Recent Global Improvements in the Representation of Women in Science Have Stalled

Julie (Sojin) Kim[1,2†*], José Ignacio Carrasco[3†], Samvardhan Vishnoi[4†], Yuqi Liang[5], Meagan Lauber[6], Julián D. Cortés[7,8], Aliakbar Akbaritabar[9], Ugofilippo Basellini[9], and Emilio Zagheni[9]

[1]Institute of Health Metrics and Evaluation, University of Washington, Seattle, WA, USA.
[2]Department of Health Metrics Sciences, University of Washington, Seattle, WA, USA.
[3]Centre on Migration, Policy and Society (COMPAS), University of Oxford, Oxford, UK.
[4]Department of Physics and Astronomy, Northwestern University, Chicago, IL, USA.
[5]Department of Sociology, University of Oxford, Oxford, UK.
[6]Boston University, Boston, MA, USA.
[7]School of Management and Business, Universidad del Rosario, Colombia.
[8]Engineering School, Universidad de Los Andes, Colombia.
[9]Department of Digital and Computational Demography, Max Planck Institute for Demographic Research, Rostock, Germany.
[*]Corresponding author, Email: sojink@uw.edu
[†]These authors contributed equally to this work. They are co-first authors.

Current version: November 1, 2023

Abstract

We leverage the Scopus bibliometric database of metadata on 33 million publications from 1996 to 2020 to assess recent trends in representation of women among published scholars. We estimated that the Gender Parity Index (number of female scholars per male scholar) increased significantly until around 2011. However, since then the trend towards higher representation of women has stagnated across the large majority of countries worldwide. Our projections indicate that, if current trends persist, gender gaps are likely to increase or to stabilize over the next decade. We identified three demographic determinants of observed trends. First, the rate at which women enter academia has decreased, relative to men; second, the rate at which women exit academia, relative to men, has been fairly stable over time; third, even within a context of enhancement of gender parity in academia, the career length of female scholars has not notably increased.

Introduction

Gender equality in science is vital to the production of innovative and impactful research (*1*). However, despite important progress and policy efforts, unequal representation of gender persists in science (*2*–*5*). Women face disadvantages in various domains of their academic career, including career duration, professional mobility, and opportunities for leadership (*6*).

Gender equality in science has been studied using a wide range of dimensions and measurements (*4*, *6*). A key concept used in the literature, gender parity, is defined as the representation of women in the scientific workforce and is measured using indicators such as the percentage of women, the Gender Parity Index (GPI) (*7*). Analyzing the population of published scholars through demographic lenses helps us understand the drivers of gender parity, as well as the likely future trajectories if recent patterns continue (*6*–*9*). In this article, we focus on three key demographic components of population dynamics, namely 'birth', 'death' and 'migration', to study gender parity in the scientific workforce. In our framework, birth can be thought of as the entry into the pool of publishing scholars; death can be thought of as the exit from this academic pool or the end of a publication career; and migration can be thought of as the movement of scholars from one country to another one, traced through change in their institutional affiliations. Recent research has shown that while the percentage of female scholars entering academia has been increasing over the last decades (*7*, *10*, *11*), their rate of retention in academia remains lower and they often struggle to achieve leadership positions (*12*, *13*). Furthermore, female scholars on average migrate less than their male counterparts (*14*), thus missing out on the career boost that comes with scholarly mobility (*15*).

While previous studies at the intersection of demography and science of science have yielded valuable insights into the issue of gender parity, these analyses have mostly focused on individual components such as entry (*4*), exit (*12*, *16*) or migration (*14*). Examining all these factors in tandem is crucial to provide a full picture of the determinants of gender representation of women in science, and to unveil their relationships. In this article, we first describe the past and current landscape of global gender parity through comprehensive country-specific GPI assessments of the scientific workforce from 2001 to 2015. Second, we utilize past trends of entry, exit and migration rates from 2001-2015 to project gender differences in the growth rate of the scientific workforce across countries. Lastly, we estimate differences in career duration for female and male scholars at a country scale from 2001 to 2015. By applying core demographic concepts to understand the population dynamics of scholars, we aim to provide a more nuanced and comprehensive picture of gender parity in the global science community. Based on suggested current and projected future gender disparity of scholars who enter, leave the academia or are internationally mobile, policymakers can develop more informed and accurate organizational or national science policies to address the issue (*6*, *8*).

Materials and Methods

Data

We used the Scopus bibliometric database on 33 million publications from 1996 to 2020. Previous studies (*17*, *18*) have shown that Scopus is one of the index and citation databases with the largest coverage of English and non-English publications in scientific journals. In addition, it offers a scholar identification number that according to Baas et al. (*18*) identifies all publications of a single author in 94.4% of cases (recall) and has a precision rate of 98.1%, which means that the records of two different authors could be merged under one id in only 1.9% of the cases. The correct identification of an author is crucial for studies such as ours that construct a scholar's publication trajectory to study entry, exit and migration.

We transformed the publication records into individual scholar activity, in order to ascertain gender information and build career trajectories for each scholar. These career trajectories laid the foundation for a demographic framework incorporating birth i.e., entry, death i.e., exit, and migration. A scholar 'enters' the publication phase in the year of their first publication, and 'exits' the publication phase in the year of their last publication. For the purposes of this study, this set of publishing scholars in Scopus-indexed journals represents the global scientific workforce.

Next, we incorporated residence and migration information into scholar career trajectories, furthering our global demographic perspective (*14*, *19*, *20*). Scopus data identify the affiliated organization of each scholar of a publication, which allowed us to infer the country of residence of scholars. When a scholar is affiliated with organizations in multiple countries for the same year, we estimated the scholar's country of residence for that year as the country of affiliation with most publications during that year for the author. Finally, we incorporated gender information of each scholar predicted using their first name. More specifically we utilized the data and approaches developed by Zhao et al. (*14*).

Two censoring issues arise in constructing scholar activities using their publication activity. First, for entry, it was not always possible to distinguish a scholar's actual entry into the workforce from entry into the Scopus database. For example, if a scholar published in 1995 and in 1997, we would only observe the publication in 1997, as the first year of our database was 1996. To limit this potential error related to entry, any information from scholars whose first publication was before 2001 were excluded from the analysis (i.e., left-truncation). Second, for exit, we could not determine if a scholar was preparing for their next publication or exited academia for those whose last publication was after 2015. Therefore, these scholars were classified as right-censored for their activity after 2015. (See Supplementary Materials 1 for additional analysis on the choice of 5-year period.) In addition, we excluded all information from scholars entered publishing pool after 2015 (i.e., right-truncation) as we could not ascertain their exit status after this year.

Demographic Indicators

Gender Parity Index (GPI). GPI for each country and year was calculated using the definition of the United Nations and commonly used in previous literature (*7*):

$$\mathrm{GPI}_{c,t} = \frac{\text{Number of female scholars}_{c,t}}{\text{Number of male scholars}_{c,t}} \tag{1}$$

Here $\mathrm{GPI}_{c,t}$ represents the number of female scholars per one male scholar in country c in year t. The GPI value can lie anywhere over or at 0, where 1 represents parity between male and female scholars.

Population Growth Rate. Using entry, exit and migration information of individual scholars, we calculated population growth rates from 2001-2015 using the following formula (*21*):

$$\text{Population Growth Rate}_{c,t,g} = \text{Entry Rate}_{c,t,g} - \text{Exit Rate}_{c,t,g} + \text{Net Migration Rate}_{c,t,g} \tag{2}$$

where

$$\begin{aligned}
\text{Entry Rate}_{c,t,g} &= \frac{B_{c,t,g}}{\frac{1}{2}(P_{c,t,g} + P_{c,t+1,g})} \\
\text{Exit Rate}_{c,t,g} &= \frac{D_{c,t,g}}{\frac{1}{2}(P_{c,t,g} + P_{c,t+1,g})} \\
\text{Net Migration Rate}_{c,t,g} &= \frac{\text{In Migration}_{c,t,g} - \text{Out Migration}_{c,t,g}}{\frac{1}{2}(P_{c,t,g} + P_{c,t+1,g})}
\end{aligned} \tag{3}$$

Here, Population Growth Rate$_{c,t,g}$ refers to population growth rate for country c, year t and gender g; $B_{c,t,g}$ refers to number of scholars identified as gender g whose first publication was in year t and their primary affiliation was country c; $D_{c,t,g}$ refers to the number of scholars identified as gender g whose last publication in the data was year t with primary affiliation in country c; In Migration$_{c,t,g}$ indicates the number of scholars identified as gender g whose primary affiliation changed to country c in year t; Out Migration$_{c,t,g}$ indicates the number of scholars identified as gender g whose primary affiliation changed from country c in year t; $P_{c,t,g}$ indicates total number of scholars with an identified gender g for a country c in year t.

Population Grwoth Rate Difference. Our primary indicator of interest was female to male difference in population growth rate (Rate Difference$_{c,t}$) calculated as following:

$$\text{Rate Difference}_{c,t} = \text{Population Growth Rate}_{c,t,female} - \text{Population Growth Rate}_{c,t,male} \tag{4}$$

We chose this as our primary indicator because the difference in growth rates over time captures progress towards gender parity in each country. For instance, a positive rate difference indicates that the female scholar population is growing at a faster rate than the male scholar population, thus progressing towards gender parity.

This indicator has multiple advantages in addressing potential unadjusted measurement errors in the data. For instance, Scopus covered more journals over time during our observation period. While the process of new journal inclusion to Scopus is expected to retrospectively include all publications in the journal prior to joining, subsequent validation with other sources is not possible. In addition, Scopus applies a scholar disambiguation algorithm (*18*) to determine the number of distinct scholars from the pool of scholars with that same name. Yet, whether this process has a differential impact over time and gender is unknown. Using female to male difference could reduce such potential of measurement errors stemming from original data, while capturing progress towards gender parity among publishing scholars.

Demographic Projection

To understand future dynamics of gender parity among actively publishing scholars, we projected female to male difference in rate of population growth. The difference in population growth rate was projected using a random walk with drift as shown in the following equation:

$$\begin{gathered}
\text{Rate Difference}_{c,t} = \text{Rate Difference}_{c,t-1} + \mu + \epsilon_i, \text{ where} \\
\epsilon_i \sim N(0, \sigma^2)
\end{gathered} \tag{5}$$

Here, μ is the average period-to-period change in the time period considered. We first estimated parameters using past data (i.e., period from 2001-2015). We then projected the future trends in difference in growth rate for years 2016 to 2040.

In order to address any potential unadjusted measurement error that has differential impact on time, we combined projections from the same model specification over different observation periods. That is, we iterated the projection process using 10 different lengths of observation period (2001-2015, 2002-2015, ..., 2010-2015). This approach is often adopted in mortality forecasting to average the effect of the fitting period on the resulting forecasts (*22*). Based on combined trajectories of projections, we calculated median, 80% and 95% prediction intervals of projected female to male difference in population growth rates.

Survival Analysis

The duration between the year of entry and the year of last observation, either right-censored or exited, was used to calculate the cumulative survival probability for cohorts that entered the publishing scholars pool between 2001 and 2005, 2006 and 2010, and 2011 and 2015. Considering the features of our data and population, we categorized migration events and the last publication after 2015 as right-censoring events. First, if a scholar moved out of a country, they were marked as right-censored in the year they migrated from their original country. Second, scholars with their last publication after 2015 were marked as right-censored, allowing a 5-year period of inactivity instead of labeling those scholars as having exited the population.

Cumulative survival probability was calculated using a Kaplan-Meier estimator to estimate survival probability by gender (*23*). For each gender g in country c, survival probability at year t from their entry was estimated using the following equation:

$$\hat{S_{c,g}}(t) = \Pi_{i:t_i \leq t}(1 - \frac{d_i}{n_i}) \tag{6}$$

Here, $S_{c,g}(t)$ refers to cumulative survival probability of gender g in country c at time t. d_i and n_i refer to the number of scholars exited at time $t = t_i$ and the number of scholars included in the publishing pool at time $t = t_i$, respectively. This analysis was conducted using 'survival' package in R (*24*). R Version 4.2.3 was used throughout the analysis of this study (*25*).

Results

Our data span more than 190 countries and over 6.5 million scholars worldwide. However, the 20 largest countries, by scholar count, cover 87% (N $\approx$ 5.7 million) of the global population of scholars. In this section, we present the results for these 20 largest countries. Estimates for all other countries are provided in the public data repository and replicability materials for this article. It should be noted, though, that for some countries with only a small number of scholars, the stochasticity of the estimates may be substantial.

For demographic projection and survival analysis, we focused on scholars who entered the the publishing pool between 2001 and 2015. For this group, among those with gender information available and affiliated with the largest 20 countries, we identified a total of 3,490,803 scholars. Of these scholars, 2,174,843 (62.3%) were classified as male and 1,315,960 (37.7%) as female. For illustrative purposes, the figures in the main article highlight four countries i.e., Brazil, Germany, Japan, and the United States, selected across four different areas of the world. Corresponding figures are provided in the Supplementary Materials 2-4 for 20 countries with the largest population of publishing scholars.

Overall, the gathered evidence points to an initial improvement in the representation of women among actively publishing scholars. This stems from higher population growth rates for female scholars as compared to their male counterparts for most countries during our period of study. There were important differences across countries in both the levels of improvement observed, and the levels required to achieve gender parity. Yet, for most countries, in recent years we also find a gradual decrease and stagnation of this improvement. Our forecasts indicate that a continuation of this trend will eventually lead to slower growth of women, as compared to men, in the near future, driving us away from the often stated goal of gender parity.

The pace of improvement in the GPI has slowed down dramatically in recent years. Figure 1 shows the GPI from 2001 to 2015. For most countries, the GPI increased steadily over time between 2001 and 2011. Starting from around 0.386 in 2001, the index for the 20 largest countries combined reached 0.528 in 2011. However, it is worth noting that the increase of gender parity has been less pronounced after 2011, as indicated by a lower rate of increase. From 2011 to 2015, the increase in GPI was 0.016, more than four times lower than a 0.074 increase between 2007 to 2011.

Brazil tended to maintain the highest gender parity across countries, starting from a relatively high level of 0.648 in 2001. Brazil also experienced a higher increase of gender parity as compared to other countries, reaching 0.931 in 2011. As with other countries, however, this increase has flattened out since 2012. In contrast to Brazil, Germany stood out as a country with relatively low GPI. Starting from 0.272 in 2001, Germany has caught up slightly with the other largest 20 countries, and reached 0.450 in 2011. In the United States, the country with the highest number of scholars included in the Scopus data, GPI increased from 0.417 in 2001 to 0.576 in 2011. Japan exhibited one of the lowest levels of gender parity. In 2001, GPI in Japan was 0.163, and it experienced a modest increase to 0.233 by 2011, followed by a decline to 0.225 in 2015. This implies that the representation of female scholars in Japan was approximately half that of the United States.

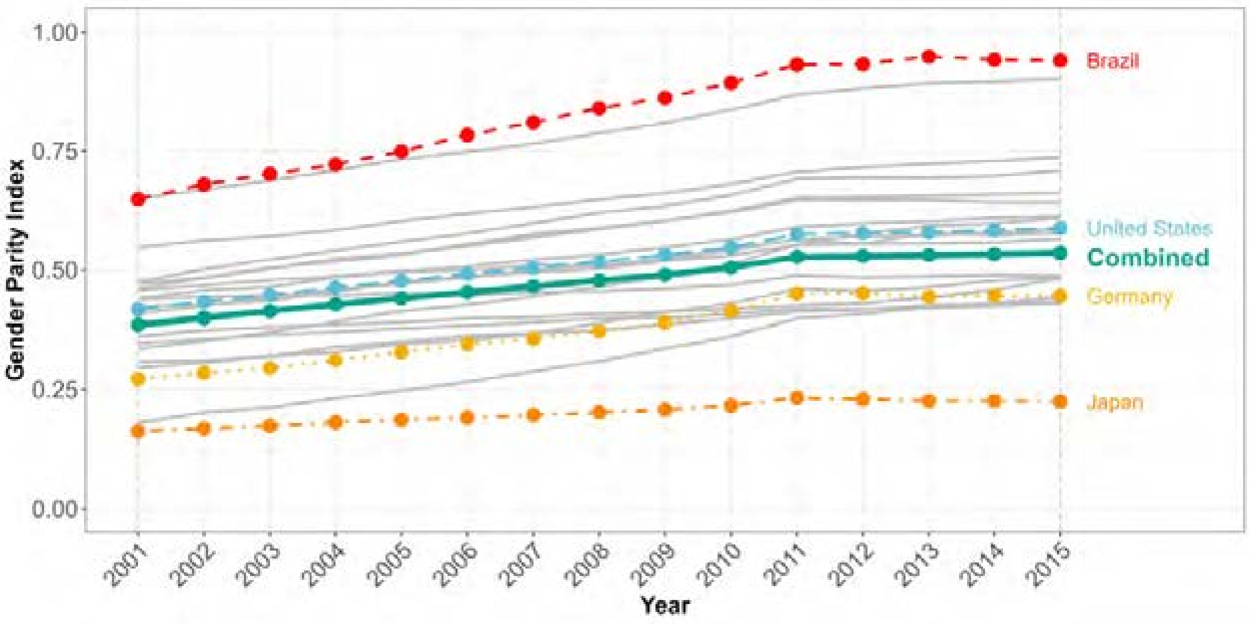


Fig. 1: Gender Parity Index (number of female scholars per male scholar) across countries with the 20 largest scholar population in Scopus data, 2001-2015

Past and future trajectories of gender difference in population growth rates suggest that female population will likely to grow slower than male population in many countries. Figure 2 presents the observed and projected growth rate difference between female and male scholars for the years 2016 to 2040. The figure shows the yearly median, 80%, and 95% prediction intervals of the 1,000 simulations obtained from 10 different time period ranges of observed data (between 2001 and 2015). The median of the simulations showed a downward trend in the growth rate difference in the four highlighted countries. Moreover, this downward trend was observed for the majority of the simulations, as shown by the 80% and 95% prediction intervals.

However, although all four trends show downward trajectories, the rates of change and variability of predictions differed between countries. The United States had the lowest rate of decrease, whereas Germany had the highest. Brazil and Japan demonstrated similar slope values. In regards to prediction ranges, Germany and Japan presented the highest variability, followed by Brazil and the United States. The median growth rate difference already turned negative in Japan (2014) and Brazil (2015), and we predict that Germany and the United States would experience the same by 2020 and 2021, respectively. (See Supplementary Materials 2 for all 20 country results.)

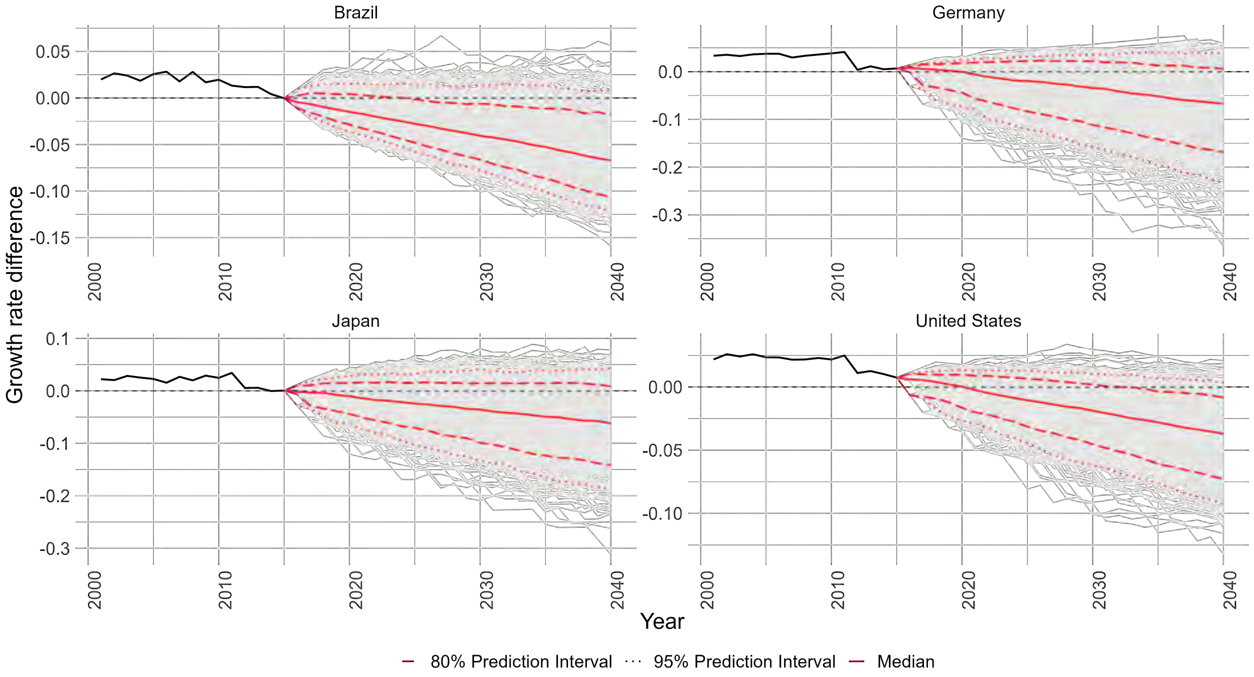


Fig. 2: Projected median (solid red lines), 80% prediction interval (red dashed lines) and 95% prediction interval (red dotted lines) of female to male difference in population growth rates for 2016-2040 from 2001-2015 (black line) estimation for Brazil (upper left), Germany (upper right), Japan (lower left) and the United States (lower right).

Gender differences in entry, exit and net migration rates show dissimilar drivers of gender disparity across countries. Figure 3 shows the differences in population growth rates, and its component entry (birth), exit (death), and net migration rate differences between female and male scholars for Brazil, Germany, Japan and the United States. In addition, Supplementary Material 3 illustrates the information for all 20 countries. Among the four highlighted countries, the female to male difference in population growth rate tended to be the highest in Germany (all-period mean: 0.028) while other three countries had relatively similar rates before 2012, the United States had all-period mean of 0.200, followed by Japan (0.019) and Brazil (0.017). This indicates that Germany was moving the fastest, among the four countries shown in the figure, towards gender parity, despite its low starting index. All four countries showed decrease in female to male difference in growth rate after 2012, explaining the slow down of the progress towards gender parity during this period. Among the 20 largest countries included in our analysis, the Netherlands tended to have the highest female to male difference in growth rate (mean across 2001-2015: 0.035), followed by Iran (0.030) and Switzerland (0.029). India recorded the lowest female to male difference in growth rate (all-period mean: 0.003), followed by China (0.008). Notably, all largest 20 countries, except for China and Russia, showed a decrease in population growth rate difference after 2012.

Decomposition of population growth rate shows dissimilar component characteristics across countries, even for countries with similar levels of difference in population growth rate. Comparing Japan and Brazil, for instance, Brazil shows relatively lower gender difference in entry rate (0.035) with the lowest gender difference in exit rate (0.016). On the other hand, Japan shows higher differences in entry rate (0.059) with the highest differences in exit rate (0.038), among all largest 20 countries.

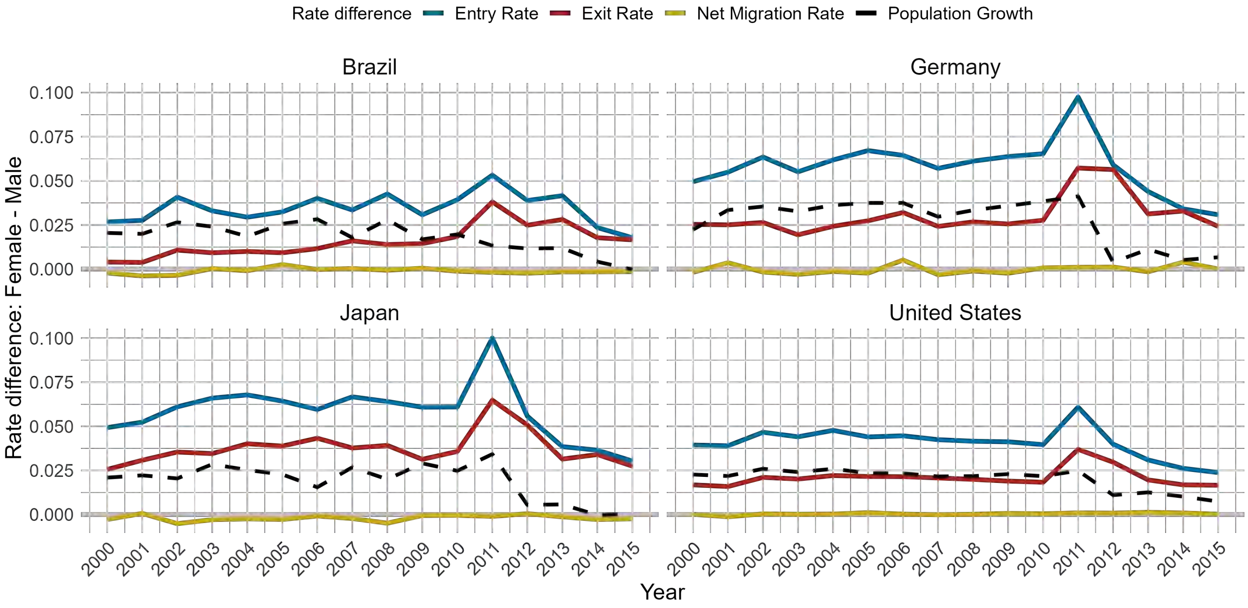


Fig. 3: Female to male difference in entry/birth (blue), exit/death (red), net migration (yellow) and population growth (black) rates for Brazil (upper left), Germany (upper right), Japan (lower left) and the United States (lower right), 2000-2015

Gender difference in career duration of publishing scholars indicates that female scholars persistently faced disadvantage with minimal progress in the past decade. Figures 4 and Supplementary Material 4 show cumulative survival probabilities of scholars from their first publication to any given year for 4 countries and all 20 countries, respectively. From this figure, we focused on four primary patterns: 1) gap between male and female survival probability in each cohort; 2) increase or decrease in the gender gap across different cohorts; 3) differences in overall level of survival probability and the rate of drop-out across countries; 4) probability of dropping out within the first year of publication.

First, among the four highlighted countries, Japan had the highest male to female gap in survival probability except for the most recent cohort, whereas the United States had the lowest difference. Second, Germany showed an increase in the gap of survival probability between cohorts that entered active publication phase in 2001-2005 and 2011-2015, indicating that more female scholars had dropped out of actively publishing than males in more recent cohorts. Third, the long-term survival probability tended to be the highest in Brazil. After 15 years of observation, for those who entered publishing scholars pool between 2001 and 2005, 45%

of female and 50% of male scholars were still actively publishing in Brazil, whereas, only 18% of female and 30% of male scholars remained after 15 years in Japan. Lastly, all largest 20 countries showed substantial drop-out of authors who published only in their first year of "life" (i.e. when they entered active publication phase) and did not publish again afterwards. These authors are captured as drop in survival probability at time $t = 0.5$ in the figures. This drop in survival probability during the first year was higher for the more recent birth cohorts. For instance, in Brazil, survival probability at year 1 was over 0.8 for both male and female scholars who entered active publication phase between 2001 and 2005. However, the same probability among those who started publishing between 2011 and 2015 was between 0.5 and 0.6. Similar trends can be observed in all largest 20 countries, with differences in survival levels.

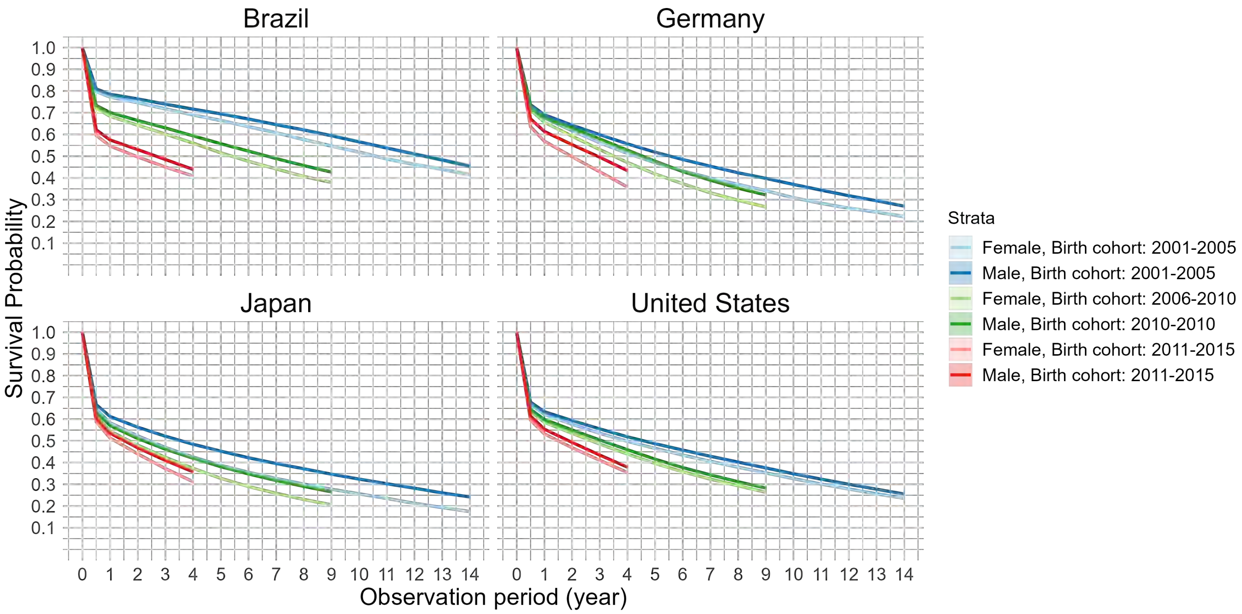


Fig. 4: Mean survival probability and 95% confidence interval by gender for cohort that started publishing in 2001-2005 (blue), 2006-2010 (green) and 2011-2015 (red) for Brazil (upper left), Germany (upper right), Japan (lower left) and the United States (lower right). Darker lines indicate male and lighter lines indicate female, 2001-2015

Discussion

Using a conceptual framework based on the entry, exit and migration components that underlie population change, we depicted, explained and projected trends in gender parity among published scholars. We focused on the gender difference in the growth of countries' scientific population, to project the future trends of gender parity in science. We relied on past and current levels of the GPI, the differences in the demographic components of population growth, and survival probabilities to explain our projections of gender parity. The results illustrate that current progress towards gender parity is stalling. In addition, there is a notable downward trend in the difference in growth rate between male and female scholars, although with considerable heterogeneity between countries.

These findings are further supported by the fact that some countries have reached negative values in the growth rate difference between gender in recent years. While the context of each country should be examined to interpret GPI, these results imply that the GPI will tend to remain at current levels or even regress back to the levels of previous years where the gender gap was larger. For instance, both Germany and Japan display a halt in the growth rate difference between male and female, which is particularly concerning because this is occurring with a stagnation at the lowest levels of gender parity compared with other countries that are within the scope of our analysis.

On the drivers contributing to gender differences in the growth rate of scientific population, our results suggest that there is a marked deceleration of such progress since 2012, despite that female population has been growing faster than the male population. The trends of gender disparity are also evident when evaluating career duration differences between men and women. While the degree of the gender differences varied across countries, we showed that all female scholars had significantly lower survival probabilities across all cohorts. Importantly, the male to female difference in career duration was not necessarily decreasing over time, and in some cases it was widening.

We have introduced a conceptual framework and methodology designed to model the factors influencing the dynamics of gender parity, which also allowed for the future projections. Studies that examined female representation in scientific community rarely accounted for all components in conjunction (*4*, *26*, *27*). Additionally, prior research employing demographic concepts has predominantly relied on data from a particular country or field (*10*, *28*–*30*), whereas our study offers a global perspective, encompassing all countries and disciplines featured in Scopus journals.

Even though this study offers a substantial contribution to the literature, we also would like to acknowledge its limitations. First, while this study seeks to examine gender equality in science, an important distinction should be made between gender parity and gender equality. In our article, parity is specifically defined as the degree to which women are equally represented in various institutions, but it is important to consider that gender parity is not always synonymous with equality as gender parity is one of the important dimensions of gender equality (*31*) and issues such a the parity paradox (*6*, *32*) remain unresolved in our study. Therefore, it should be noted that many of the metrics used in this study are measures of gender parity although they may fail to capture more complex dynamics in gender inequality in the academic realm.

In addition, it should be considered that our entry, exit and migration definitions are not without drawbacks. Particularly, the right- and left-truncation issues that we face have shortened the time period of our analysis. Moreover, because our measurement of entry into academia is based on the country corresponding to the institutional affiliation, our results regarding scholars' entry origins do not reflect birth nor nationality. Lastly, similarly to previous studies (*14*), we were less capable of measuring gender disparity in China as the gender detection algorithm is less accurate in detecting genders from transliterated Chinese names. In addition, we note that the gender inference using first names to binary categories is a limitation of algorithmic labeling. It neglects the diversity in gender identities beyond this dichotomy which is necessary in such an analysis. Nevertheless, obtaining such diversity needs further metadata beyond names that would involve surveying scholars and was beyond the scope of our analysis.

Because of the three demographic components, we are able to provide a much more detailed understanding of gender parity in science compared to other research. In this way, our study enables policy makers to identify country-specific ways of facilitating gender parity, e.g., by enabling female scholars to stay in the academic sector longer. Thus, our results involve actionable insights for public policy. More broadly, policymakers and researchers should not only consider particular aspects of gender parity (such as entry or participation). Instead, it is essential to consider a combination of factors and their inter-relationships to obtain a more comprehensive picture of gender parity in science. For example, some countries, such as Brazil, are leading the way with near parity in gender distribution of new scholars, but female scholars still have shorter careers on average than their male counterparts. This fact highlights the role of survival and exit of scholars. Namely, the end of a scholar's academic career can be greatly influenced by their gender.

In this study, we focused on the demographic components underlying past and future trends in gender parity, rather than the individual mechanisms that may be explaining these trends. As one of the explanations, researchers have provided evidence that stereotypes and discrimination may prevent women from moving up the career ladder in science, or even cause them to leave the system altogether (*33*–*35*). Likewise, we emphasise that attracting female scholars from abroad and reducing their likelihood to leave are also important drivers of gender parity. These would likely interact with the institutional characteristics of the countries' scientific environments.

Acknowledgements

We thank our colleagues Tom Theile and Xinyi Zhao for sharing some replication scripts with us.

Funding: This study has received access to the bibliometric data through the project "Kompetenzzentrum Bibliometrie" (*36*) and we acknowledge their funder Bundesministerium für Bildung und Forschung (grant number 16WIK2101A).

Authors' contributions:

- Conceptualization: A.A.; U.B.; E.Z.; J.K.; J.C.; S.V.
- Data curation: A.A.; I.C.; S.V.; J.K.
- Formal Analysis: I.C.; S.V.; J.K.
- Funding acquisition: E.Z.
- Methodology: A.A.; U.B.; E.Z.; J.K.
- Project administration: A.A.
- Software: S.V.; J.K.; I.C.
- Supervision: A.A.; U.B.; E.Z.
- Visualization: I.C.; S.V.; J.K.
- Writing - original draft: I.C.; Y.L.; J.K.; S.V.
- Writing - review & editing: I.C.; Y.L.; J.K.; S.V.; M.L.; J.C.; A.A.; U.B.; E.Z.

Competing interests: The authors declare that they have no competing interests.

Availability of data and materials: We publicly share aggregate data and scripts that allow reconstructing our analysis in this GitHub repository (**will be added upon acceptance; anonymized link is available for peer review purposes**).

References

1. Y. Yang, T. Y. Tian, T. K. Woodruff, B. F. Jones, B. Uzzi, *Proceedings of the National Academy of Sciences of the United States of America* **119**, e2200841119, ISSN: 10916490, (https://www.pnas.org/doi/abs/10.1073/pnas.2200841119) (36 Sept. 2022).
2. K. Christian, *Gender inequality in sciences: why is it still with us?*, (https://blogs.nature.com/naturejobs/2018/04/09/gender-inequality-in-the-sciences-why-is-it-still-with-us/) (2018).
3. Directorate-General for Research and Innovation (European Commission), *She Figures 2018* (Publications Office of the European Union, LU, 2019), ISBN: 978-92-79-86715-6.
4. V. Larivière, C. Ni, Y. Gingras, B. Cronin, C. R. Sugimoto, *Nature* **504**, 211–213, ISSN: 1476-4687 (Dec. 2013).
5. B. Macaluso, V. Larivière, T. Sugimoto, C. R. Sugimoto, *Academic Medicine: Journal of the Association of American Medical Colleges* **91**, 1136–1142, ISSN: 1938-808X (Aug. 2016).
6. C. R. Sugimoto, V. Larivière, *Equity for women in science: dismantling systemic barriers to advancement* (Harvard University Press, Cambridge, Massachusetts, 2023), p. 256, ISBN: 9780674919297, (https://www.hup.harvard.edu/catalog.php?isbn=9780674919297).
7. V. Gal´n-Muros, M. Bouckaert, J. Roser, "The Representation of Women in Academia and Higher Education Management Positions" (UNESCO, 2023), (https://www.iesalc.unesco.org/wp-content/uploads/2023/03/PB-Gender_final_EN.pdf).
8. D. of Economic, U. N. Social Affairs, "The sustainable development goals", (https://sdgs.un.org/goals).
9. F. Petropoulos *et al.*, *International Journal of Forecasting* **38**, 705–871, ISSN: 01692070, arXiv: 2012.03854 (July 2022).
10. S.-N. C. Liu, S. E. Brown, I. E. Sabat, *Archives of Scientific Psychology* **7**, 32 (2019).
11. M. C. Maphalala, N. Mpofu, *Gender and Behaviour* **15**, 9216–9224 (2017).
12. D. Kaminski, C. Geisler, *Science* **335**, 864–866, ISSN: 10959203, (https://doi.org/10.1126/science.1214844) (Feb. 2012).
13. J. M. Box-Steffensmeier *et al.*, *PLOS ONE* **10**, e0143093, ISSN: 1932-6203, (https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0143093) (11 Nov. 2015).
14. X. Zhao, A. Akbaritabar, R. Kashyap, E. Zagheni, *Proceedings of the National Academy of Sciences* **120**, e2214664120, (2023; https://www.pnas.org/doi/10.1073/pnas.2214664120) (Mar. 2023).
15. C. Franzoni, G. Scellato, P. Stephan, en, *Economics Letters* **122**, 89–93, ISSN: 0165-1765, (2020; http://www.sciencedirect.com/science/article/pii/S0165176513004874) (Jan. 2014).
16. K. Spoon *et al.*, *Science Advances* **9**, Publisher: American Association for the Advancement of Science, eadi2205, (2023; https://www.science.org/doi/10.1126/sciadv.adi2205) (Oct. 20, 2023).
17. M. Visser, N. J. van Eck, L. Waltman, *Quantitative Science Studies* **2**, 20–41, ISSN: 2641-3337 (Apr. 2021).
18. J. Baas, M. Schotten, A. Plume, G. Côté, R. Karimi, *Quantitative Science Studies* **1**, 377–386, ISSN: 2641-3337, (2021; https://doi.org/10.1162/qss_a_00019) (2020).

19. X. Zhao, S. Aref, E. Zagheni, G. Stecklov, *Scientometrics* **127**, 7707–7729 (2022).

20. E. Sanliturk, E. Zagheni, M. J. Dańko, T. Theile, A. Akbaritabar, *Proceedings of the National Academy of Sciences* **120**, e2217937120 (Jan. 2023).

21. S. Preston, P. Heuveline, M. Guillot, *Demography: Measuring and Modeling Population Processes* (Wiley, Oct. 2000), ISBN: 978-1-55786-451-2.

22. U. Basellini, C. G. Camarda, H. Booth, *International Journal of Forecasting*, ISSN: 0169-2070 (Dec. 2022).

23. E. L. Kaplan, P. Meier, eng, *Journal of the American Statistical Association* **53**, 457–481, ISSN: 0162-1459 (1958).

24. T. M. Therneau, *A Package for Survival Analysis in R*, R package version 3.5-7, (`https://CRAN.R-project.org/package=survival`).

25. R Core Team, *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing (Vienna, Austria, 2023), (`https://www.R-project.org/`).

26. D. C. Sing, D. Jain, D. Ouyang, *The Spine Journal* **17**, 1749–1754 (2017).

27. A. Akbaritabar, F. Squazzoni, *Science, Technology, & Human Values*, ISSN: 0162-2439 (July 2020).

28. J. Clark Blickenstaff*, *Gender and education* **17**, 369–386 (2005).

29. C. E. Gasser, K. S. Shaffer (2014).

30. M. Schaer, J. Dahinden, A. Toader, *Journal of Ethnic and Migration Studies* **43**, Publisher: Routledge _eprint: https://doi.org/10.1080/1369183X.2017.1300254, 1292–1307, ISSN: 1369-183X, (2021; `https://doi.org/10.1080/1369183X.2017.1300254`) (June 2017).

31. J. S. Chafetz, M. F. Fox, *Handbook of the Sociology of Gender*, 441–457 (2006).

32. G. Stoet, D. C. Geary, *Psychological science* **29**, 581–593 (2018).

33. D. I. Miller, A. H. Eagly, M. C. Linn, *Journal of Educational Psychology* **107**, 631 (2015).

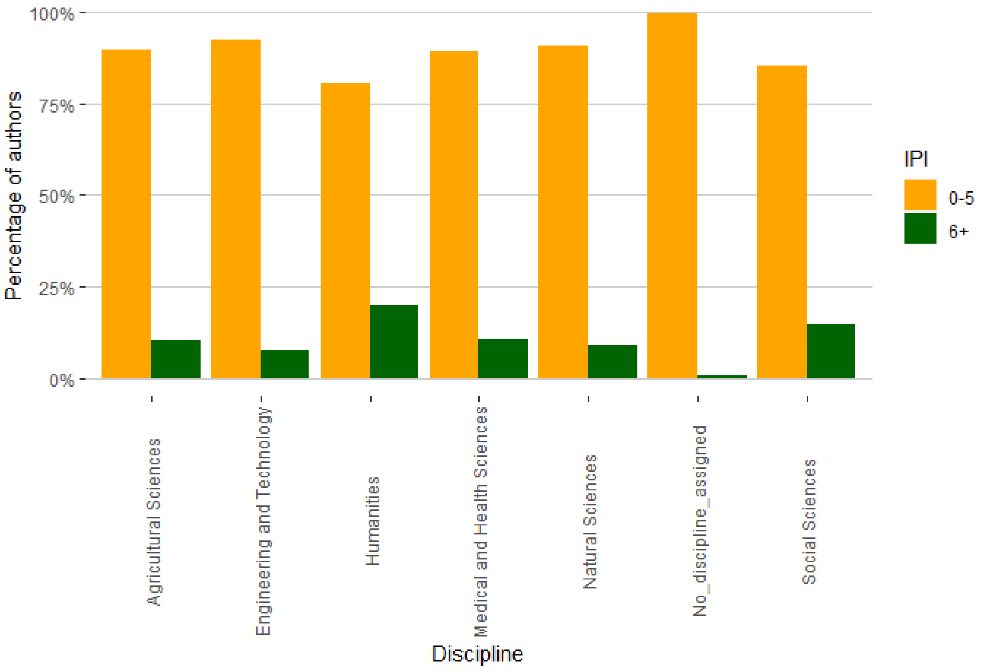
34. S. Cheryan, V. C. Plaut, C. Handron, L. Hudson, *Sex roles* **69**, 58–71 (2013).

35. D. Knights, W. Richards, *Gender, Work & Organization* **10**, 213–238 (2003).

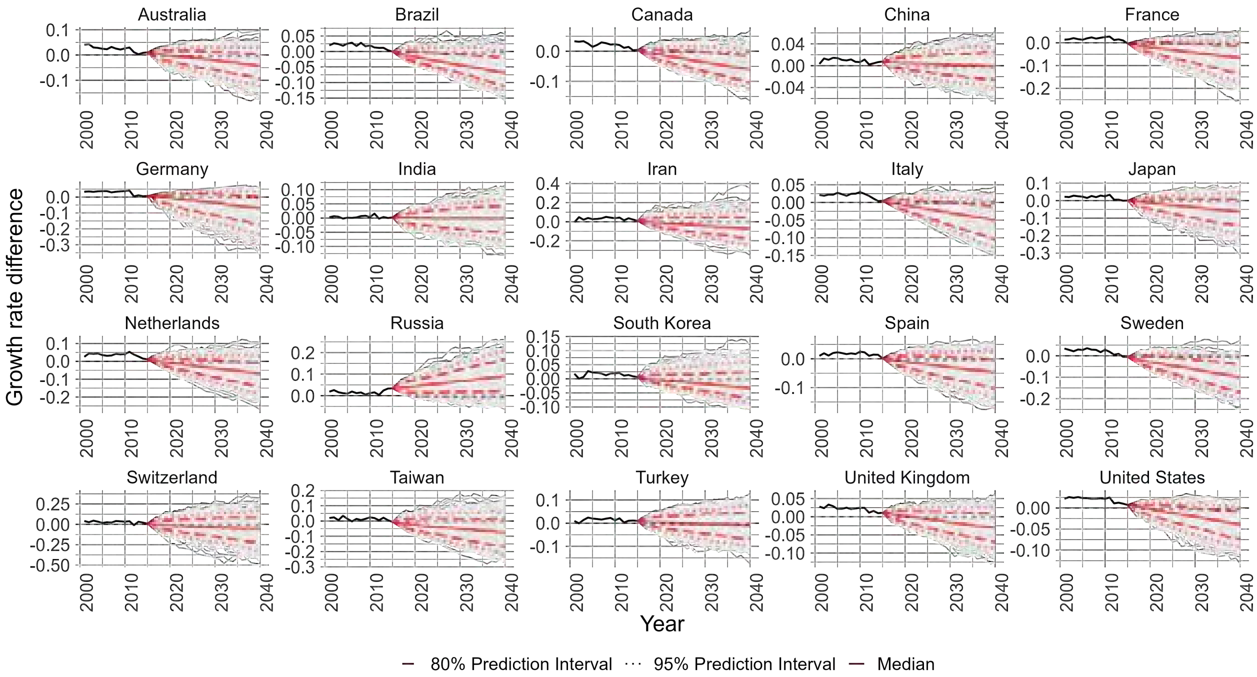
36. German Competence Network for Bibliometrics, *Kompetenzzentrum Bibliometrie (KB)*, 2021, (2021; `https://bibliometrie.info/`).

Supplementary Materials

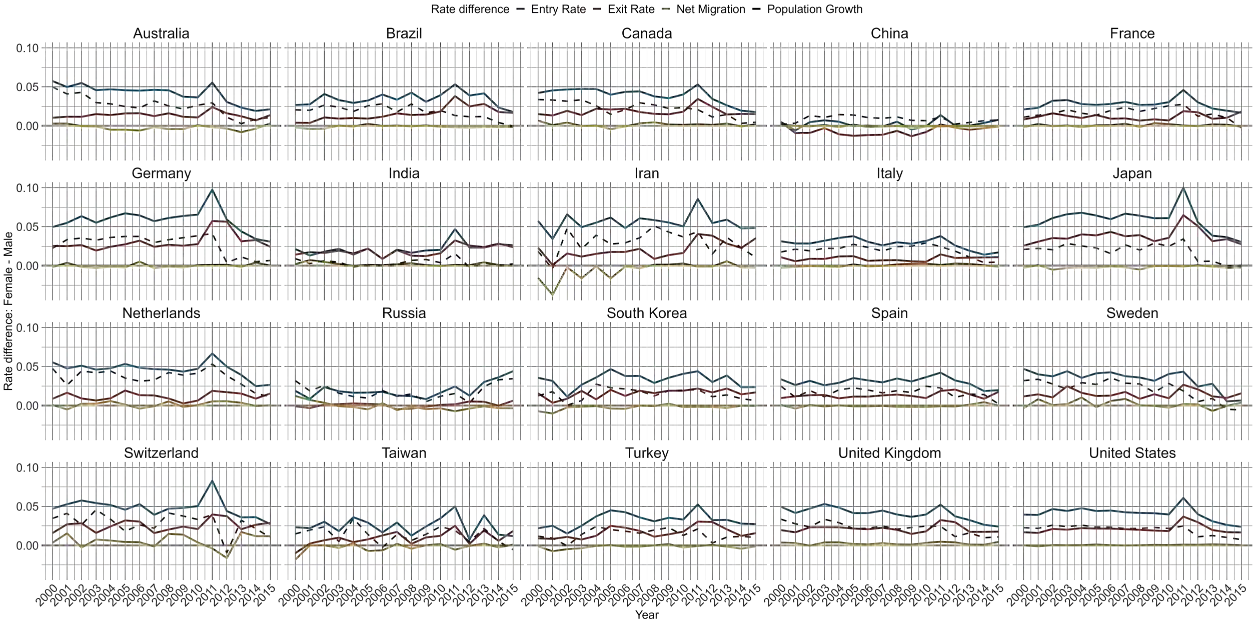
Defining 5-year period of inactivity for right-censoring. We conducted a complementary analysis to set a threshold of period for a scholar's publication trajectory to determine if the scholar was in inter-publication phase or exited. For each discipline, we computed the share of scholars published after five or more years of inactivity. As an example, consider a multi-paper scholar with a set of publication years, represented as: $P = \{p_1, p_2, \ldots, p_n\}$, where p_i represents the year of the i^{th} publication and n is the total number of publications. We calculate the Inter Publication Interval (IPI) as $IPI = \{p_2 - p_1, p_3 - p_2, \ldots, p_n - p_{n-1}\}$. The longest IPI can then be represented as $max(IPI)$. For instance, a scholar with IPI of 6 is an author whose longest period of IPI between an article or review published and the next one was 6 years. Over 92% of the scholars in miscellaneous and engineering and technology had an IPI of 0-5; for natural and agricultural sciences, it was 89%; and for social science and humanities was 85% and 80%, respectively. Therefore, an IPI of 0-5 is a reasonable time window between publications that did not have differential impact across disciplines.



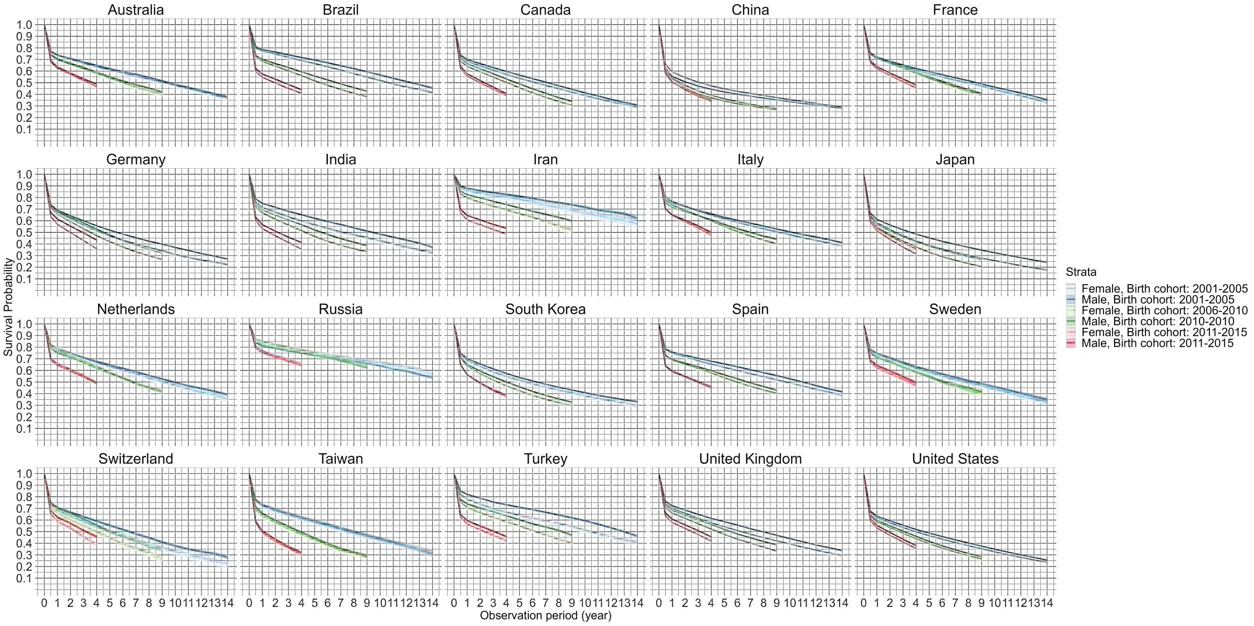
Supplementary Material 1: Percentage of authors grouped by IPI 0-5 or 6+ by disciplines



Supplementary Material 2: Projected median (solid red lines), 80% prediction interval (red dashed lines) and 95% prediction interval (red dotted lines) of female to male difference in population growth rates for 2016-2040 from 2001-2015 (black line) estimation in all 20 countries with the largest scholar population.



Supplementary Material 3: Female to male differences in entry rates (blue), exit (red), net migration (yellow) and population growth (black) rates, 2000-2015 in all 20 countries with the largest scholar population.



Supplementary Material 4: Survival Probability by gender for cohort that started publishing in 2001-2005 (blue), 2006-2010 (green) and 2011-2015 (red), darker lines indicate male, lighter lines indicate female, 2001-2015 in all 20 largest countries.