9^{th} International Conference on Computational Social Science IC^2S^2
July 17-20, 2023, Copenhagen, Denmark

Gender differences in self-promotion on Twitter among early-career researchers: Are men more likely to disseminate their first publication and benefit more?

Keywords: Gender differences; self-promotion; Twitter; Altmetrics; Science of science

Extended Abstract

Over the past decades, Twitter, Facebook, LinkedIn and other social media platforms have been widely leveraged by researchers for scholarly purposes. The widespread availability of social media has significantly changed scholarly communication, enabled researchers to more widely disseminate their research, and facilitated interaction among researchers worldwide. For early-career researchers in particular, promoting themselves in the academic community is particularly important since being visible is the first step for distributing their findings and leading to a growing network, which is essential for a long-term career in academia.

Social media offers a promising opportunity for early-career academics to engage in dissemination of their research in a faster and easier way [Vásárhelyi et al., 2021, Yammine et al., 2018]. In addition, social media has the potential to serve an important role in the movement toward increased equity, diversity, and inclusion within academia as it provides a widely available and readily accessible platform to scholars, including underrepresented groups, largely reducing the time and financial cost for participation [Yammine et al., 2018]. However, despite this potential for equalizing participation, recent research has shown that social inequalities are often reproduced online. Research has shown that female researchers receive less online visibility than their male counterparts. Partly because women are significantly less likely than men to self-promote their papers [Vásárhelyi et al., 2021, Peng et al., 2022]. Whether such gender gaps are visible already within the early stages of an academic career and the subsequent impacts of self-promotion in early career stages however are less well understood.

To fill this gap, this study uses large-scale bibliometric data from Scopus combined with Altmetrics as well as the Twitter's public API. Our analysis proceeds in two steps: (1) we first estimate gender differences in the probability to promote author's first publication on social media; (2) we compare the subsequent impact of this self-promotion in early stages of the career on citation counts. We first identified all authors who published their first publication during the period 2012-2016 and received mentions on their first publication on Twitter. Furthermore, we judged whether they had self-promoted their first publication by comparing their names retrieved from Scopus and the names of the tweeters who mentioned the publication (see Fig. 1 for the matching workflow). Among 1,232,080 Scopus-published researchers from the cohort of 2012-2016 (defined using the publication year of the first paper), 370,560 (30.08%) researchers were detected to receive Twitter mentions on their first publication and 10,183 (0.83%) researchers had promoted their first publication themselves. By using a systematic process composed of name-gender dictionaries and deep learning methods and a category of six macro research fields of speciality, we identified the genders and disciplines of these researchers. Descriptive statistics on the sample and self-promotion rates are shown in Table. 1.

To model the probability of authors promoting their first publication online and to examine whether it differs by gender, we employed a mixed-effects logistic regression, incorporating

random effects for country variability in the intercept and slope. Our main variable of interest is gender, and its interactions with other covariates to explore heterogeneity in gender differences. Our models include the author's cohort, the field of speciality, the number of authors in the first publication, the author order (first, middle, or last), the ranking quantile of the published journal in the subject area, as well as whether their publication has been mentioned on Twitter first by others as control variables.

Fig. 2 shows the marginal effects of gender (1 = male, 0 = female) in the probabilities of self-promoting the first publication among researchers from the estimated model, by cohort, discipline, journal rank and author order. The marginal effects can be interpreted as the gender gap, i.e. the difference between male and female predicted probabilities of self-promotion. Across all categories, male researchers have higher probability to promote their first publication. The gender gap in self-promotion increases across cohorts, from 0.2 percentage points (p.p) in the cohort of 2012 to 0.35 p.p in 2016. Disaggregating by six macro fields of speciality, the gender gap in self-promotion is larger for those working in Social Sciences and Humanities compared to those in Engineering and Technology. Overall levels of self-promotion are also higher in the Social Sciences and Humanities across both genders compared with STEM fields. We can also see that male researchers are more likely to promote their first paper when the paper was published in a Q1 (top 25%) journal or when they were the first author.

To explore the impact arising from self-promotion among early-career researchers on the publication's citation count, we used propensity score matching (psm) to match researchers who promoted their first publication with those who did not. As factors in the matching process, we considered gender, cohort, discipline, the journal rank of their first publication, affiliated country, the authors' order in the publication, and the number of authors. This gives us 10,179 pairs of researchers (20,358 individuals). We further confined each pair of researchers who have the same gender, are from the same cohort, do research in the same speciality and publish their first article in the journals ranking at the same quantile. And we excluded the publications co-authored by over 15 authors, and get 7,880 pairs (15,760 individuals). Among them, 2,996 pairs (5,992 individuals, 38%) are females. We compared the impact of self-promotion on the citations they received on their first publication, shown in Fig. 3. Self-promotion increases the citation count of the first publication for both female and male researchers. This positive effect on citation is larger for researchers who are from the earlier cohorts (i.e. 2012-2014), those in the field of Natural Sciences and those who published their first paper in a Q1 journal. Despite the significant increase in citation for both female and male researchers, the payoff of self-promotion is higher for male researchers. These results suggest a process of a cumulative disadvantage starting early in scientific careers: first, we find men are more likely to promote their first publication, with subsequent citation impacts also being larger for men.

References

H. Peng, M. Teplitskiy, D. M. Romero, and E.-A. Horvát. The gender gap in scholarly self-promotion on social media, 2022. URL `https://arxiv.org/abs/2206.05330`.

O. Vásárhelyi, I. Zakhlebin, S. Milojevic, and E.-A. Horvát. Gender inequities in the online dissemination of scholars' work. *Proceedings of the National Academy of Sciences*, 118:e2102945118, 09 2021. doi: 10.1073/pnas.2102945118.

S. Z. Yammine, C. Liu, P. B. Jarreau, and I. R. Coe. Social media for social change in science. *Science*, 360 (6385):162–163, 2018. doi: 10.1126/science.aat7303. URL `https://www.science.org/doi/abs/10.1126/science.aat7303`.

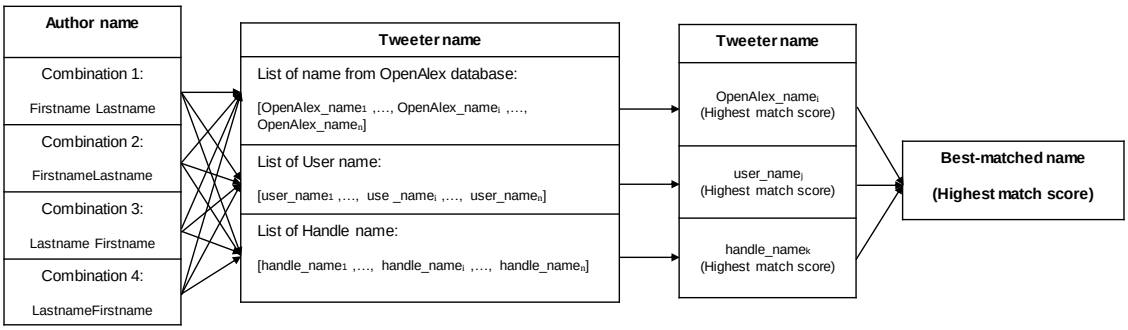

Figure 1: The workflow of matching Scopus-published researchers and Tweeters. The OpenAlex name is acquired from: Mongeon, P., Bowman, T. D., & Costas, R. (2022). An open dataset of scholars on Twitter (arXiv:2208.11065).

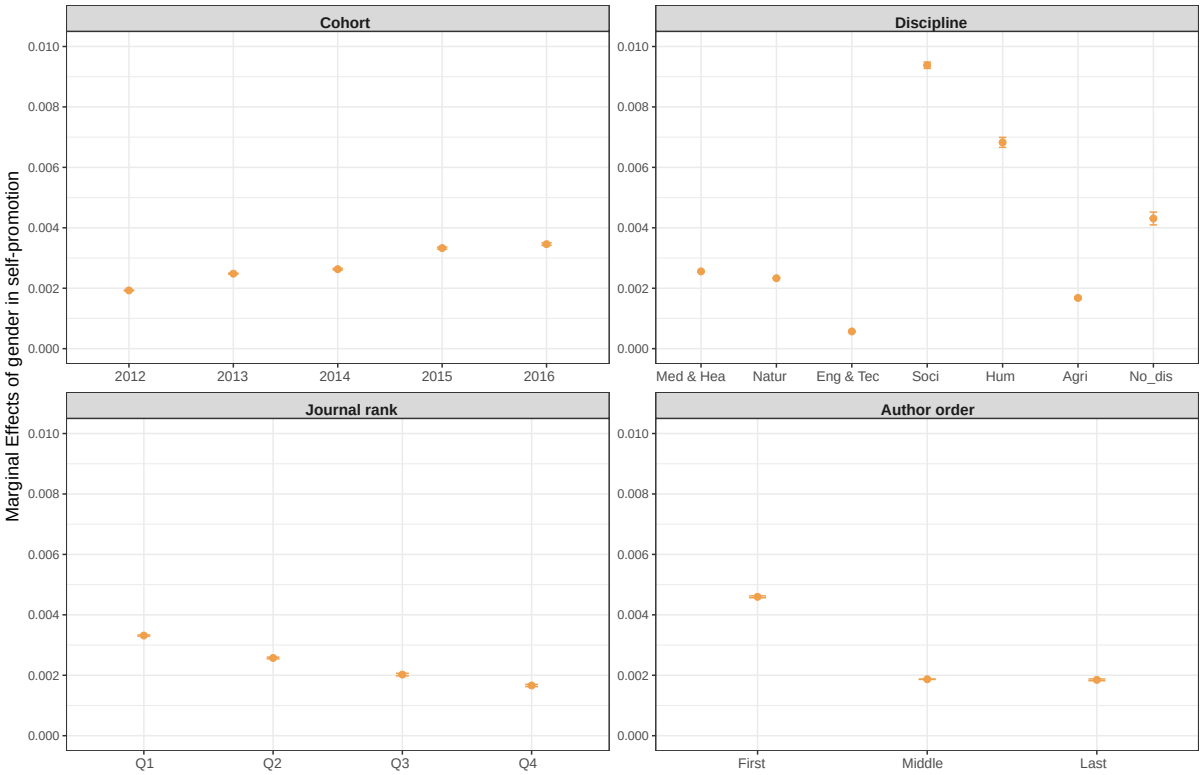

Figure 2: Marginal effects of gender (gender gap, i.e. difference between male and female probabilities) on self-promotion of the first publication, by cohort, discipline, journal rank and author names' order. The prediction of self-promotion is based on a mixed-effects logistic regression model with this formula: $logit(P(Y_{ij}=1)) = \beta_0 + \beta_{0j} + (\beta_1 + \beta_{1j} \times X_i) \times Country_j + e_{ij}$, where $X_i = Gender_i + Z_i + Gender_i \times Z_i$. $logit(P(y_{ij}=1))$ indicates the probability of self-promoting the first publication for the researcher i in the country j, by controlling the variations in country ($Country_j$). In addition to the variable of interest, i.e. gender ($Gender_i$), we also consider these control variables: author's cohort (2021 is the reference level), the field of speciality (without a discipline assignment is the reference level), the number of authors in the first publication, the authors names' order (first, middle, or last, the middle position is the reference level), the ranking quantile of the published journal in the subject area (the Q4 is the reference level), and whether their publication has been mentioned on Twitter first by others (not being tweeted is the reference level). We also included interactions between gender and other control variables in the model.

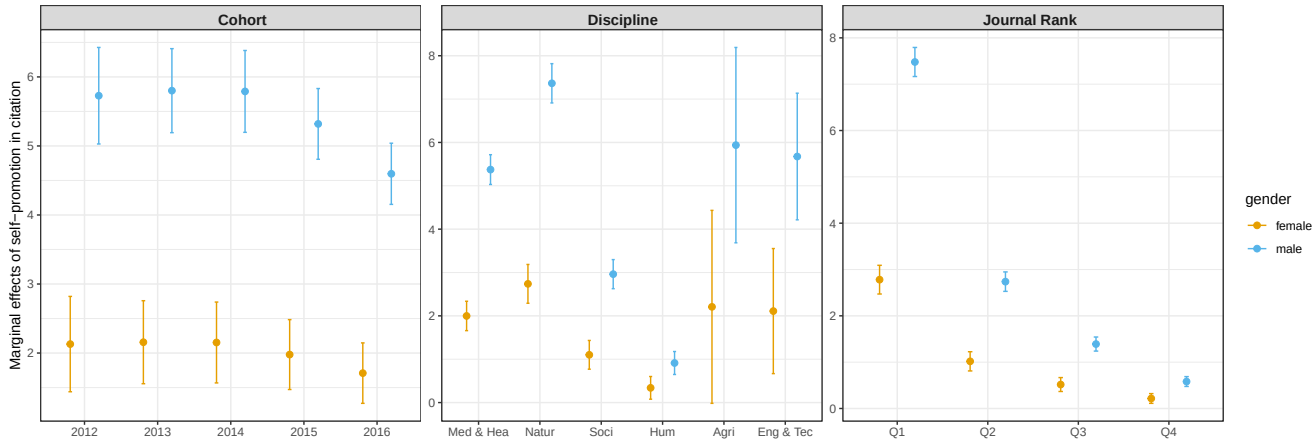

Figure 3: Marginal effects of self-promotion in the citations count by gender (gender gap, i.e. difference between male and female probabilities), at the categories of cohort, discipline, and the journal rank. The prediction of the citation counts by gender is based on a negative binomial regression model with this formula: $Y_i = \beta_0 + \beta_{gender} \times Gender_i + \beta_{pro} \times \text{Self promotion}_i + \beta_{gender \times pro} \times Gender_i \times \text{Self promotion}_i + \beta_z \times Z_i + e_i$. Where Y_i is the citations count for the researcher's first publication, and Z_i represents the control variables including author's cohort (2021 is the reference level), the field of speciality (without a discipline assignment is the reference level), the ranking quantile of the published journal in the subject area (the Q4 is the reference level) and whether being mentioned by others on Twitter (no others' mention is the reference level).

Table 1: The count and percentage of the published researchers and those who self-promoted their first publication by cohort (2012-2016), gender and discipline. Note: the few cases where female researchers have the lowest difference or outnumber male researchers are indicated with a **bold** font.

		Number of newly-published authors		**Number of self-promoted authors** (% among published authors)	
Cohort	**Discipline**	**Female**	**Male**	**Female**	**Male**
2012	Agricultural Sciences	3,691	4,468	12 (0.33%)	10 (0.22%)
	Engineering and Technology	5,910	17,959	8 (0.14%)	22 (0.12%)
	Humanities	1,503	1,855	21 (1.4%)	37 (1.99%)
	Medical and Health Sciences	50,568	55,554	262 (0.52%)	316 (0.57%)
	Natural Sciences	43,786	80,293	143 (0.33%)	372 (0.46%)
	Social Sciences	**8,233**	8,035	110 (1.34%)	176 (2.19%)
	No discipline assigned	425	757	6 (1.41%)	10 (1.32%)
	Total	114,116	168,921	562 (0.49%)	943 (0.56%)
2013	Agricultural Sciences	3,556	3,958	16 (0.45%)	20 (0.51%)
	Engineering and Technology	5,612	16,913	9 (0.16%)	17 (0.1%)
	Humanities	1,422	1,843	20 (1.41%)	41 (2.22%)
	Medical and Health Sciences	47,014	54,037	266 (0.57%)	381 (0.71%)
	Natural Sciences	40,265	73,909	196 (0.49%)	431 (0.58%)
	Social Sciences	7,071	7,421	154 (2.18%)	231 (3.11%)
	No discipline assigned	609	991	**14 (2.3%)**	8 (0.81%)
	Total	105,549	159,072	675 (0.64%)	1,129 (0.71%)
2014	Agricultural Sciences	2,624	3,392	11 (0.42%)	16 (0.47%)
	Engineering and Technology	5,649	17,101	7 (0.12%)	36 (0.21%)
	Humanities	1,358	1,713	38 (2.8%)	49 (2.86%)
	Medical and Health Sciences	41,769	50,910	318 (0.76%)	416 (0.82%)
	Natural Sciences	36,226	68,883	214 (0.59%)	477 (0.69%)
	Social Sciences	**6,770**	6,773	178 (2.63%)	224 (3.31%)
	No discipline assigned	410	786	6 (1.46%)	14 (1.78%)
	Total	94,806	149,558	772 (0.81%)	1,232 (0.82%)
2015	Agricultural Sciences	2,531	3,264	12 (0.47%)	20 (0.61%)
	Engineering and Technology	6,112	17,178	7 (0.11%)	38 (0.22%)
	Humanities	1,289	1,583	43 (3.34%)	70 (4.42%)
	Medical and Health Sciences	37,284	45,581	353 (0.95%)	436 (0.96%)
	Natural Sciences	33,009	62,639	246 (0.75%)	534 (0.85%)
	Social Sciences	6,504	6,737	213 (3.27%)	278 (4.13%)
	No discipline assigned	517	883	8 (1.55%)	18 (2.04%)
	Total	87,246	137,865	882 (0.81%)	1,394 (1.01%)
2016	**Agricultural Sciences**	2,392	3,132	**24 (1%)**	20 (0.64%)
	Engineering and Technology	5,963	16,969	14 (0.23%)	47 (0.28%)
	Humanities	1,271	1,538	49 (3.86%)	79 (5.14%)
	Medical and Health Sciences	35,291	43,381	441 (1.25%)	543 (1.25%)
	Natural Sciences	31,600	59,310	299 (0.95%)	576 (0.97%)
	Social Sciences	5,787	6,068	205 (3.54%)	265 (4.37%)
	No discipline assigned	815	1,430	7 (0.86%)	25 (1.75%)
	Total	83,119	131,828	1,039 (1.25%)	1,555 (1.18%)