

## Chapter 3

# The gender publication gap in mathematics: A bibliometric analysis of zbMATH data

Helena Mihaljević, Lucía Santamaría

The achievement of *gender equality and empowerment of all women and girls* is one of the 17 goals listed by the United Nations' 2030 Agenda for Sustainable Development towards a more peaceful, inclusive, equal, prosperous and sustainable world. According to the most recent Global Gender Gap Report 2021 of the World Economic Forum, *the COVID-19 crisis has increased pre-existing gender inequalities*, meaning that “another generation of women will have to wait for gender parity”.<sup>1</sup> Recent investigations have collected evidence that the pandemic has affected female academics in STEMM fields (science, technology, engineering, mathematics, and medicine) particularly hard along multiple dimensions, such as productivity, boundary setting and control, and the ability to engage actively in collaborations and network building.<sup>2</sup> In order to fully understand the gender gap in academia and its development, for instance to assess and counteract the effects of crises such as pandemics, fine-grained data are needed. These typically need to go beyond the often-employed high-level statistics such as those measured by the Global Gender Gap Index applied in the WEF Report.

## 1 The Gender Gap in Science Project

In 2017 eleven scientific organizations, led by the International Mathematical Union (IMU) and the International Union of Pure and Applied Chemistry (IUPAC), joined efforts to conduct an interdisciplinary, cross-national project to gather and analyse comprehensive data on the situation of women in mathematics, computing and natural sciences. The project “A Global Approach to the Gender Gap in Mathematical, Computing, and Natural Sciences: How to Measure It, How to Reduce It?”<sup>3</sup> was funded for the period 2017–2020 by the International Science Council (ISC). Annual

---

<sup>1</sup>[https://www3.weforum.org/docs/WEF\\_GGGR\\_2021.pdf](https://www3.weforum.org/docs/WEF_GGGR_2021.pdf)

<sup>2</sup>E. Higginbotham and M. Lund Dahlberg (eds.), *The impact of COVID-19 on the careers of women in academic sciences, engineering, and medicine*. A Consensus Study Report of the National Academies of Sciences, Engineering, and Medicine. The National Academies Press, Washington, DC, 2021 <https://doi.org/10.17226/26061>

<sup>3</sup><https://gender-gap-in-science.org>

coordination meetings were held by partners to discuss goals, approaches and methodology. A well-attended final conference was organized in November 2019 at the ICTP in Trieste, after which the project's final report was made public.<sup>4</sup>

The Gender Gap project was articulated around three central themes. Besides a Global Survey of Scientists and the creation of a Database of Good Practices, the third working package consisted of the *examination of the situation of academic authors and their publication practices in different academic fields across world countries and regions with respect to the scientists' gender*. This type of analysis makes it possible to identify common, and discipline-specific issues that might require interventions in view of the measured gender gap.

The reason for a focus on publishing practices lay in the importance of publications for academic careers. Scientific publications are not only the major outlet for scholarly communication, they are regarded as a proxy for a researcher's scientific credo and play a key role in achieving and maintaining a successful career in academia. Decisions on tenure and other academic promotions are mostly based on evaluations of the candidate's research portfolio that pay special attention to research publications like journal articles, in addition to grants, conference presentations, and how visible or well-recognized a scholar is. Thus, *the understanding of publication practices, obtained through measurable data on research output, is of great interest to academic institutions, science policymakers, and researchers alike*.

Multiple studies based on bibliometric data have concentrated on the variable of gender. The literature also comprises discipline-specific findings from the area of mathematics, albeit in small numbers. Much of the existing scientometric research builds on cross-discipline corpora such as Scopus and, accordingly, focuses less on individual fields. Research directed to a specific discipline or subfield, in turn, typically examines a limited selection of journals or conferences or a narrow time period. In the aforementioned Gender Gap project, we built on existing results and focused on data sources managed by community organisations and curated by experts, encompassing the respective disciplines as comprehensively as possible in terms of content and temporal coverage. *The analyses of publication behavior in mathematics were performed on zbMATH data, made available to us at regular intervals in order to provide the most up-to-date status of the additional information gathered by the zbMATH office, such as improved author profiles or extracted geo-entities*.

Below we present various key findings from the Gender Gap in Science project related to mathematics. Further results related to gender, as well as to mathematical

---

<sup>4</sup><https://gender-gap-in-science.org/2019/11/09/celebration-of-the-conference-on-global-approach-to-the-gender-gap-in-mathematical-computing-and-natural-sciences-how-to-measure-it-how-to-reduce-it>

publishing in general, plus additional context information e.g. on the data processing algorithms that were employed, can be found in the final project report.<sup>5</sup>

## 2 The Gender Gap in mathematical publications: Cohorts and gender analyses

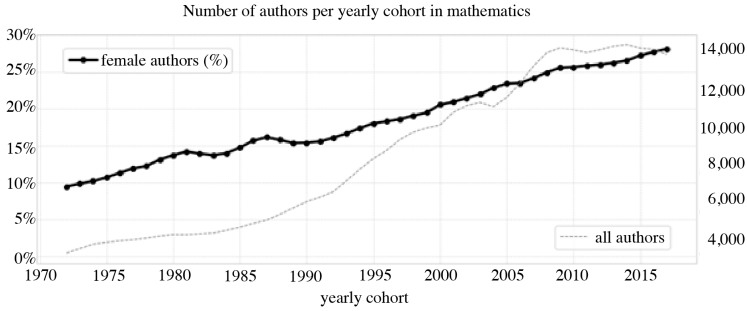
We analysed the full collection of publications by scientists with a main research focus in mathematics (“core mathematicians”) from 1970 until July 2019. This *data set comprises more than 3 million documents corresponding to more than 5.2 million authorships* (pairs of author and document), yielding an average of 1.7 authors per article. We inferred the gender of these authorships from the authors’ names via various statistical name-gender databases and services, resulting in approximately 3.6 million being assigned to men, 0.5 million to women, and 1.2 million that could not be matched to any gender. Omitting authors for which our gender assignment procedure led to no reliable result, authorships of women accounted for about 12% of the total. These<sup>6</sup> in turn belong to ca. 65,000 authors labeled as women and ca. 260,000 authors labeled as men, which yields around 21% women among all recorded authors in zbMATH in the mentioned time span. Figure 1 shows the number of authors according to the year of their first publication (“cohort”), and the percentage of women among them. The proportion of women has increased steadily, growing from less than 10% in the 1970s to over 27% after 50 years. Moreover, nowadays, *more than 14,000 new mathematicians start publishing per year, corresponding to 4,000 women that enter the field of mathematics annually.*

While more and more women become part of academic mathematical research, the question arises how many of them continue to pursue scientific careers in the field several years later. After all, numerous studies show that the percentage of women decreases drastically the higher one looks up the career ladder. Therefore we analysed how many authors “drop out” after a given number of years: we checked, per author and time span, whether each author still appears in zbMATH a number of years after their first publication. Figure 2 visualizes the proportions grouped by cohort and gender for all authors who had been initially active for five years. The assumption of an initial period of five years of activity serves as a proxy for the post-doctoral stage,

---

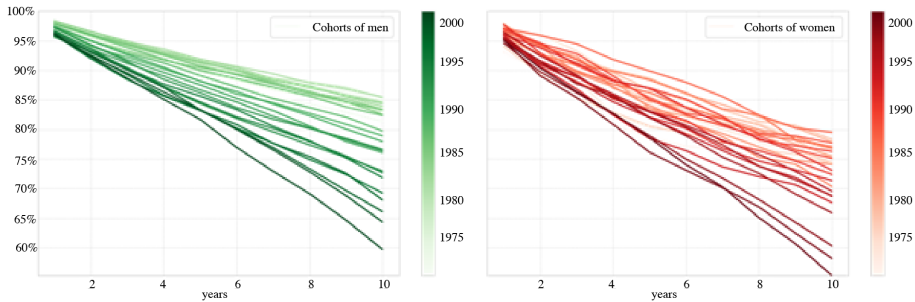
<sup>5</sup>M.-F. Roy, C. Guillopé, M. Cesa, R. Ivie, S. White, H. Mihaljevic, L. Santamaría, R. Kelly, M. Goos, S. Ponce Dawson, I. Gledhill, and M.-H. Chiu, A global approach to the gender gap in mathematical, computing, and natural sciences: How to measure it, how to reduce it? International Mathematical Union (2020) <https://doi.org/10.5281/zenodo.3882609>

<sup>6</sup>Not all authorships can be assigned to a unique author, in particular if the author’s name is frequent.



**Figure 1.** (Dotted grey line; right axis) Defining a zbMATH author’s cohort as the year of their first publication, number of authors found in the database per cohort from 1970 until 2017. (Solid black line; left axis) Percentage of all authors that could be algorithmically assigned as female.

thus the figure implies the following: the number of authors that stay in academia further 6 to 10 years has reduced enormously when comparing the 1970s cohorts with those from the 2000s. If we associate the subsequent 10-year period with the time when a permanent academic position is secured, then around 60% of the male “post-docs” from the most recent cohorts manage to achieve such a career milestone. For women, the percentages have been, and continue being, lower than for men. However, the differences between women and men have reduced over time. Likely, this is mainly due to the fact that the number of PhD students and post-docs has grown much faster than the available permanent positions in mathematical research.



**Figure 2.** Percentage of male (left) and female (right) mathematicians that continue publishing for another 1 to 10 years after having been active for 5 years. The colors indicate cohorts, with dark colors indicating the most recent ones. The figure exposes a “publishing drop-out rate” in mathematics throughout the past four decades.

### 3 The Gender Gap in renowned mathematical journals

As already mentioned, scholarly journals are a crucial vehicle for the forging of academic careers in STEMM, as decisions on tenure, funding, and promotions strongly depend on the researchers' publication record. Moreover, it is not just the number of articles a scientist writes that matters, but also the venue where they appear. Publishing in highly renowned journals in one's discipline is a powerful determinant of tenure in many STEMM fields including mathematics, and an important predictor of professional success. Thus, any bibliometric study on publication practices ought to take into account their impact in the making of academic careers.

In previous research,<sup>7</sup> also based on zbMATH data, we had already demonstrated that authorships by women are vastly underrepresented in journals with a high reputation in terms of two common ranking methods, the manually compiled Australian ERA indicator and the journal impact factor (JIF). In this project, we intended to offer the scientific community the opportunity of examining gender distributions in journals of particular relevance to them or their subfield. We made this possible via a dedicated web interface that allows readers to filter specific publication venues of their interest.

Additionally, we have taken *a close look at selected journals published by mathematical societies as well as those particularly renowned in individual topical subfields*. Figure 3 illustrates that the percentages of authorships from women in said selected journals are predominantly constrained below 20%. Around half of the society journals show a rising tendency over the past decades. The *Bulletin de la Société Mathématique de France* shows a rather noisy behavior and no clear chronological trend, with close to no publications by women at all in various years. The average share is around 10%, similar to the *Journal of the European Mathematical Society*. The lowest percentages are found in the *Journal of the American Mathematical Society*, where the proportion of women is around 5% or less, and shows no noticeable increase over time. The bottom three topical journals on the right-hand column, which mainly feature works in areas of applied mathematics, display a rising development over time with shares above 10% in recent years. Except for the *Journal of Differential Geometry*, all journals reveal a slight positive trend. The renowned journals *Inventiones Mathematicae* and *Annals of Mathematics*, which for the most part publish work in pure mathematics, stand out with percentages of women authorships predominantly in the single-digit range.<sup>8</sup>

---

<sup>7</sup>H. Mihaljević–Brandt, L. Santamaría, and M. Tullney, The effect of gender in the publication patterns in mathematics. PLOS ONE 11 (10): e0165367 (2016) <https://doi.org/10.1371/journal.pone.0165367>

<sup>8</sup>For more details, see H. Mihaljević and L. Santamaría, Authorship in top-ranked mathematical and physical journals: Role of gender on self-perceptions and bibliographic evidence. Quantitative Science Studies 1 (4): 1468–1492 (2020) [https://doi.org/10.1162/qss\\_a\\_00090](https://doi.org/10.1162/qss_a_00090)

There may be several potential causes for the measured underrepresentation, but these cannot be determined from the bibliographic data. As an alternative data source we have leveraged the 2018 Global Survey of Mathematical, Natural, and Computing Scientists that was conducted as another working package of the project to obtain answers from almost 10,000 mathematicians, physicists, and astronomers about their submission practices to top-ranked journals in their disciplines. More precisely, we asked the following question: “*During the last five years, how many articles have you submitted to journals that are top-ranked in your field?*” Respondents were expected to provide a number between 0 and 30; larger values were clustered together. According to the obtained responses, women and men self-report to have submitted similar numbers of articles in the past 5 years, with no major statistically significant differences in subgroup analyses broken down by disciplines or world regions. What matters much more than gender in the computed model is strong research activity, a professional network, and overall academic success.

The reported perceived submission practices do not support the hypothesis that the underrepresentation of women in prestigious journals is mainly rooted in them submitting less manuscripts for consideration than men. Considering the importance of publishing in renowned journals on the one hand and the conflicting bibliographic analysis on the other, this begs the question on the role of peer review. We observe that the refereeing system in mathematics lacks homogeneity and relies substantially on the authors’ credit and the level of trust between editors and reviewer(s). In this regard, we stress that *there are hardly any systematic studies on the peer review process in mathematics*,<sup>9</sup> a need that very much ought to be addressed.

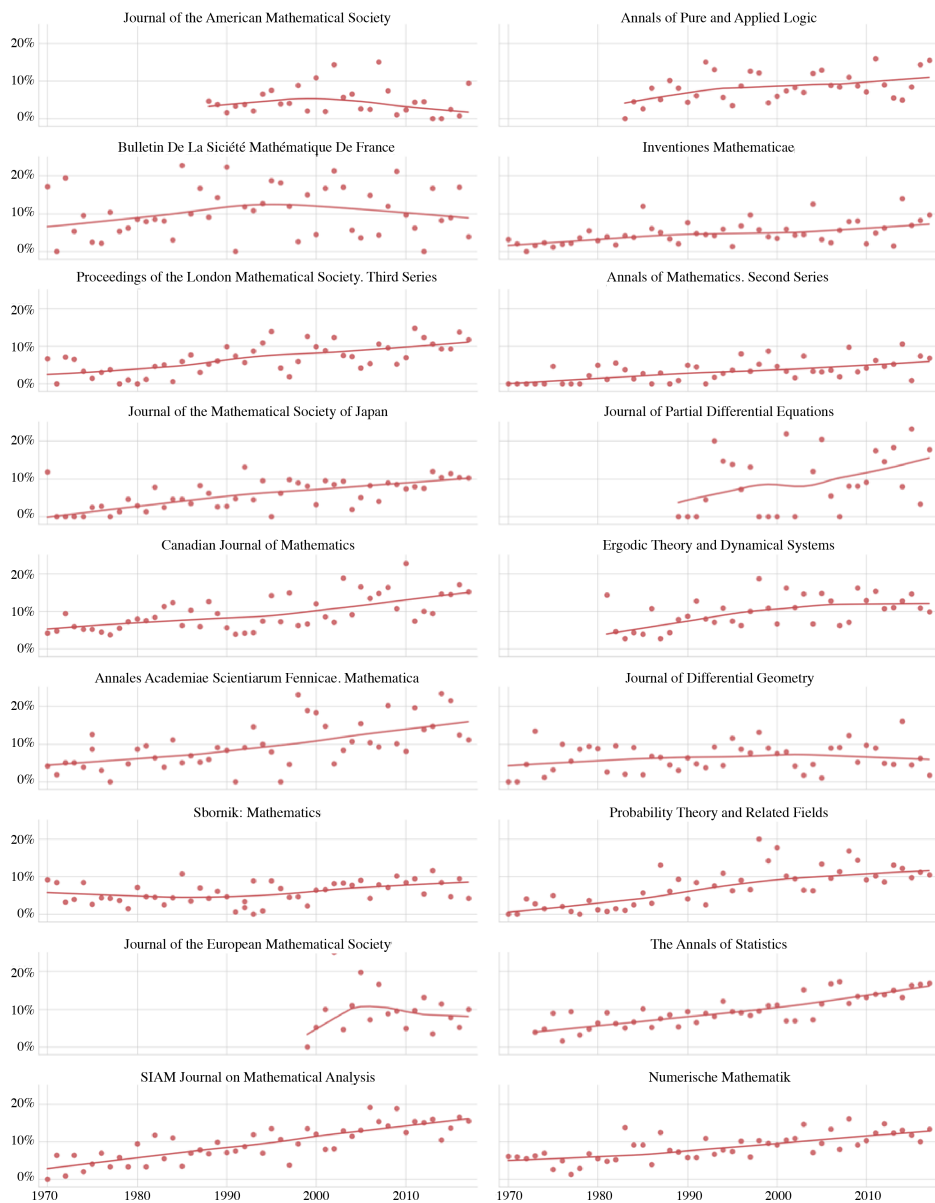
#### **4 Learnings and perspectives on the Gender Gap in mathematical publications**

Inspired by the UN’s agenda to reach gender equality and empowerment of all women and girls within the next decade, we set out to investigate the existence and characteristics of a particular gender gap: the *underrepresentation of female authors in academic publishing in mathematics with respect to their male counterparts*. The comprehensive data collection from zbMATH as well as our usage of algorithmic methods at scale make this bibliometric analysis feasible.

There are various aspects to consider when speaking of a gender gap. We have provided insights on the gap defined by the proportional presence of women as authors

---

<sup>9</sup>C. Geist, B. Löwe and B. Van Kerkhove, Peer review and knowledge by testimony in mathematics. In *PhiMSAMP: Philosophy of mathematics: Sociological aspects and mathematical practice*, pp. 155–178. London, College Publications, 2010.



**Figure 3.** Percentage of authorships from women in renowned mathematics journals per year between 1970 and 2017.

of core mathematics publications; we have also investigated whether there is a gender gap in the dropout rates that affect the length of mathematicians' academic careers; finally, we have focused on the gender gap in renowned, high-impact mathematical journals.

Consistent with the global trend in higher education, we observe *increasing proportions of women entering the field of mathematics with each passing year*. The understanding of the extent to which those newcomers will progressively attain senior academic positions is crucial to address the “leaky pipeline” phenomenon. Thanks to our cohort analysis based on zbMATH publication data, we are able to provide insights on this issue. We show that dropout rates of mathematicians after their post-doctoral stage, which used to be higher for women, are converging on similar figures for both genders. These data certainly offer optimistic prospects regarding the eventual closure of this particular aspect of the gender gap.

On the other hand, our analysis of women's presence in renowned journals is a good measure of the gender gap in relation to achieving a prestigious academic career. In this regard, a non-negligible number of the *prestigious mathematical journals under consideration show a meager representation of women among their authors*. All other factors being equal, the expectation is that the proportion of women among all authors should roughly resemble the percentage of established female mathematicians in the profession, a number that has been steadily growing and that is estimated to be currently around 25%. Remarkably, several of the analysed journals publish very few articles authored by women and exhibit no signs of turnaround over the last couple of decades. An explanation for this fact might lie in the characteristics of the peer review process in mathematics, which favors close interactions and trust relationships between editors and reviewers and opens the door to conscious and unconscious biases. Regarding subfields, applied areas display a better situation for women than pure ones, which in itself introduces a series of discussion points regarding the intrinsic differences among subfields of mathematics.

The above remarks provide a compelling starting point for future research questions. Is the increasing number of young female mathematicians enough to stop the pipeline from leaking? Which factor in the retention of women in academia is played by the professional atmosphere in pure versus applied mathematics? What is the importance of informal academic networks to make a mathematician's career thrive? Is the lack of double-blindness in peer review hindering women and other underrepresented groups in mathematics? *It would be excellent to discuss our data-backed findings with experts from the respective subfields in the mathematical community*, with the goal of formulating plausible hypotheses that could explain the observations found by our work in the Gender Gap project.