Gender Differentials in High-Skilled Mobility in Europe via LinkedIn

Daniela Perrotta[1], Elizabeth Marie Jacobs[1], Tom Theile[1], and Emilio Zagheni[1]

[1]Max Planck Institute for Demographic Research
Corresponding author: perrotta@demogr.mpg.de

Abstract

International high-skilled migration represents an increasingly important component of migration streams with a significant impact on the global flow of skills and on migration policies. Understanding the factors that explain why highly skilled workers move, and where they go, is of paramount importance in migration research, but generally difficult to measure and model. In this study, we use a largely untapped data source, the LinkedIn advertising platform, to retrieve the mobility patterns of professionals across countries in Europe and characterize the gender disparities among professionals migrating within Europe. Here we present some descriptive statistics to highlight the potential utility and challenges of employing these data to complement traditional survey data. Our findings reveal substantial variations in gender composition among different countries and age cohorts, with older high-skilled migrants showing a predominant male presence.

1 Introduction

International high-skilled migration represents an increasingly large component of global migration streams with a significant impact on the global economy, gender imbalance in employment, and migration policies. Timely, accurate, and comparative data on international migration are therefore of paramount importance for demographic research and policy implementation. However, such data is often scarce and not easily accessible. Traditional data, like censuses, is usually expensive, non-scalable, and often lack spatial and temporal granularity as well as consistent definitions and data collection standards across countries. In this context, the recent availability of large geo-located datasets has rapidly fostered new research in this field, as evidenced by recent prominent examples leveraging various sources of data, such as Facebook advertising data [4], LinkedIn advertising and recruiting platforms [2, 3], as well as bibliometric databases [5, 6]. These digital data sources hold the potential to provide valuable insights into various facets of high-skilled migration, addressing some of the limitations associated with traditional data collection methods.

In this study, we aim to reduce the existing data gap and offer new insights into the gender differentials in high-skilled migration across countries in Europe. For this, we leverage social media data we collected via the LinkedIn advertising platform. This research extends our prior work, which also utilized this dataset to characterize the mobility patterns of professionals in Europe and assess the relative attractiveness of European countries for high-educated workers. In this paper, we aim to further analyze the gender differences in the high-skilled mobility

patterns, exploring variations by sending and receiving countries, age cohorts, and industrial categories.

2 Data and Methods

2.1 LinkedIn dataset

We collected data from the LinkedIn advertising platform. This is a tool for advertisers to create ads and content aimed at a specific audience of LinkedIn users who can be reached by specifying certain targeting criteria (e.g. gender, age, industry). Given these options, LinkedIn returns an estimate of the LinkedIn users (i.e. audience size) who meet those criteria and can potentially be reached by the advertisement. In this work, we leverage this feature in order to capture migration events and retrieve the estimates of highly educated LinkedIn users who migrated across countries in Europe for job purposes. For this, we assume that the latest update on LinkedIn users' profile indicates their current job location and thus the country whey they are currently located. To obviate the lack of country of birth or previous country of residence on LinkedIn, we use the country of education as a proxy for a migration event from the country where Linkedin users studied to another where they are currently employed. To do this, we used the list of European universities from UniRank (`https://www.4icu.org/`) to select those LinkedIn users having at least a BA, MA or PhD.

We collected data in January 2023 by various targeting variables, including: 1) country of education, 2) country of current employment, 3) gender, 4) age, and 5) industrial category. The data collection process involves a separate search query for each combination of the targeting variables and returns the number of LinkedIn users meeting the given criteria. As an illustrative example, a single search would collect the number of female LinkedIn users aged 35-54 who studied at Oxford University in the United Kingdom and are now employed in a high-tech company in Germany. Note that the targeting options provided by the LinkedIn advertising platform are standardized. It is important to note that we collected only anonymous, aggregate-level data, from which identification of individuals is impossible. The data were collected purely for scientific purposes using the LinkedIn Ads platform.

2.2 Methods

In this work, we focus on Europe and consider countries according to the UN geoscheme. We dropped a few countries because of the small data amount (e.g. Monaco, Malta) or because of ongoing situations that could potentially influence the data quality (e.g. Russia, Ukraine). Given our dataset, we can reconstruct an origin-destination matrix of the high-educated LinkedIn users n_{ij} who are currently employed in country j and studied in country i. We can therefore define three categories of LinkedIn users in our dataset: i) emigrants, ii) immigrants, and iii) non-movers or returning migrants. This is when the current job location coincides with the education place. Since we cannot look at users' trajectories, this might correspond to users who didn't migrate abroad for their job or who returned to the country where they obtained their education.

We compute a female-to-male migrant ratio that measures the immigrant and emigrant gender ratio among migrant professionals. This extends Kashyap and Verkroost's previous work [1] estimating overall population gender gaps on the LinkedIn advertising platform. Values of gender ratio below one indicate that women are underrepresented relative to men in the migrant population, while values above one indicate a higher proportion of women relative to men.

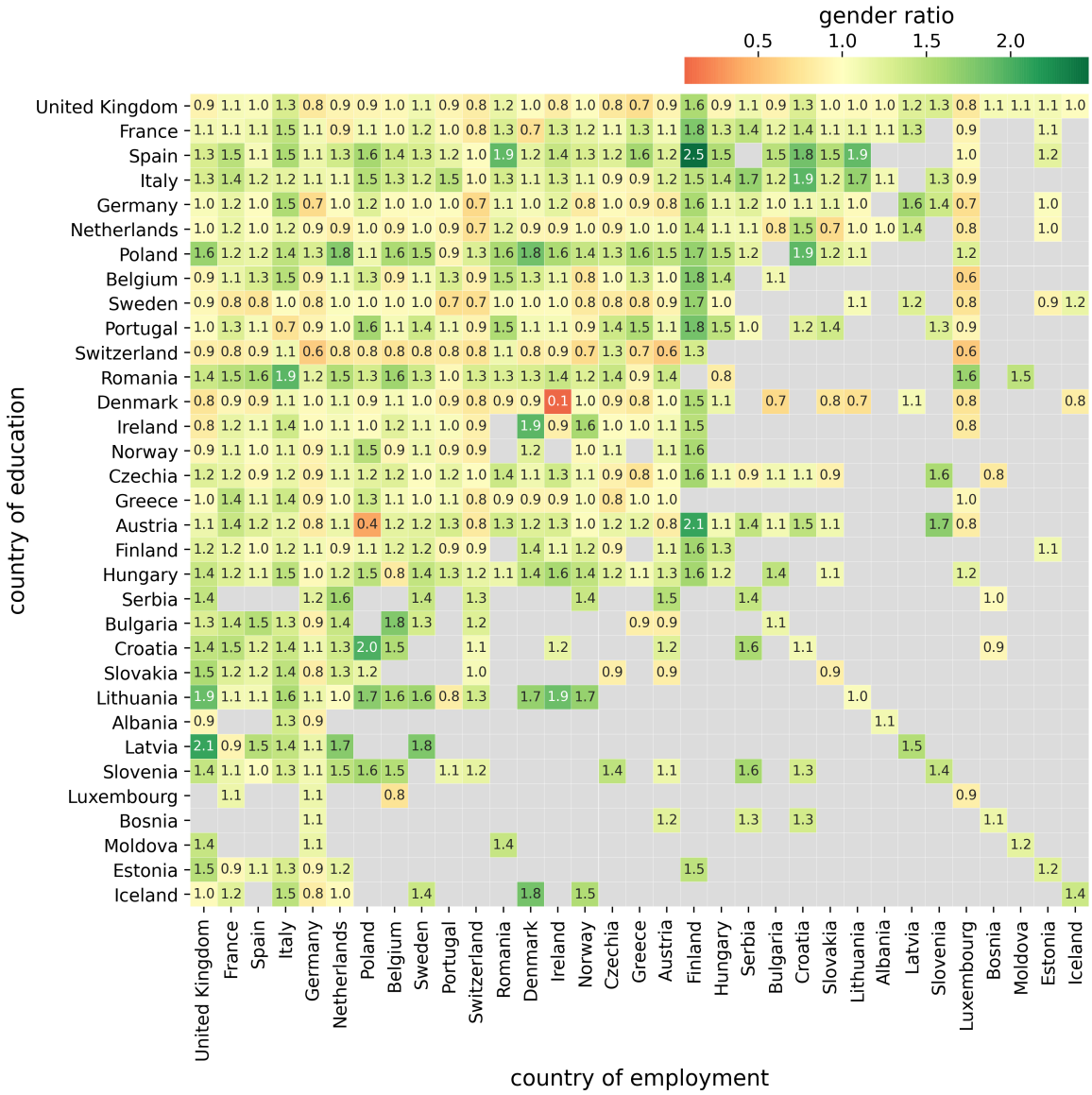

Fig. 1. Gender ratio of high-skilled migration in Europe. The heatmap shows the gender ratios relative to the high-skilled mobility patterns of LinkedIn users who moved from the country where they studied (on the y-axis) to the country where they are currently employed (on the x-axis). The color code represents the female-to-male migrant ratio (grey indicates missing values or zeros). Countries are sorted according to LinkedIn population size.

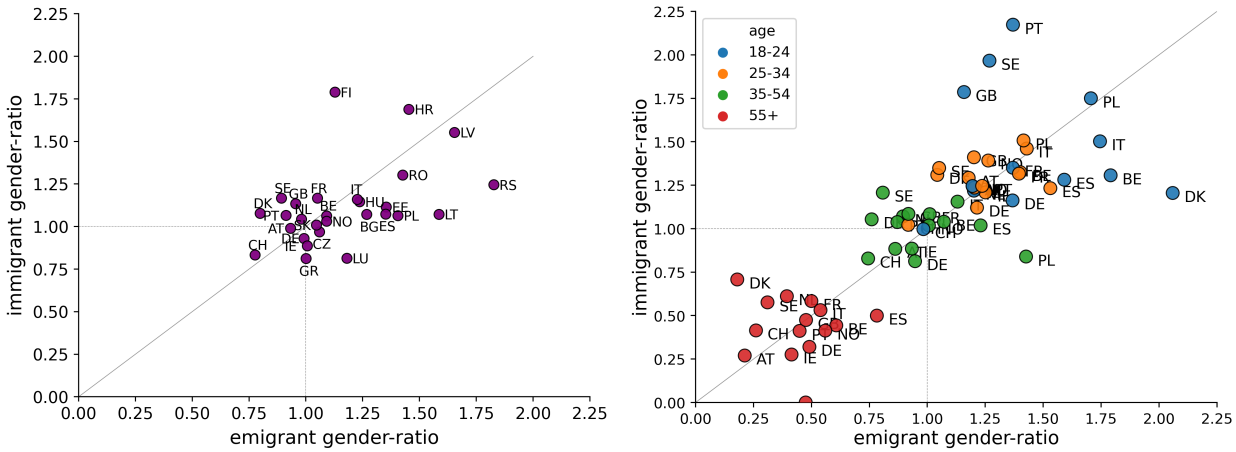

Fig. 2. Immigrant vs. emigrant gender ratio. The scatterplots show the relationship between the gender ratio of immigrants (on the y-axis) and the gender ratio of emigrants (on the x-axis) across countries in Europe. The scatterplot on the right shows the breakdown by age groups. The solid line $x = y$ is a guide to the eye and separates countries where more women enter the country compared to those who emigrated. The dashed lines indicates gender parity with the female-to-male migrant ratio equal to 1.

3 Preliminary results

Figure 1 shows the resulting origin-destination matrix of the gender ratios relative to the high-skilled mobility patterns of LinkedIn users who moved from the country where they studied to the country where they are currently employed. We observe that in certain countries, high-skilled emigration patterns are characterized by a predominance of women (e.g. Spain, Italy) or men (e.g. Switzerland). Similarly, certain countries tend to receive more women (e.g. Finland) or men (e.g. Luxembourg). Also, the matrix is not symmetrical, thus indicating countries that may act as a source or sink for high-skilled migration which can be predominantly male or female. This is more evident in Figure 2 that shows the relationship between the gender ratio of immigrants vs. emigrants across countries in Europe. Here we observe countries having values of gender ratios below 1 (e.g. Switzerland) as well as countries where the immigration of women outnumbers the emigration, and viceversa. The scatterplot on the right further shows this relationship by age groups. Here we observe that the gender ratios decrease with age almost consistently across all countries in Europe. Notably, high-skilled migrants above 55 years old are predominantly male in all countries in Europe.

4 Discussion and future work

Non-traditional data derived from social media offers a cost-effective resource to analyze migration patterns at an unprecedented scale. In this context, here we presented the potential of using LinkedIn Ads data to characterize international migrations of highly skilled workers in Europe. This approach enables us to provide new insights into the gender disparities among professionals migrating within Europe. Compared to traditional survey data, the LinkedIn data that we employ here offers the advantages that (1) are continuously available, (2) have consistently defined variables across countries and languages, and (3) provide a global snapshot of high-skilled migration as recent as the latest update to a person's LinkedIn profile.

These advantages notwithstanding, there are also some systematic biases to take into ac-

count. First, multiple counts of the same user can occur whenever they moved multiple times during their career. For example, if a LinkedIn user obtained their BA from the Sorbonne in France and their MA from the Imperial College London in the United Kingdom and then moved to Germany for work reasons, then the user is counted twice, respectively in the audience size of LinkedIn users who have studied in the United Kingdom and currently work in Germany, and in the audience size of LinkedIn users who have studied in France and currently work in Germany. Second, the audience sizes of LinkedIn users by age and gender are systematically smaller due to the fact that age and gender are inferred from the users' profile information and may not be available for all users. Third, users' nationality is unknown, thus limiting the interpretation of the direction of migration patterns, if towards the home country or towards a foreign country. Finally, the use of LinkedIn and, therefore, users' self-selection bias are heterogeneous across countries. Hence, migration flows must be adequately weighted to take into account the different LinkedIn penetration rates.

In this study, we aim to address these issues to eventually identify the utility and limitations of using LinkedIn Ads data to characterize gender differences in high-skilled mobility patterns. The final paper will include a detailed description of the data collection, processing and validation. Moreover, we will include further analysis on the variation of gender differences by industry categories.

References

[1] Kashyap, R.; and Verkroost, F. C. 2021. Analysing global professional gender gaps using LinkedIn advertising data. EPJ Data Science, 10(1): 39.

[2] Perrotta, D.; Johnson, S. C.; Theile, T.; Grow, A.; de Valk, H.; and Zagheni, E. 2022. Openness to migrate internationally for a job: Evidence from LinkedIn data in Europe. In Proceedings of the International AAAI Conference on Web and Social Media, volume 16, 759–769.

[3] Rodriguez, M.; Helbing, D.; Zagheni, E.; et al. 2014. Migration of professionals to the us. In International conference on social informatics, 531–543. Springer.

[4] Zagheni, E.; Weber, I.; and Gummadi, K. 2017. Leveraging Facebook's advertising platform to monitor stocks of migrants. Population and Development Review, 721–734.

[5] Zhao, X.; Akbaritabar, A.; Kashyap, R.; and Zagheni, E. 2023. A gender perspective on the global migration of scholars. Proceedings of the National Academy of Sciences, 120(10): e2214664120.

[6] 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, 1–23.