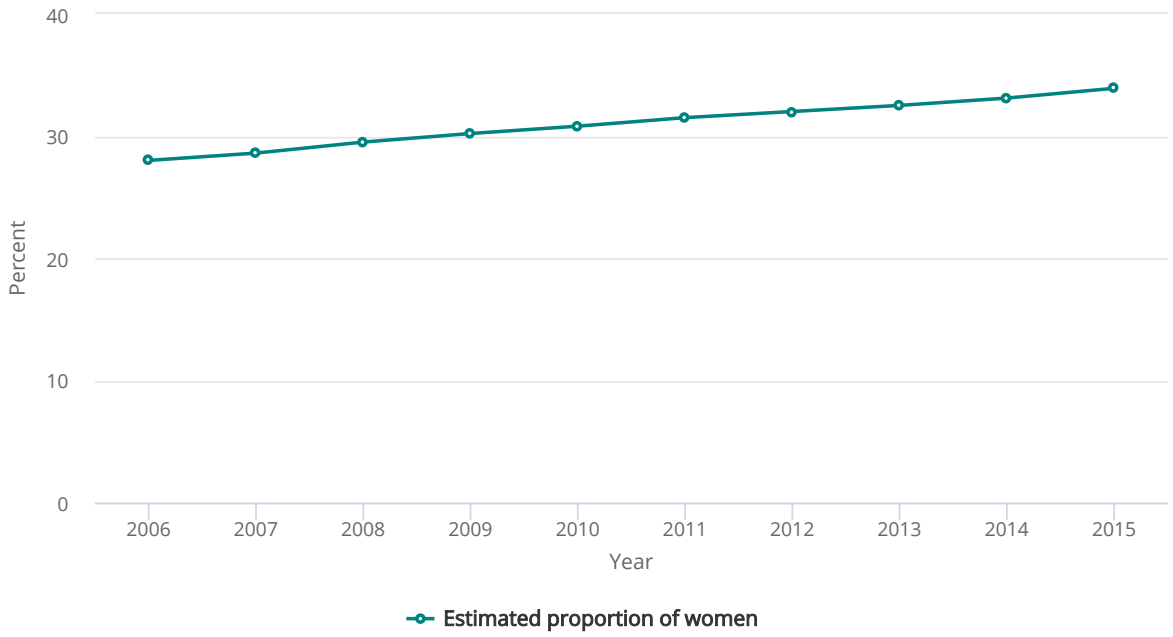
SIDEBAR

S&E Publication Patterns, by Gender

Recently, researchers have begun matching publication authorship data with name databases containing gender identifiers such as Social Security (Lariviére et al. 2013) or Genderize.io (Elsevier 2017). This sidebar summarizes research from Science-Metrix using Scopus bibliometric data and Namsor for estimating author gender (Science-Metrix 2017b).* Because of the various sources of uncertainty associated with matching name with gender, the analyses presented in this sidebar should be seen as exploratory research and interpreted with caution. Researchers are finding that although male authors historically comprised a larger share of peer-reviewed scientific publications, female authorship is growing. Science-Metrix estimates that from 2006 to 2015, female scientific authorship increased over 20% and reached nearly 34% in 2015 (Figure 5-G). Other researchers (Lariviére et al. 2013) have coined the term “productivity paradox” to discuss the current phenomenon where men publish more papers on average, even though there are more female than male undergraduate and graduate students in many countries. Gender balance is said to occur when women make up 40%–60% of any group (European Commission 2015).

FIGURE 5-G

Trends in the proportion of female authors of S&E publications in Scopus: 2006–15



Note(s)

Data are presented according to publication year. The 2015 data are preliminary and do not represent total 2015 publications.

Source(s)

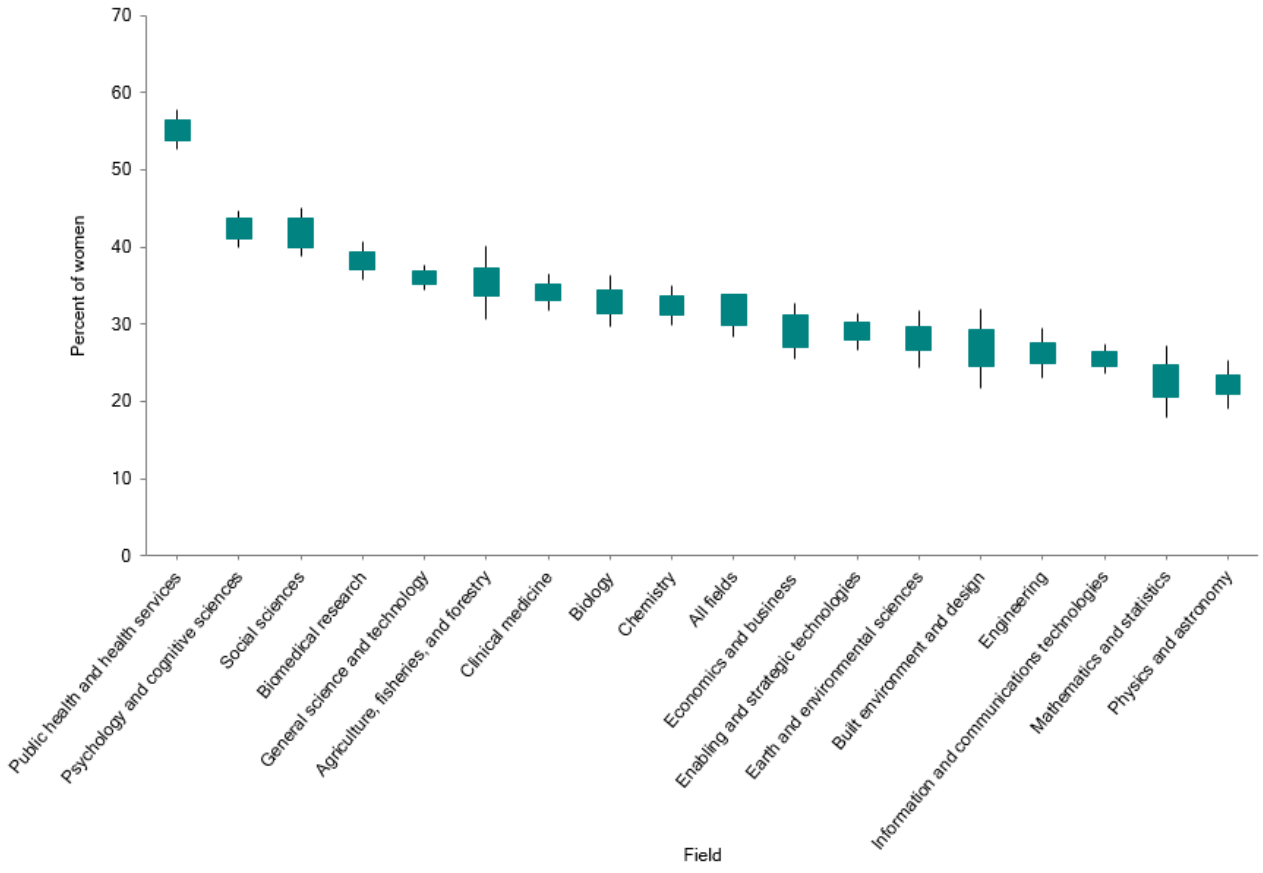
Science-Metrix; NamSor, accessed October 2016; Elsevier, Scopus abstract and citation database, accessed August 2016.

Science and Engineering Indicators 2018

The gender of authors differs across research fields.[†] Science-Metrix finds that over the last decade, the research field with the highest proportion of female authorship is public health and health services; this is also the only field where women represent a larger share of authorship than men (Figure 5-H). Psychology and cognitive sciences and social sciences are all at or above the 40% mark. Six other fields are above 30%, and the remaining 8 fields are below 30% (which is the overall average across Scopus for the 2006–15 period).

FIGURE 5-H

Proportion of female authors of S&E publications, by field: 2006–15



Note(s)

Rectangle is the 95% confidence interval for NamSor only. The line is the 95% confidence interval due to the sampling error.

Source(s)

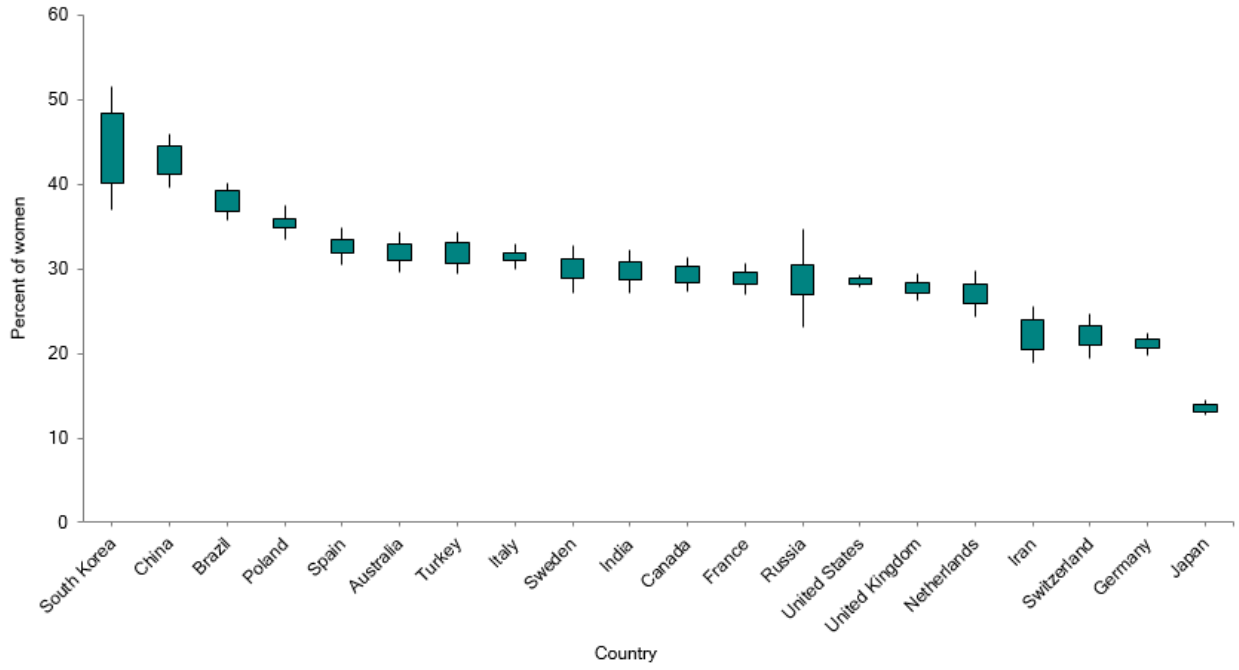
Science-Metrix; NamSor, accessed October 2016; Elsevier, Scopus abstract and citation database, accessed August 2016.

Science and Engineering Indicators 2018

Among the 20 countries that produce the most academic articles, female authorship in Scopus (over the 2006–15 period) measured highest for South Korea and China, which all scored above the 40% mark (Figure 5-I). In the United States, women account for just under 30% of U.S. authorship. Two countries, South Korea and China, achieve gender parity with 40% or more of women in authorship, 6 more are near parity with above 30%, 11 more (including the United States) are above 20%, and Japan is around 13%.

FIGURE 5-I

Proportion of female authors of S&E publications, by country: 2006–15



Note(s)

The rectangle is the 95% confidence interval for NamSor only. The line is the 95% confidence interval due to the sampling error. Twenty leading academic article-producing countries are included.

Source(s)

Science-Metrix; NamSor, accessed October 2016; Elsevier, Scopus abstract and citation database, accessed August 2016.

Science and Engineering Indicators 2018

These results summarize Science-Metrix research using a name-based gender identification tool in conjunction with the Scopus database to estimate a probability for correctly identifying gender, based on an author's full name (given name and family name, though given name is usually more informative). For example, Helen is clearly female, but Riley is 40% male. The name-gender matched data set therefore introduces some uncertainties into the analysis because some names are not gender specific, especially for countries in East Asia, and those authorships that cannot be tagged with a high degree of certainty are removed from the analyzed population.[‡]

Further uncertainties exist in the analysis because Scopus has some coverage limits in non-English speaking countries (e.g., Russia), and some fields or editorial policies promote the use of initials rather than an author's given name (astronomy). Over time, the availability of full first names (as opposed to initials only) has risen in the Scopus database, approaching 70% in 2002 and 2003, and has been consistently close to 80% since 2013.

[*] The Namsor tool is used to code gender and provides a probability for correctly identifying gender (http://www.namsor.com/).

[†] Note that definitions of fields and countries are those of Science-Metrix, not those used elsewhere in *Indicators.*

[‡] To create estimates of sampling error for these proportions, Science-Metrix assumes that the authors who are identified by gender in the database are fully representative of the author population as a whole, including those for whom gender

cannot be determined. Confidence interval estimates have been calculated under this assumption, but major differences between the gender distribution of those whose gender can be identified and those that are of unknown gender could result in other sources of uncertainty.