

Extended Abstract

Theoretical focus

In the current era of knowledge-based economies, highly skilled people are among the most mobile population groups worldwide (Schiller & Cordes, 2016). In OECD countries, highly qualified individuals, including researchers, make up one-third of the immigrant population (Schiller & Cordes, 2016). The international mobility of researchers facilitates the exchange of knowledge, ideas, and skills, and thus contributes to the dynamic development of the global knowledge production system (Bauder, 2015; Franzoni et al., 2015; Netz et al., 2020). As Germany is a science powerhouse that hosts more non-EU researchers than any other EU country (Guthrie et al., 2017; Aman, 2018), Germany should be able to attract and retain international researchers with high citation performance. A lack of fine-grained data on the international migration of German-affiliated researchers and their citation performance makes it difficult to understand the in- and out-flow patterns associated with this kind of mobility. This study aims to utilize over 8 million individual-level publication data from Scopus to document and analyze international migration to and from Germany among researchers over the past 24 years. The cleaned and pre-processed bibliometric data provides a unique perspective on the international migration patterns and geographical trajectories of mobile researchers. The results help to clarify the position of Germany in the global science system. In addition, the analysis evaluates the interplay between migration and citation performance across different disciplines. This analysis adds a demographic dimension to the science of science literature by providing several in-depth statistics related to academic mobility, disaggregated by gender.

Data

This study relies on large-scale bibliometric data from over 8 million Scopus publications. This includes an exhaustive set of over a million researchers who have published with an affiliation address in Germany at some point during the 1996-2020 period in Scopus-indexed publication venues. The idea of using the historical records of researchers to follow their geographical movements can be traced back to a study by Rosenfeld and Jones (1987) on the movements of psychologists in the US. The digital revolution and the advent of digitized sources of bibliometric information enables us to expand this simple idea to cover a large number of data points with a flexible level of granularity that is suitable to our research objectives. Previous studies that used similar applications of bibliometric data have demonstrated that it is possible to map academic mobility among countries, and have shed light on the causes and consequences of academic mobility (Moed & Halevi, 2014). Recent methodological innovations for re-purposing bibliometric data to study internal migration within country boundaries have made the process of inferring migration from affiliation addresses more reliable (Miranda-González et al., 2020). The enriched records of individual scientific publications allow us to analyze the mobility of researchers to and from Germany over the last 24 years. Furthermore, the data can be linked with the comprehensive citation information and disciplines of publication venues offered by Scopus to evaluate and compare the academic performance of mobile researchers and non-movers in all fields of scholarship.

Research methods

This study firstly pre-processes the raw bibliometric data to ensure that the analysis of mobility events based on the changes in affiliation addresses are reliable. The pre-processing steps involve (1) an unsupervised machine learning algorithm for disambiguating authors, (2) a neural network algorithm for handling missing values, (3) a statistical method for detecting the main disciplines of researchers, and (4) a method for systematically inferring the genders of researchers from their first names (Figure 1).

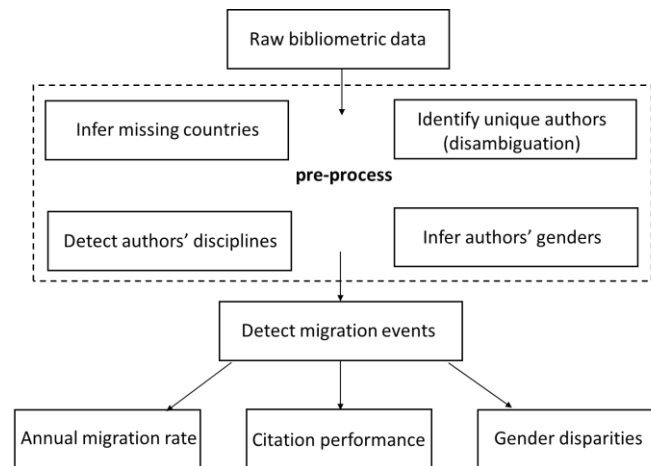


Figure 1. Workflow for pre-processing and analyzing bibliometric data for studying the migration of researchers

The international mobility of researchers is determined by identifying the changes in the affiliation addresses of authors across different publications over time. To detect migration events more reliably, the most frequent (mode) country(ies) of affiliation is extracted for each researcher in each year. A migration event is considered to have happened only if the mode country of affiliation changes for the researcher across different years such that the previous mode country disappears. By aggregating the movements for each pair of origin and destination countries, we can estimate the international migration flows among researchers. Accordingly, the country of academic origin is defined as the mode country during the first year of publishing. Similarly, the country of destination is defined as the mode country of the most recent year of publishing. Based on their migration events or lack thereof, and on their academic origins and destinations, researchers are assigned to one of the following six categories (mobility types) from the perspective of the German science system:

- (1) Single-paper author (having only one publication);
- (2) Non-mover (having multiple publications and Germany as the only mode country);
- (3) Immigrant (origin: not Germany; destination: Germany);
- (4) Emigrant (origin: Germany; destination: not Germany);
- (5) Returnee (origin: Germany; destination: Germany, with international mobility); and
- (6) Transient (origin: not Germany; destination: not Germany, but with Germany being among the researcher's mode countries at some point).

The annual net migration rate (NMR) is calculated based on the difference between incoming and outgoing flows in a year divided by the population of scholars engaged in active research in that year, and

then expressed as a per-thousand rate. The size of the population of scholars conducting active research in a given year is based on the number of researchers who have listed Germany as their country of affiliation on publications within a two-year vicinity of that year. A positive NMR value indicates that more researchers have entered than those who have left the country under analysis in the respective year.

We define the *academic age* of researchers as the number of years since their first publication (as of 2020). We calculate the average annual citation rate for each researcher by dividing their total number of citations (as of April 2020) by their academic age. This measure allows us to compare the citation performance of researchers who have different levels of experience, but are in the same field. When comparing the citations of migrant researchers from different fields, we also apply a second normalization by dividing the annual citation rate by the average annual citation rate of migrant researchers in that field.

Findings

By examining the publications of German-affiliated researchers during the 1996-2020 period, we identify the migration researchers of different migration types and calculate the annual net migration rates of researchers for Germany. We also provide descriptive results on common countries of origin and destination. Furthermore, the discipline-normalized citation performance of the researchers is evaluated and compared for the common origin and destination countries. Finally, we provide detailed results on the gender ratios of researchers by discipline, and compare the gender ratios among mobile researchers and all researchers.

For the past 24 years, we find that internationally mobile researchers accounted for nearly 14% of the population of published researchers with ties to Germany. Among them, 3.69% are immigrants, 4.54% are emigrants, 2.67% and 3.01% are returnees and transients respectively. It shows a shifting and mostly negative net migration rate for nearly the entire two decades, which suggests that Germany has gone beyond reciprocating these levels of mobility, with higher numbers of researchers emigrating from Germany to other countries. By differentiating the population of these mobile researchers according to their countries of origin and destination, we demonstrate that the US, the UK, and Switzerland are the three largest origin countries of researchers immigrating to Germany. While these three countries are the primary destinations and accounted for 58% of all scholarly migration in this period, they are not necessarily the most significant ones in terms of their scientific impact. Researchers who emigrate to Denmark, Sweden, and Austria and researchers who immigrate from South Korea and the Netherlands are more likely to outperform in citation impact than their counterparts with ties to these three countries.

We also examine the gender ratios among the migrant researchers of each discipline, and then compare them to the corresponding ratios among all German-affiliated researchers. We observe that the gender ratios (males to females) are nearly equal in disciplines such as veterinary sciences and psychology, whereas in fields such as engineering, physics, and astronomy, there are six to eight times more men than women. The most striking finding is that in disciplines that have a gender ratio above four among all researchers, the gender disparity is less severe among the migrant researchers. In contrast, for most disciplines in the life sciences, the male-to-female ratio is more balanced than it is in the physical sciences, and it is slightly higher among migrant researchers.

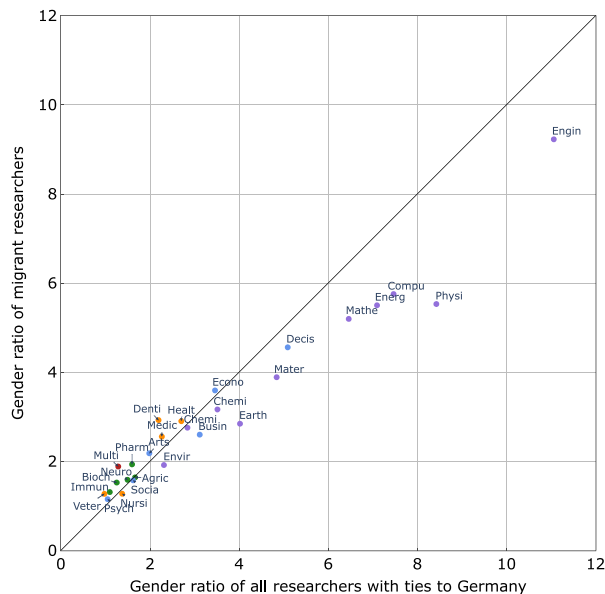


Figure 2. Male-to-female gender ratios of migrants and all researchers

Building on this work, we intend to gain further insights into the relationship between migration and scientific performance from a micro-level perspective. This approach will be particularly useful for studying how women are affected, especially in fields where they are severely under-represented, and for examining the impact of specific national policies in Germany that aim to improve the gender balance in academic fields. Ultimately, by understanding the role of migration both into and out of Germany in each field and discipline, we can provide additional insights into researchers' characteristics, mobility behavior, and performance, and thus represent a key step toward gaining a better understanding of migration among scholars.

Reference

- Aman, V. (2018). Does the Scopus author ID suffice to track scientific inter- national mobility? A case study based on Leibniz laureates. *Scientometrics*, 117(2), 705-720.
- Bauder, H. (2015). The international mobility of academics: A labour market perspective. *International Migration*, 53(1), 83-96.
- Franzoni, C., Scellato, G., & Stephan, P. (2015). Chapter 2 - International mobility of research scientists: Lessons from GlobSci. In Geuna, A. (Eds), *Global Mobility of Research Scientists* (pp. 35-65). Academic Press.
- Guthrie, S., Lichten, C. A., Harte, E., Parks, S., & Wooding, S. (2017). *International mobility of researchers: A survey of researchers in the UK*. RAND Corporation.
- Miranda-González, A., Aref, S., Theile, T., & Zagheni, E. (2020). Scholarly migration within Mexico: Analyzing internal migration among researchers using Scopus longitudinal bibliometric data. *EPJ Data Science*. 9(1), 1-26
- Moed, H. F. & Halevi, G. (2014). A bibliometric approach to tracking inter- national scientific migration. *Scientometrics*, 101(3), 1987-2001.
- Netz, N., Hampel, S., & Aman, V. (2020). What effects does international mobility have on scientists' careers? a systematic review. *Research Evaluation*, 29(3), 327-351.
- Rosenfeld, R. A. & Jones, J. A. (1987). Patterns and effects of geographic mobility for academic women and men. *The Journal of Higher Education*, 58(5), 493-515.
- Schiller, D. & Cordes, A. (2016). Measuring researcher mobility. In *OECD Blue Sky Forum*, pp. 24.