RESEARCH

Migration and collaboration of scholars worldwide: A word embedding representation

Erin Walk[1*] and Aliakbar Akbaritabar[2]

*Correspondence: ewalk@mit.edu
[1]Massachusetts Institute of Technology (MIT) Social and Engineering Systems, 77 Massachusetts Avenue, Cambridge, USA
Full list of author information is available at the end of the article

Abstract

Motivation: Migration of academics is a core focus of the recent literature. Factors affecting academics' decision to migrate, such as scholarly collaboration, are understudied and those studies which do consider scholarly migration and collaboration in tandem report paradoxical findings. We take a two pronged approach, considering both how spatial representations of mobility and collaboration networks compare and how collaboration patterns of mobile authors relate to their movement between institutions.

Data: We selected a random sample of authors worldwide from Scopus 2020 data based on number of publications, corresponding authorship, publication in top-ranked journals, and mobility status to construct control/observation groups.

Results: We find that collaboration and mobility trajectories are highly similar across groups, though embedding representations of collaboration are more densely packed than those for mobility. Authors who are mobile or talents (top 1% based on our selection criteria) are more likely to have a high number of collaborators. Furthermore, though few authors collaborate only before or after a movement event, collaboration increases leading up to an affiliation change and the majority of publications with a target institution are published prior to the initial move there.

Conclusions: We use a novel spatial approach to model mobility and collaboration trajectories, enabling us to compare ways in which these phenomena differ. Then, we look at author patterns of collaboration pre- and post- mobility events. Our methodological framework opens up promising avenues for future research on individual level forecasting of scholarly migration and on global dynamics of academic talent circulation.

Keywords: Scholarly migration; Scientific collaboration; Bibliometric data; Word embedding vectors

Introduction

There is a global competition to attract the highly-skilled and talented [1, 2], as they are considered innovation powerhouses [3, 4]. Academics as a subset of the highly-skilled population are highly mobile, even called *globetrotters* [5, 6], making their mobility experience the focus of much recent literature [e.g., 7–14]. Migrant academics contribute both to the innovation in home [15, 16] and host [17] countries. Modeling past trajectories of mobility and factors affecting it enables both an explanation and speculation for forecasting future mobility events [18].

Different factors affect the decision to emigrate [19]. In the case of academics, one important and influential factor is scientific collaboration [20]. Through collaboration scholars share their knowledge, complement their skills, and network, connecting them wi opportunities to move and further their scientific career. Scholars can form collaboration ties *before*, or *during* their mobility experience [21]. Some of these formed ties *persist* [22], even after the mobility event. Nevertheless, the sequence of events does not always follow a defined order of mobility and collaboration. A theoretical framework considering the effect of *network tie formation* [23] in migration would help in identifying the influence of collaboration ties on scholarly mobility. Further, an intertwined study of scholarly migration and collaboration is necessary to disentangle this sequence of events, but has been rarely done [20].

Most of the previous works, while proposing innovative representations of mobility, lack an integrated investigation of "mobility" with "collaboration". For instance, Block et al. [24] use extensions to Exponential Random Graph Models (ERGMs) to model mobility based on *social selection* —one's socioeconomic status and attributes affect the mobility decision and trajectory in the form of self-selection to migrate— versus *social influence* —having other mobile people in one's network influences an individual's mobility decision and trajectory, i.e., to be more mobile— and they simultaneously test these competing hypotheses on social selection and influence. But, this innovative work, which also considers immobile individuals, does not speculate on "why" someone moves and others do not. It does not consider the effects of collaboration on mobility. Another possible representation of mobility in conjunction with collaboration has been proposed by Boekhout et al. [21]. They use affiliation as an attribute of each author to check what proportion of an author's current collaborators are in the same institute or country versus abroad.

When two former collaborators from within a country i.e., a national tie, since both authors affiliate to institutions inside a country, continue to collaborate as an international tie between two countries, they infer that a mobility event must have happened for one of the authors. Nevertheless, this innovative setup only "infers" mobility without identifying who has actually moved. This type of setup is prone to ignore an individual's specific mobility trajectory —which is dependent on their attributes, e.g., gender, scientific discipline, country of origin etc.— and mix it with their collaborators' mobility trajectory. It also ignores the direction of the migration if it is an immigration, emigration or return move. Another approach consisted of aggregating individual author's mobility trajectories to "institution", "city", "region" and "country" level, representing these aggregated entities as nodes in a network and recording a mobility event when authors move between aggregated nodes [e.g., 11, 12, 25–27]. This type of aggregation neglects the effect of an author's attributes, trajectory and decisions (e.g., former/future collaborations) in shaping mobility patterns and does not consider immobility. Immobility and the rationale behind it is important to indicate if lack of mobility is strategic or involuntary [28]. In addition, because some scholars form international collaboration ties [29, 30] or host other mobile scholars [31] in order to circumvent immobility's disadvantages [32] and partially prevent a loss of mobility's advantages e.g., in productivity or citations, since mobile scholars are more productive and receive higher citations [7, 8].

Once mobility and collaboration are investigated alongside each other, previous studies find paradoxical results regarding the direction of the effect between the scholarly collaboration and migration. Some have advocated for an effect of mobility on collaboration [33], while others have found a bidirectional [21, 34] or an inconclusive [35] effect between the two. Hence, the direction of the effect and causation is still unclear. Adopting a spatial approach by representing the trajectory of mobility and collaboration using word embedding vectors [36] might help to resolve this paradox by enabling the consideration of multiple vectors, e.g., for mobility, collaboration, and similar other factors such as geographic distance or linguistic and cultural similarities, in a unified framework to compare the communities formed through each process. Though Murray et al. [36] consider how affiliation trajectories reflect geographical and language differences between countries, we additionally

consider how collaboration and affiliation trajectories reflect differences in mobility and collaboration patterns between organizations and authors. Specifically, we compare how an author's collaboration and mobility trajectories are similar or different, as well as in what countries are more insular in terms of collaboration or movement. Following this, we model collaboration over time with a target institution to assess how a mobility event impacts collaboration.

To address the described lack of comprehensive intertwined studies on scholarly collaboration and migration, and to determine the direction of causation between the two, we chose a random sample of more than 10 thousand Scopus authors worldwide. We used bibliometric criteria, that were extensively validated in previous research [37], on a) productivity, in terms of the number of publications, b) having a principal role in the publication through being the corresponding author, and c) publishing in top-ranked Q1 journals to select the sample of authors. Using these three criteria, we identify a group of *potential talents*, who are at the top 1% in these criteria, as our observation group. We identify also a control group of authors who are at the top 5-10% of the selection criteria. In addition, and to differentiate between mobile and non-mobile scholars in our investigation, we considered two types of mobility [11], namely, internal mobility inside one country or between sub-national regions, versus international mobility between multiple countries. Furthermore, we considered a group of authors who had never experienced mobility, i.e., immobile scholars. These groups formed our observation and control sets. Haunschild and Bornmann [37] have shown that these criteria are reliable in identifying a group of talents who tend to be more successful in their scientific career. Our study carries out a comparison of these talents with a control group in terms of collaboration and mobility and contributes to the literature by resolving the mobility-collaboration paradox outlined above. In the following sections, since we identify a mobility event based on a change in affiliation addresses [13, 14], we use "affiliation" and "mobility"/"migration" trajectory interchangeably.

Results

Fig. 1 shows the geographical distribution of the chosen 10,963 authors worldwide. Although the number of selected authors for the observation group (3,564 authors, top panel in Fig. 1) was smaller than the number chosen for the control group (7,399

authors, bottom panel), the geographical distribution of these authors is consistent with only a few exceptions where a country is only represented in one of the two groups. Table S1 in the appendix presents more detail on the bibliometric criteria used [37] and the chosen number of authors. These criteria allowed us to identify a group of **potential talents** among authors, top 1%, and a control group of top 5-10%. Based on the mobility status [11], we chose 2k immobile authors, 3k with internal, 3k with international and 2,963 authors who had experienced both types of mobility.

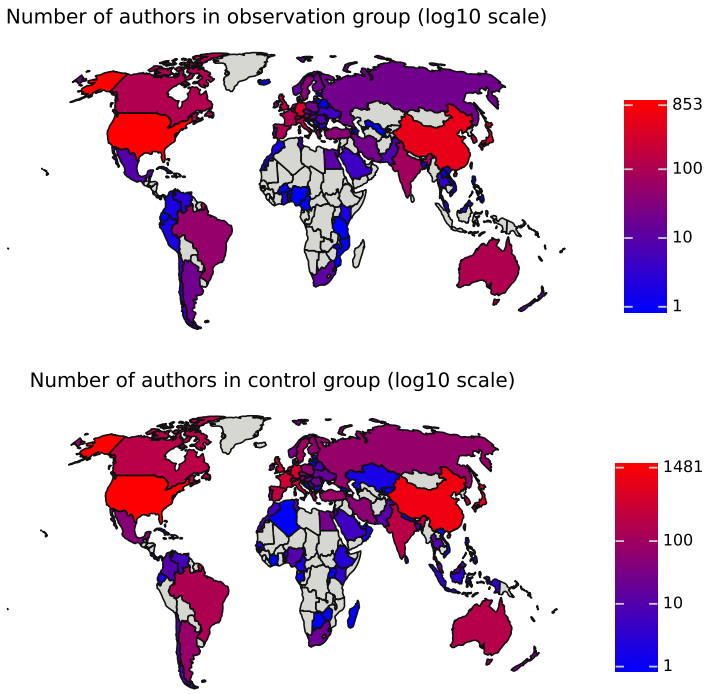


Figure 1: **Geographical distribution of authors in observation (top) and control (bottom) groups.** Selected authors are affiliated with similar countries worldwide (in the case of most countries, and especially in larger well-established science systems) with a few countries missing from either observation or control group (e.g., see the cases of Peru, Algeria, Kazakhstan and Indonesia which are only present in the control group).

In Figures 2 and 3 we show the clustering of countries based upon the mobility and collaboration between individuals at organizations within those countries. In the affiliation heat map which is based on scholars' mobility trajectories, Figure 2, we

note that the Nordic countries have their own cluster, and are closer to the cluster for Western Europe and that for The UK, Ireland, Australia and New Zealand, Israel. The second large grouping, as seen in the dendrogram (on the left), includes the clusters for eastern and south eastern Asian countries, Middle East and North Africa as well as the Americas and Eastern Europe. When we consider collaboration, Figure 3, the clusters become far denser as indicated by larger swatches of dark squares within the heat map. This is especially true for Europe, where individual clusters become far less clear, indicating a high level of collaboration amongst institutions in these countries, confirming [29]'s findings, which indicate the emergence of an "integrated European Research Area" [38, 39].

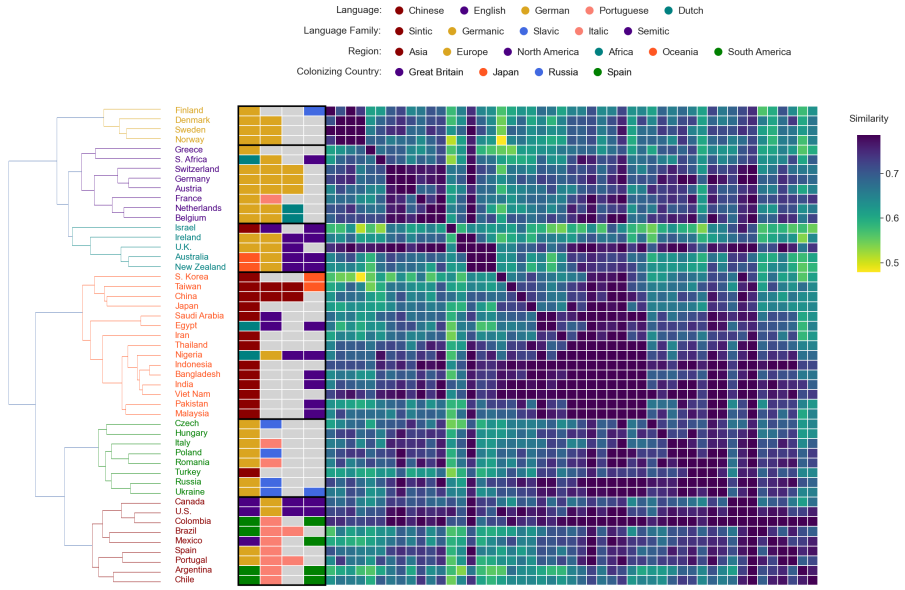


Figure 2: Clustering based on affiliation trajectories of individuals and mobility between institutions in the given countries. Colors in the heatmap represent the cosine similarity between country vectors. Boxes on the left indicate, from left to right, region, language family, language, and colonizing country.

Collaboration, in general, seems to be more regionally driven than affiliation changes and mobility, and we find empirically that high-level clustering for both mobility and collaboration is related to region and the colonizing country, while at the lower levels, it is language based. A graphical representation of this is available in Appendix Figure S2. In addition to the heat maps shown here, Appendix Figure S1 presents the clustering with each organization as a single point, showing that

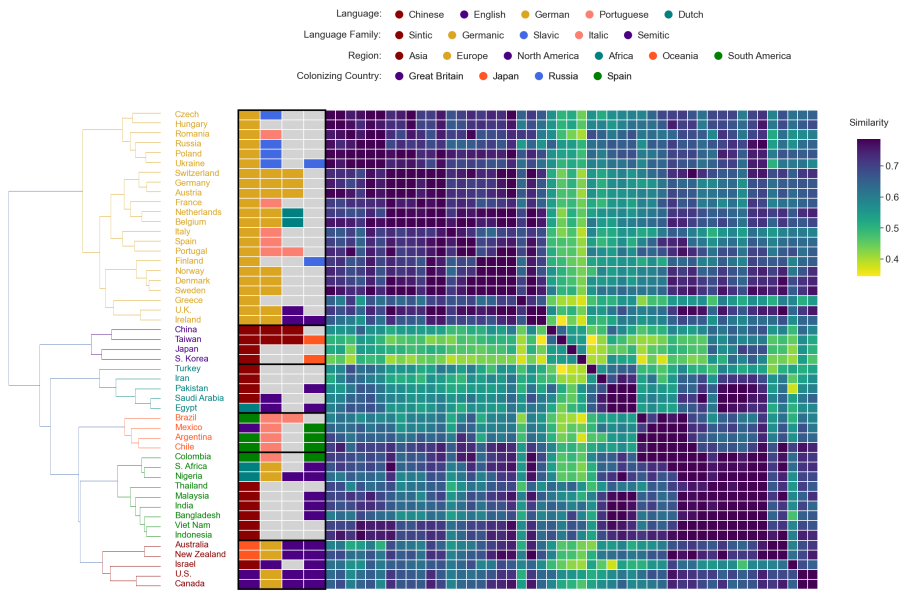


Figure 3: Clustering based on collaboration trajectories of individuals between institutions in the given countries. Colors in the heatmap represent the cosine similarity between country vectors. Boxes on the left indicate, from left to right, region, language family, language, and colonizing country.

most organizations group into country specific clusters, a phenomenon noted by [36]. Pairwise comparisons of organizations can be somewhat difficult to interpret due to the variety of institution types present. However, we note that the vast majority of organizations with a large difference between collaboration and mobility, as defined by a 0.5 or greater difference in vector similarities, have high levels of collaboration and low levels of mobility between them.

Table 1: Mean similarity between mobility and collaboration trajectories for authors in the dataset separated by movement status, top, and talent status, bottom, as well as the average number of institutions collaborated with.

Group	Similarity	Avg. # Collabs
non-mobile (2k)	0.93	39.32
mobile internal (3k)	0.94	99.00
mobile international (3k)	0.92	129.12
mobile both (2,963)	0.94	121.00
Control group	0.94	62.71
Potential talents	0.92	184.50

In order to understand the connection between collaboration and mobility, we look at the similarities between these trajectories for each author. The mean similarity value for each group is in Table 1, and is very high for all of the groups. Though this does not imply any directionality, it does show that collaboration and mobility are highly associated, and that two organizations directly next to each other in a collaboration trajectory are likely also co-located in the affiliations and mobility trajectory. The mean number of collaborations is far higher in all of the mobile groups than it is in the non-mobile group. Authors who have moved between countries, as indicated by *mobile both* and *mobile international*, also have more collaborations than their internal counterparts. International movement especially has a close association with collaborations. In addition, the Potential talents have far more collaborations than the Control group and their collaboration and mobility trajectories are slightly less similar.

The spatial representation of collaboration and mobility provides information about how close two organizations are, but does not provide a time component because word embeddings are not context dependent. In order to add these considerations to our analysis, we analyze the collaboration behavior of authors before and after moving to a target institution. We find that the majority of collaborative publications are published prior to moving to a target institution, as can be seen in Table 2. This is true for all groups. In addition, the number of moves and the length of the moves remain similar. One possible confounder is that we only explore the first move in each author-institution pair, so it may be that certain authors are more likely to return in the future to an institution which they have already been affiliated with. The average year for first collaborative publication with the target institution is also further from the movement than the last publication afterwards, suggesting that collaboration is more beneficial as a way to create ties with an institution than as a way to maintain ties after leaving.

Table 2: Summary statistics of movement behavior by group.

Group	Number Moves	Move Length (Years)	1st Pub (Years Pre)	Last Pub (Years Post)	Pre-Move (Pub #)	Post-Move (Pub #)	At Inst. (Pub #)
All authors	3	1.5	3.2	1.7	5	2	2
Control group	3	1.7	2.9	1.5	3	1	2
Potential talents	4	1.2	3.6	1.9	7	3	2
Internal mobility	3	1.6	3.3	1.9	6	2	2
International mobility	3	1.5	2.9	1.7	6	2	2
Both	4	1.3	3.1	1.3	4	1	1

When we consider all move related publications, we also see that the vast majority are published prior to the move, as can be seen in Table 3. In addition, though the majority of authors collaborate both before and after their affiliation change, it is much more common to collaborate *only before* and at the institution than it is to collaborate only at the institution and after leaving. Table 4 shows that nearly 40% of moves are only preceded by collaboration, while between 10 and 20% of moves, depending on the group, are only followed by it.

Table 3: All publications by mobile authors in which they collaborate with a target institution and the percentage which occur before, during, and after affiliation with the institution.

Group	Total Authors	Move Related Pubs	Pre Move	Post Move	At Inst
All	8,836	54,782	41%	28%	31%
Control group	5,310	28,663	41%	25%	35%
Potential talents	3,526	26,119	42%	31%	27%

Finally, the outcomes from our regression analysis indicate that collaboration with the target institution increases leading up to the move, then starts to decrease following the move as in Figure 4. The increase directly post-move may be related to the individual still being at the target institution or the delay between starting a paper and publication [40]. Furthermore, there is much higher variance in publication post-move than pre-move, likely because there are also fewer datapoints. These results indicate that though in many cases collaboration with the target institution both precedes and follows affiliation changes, it is more common for collaboration to precede movement behavior, rather than the other way around.

Table 4: All moves by mobile authors and percentage where collaboration occurred only before (and during) or only after (and during) the affiliation change.

Group	Moves	Only Pre	Only Post
All	34,064	39%	17%
Control group	18,154	40%	10%
Potential talents	15,910	37%	18%

Discussion

In this study, we implemented a comprehensive intertwined framework to consider scholarly collaboration and migration simultaneously. Our results indicate that col-

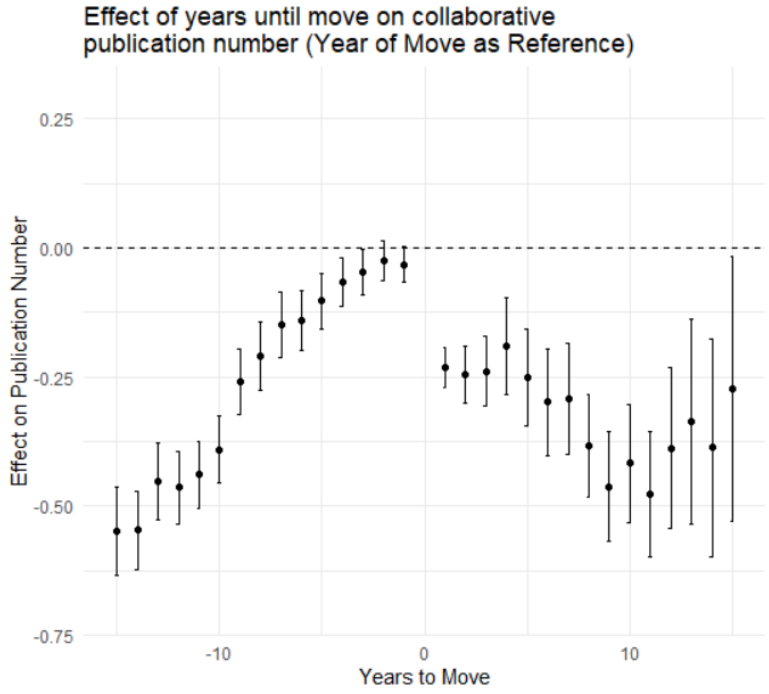


Figure 4: Effect of years to move on publication number in collaboration with target institution. Reference level is number of collaborative publications in year 0 (year of move). Intercept is 1.2. Standard Errors (SEs) are clustered at the author level (see Table S2 for regression coefficients).

laboration and mobility are highly connected. Likely due to the lower barrier to entry for collaboration versus moving between institutions, when considered spatially the clusters in the affiliations embedding are far more spread and distinct than in the collaborations embedding. In addition, alliances such as the European Union and defense agreements impact collaboration by creating ties between countries which may be geographically far apart. Overall, we find that in cases where collaboration and affiliation are not highly similar between two organizations collaboration is the closer tie, as supported by both visualization and pairwise organization comparisons. This could signal a phenomenon reported in literature as *helicopter research* [41–43], where researchers from the Global North countries collaborate with those from the Global South, but the most benefits and credit are attributed to Global North [44] and the mobility of scholars between these institutions are not both-ways, perhaps due to prestige unbalance [27], which needs further probes in our sample.

In terms of our understanding of the relationship between collaboration and mobility, we find that for all authors collaboration and affiliation trajectories are highly similar. This is in some ways surprising, as one might expect far more collaborations than affiliations, due to costs associated with mobility [32, 45, 46]. In addition, both authors who move between institutions and those deemed 'Potential talents' [37] in our dataset have a far larger number of collaborators. This indicates a symbiotic relationship between collaboration and changing institutions, though directionality remains unclear.

Moving from the spatial results to directional results, we find that potential talents and those with internal and international mobility publish more papers in collaboration with an institution prior to a move versus after. In general, collaboration before a mobility event is more common than collaboration after, or even at the institution, with close to half of papers in collaboration with a target institution being published prior to the move. In keeping with this, collaborative papers increase in number leading up to a mobility event, and then decrease following the event. Future work may explore the effect of organization type (e.g., education, healthcare, government, NGOs, etc.) and scientific discipline of the authors on the observed results.

Our results have certain implications for global academic talent circulation and forecasting scholars' migration. Here we have shown that scholars have a higher tendency to form collaboration ties with an institution before they move there, and publication with the target institution increases leading up to the mobility event. Using these results would make it possible to identify mobility tendencies and forecast where academic talents are going to move next. This can inform policies on brain circulation worldwide.

Materials and methods

We selected and used a random sample of 10,963 authors from a 2020 snapshot of Elsevier's Scopus data provided to us by the German Competence Network for Bibliometrics [47] through Max Planck Digital Library (MPDL). These authors were chosen based on multiple criteria proposed and extensively validated by Haunschild and Bornmann [37] (N. of publications (o), N. as corresponding author (c), N. in Q1 (top ranked) journals (q1)) to construct a *control* and *observation* group. The

observation group includes 3,564 authors which are considered potential talents based on the used criteria and the control group includes 7,399 authors. In our selection process, we also considered the mobility status of these authors in terms of internal mobility (between sub-national regions of one country), international mobility between multiple countries, both types of mobility, and immobility. Table S1 in the appendix presents more detail on the selected random sample of authors.

We modeled the distance between organizations using word2vec word embeddings [48], where each "sentence" was an individual author's affiliation (mobility) or collaboration trajectory. As in [36], in instances where an individual had multiple affiliations or collaborations in a given year we updated the model five times incorporating random shuffling of the organizations to prevent order from impacting our results. We used UMAP reduction from [49] with cosine similarity as the metric to visualize the results in 2D. Cosine similarity is the most common measurement for comparison between words in word2vec, and measures the difference between the vectors of the words. A high cosine similarity means that the words are either used in the same context (interchangeable) such as 'cat' and 'lion' or frequently co-located such as 'cat' and 'meow.'

In addition to visualizing the whole embedding space as described above, we also ran pairwise cosine similarity between institutions in order to consider whether institutions were spaced differently in the collaboration and affiliation models. To better understand how these choices affect authors, and as a first step to understanding the relationship between collaboration and affiliation, we computed the similarity between each author's collaboration and affiliation vectors. In order to do this, we made a vector of the features within each trajectory and divided it by the total number of features so that length did not penalize similarity. We then computed the similarity between the two vectors. Differences in author affiliation and collaboration paths were considered in regards to their overall mobility and their success.

For the time varying collaboration component we consider all moves by each author in the dataset. In cases where an author came back to an institution they had previously been affiliated with after leaving it only the first time they moved to that given institution is considered. We find all publications in which the author collaborated with the target institution, and assign each a year index based on

when the paper was published in relation to the year of the affiliation change. For example, an index of -15 indicates that it was published 15 years prior to the move, and one of 10 indicates that it was published 10 years after the move. All post-move indices are positive regardless of whether the author is still at the target institution or has moved to a new institution at the time of publication. We create a new variable, all_works, which is the sum of all publications for an author in a given year index for a given target institution. Since there are naturally years in which an author is active but either does not have publications or does not have publications with the target institution we fill all 'known active' years, or years with no entries that are between the author's first and last publications, with 0 in the all_works column. We run a regression with publication number as the outcome, the index factor as the independent variable, and covariates for author group and mobility type. Standard errors are clustered at the author level.

Competing interests

The authors declare that they have no competing interests.

Author's contributions

EW: Conceptualization, Formal Analysis, Methodology, Software, Visualization, Writing - original draft, Writing - review & editing. AA: Conceptualization, Data curation, Methodology, Visualization, Writing - original draft, Writing - review & editing.

Acknowledgements

We would like to thank Murray et. al [36] for their generosity in sharing their replication scripts.

Funding

This study has received access to the bibliometric data through the project "Kompetenzzentrum Bibliometrie" and we acknowledge their funder Bundesministerium für Bildung und Forschung (grant number 16WIK2101A). It is also supported by the NSF Graduate Research Fellowship Program (ID number 2021311265).

Availability of data and materials

The data that support the findings of this study are available from the German Competence Network for Bibliometrics [47] but restrictions apply to the availability of these data, which were used under license for the current study, and so are not publicly available and require the permission of the the German Competence Network for Bibliometrics [47]. Nevertheless, we publicly share replication materials that allow for replicating our analysis.

Author details

[1]Massachusetts Institute of Technology (MIT) Social and Engineering Systems, 77 Massachusetts Avenue, Cambridge, USA. [2]Max Planck Institute for Demographic Research (MPIDR), Konrad-Zuse Strasse 1, 18057 Rostock, Germany.

References

1. OECD: The Global Competition for Talent Mobility of the Highly Skilled: Mobility of The Highly Skilled. OECD Publishing, ??? (2008)
2. Laudel, G.: Studying the brain drain: Can bibliometric methods help? Scientometrics **57**(2), 215–237 (2003). doi:10.1023/A:1024137718393. Accessed 2021-04-29

3. Hunt, J.: Immigrant patents boost growth. Science **356**(6339), 697–697 (2017). doi:10.1126/science.aan2468. Accessed 2022-09-09
4. Kapur, D., McHale, J.: Give Us Your Best and Brightest: The Global Hunt for Talent and Its Impact On the Developing World. Center for Global Development, ??? (2005)
5. Bohannon, J.: Restless minds. Science **356**(6339), 690–692 (2017). doi:10.1126/science.356.6339.690. Accessed 2021-04-06
6. Mahroum, S.: Highly skilled globetrotters: mapping the international migration of human capital. R&D Management **30**(1), 23–32 (2000). doi:10.1111/1467-9310.00154. Accessed 2022-09-21
7. Scellato, G., Franzoni, C., Stephan, P.: A mobility boost for research. Science **356**(6339), 694–694 (2017). doi:10.1126/science.aan4052. Accessed 2021-07-22
8. Sugimoto, C.R., Robinson-Garcia, N., Murray, D.S., Yegros-Yegros, A., Costas, R., Larivière, V.: Scientists have most impact when they're free to move. Nature **550**(7674), 29–31 (2017). doi:10.1038/550029a. Accessed 2020-04-16
9. Robinson-Garcia, N., Sugimoto, C.R., Murray, D., Yegros-Yegros, A., Larivière, V., Costas, R.: The many faces of mobility: Using bibliometric data to measure the movement of scientists. Journal of Informetrics **13**(1), 50–63 (2019). doi:10.1016/j.joi.2018.11.002. Accessed 2021-07-29
10. Momeni, F., Karimi, F., Mayr, P., Peters, I., Dietze, S.: The many facets of academic mobility and its impact on scholars' career. Journal of Informetrics **16**(2), 101280 (2022). doi:10.1016/j.joi.2022.101280. Accessed 2022-06-16
11. Akbaritabar, A., Zhao, X., Zagheni, E.: Internal versus international scholarly mobility and migration worldwide, International Union for the Scientific Study of Population (IUSSP) (2021). https://ipc2021.popconf.org/abstracts/211016 Accessed 2022-04-19
12. Czaika, M., Orazbayev, S.: The globalisation of scientific mobility, 1970–2014. Applied Geography **96**, 1–10 (2018). doi:10.1016/j.apgeog.2018.04.017. Accessed 2021-07-14
13. Sanliturk, E., Zagheni, E., Dańko, M.J., Theile, T., Akbaritabar, A.: Global patterns of migration of scholars with economic development. Proceedings of the National Academy of Sciences **120**(4), 2217937120 (2023). doi:10.1073/pnas.2217937120. Accessed 2023-01-18
14. Zhao, X., Akbaritabar, A., Kashyap, R., Zagheni, E.: A gender perspective on the global migration of scholars. Proceedings of the National Academy of Sciences **120**(10), 2214664120 (2023). doi:10.1073/pnas.2214664120. Accessed 2023-02-27
15. Sbalchiero, S., Tuzzi, A.: Italian Scientists Abroad in Europe's Scientific Research Scenario: High skill migration as a resource for development in Italy. International Migration **55**(4), 171–187 (2017). doi:10.1111/imig.12340. Accessed 2021-08-24
16. Saint-Blancat, C.: Making Sense of Scientific Mobility: How Italian Scientists Look Back on Their Trajectories of Mobility in the EU. Higher Education Policy **31**(1), 37–54 (2018). doi:10.1057/s41307-017-0042-z. Accessed 2022-03-03
17. Scellato, G., Franzoni, C., Stephan, P.: Migrant scientists and international networks. Research Policy **44**(1), 108–120 (2015). doi:10.1016/j.respol.2014.07.014. Accessed 2020-08-31
18. Vaccario, G., Verginer, L., Schweitzer, F.: Reproducing scientists' mobility: a data-driven model. Scientific Reports **11**(1), 10733 (2021). doi:10.1038/s41598-021-90281-9. Accessed 2022-09-30
19. Czaika, M., Bijak, J., Prike, T.: Migration Decision-Making and its key Dimensions, 16 (2021)
20. Appelt, S., van Beuzekom, B., Galindo-Rueda, F., de Pinho, R.: Chapter 7 - Which Factors Influence the International Mobility of Research Scientists? In: Geuna, A. (ed.) Global Mobility of Research Scientists, pp. 177–213. Academic Press, San Diego (2015). doi:10.1016/B978-0-12-801396-0.00007-7. https://www.sciencedirect.com/science/article/pii/B9780128013960000077 Accessed 2022-08-12
21. Boekhout, H.D., Traag, V.A., Takes, F.W.: Investigating scientific mobility in co-authorship networks using multilayer temporal motifs. Network Science **9**(3), 354–386 (2021). doi:10.1017/nws.2021.12. Accessed 2021-11-14
22. Dahlander, L., McFarland, D.A.: Ties That Last: Tie Formation and Persistence in Research Collaborations over Time. Administrative Science Quarterly **58**(1), 69–110 (2013). doi:10.1177/0001839212474272. Accessed 2020-04-16

23. Massey, D.S., Arango, J., Hugo, G., Kouaouci, A., Pellegrino, A., Taylor, J.E.: Theories of International Migration: A Review and Appraisal. Population and Development Review **19**(3), 431 (1993). doi:10.2307/2938462. Accessed 2020-04-26
24. Block, P., Stadtfeld, C., Robins, G.: A statistical model for the analysis of mobility tables as weighted networks with an application to faculty hiring networks. Social Networks **68**, 264–278 (2022). doi:10.1016/j.socnet.2021.08.003. Accessed 2021-09-29
25. Vaccario, G., Verginer, L., Schweitzer, F.: The mobility network of scientists: analyzing temporal correlations in scientific careers. Applied Network Science **5**(1), 1–14 (2020). doi:10.1007/s41109-020-00279-x. Accessed 2021-04-09
26. Windzio, M.: The network of global migration 1990–2013: Using ERGMs to test theories of migration between countries. Social Networks **53**, 20–29 (2018). doi:10.1016/j.socnet.2017.08.006. Accessed 2021-09-27
27. Clauset, A., Arbesman, S., Larremore, D.B.: Systematic inequality and hierarchy in faculty hiring networks. Science Advances **1**(1), 1400005 (2015). doi:10.1126/sciadv.1400005. Accessed 2020-04-17
28. Carling, J.: Migration in the age of involuntary immobility: Theoretical reflections and Cape Verdean experiences. Journal of Ethnic and Migration Studies **28**(1), 5–42 (2002). doi:10.1080/13691830120103912
29. Akbaritabar, A., Barbato, G.: An internationalised Europe and regionally focused Americas: A network analysis of higher education studies. European Journal of Education **56**(2), 219–234 (2021). doi:10.1111/ejed.12446. Accessed 2022-01-29
30. Akbaritabar, A., Traag, V.A., Caimo, A., Squazzoni, F.: Italian Sociologists: a Community of Disconnected Groups. Scientometrics (2020). doi:10.1007/s11192-020-03555-w. Accessed 2020-06-14
31. Schaer, M., Jacot, C., Dahinden, J.: Transnational mobility networks and academic social capital among early-career academics: beyond common-sense assumptions. Global Networks **n/a**(n/a) (2020). doi:10.1111/glob.12304. Accessed 2021-05-28
32. Aksnes, D.W., Rørstad, K., Piro, F.N., Sivertsen, G.: Are mobile researchers more productive and cited than non-mobile researchers? A large-scale study of Norwegian scientists. Research Evaluation **22**(4), 215–223 (2013). doi:10.1093/reseval/rvt012. Accessed 2021-11-02
33. Kato, M., Ando, A.: National ties of international scientific collaboration and researcher mobility found in Nature and Science. Scientometrics **110**(2), 673–694 (2017). doi:10.1007/s11192-016-2183-z. Accessed 2021-07-26
34. Chinchilla-Rodríguez, Z., Miao, L., Murray, D., Robinson-García, N., Costas, R., Sugimoto, C.R.: A Global Comparison of Scientific Mobility and Collaboration According to National Scientific Capacities. Frontiers in Research Metrics and Analytics **3** (2018). doi:10.3389/frma.2018.00017. Accessed 2020-06-07
35. Netz, N., Hampel, S., Aman, V.: What effects does international mobility have on scientists' careers? A systematic review. Research Evaluation **29**(3), 327–351 (2020). doi:10.1093/reseval/rvaa007. Accessed 2020-08-18
36. Murray, D., Yoon, J., Kojaku, S., Costas, R., Jung, W.-S., Milojević, S., Ahn, Y.-Y.: Unsupervised embedding of trajectories captures the latent structure of mobility. arXiv:2012.02785 [physics] (2021). Accessed 2021-07-10
37. Haunschild, R., Bornmann, L.: Identification of potential young talented individuals in the natural and life sciences: A bibliometric approach. Journal of Informetrics **17**(3), 101394 (2023). doi:10.1016/j.joi.2023.101394. Accessed 2023-03-12
38. Hoenig, B.: Europe's New Scientific Elite: Social Mechanisms of Science in the European Research Area. Routledge, ??? (2017)
39. Hoekman, J., Frenken, K., Tijssen, R.J.W.: Research collaboration at a distance: Changing spatial patterns of scientific collaboration within Europe. Research Policy **39**(5), 662–673 (2010). doi:10.1016/j.respol.2010.01.012. Accessed 2020-05-10
40. Björk, B.-C., Solomon, D.: The publishing delay in scholarly peer-reviewed journals. Journal of Informetrics **7**(4), 914–923 (2013). doi:10.1016/j.joi.2013.09.001. Accessed 2022-03-02
41. Adame, F.: Meaningful collaborations can end 'helicopter research'. Nature (2021). doi:10.1038/d41586-021-01795-1. Accessed 2022-07-14
42. Haelewaters, D., Hofmann, T.A., Romero-Olivares, A.L.: Ten simple rules for Global North researchers to stop perpetuating helicopter research in the Global South. PLOS Computational Biology **17**(8), 1009277 (2021).

doi:10.1371/journal.pcbi.1009277. Accessed 2022-07-14

43. Boshoff, N.: Neo-colonialism and research collaboration in Central Africa. Scientometrics **81**(2), 413 (2009). doi:10.1007/s11192-008-2211-8. Accessed 2021-01-14
44. Aasli Nur, A., Akbaritabar, A., Vignau Loria, M., Coolman, N., Xinguang, F., Curran, S.R.: The diffusion of scientific knowledge on family planning behavior: when, where, how and by whom is research being done?, International Union for the Scientific Study of Population (IUSSP) (2021). https://ipc2021.popconf.org/abstracts/211073 Accessed 2022-10-03
45. Zhao, Z., Bu, Y., Li, J.: Does the mobility of scientists disrupt their collaboration stability? Journal of Information Science, 016555152094874 (2020). doi:10.1177/0165551520948744. Accessed 2020-08-24
46. Baruffaldi, S.H., Landoni, P.: Return mobility and scientific productivity of researchers working abroad: The role of home country linkages. Research Policy **41**(9), 1655–1665 (2012). doi:10.1016/j.respol.2012.04.005. Accessed 2021-07-31
47. German Competence Network for Bibliometrics: Kompetenzzentrum Bibliometrie (KB) (2021). https://bibliometrie.info/ Accessed 2021-11-02
48. Mikolov, T., Chen, K., Corrado, G., Dean, J.: Efficient Estimation of Word Representations in Vector Space. arXiv (2013). doi:10.48550/ARXIV.1301.3781. https://arxiv.org/abs/1301.3781
49. McInnes, L., Healy, J., Saul, N., Großberger, L.: Umap: Uniform manifold approximation and projection. Journal of Open Source Software **3**(29), 861 (2018). doi:10.21105/joss.00861

Appendix

Table S1: The random sample of 10,963 authors with composition of the control (7,399 authors) and observation (3,564) groups (numbers printed in the mobility status column are the sum of authors per mobility type over control and observation groups. N. of publications = o, N. as corresponding author = c, N. of publications in Q1 (top ranked) journals = q1. Interaction between these criteria is indicated with an x).

Group	Mobility status	Indicator combination	Count of unique authors
Control group	non-mobile (total: 2k)	top5-10% c	500
		top5-10% o	499
		top5-10% oxc	490
		top5-10% oxq1	496
		top5-10% oxq1xc	15
Control group	mobile internal (3k)	top5-10% c	300
		top5-10% o	296
		top5-10% oxc	277
		top5-10% oxq1	287
		top5-10% oxq1xc	211
		top5-10% q1	264
		top5-10% q1xc	196
Control group	mobile international (3k)	top5-10% c	300
		top5-10% o	298
		top5-10% oxc	283
		top5-10% oxq1	275
		top5-10% oxq1xc	184
		top5-10% q1	270
		top5-10% q1xc	177
Control group	mobile both (2,963)	top5-10% c	686
		top5-10% o	774
		top5-10% oxq1	98
		top5-10% q1	223
Potential talents	mobile internal (3k)	top1% c	288
		top1% o	284
		top1% oxc	253
		top1% oxq1	253
		top1% oxq1xc	91
Potential talents	mobile international (3k)	top1% c	284
		top1% o	277
		top1% oxc	235
		top1% oxq1	243
		top1% oxq1xc	174
Potential talents	mobile both (2,963)	top1% c	572
		top1% o	151
		top1% oxq1	153
		top1% q1	306

	Model 1
(Intercept)	1.06^{***}
	(0.03)
Years before move -15	-0.55^{***}
	(0.04)
Years before move -14	-0.55^{***}
	(0.04)
Years before move -13	-0.45^{***}
	(0.04)
Years before move -12	-0.46^{***}
	(0.04)
Years before move -11	-0.44^{***}
	(0.03)
Years before move -10	-0.39^{***}
	(0.03)
Years before move -9	-0.26^{***}
	(0.03)
Years before move -8	-0.21^{***}
	(0.03)
Years before move -7	-0.15^{***}
	(0.03)
Years before move -6	-0.14^{***}
	(0.03)
Years before move -5	-0.10^{***}
	(0.03)
Years before move -4	-0.07^{**}
	(0.02)
Years before move -3	-0.05^{*}
	(0.02)
Years before move -2	-0.02
	(0.02)
Years before move -1	-0.03
	(0.02)
Years after move 1	-0.23^{***}
	(0.02)
Years after move 2	-0.25^{***}
	(0.03)
Years after move 3	-0.24^{***}
	(0.03)
Years after move 4	-0.19^{***}
	(0.05)
Years after move 5	-0.25^{***}
	(0.05)
Years after move 6	-0.30^{***}
	(0.05)
Years after move 7	-0.29^{***}
	(0.05)
Years after move 8	-0.38^{***}
	(0.05)
Years after move 9	-0.46^{***}
	(0.05)
Years after move 10	-0.42^{***}
	(0.06)
Years after move 11	-0.48^{***}
	(0.06)
Years after move 12	-0.39^{***}
	(0.08)
Years after move 13	-0.34^{**}
	(0.10)
Years after move 14	-0.39^{***}
	(0.11)
Years after move 15	-0.27^{*}
	(0.13)
Mobile Internal (3k)	0.37^{***}
	(0.05)
Mobile International (3k)	0.44^{***}
	(0.05)
Potential talents	0.64^{***}
	(0.04)
R^2	0.02
Adj. R^2	0.02
Statistic	28.21
P Value	0.00
DF Resid.	4914.40
N observations	222248

$^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$

Table S2: Regression coefficients for the mobility results presented in Figure 4.

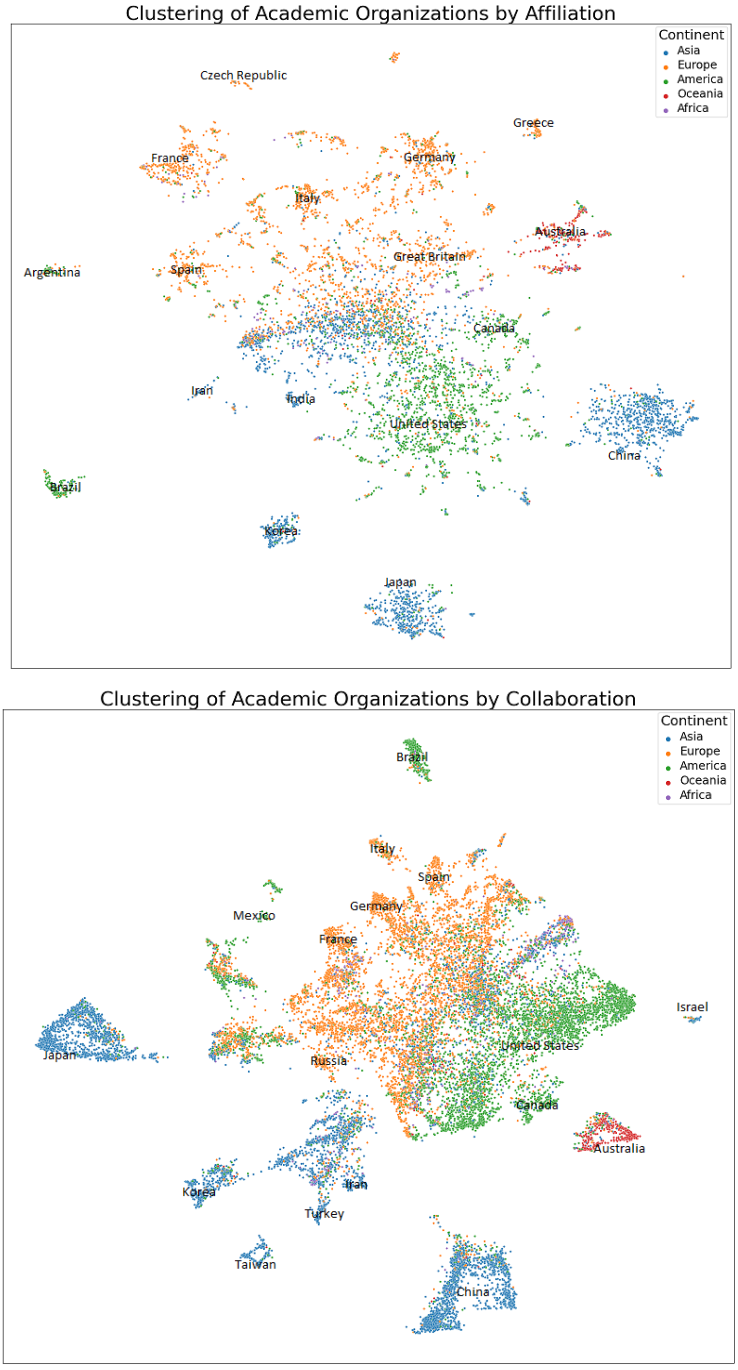
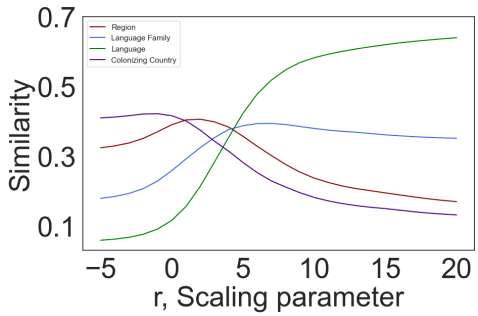
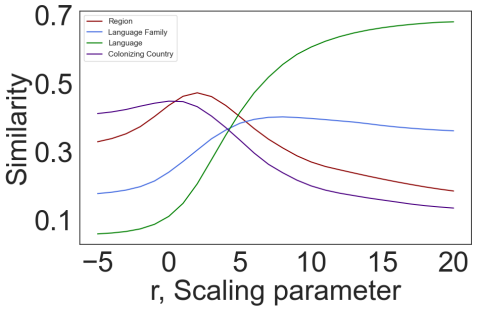


Fig. S1: **Clustering of organizations based on affiliation trajectories (top) and collaboration trajectories (bottom).** The collaboration graph is much more dense, and many of the European countries lose their distinct clusters.



(a) Affiliation



(b) Collaboration

Fig. S2: Factors influencing hierarchical clustering at lower and higher levels for affiliation and mobility of scholars, left, and collaboration, right. Colonizing country is the most influential at high levels, while language is the most influential at lower levels.