed accordance between both data sources was 78% on average. There was 60% full compliance for countries according to the Scopus data and the CVs (code 1). For an average of 18%, there was a match but with additional information in Scopus (14%) or the CV (3%) or both (1%). This suggests that there may be a merge author error, because there is no documentation in whatever form of the abroad stay in the publications but not in the CV., vice versa would be very unusual. To identify the number of merge authors, we searched for different criteria like publication rate, institution etc. We did this with all (27) authors that had a code 2 mismatch in at least one year. These authors show a relatively high number of affiliations, 17.59 on average with values ranging from 7 to 49. However, the number of fields in which these authors were active did not shed new light on this issue, because every one of the 27 authors had only published in 1 field according to the Scopus data. The number of publications ranged from 30 to 408, where the maximum value was able to be confirmed by the CV of that specific author.
It is interesting that there were cases in which one author was assigned codes 2, 3 and 4 in different years. Only 22% of the cases received code 1 throughout 2001-2010.
Here, it must be noted that we considered only a very small random sample and random effects are not excluded. Nevertheless, the results show that a scientist to whom a foreign country is assigned in Scopus also had an approximately 63% chance of being present in the country in the corresponding year and thus deductions for the mobility of researchers are possible.

Results

H1: Only a small percentage of German scientists leave GER for a long time or permanently.

Table 1 provides an overview of the data set. Of the 64,925 identified authors, on average 83% published between 2001 and 2010 only with German addresses, 8% with German and an additional foreign address and 9% published unambiguously with only a foreign address (migratory scientists). Figure 2 offers a more detailed overview.

Table 1: Description of data set (2001-2010)

	Total data set	GER address	GER and foreign address	Foreign address
Average number of authors	64,925 (100%)	53,854 (83%)	5,443 (8%)	5,628 (9%)
Number of Publications				
1-5	25%	37%	4%	14%
6-10	13%	17%	7%	11%
11-20	17%	18%	15%	17%
21-50	26%	21%	35%	29%
51+	19%	8%	39%	28%

Research fields				
-Medicine	33%	34%	34%	26%
-Physics and Astronomy	16%	12%	22%	24%
-Biochemistry, Genetics and Molecular Biology	11%	10%	12%	13%
-Chemistry	8%	8%	7%	7%
-Engineering	7%	8%	5%	6%
-Materials Science	5%	5%	3%	4%
-Earth and Planetary Sciences	4%	3%	4%	5%
-Agricultural and Biological Sciences	3%	3%	3%	3%
-Neuroscience	2%	2%	2%	2%
-Mathematics	2%	2%	1%	2%
-Environmental Science	2%	2%	1%	1%
-Computer Science	2%	2%	1%	2%
-Immunology and Microbiology	2%	2%	1%	2%
-Pharmacology, Toxicology and Pharmaceutics	1%	1%	1%	1%
-Chemical Engineering	1%	1%	0%	0%
-Veterinary	1%	1%	1%	0%
-Psychology	1%	1%	0%	1%

H2: Migrating scientists publish more than scientists who remain in GER.

In the following, we compare the three groups by their publication behavior and research fields. We count the number of publications for each of the three groups and in total. We show values as percentages to compare them. The authors publishing with foreign addresses (even if they keep their German address) are more likely to publish more than 21 publications in 10 years than those who publish with only a German address. However, at this stage of the analysis, interpretations are a chicken and egg problem: Either moving abroad increases the publication rate per author or having a higher publication output increases the chance of finding a position abroad (e.g. also because it increases the chance of getting a scholarship).

To obtain an impression of differences between research fields, we count the movements per field of the three groups and in total. We use Scopus classifications to assign authors to research fields. There is no huge difference between the three groups, except for medicine and physics and astronomy. There are fewer authors in medicine who publish with a foreign address (26%) than publish with a German address (34%). In physics and astronomy, there are more authors who publish with a foreign address (24%) than publish with a German address (12%).

H3: The share of scientists publishing abroad is increasing over time.

Figure 2 shows that the number of scientists publishing with a foreign address (no German address) increased during the time period; in 2004, it even exceeds the number of scientists with a German and a foreign address. This means that especially scientists with double addresses (German and foreign) gradually give up their German address.

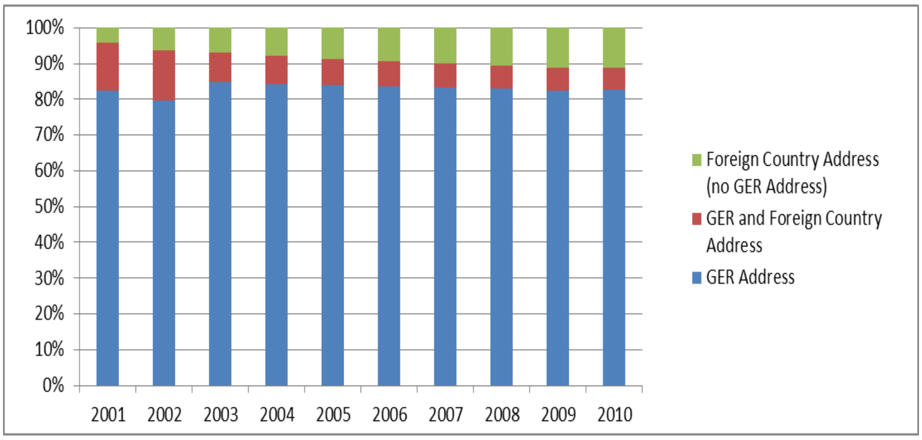

Figure 2: Author addresses 2000-2010.

For the further analyses, the scientists with German and German or foreign addresses are grouped together. These are regarded as the comparison group to the movers who publish only with a foreign address.

H4: If migrating scientists have not returned after two years, they will stay away for a longer period or permanently.

Those migratory scientists (2,698 in 2001 and up to 7,377 in 2010) are now considered in more detail in order to investigate whether a brain drain exists in GER.
Figure 3 provides an overview of the stations of migratory scientists.

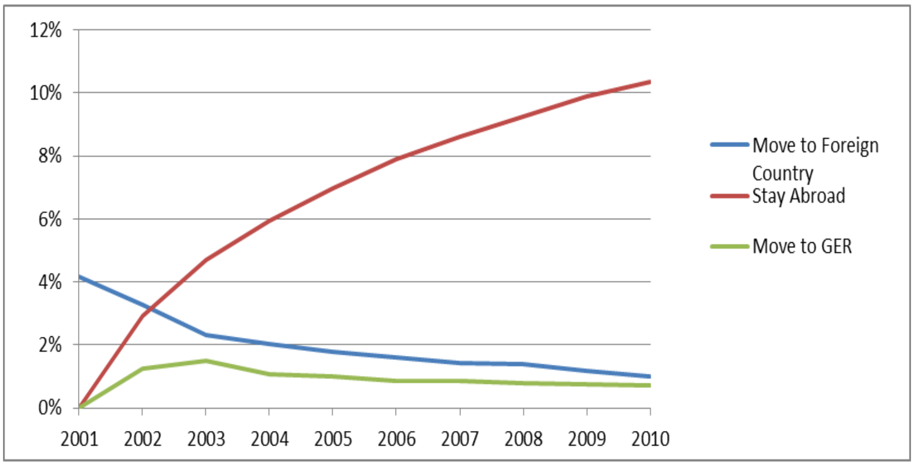

Figure 3: Migration stations of German scientists 2000-2010.

The analysis started in 2000. Two years later, in 2002, 93% of the scientists are still in GER, 3% had moved abroad, a further 3% left GER and the first 1% return to GER. Each year fewer scientists moved abroad and fewer returned, so that by 2010 most of the migrating scientists were staying abroad. This confirms the assumption that these movers without at least one German address are more likely to remain abroad than move back. Most scientists tend to migrate in the early phase (until 2003) of the observation period, and then this share drops rapidly.
We analyzed how long scientists remained abroad. We count migratory scientists by categorizing them after how many years they return to GER. It could be that one scientist moved abroad and back to GER several times within the period analyzed. We capture and evaluate only the maximum stay abroad. Also, for researchers who were resident abroad in the final year of observation, it is not clear how long the total time abroad will last. Mixing

these numbers (69% of all migrants) with those of already finished stays would distort the results. Therefore, we excluded all the researchers who were still living abroad at the end of the observation period except those that had been abroad since 2001 and thus were unlikely to return. Figure 4 shows the distribution of the period spent abroad for this set of researchers.

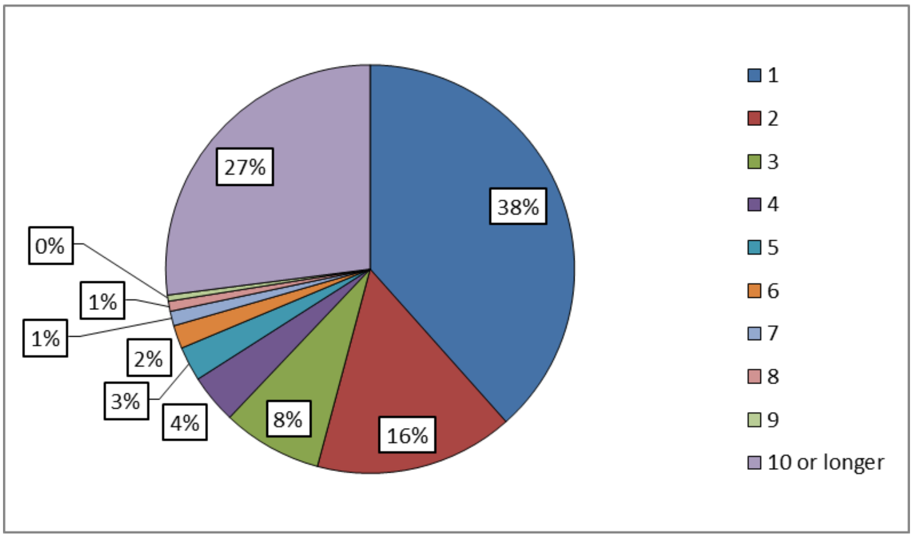

Figure 4: Duration of period spent abroad in years (residences of less than 10 years which were not finished in 2010 were excluded).

A relatively high proportion (27%) of the movers[1] in the observation period did not return to GER within 10 years and are thus unlikely to do so in the future. The other movers left their German address only temporarily and mostly returned after one (38%) or two years (16%). The probability of returning diminishes for longer periods abroad, but also because the possibility of staying for 9 years during a 10 year observation period is smaller than for shorter periods. However, merely 9% of the researchers from our initial set had not finished their longest stay abroad in 2010. To summarize, we can conclude that the actual brain drain accounts for approximately 30% of the researchers who migrated (9% of the data set) and who are unlikely to return to GER.

H5: The most popular countries of migration are the US, the UK and Switzerland.

The question arises as to which countries are particularly attractive to German migratory scientists. Data were selected for each year of observation for which a migration from GER took place. Figure 5 shows the top 10 most popular migration countries for German scientists in the period between 2000 and 2010.

[1] Movers are scientists who published 1999/98 and 2000 in GER and then moved to other countries

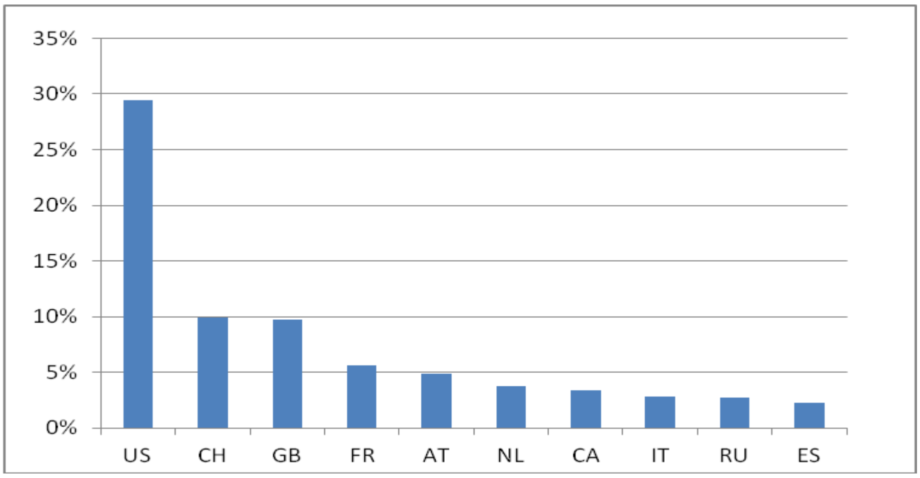

Figure 5: Top 10 migration countries.

The United States is clearly the most popular country, because it is the target of 29% of movements. The values for the UK and Switzerland are similar (around 10 %), but with a large gap to the US. France, Austria, the Netherlands, Canada, Italy, Russia and Spain are on approximately the same level (between 2% and 6% of movements) with the ranks 4 to 10.
More scientists remain abroad than return after moving from GER (Figure 3). Now the countries from which the scientists return are analyzed. Analogous to the above procedure, the data were selected so that returnees can be mapped.
In 2001, the first migration was observed and thus the first returnees were due in 2002. Figure 6 shows the proportions of returnees from the respective countries. The top 10 countries account for approximately three fourths of the returnees each year. Rates of returnees from the remaining countries reach a maximum of 3% per year and country.
Most German scientists per year, up to 38% in 2004, returned from the US. This seems plausible, since the US is also the country to which most German scientists moved. Especially in 2003, there was a peak in the returns from almost all countries, particularly from the UK and Russia. This peak could be explained by the large number of movers in 2001. For Switzerland, there seems to be a cycle of returnees every two years, which is relatively constant. The US, which ranked first at the beginning of our analysis, is now gradually losing shares, while Switzerland and the UK, which were the second and third most popular choice of German emigrants according to Figure 5, increase their shares over time.

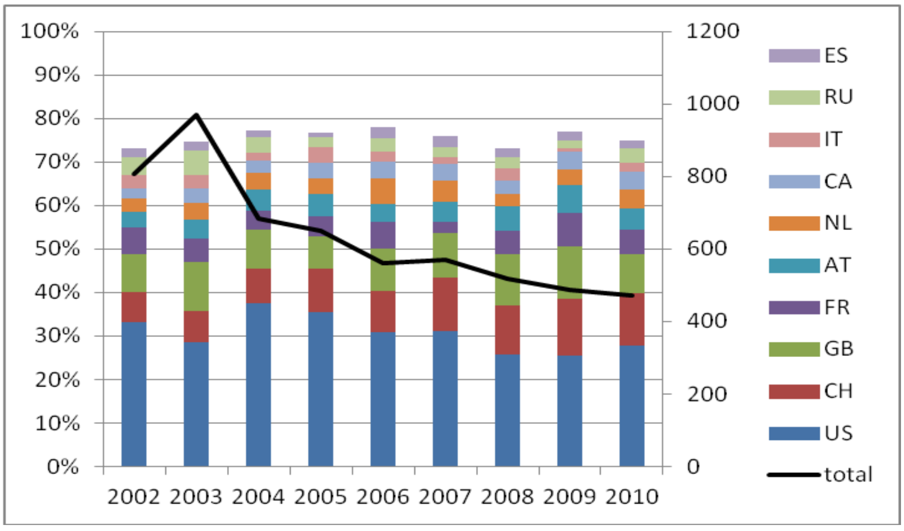

Figure 6: Top 10 countries from which German scientists returned.

Figure 7 shows the proportion of returnees in relation to the total number of migrants to the top 10 migration countries. As the most popular migration destination turns out to be Italy, the share of migrants returning is relatively small at 37%, closely followed by Spain (38%). In contrast, most scientists come back from Russia (55%). The other countries share the middle field.

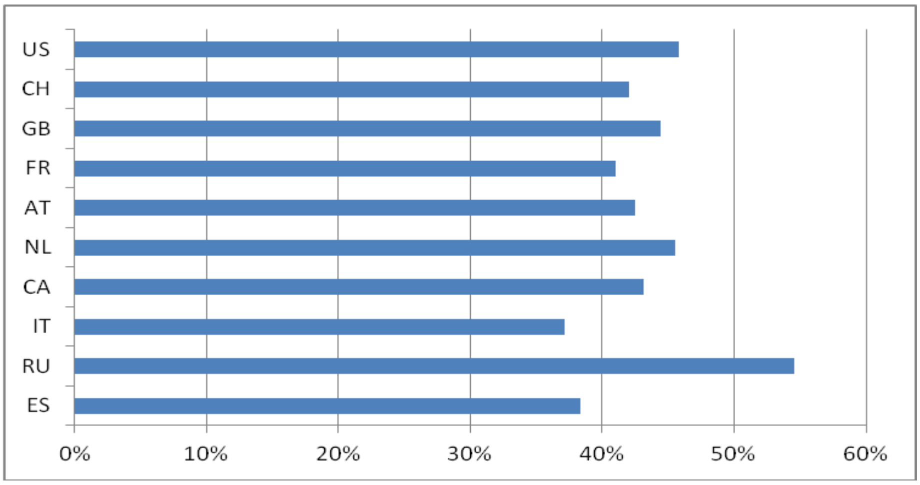

Figure 7: Rate of return of migratory scientists (top 10 countries of migration)

At the current stage of research, it is difficult to say whether the migrants to those countries with exceptionally high return rates – i.e. US, CH, UK, RU - have more incentives for returning or planned shorter stays abroad from the outset. This matter will be investigated in an online survey.

Discussion and Conclusions

As the results show on average 17% of German scientists migrate according to Scopus data. To detect whether a brain drain is taking place, it was necessary to look at those scientists who give up their German addresses and never come back; these amounted to 9% of all migrations between 2000 and 2010. We cannot be sure that all these scientists are part of a brain drain movement, however, because there are still other factors to be considered. On the one hand, it can only be assumed how many scientists leave the country "forever" based on a finite data set. However, an online survey of the respective authors was able to corroborate our finding that a return is very unlikely for at least 27%. On the other hand, there seem to be split and/or merge author errors in the Scopus data, but their extent is unclear as we have only been able to test a small random sample manually so far. To obtain representative data about the extent of errors, we would have to enlarge the random sample. In particular, our data suggests there are very few matches between the CVs and the Scopus data in 2001 and 2002, which should be investigated in more detail.

The results also show that a large number of scientists give up their home address for one year or two years and publish using a foreign address even during only short periods abroad.

Remarkably, scientists who published with a foreign address published a lot more than those with a German address. This could mean that scientists who move abroad are stimulated to publish more than those who stay in Germany, but it could also be that scientists with a higher publication output have a greater chance of finding a position abroad. In addition, it is not possible to exclude an effect of the merge author error here.

Especially in the research field of physics and astronomy the share of migrating scientists is much higher than the share of scientists remaining in GER.

The favorite migration countries, the US in first place followed by Switzerland and the UK, are also the top countries from which German scientists returned. A significant proportion of scientists moved repeatedly during our observation period. There seems to be a two-year rhythm of migration and return, which needs to be explored in more detail. Mapping migration destinations could shed new light on the motivations of migrants.

In our next project, we plan to use the author data from Scopus and the included e-mail addresses as the starting point for an online survey. The aim of the survey among migrants is to examine the motivations of scientists and especially which specific institutions are desirable targets for researchers. At the same time, the Scopus data can be validated on a larger scale using the answers in the survey.

To compare our brain drain results for Germany with other countries, we plan to expand the study to more countries like the US, Austria, the UK and South Africa.

A regression analysis of publication and citation data could provide evidence of a causal relationship between the migrating and the permanent scientists of a country.

Furthermore, it would be possible to measure the research success of migratory and non-migratory scientists based on bibliometric indicators.

References

Allmendinger, J.; Eickmeier, A. (2003): Brain Drain. Ursachen für die Auswanderung akademischer Leistungseliten in die USA. Beiträge zur Hoschschulforschung, Heft 2, 25, p. 26-35.

Backhaus, B.; Ninke, L. ; Over, A. (2002): Brain Drain – Brain Gain – Eine Untersuchung über internationale berufskarrieren. Stifterverband für deutsche Wissenschaft

Büchtemann, C. F. (2001): Deutsche Nachwuchswissenschaftler in den USA. Perspektiven der Hochschul- und Wissenschaftspolitik, Bonn.

Grubel, H.G. (1994): "Brain Drain, Economics of", in T. Husen and T. Neville Postlethwaite (eds.), The International Encyclopedia of Education, Vol. I, Oxford, pp. 554-561.

Hunger, U. (2003): Vom Brain Drain zum Brain Gain. Die Auswirkungen der Migration von Hochqualifizierten auf Abgabe- und Aufnahmeländer. Bonn : Wirtschafts- und sozialpolitisches Forschungs- und Beratungszentrum der Friedrich-Ebert-Stiftung, Abteilung Arbeit und Sozialpolitik, http://www.fes.de/aktuell/focus_interkulturelles/focus_1/documents/17_000.pdf

Janson, K.; Schomburg, H.; Teichler, U. (2006): Wissenschaftliche Wege zur Professor oder ins Abseits? Strukturinformationen zu Arbeitsmarkt und Beschäftigung an Hochschulen in Deutschland und den USA. Kassel: INCHER

Kahn, M.J, Blankley, W., Reddy, V., Pogue, T., Maharajh, R. (2004): Flight of the Flamingos: A study on the Mobility of R&D Workers. Pretoria: HSRC.

Lowell, B.L. (2002): Skilled Labour Migration from Developing Countries: Annotated bibliography. International Migration Papers, 56.

Moed, H.F., Aisati, M., Plume, A. (2013): Studying scientific migration in Scopus. Scientometrics. March 2013, Volume 94, Issue 3, pp 929-942.

OECD: International Mobility of the Highly Skilled. Paris: OECD 2001. http://www.oecd.org/science/innovationinsciencetechnologyandindustry/1950028.pdf

Roberge, G. ; Campbell D. (2012) Canadian Researchers Migration Analysis based on Scopus Author IDs. STI 2012 - 17th International Conference on Science and Technology Indicators. Montréal. p. 884-885. http://sticonference.org/Proceedings/vol2/Roberge_Canadian_884.pdf