nature neuroscience

SUPPLEMENTARY INFORMATION

https://doi.org/10.1038/s41593-020-0658-y

In the format provided by the authors and unedited.

# The extent and drivers of gender imbalance in neuroscience reference lists

**Jordan D. Dworkin[1], Kristin A. Linn[1], Erin G. Teich[2], Perry Zurn[3], Russell T. Shinohara[1] and Danielle S. Bassett[2,4,5,6,7,8]✉**

[1]Penn Statistics in Imaging and Visualization Center, Department of Biostatistics, Epidemiology and Informatics, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, USA. [2]Department of Bioengineering, School of Engineering and Applied Science, University of Pennsylvania, Philadelphia, PA, USA. [3]Department of Philosophy and Religion, American University, Washington, DC, USA. [4]Department of Physics & Astronomy, College of Arts and Sciences, University of Pennsylvania, Philadelphia, PA, USA. [5]Department of Electrical & Systems Engineering, School of Engineering and Applied Science, University of Pennsylvania, Philadelphia, PA, USA. [6]Department of Neurology, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, USA. [7]Department of Psychiatry, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, USA. [8]Santa Fe Institute, Santa Fe, NM, USA.
✉e-mail: dsb@seas.upenn.edu

# Supplementary Information for "The extent and drivers of gender imbalance in neuroscience reference lists"

Jordan D. Dworkin[1], Kristin A. Linn[1], Erin G. Teich[2], Perry Zurn[3], Russell T. Shinohara[1] & Danielle S. Bassett[2,4,5,6,7,8]

[1]Penn Statistics in Imaging and Visualization Center, Department of Biostatistics, Epidemiology, and Informatics,Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, USA
[2]Department of Bioengineering, School of Engineering and Applied Science, University of Pennsylvania, Philadelphia, PA, USA
[3]Department of Philosophy and Religion, American University, Washington DC, USA
[4]Department of Physics & Astronomy, College of Arts and Sciences, University of Pennsylvania, Philadelphia, PA, USA
[5]Department of Electrical & Systems Engineering, School of Engineering and Applied Science, University of Pennsylvania, Philadelphia, PA, USA
[6]Department of Neurology, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, USA
[7]Department of Psychiatry, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, USA
[8]Santa Fe Institute, Santa Fe, NM, USA

***Validation of gender determination.*** Because of the degree to which the results presented in this work rely on accurate estimation of author genders, it was vital to understand the performance of the gender determination methods in this data. To test the performance of the gender assignments, a random sample of 100 papers were drawn from the full data, yielding 200 first and last author names. For each of these 200 authors, manual gender determination was conducted by assigning inferred gender via publicly available pronouns (e.g., on lab websites or social media pages). Although authors who use 'she'/'her'('he'/'him') may identify as any number of genders, their gender is most likely to be 'woman'('man') and least likely to be 'man'('woman'). For authors without publicly available pronouns, a gender determination was conducted by assigning inferred gender via publicly available photographs. Although authors whose gender expression most approximates 'woman'('man') may identify as any number of genders, their gender is most likely to be 'woman'('man') and least likely to be 'man'('woman'). Authors for whom photographs were ambiguous or unavailable, or whose names could not be reliably matched to one scholar, were marked as 'unknown.' Relative to this manual assignment, the accuracy of the automated gender assessment method can be seen at the author level in Table S1, and the article level in Table S2.

| | Manual assessment | | |
|---|---|---|---|
| Automated assessment | Man | Woman | Unknown |
| Man | 140 | 5 | 7 |
| Woman | 2 | 42 | 4 |

**Table S1. Accuracy of automated gender determination technique in a random sample of 200 authors**.

| | Manual assessment | | | | | |
|---|---|---|---|---|---|---|
| Automated assessment | MM | WM | MW | WW | Unknown+M | Unknown+W |
| MM | 51 | 1 | 1 | 0 | 6 | 0 |
| WM | 0 | 18 | 0 | 1 | 3 | 0 |
| MW | 1 | 0 | 8 | 2 | 0 | 1 |
| WW | 0 | 0 | 1 | 5 | 0 | 1 |

**Table S2. Accuracy of automated gender determination technique for assignment of papers to one of four gender categories in a random sample of 100 papers**.

***Imputation of missing data under the null.*** The analyses presented in the main text were conducted on the 88% of papers ($n = 54,226$) for which gender could be reliably assigned to both the first and last author. The remaining 12% of papers, for which the gender of at least one author was unknown, were excluded from the primary analyses. We conducted a secondary analysis in which missing data were imputed and estimates were computed on the full data.

To impute the missing author genders, the gender probabilities obtained from the GAM were used. This model estimated probabilities that a given paper was MM, WM, MW, or WW based on its year of publication, the number of authors, the seniority of the authors, the journal in which it was published, and whether it was a review article. If neither of a paper's authors had an assigned gender, a gender category was randomly assigned to it using the four GAM-estimated probabilities. If one of the paper's authors had an assigned gender, one of the two possible gender categories was randomly assigned to it using the two relevant GAM-estimated probabilities (e.g., if the first author was a woman and the gender of the last author was unknown, either WM or WW would be assigned based on their relative GAM-estimated probabilities). Importantly, this procedure inherently assumes the null of 'no effect of author gender on citation patterns' for all imputed data. Thus, the estimates obtained in this sensitivity analysis are likely conservative.

This imputation process, combined with a bootstrapping procedure, was carried out 1,000 times. Approximations of the mean effect estimates and their 95% confidence intervals under this conservative missing data scheme were drawn from the resulting 1,000 estimates. The results are shown in Table S3, and can be compared to the main paper results found in Table S4.

| | MM teams | | W∪W teams | | WM teams | | MW teams | | WW teams | |
|---|---|---|---|---|---|---|---|---|---|---|
| Effect | Est. | 95% CI | Est. | 95% CI | Est. | 95% CI | Est. | 95% CI | Est. | 95% CI |
| MM paper citation rate | +7.7% | [7.1, 8.3] | +2.8% | [2.2, 3.4] | +4.1% | [3.5, 4.8] | +2.9% | [1.8, 3.9] | -1.5% | [-2.8, -0.3] |
| WM paper citation rate | -9.0% | [-10.1, -7.8] | -4.8% | [-5.9, -3.7] | -5.2% | [-6.6, -3.9] | -6.2% | [-8.2, -4.2] | -1.9% | [-4.1, 0.2] |
| MW paper citation rate | -7.8% | [-10.0, -5.8] | +0.0% | [-2.2, 2.0] | -3.2% | [-5.8, -0.5] | +1.3% | [-2.1, 4.7] | +7.4% | [3.4, 11.6] |
| WW paper citation rate | -22.6% | [-24.8, -20.3] | -5.5% | [-7.9, -3.0] | -11.5% | [-14.4, -8.6] | -3.5% | [-7.9, 1.2] | +9.5% | [4.0, 15.0] |
| MM overcitation trend (perc. points per year) | 0.49 | [0.39, 0.60] | 0.27 | [0.16, 0.39] | 0.31 | [0.16, 0.46] | 0.24 | [0.02, 0.45] | 0.26 | [0.01, 0.52] |
| Unconditional MM overcitation (perc. points) | 5.2 | [4.8, 5.7] | - | - | 3.1 | [2.6, 3.7] | 2.5 | [1.6, 3.3] | -0.4 | [-1.4, 0.5] |
| MM overcitation given network (perc. points) | 3.5 | [3.0, 4.0] | - | - | 2.1 | [1.5, 2.7] | 1.9 | [1.0, 2.7] | -0.3 | [-1.2, 0.7] |

**Table S3. Table of main results after conservative imputation of missing data (no assumed gender imbalance within missing entries)**. Results are calculated from 461,674 total citations. Confidence intervals are calculated from 1,000 bootstrap resampling iterations.

| | MM teams | | W∪W teams | | WM teams | | MW teams | | WW teams | |
|---|---|---|---|---|---|---|---|---|---|---|
| Effect | Est. | 95% CI | Est. | 95% CI | Est. | 95% CI | Est. | 95% CI | Est. | 95% CI |
| MM paper citation rate | +8.0% | [7.6, 8.5] | +2.5% | [2.0, 3.1] | +4.0% | [3.3, 4.7] | +2.7% | [1.6, 3.7] | -2.0% | [-3.3, -0.8] |
| WM paper citation rate | -9.3% | [-10.2, -8.3] | -4.6% | [-5.6, -3.7] | -4.8% | [-6.0, -3.5] | -6.6% | [-8.5, -4.6] | -1.8% | [-4.0, 0.5] |
| MW paper citation rate | -9.0% | [-10.6, -7.4] | -0.1% | [-2.0, 2.0] | -4.0% | [-6.3, -1.6] | +2.1% | [-1.4, 5.7] | +8.2% | [3.6, 12.5] |
| WW paper citation rate | -23.4% | [-25.4, -21.5] | -4.2% | [-6.4, -2.0] | -11.4% | [-14.2, -8.4] | -1.1% | [-6.2, 3.6] | +12.2% | [6.7, 17.8] |
| MM overcitation trend (perc. points per year) | 0.53 | [0.42, 0.63] | 0.29 | [0.17, 0.40] | 0.34 | [0.19, 0.50] | 0.23 | [0.01, 0.48] | 0.25 | [-0.03, 0.53] |
| Unconditional MM overcitation (perc. points) | 5.5 | [5.1, 5.9] | - | - | 3.1 | [2.6, 3.6] | 2.3 | [1.6, 2.9] | -0.7 | [-1.6, 0.3] |
| MM overcitation given network (perc. points) | 3.5 | [3.1, 4.0] | - | - | 2.0 | [1.5, 2.6] | 1.6 | [0.8, 2.3] | -0.3 | [-1.0, 0.6] |

**Table S4. Table of main results under primary analysis strategy**. Results are calculated from 303,886 total citations. Confidence intervals are calculated from 1,000 bootstrap resampling iterations.

***Rationale and impact of self-citation exclusion.*** For all analyses reported in the primary manuscript, self-citations were removed from consideration. Although the potential for gendered patterns in self-citations is potentially interesting, their removal in this study allowed us to isolate more comparable aspects of citation behavior between men and women in the field. Here we show additional results related to self-citation behavior. First, to assess the impact that self-citations have on gendered citation behavior, we conducted the primary analyses of over/undercitation without removing instances of self-citation. Here, we find that the inclusion of self-citations does not make a meaningful difference in the overall over/undercitation patterns. Their inclusion does, however, create even larger differences between the citation behavior of MM and W∪W teams (Figure S1), as self-citations necessarily increase the citation rate of men-led papers within MM reference lists, and increase the citation rate of women-led papers within W∪W reference lists. To ensure that missed self-citations were not driving the observed differences, we removed all citations in which the cited first or last author had the same last name as either the first or last author of the citing paper (i.e., a very conservative removal of potential self-citations). This more expansive removal led to negligible change in the primary results, suggesting that missed instances of self-citation were rare (Table S5).

Additionally, we examined potential gendered differences in the rate of self-citations across author genders. While a more thorough version of this analysis was carried out in King et al., 2017 (1), it may be of interest to the field to see the results for neuroscience. We found that as a proportion of reference list length, MM and WM teams tended to self-cite at higher rates than MW and WW teams. However, as a proportion of potential self-citations (i.e., count of authors' previous citable papers), rates of self-citation were relatively similar across groups (though still slightly higher in MM than WW; Table S6).

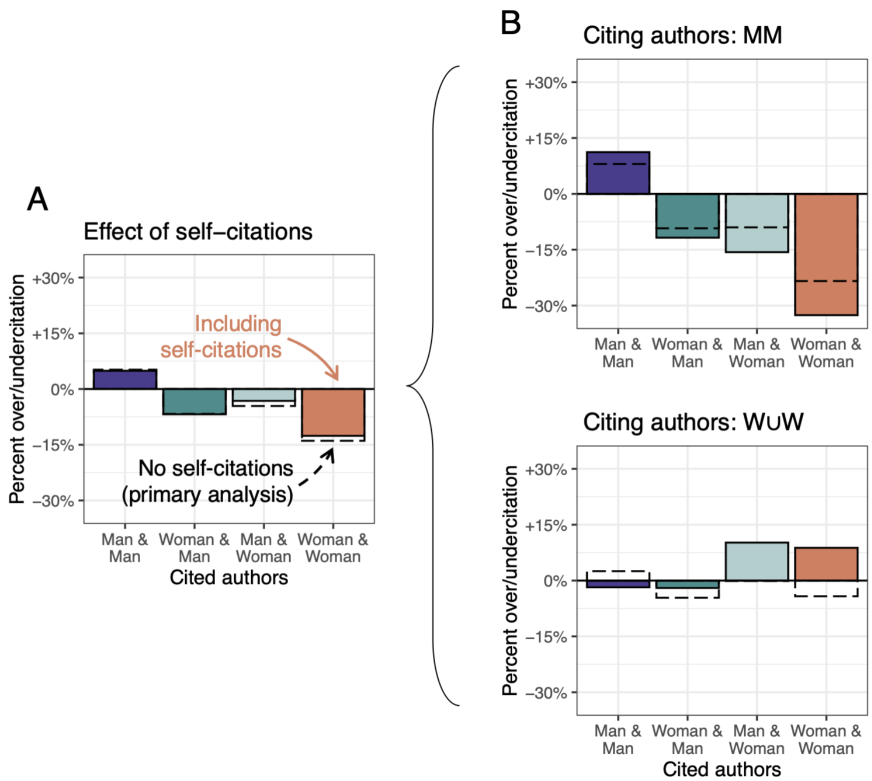


**Fig. S1. Effect of self-citations on the distribution of citations across gender categories.** *(A)* Overall over/undercitation of author gender categories excluding self-citations (dotted) and including self-citations (colored). Values are calculated from 303,886 total citations. *(B)* Over/undercitation of author gender categories excluding and including self-citations within MM reference lists (top) and W∪W reference lists (bottom). Values are calculated from 277,030 total citations (146,077 within MM reference lists, and 130,953 within W∪W reference lists)

| | Overall | | MM teams | | W∪W teams | |
|---|---|---|---|---|---|---|
| Effect | Primary | Strict | Primary | Strict | Primary | Strict |
| MM paper citation rate | +5.21% | +5.21% | +8.03% | +8.01% | +2.53% | +2.56% |
| WM paper citation rate | -6.69% | -6.69% | -9.26% | -9.22% | -4.63% | -4.65% |
| MW paper citation rate | -4.57% | -4.59% | -9.00% | -8.96% | -0.12% | -0.16% |
| WW paper citation rate | -13.92% | -13.95% | -23.43% | -23.39% | -4.23% | -4.36% |

**Table S5. Potential impact of missed instances of self-citation**. Values are calculated from 303,886 total citations (in column "Overall"), 146,077 citations within MM reference lists (in column "MM teams"), and 130,953 citations within W∪W reference lists (in column "W∪W teams").

| | MM teams | | WM teams | | MW teams | | WW teams | |
|---|---|---|---|---|---|---|---|---|
| Effect | Est. | 95% CI | Est. | 95% CI | Est. | 95% CI | Est. | 95% CI |
| Mean self-citations by reference list length | 13.5% | [13.2, 13.7] | 13.6% | [13.2, 14.0] | 12.1% | [11.5, 12.6] | 10.8% | [10.2, 11.3] |
| Median self-citations by reference list length | 9.1% | [8.7, 9.1] | 8.5% | [8.3, 9.1] | 7.2% | [6.7, 7.7] | 6.3% | [5.6, 7.1] |
| Mean self-citations by num. previous papers | 17.0% | [16.7, 17.3] | 17.0% | [16.6, 17.4] | 17.8% | [17.2, 18.5] | 17.1% | [16.3, 17.8] |
| Median self-citations by num. previous papers | 12.5% | [11.8, 12.5] | 11.7% | [11.1, 12.5] | 13.0% | [12.1, 14.3] | 10.9% | [9.1, 12.5] |

**Table S6. Comparison of self-citation rates across author gender categories, relative to both the length of reference lists and the number of potential self-citations**. Values are calculated across 27,636 author teams (14,151 MM teams, 7,711 WM teams, 3,106 MW teams, and 2,668 WW teams). Confidence intervals are calculated from 1,000 bootstrap resampling iterations.

***Effect of varying definitions of self-citation.*** For the primary analyses, we defined self-citations as papers for which either the cited first or last author was the first or last author on the citing paper (according to the previously described author name disambiguation procedure). Although this is a somewhat restrictive definition, it is the only type of self-citation for which the author gender of the cited paper is necessarily determined by the author gender of the citing paper. Here, we determine whether using broader definitions of self-citation has a meaningful effect on the results. We consider two alternate definitions of self-citation. *Broad - Citing* defines self-citations as papers for which either the cited first or last author was a co-author (not exclusively first/last) on the citing paper. *Broad - Cited* defines self-citations as cited papers for which any co-authors (not exclusively first/last) were first or last author on the citing paper. Tables S7 and S8 illustrate that all possible definitions yielded highly similar results.

| | Overall | | MM teams | | W∪W teams | |
|---|---|---|---|---|---|---|
| Effect | Primary | Broad - Citing | Primary | Broad - Citing | Primary | Broad - Citing |
| MM paper citation rate | +5.21% | +5.18% | +8.03% | +8.01% | +2.53% | +2.48% |
| WM paper citation rate | -6.69% | -6.83% | -9.26% | -9.32% | -4.63% | -4.87% |
| MW paper citation rate | -4.57% | -4.26% | -9.00% | -8.89% | -0.12% | +0.52% |
| WW paper citation rate | -13.92% | -13.71% | -23.43% | -23.35% | -4.23% | -3.90% |

**Table S7. Effect of expanding the definition of self-citation to include additional papers**. Primary analysis defined self-citation as cited papers for which the first or last author is the first or last author on the citing paper. "Broad - Citing" analysis here refers to the process of defining self-citations as any papers for which the *cited* first or last author is a co-author on the *citing* paper. Values are calculated from 303,886 total citations (in column "Overall"), 146,077 citations within MM reference lists (in column "MM teams"), and 130,953 citations within W∪W reference lists (in column "W∪W teams").

| | Overall | | MM teams | | W∪W teams | |
|---|---|---|---|---|---|---|
| Effect | Primary | Broad - Cited | Primary | Broad - Cited | Primary | Broad - Cited |
| MM paper citation rate | +5.21% | +5.25% | +8.03% | +8.05% | +2.53% | +2.64% |
| WM paper citation rate | -6.69% | -6.83% | -9.26% | -9.27% | -4.63% | -5.00% |
| MW paper citation rate | -4.57% | -4.49% | -9.00% | -9.11% | -0.12% | +0.27% |
| WW paper citation rate | -13.92% | -14.14% | -23.43% | -23.77% | -4.23% | -4.42% |

**Table S8. Effect of expanding the definition of self-citation to include additional papers**. Primary analysis defined self-citation as cited papers for which the first or last author is the first or last author on the citing paper. "Broad - Cited" analysis here refers to the process of defining self-citations as any papers for which the *citing* first or last author is a co-author on the *cited* paper. Values are calculated from 303,886 total citations (in column "Overall"), 146,077 citations within MM reference lists (in column "MM teams"), and 130,953 citations within W∪W reference lists (in column "W∪W teams").

***Impact of weighting articles by length of reference list.*** The primary analyses conducted in this paper are conducted by either passively or actively weighting articles' contributions to estimates by the number of candidate citations present in their reference lists. For the first sections, in which percent over/undercitation is calculated, this weighting is done passively by summing over all citations across citing papers. In the section discussing social network effects, this weighting is done actively by running a quantile regression weighted by the number of citations in an article. Although this weighting was done to stabilize the estimates in the face of articles with very few citations (i.e., percent over/undercitation measures at an individual paper level are highly variable), it was of interest to ensure that the results were robust to this decision. To test for robustness, we re-ran the primary analyses in the paper using unweighted article-level measures of over/undercitation, as opposed to measures that relied on weighting or collapsing across all citations. These results are shown in Table S9, and can be compared to the main paper results found in Table S4.

| | MM teams | | W∪W teams | | WM teams | | MW teams | | WW teams | |
|---|---|---|---|---|---|---|---|---|---|---|
| Effect | Est. | 95% CI | Est. | 95% CI | Est. | 95% CI | Est. | 95% CI | Est. | 95% CI |
| MM paper citation rate | +7.1% | [6.4, 7.7] | +1.5% | [0.8, 2.2] | +2.9% | [2.0, 3.8] | +2.2% | [0.8, 3.6] | -3.0% | [-4.6, -1.4] |
| WM paper citation rate | -8.8% | [-10.0, -7.6] | -3.2% | [-4.5, -1.9] | -3.5% | [-5.2, -1.8] | -6.1% | [-8.6, -3.5] | +1.0% | [-2.1, 4.0] |
| MW paper citation rate | -6.8% | [-9.0, -4.5] | +0.2% | [-1.9, 2.6] | -2.8% | [-5.9, 0.5] | +4.2% | [-0.9, 9.1] | +4.3% | [-1.0, 9.5] |
| WW paper citation rate | -21.9% | [-24.5, -19.3] | -4.7% | [-7.4, -2.0] | -11.7% | [-15.1, -8.1] | -0.4% | [-6.5, 5.4] | +10.4% | [4.2, 16.7] |
| MM overcitation trend (perc. points per year) | 0.39 | [0.25, 0.52] | 0.21 | [0.07, 0.35] | 0.17 | [-0.03, 0.35] | 0.26 | [-0.04, 0.56] | 0.25 | [-0.08, 0.58] |
| Unconditional MM overcitation (perc. points) | 5.1 | [4.7, 5.5] | - | - | 2.5 | [1.9, 3.0] | 2.2 | [1.4, 2.9] | -0.8 | [-1.7, 0.0] |
| MM overcitation given network (perc. points) | 3.3 | [2.8, 3.7] | - | - | 1.6 | [1.0, 2.2] | 1.8 | [1.0, 2.5] | -0.4 | [-1.3, 0.5] |

**Table S9. Table of main results using unweighted article-level assessments of over/under citation**. Results are calculated from 303,886 total citations. Confidence intervals are calculated from 1,000 bootstrap resampling iterations.

***Specification of the generalized additive model (GAM).*** Many results in this study rely on expected citation rates across author gender categories that account for other characteristics. These rates were assessed by modeling the relationship between gender category and relevant paper characteristics as a generalized additive model (GAM) on the multinomial outcome {MM, WM, MW, WW}, in which the model's features were 1) month and year of publication, 2) combined number of publications by the first and last authors, 3) number of total authors on the paper, 4) the journal in which it was published, and 5) whether it was a review paper. Smooth terms were fit for publication date, author experience, and team size. In the primary analysis, these smooth terms were fit using univariate thin plate splines (i.e., assuming additive univariate functions, and no interactions in two- or three-dimensional variable space). Similarly, no interactions were modeled within the journal and review article variables, nor between these variables and the smooth terms. To assess the extent to which this simplified model accurately reflects the relevant relationships between these variables and author gender, we conducted a sensitivity analysis that used a more complex specification of the GAM. In this sensitivity analysis, the model was formulated as follows: separate smooth terms for publication date were fit for each journal, a bivariate thin plate spline was used to fit a joint smooth term for author experience and team size, and an interaction term was fit between journal and review article status. Figure S2 visualizes the primary results of the study under both the primary (dashed lines) and sensitivity (colored bars) specifications of the GAM. The results suggest that the relationships between the paper characteristics and author gender are largely captured by univariate additive terms, and that specifying a more complex model does not meaningfully change the results of the paper.

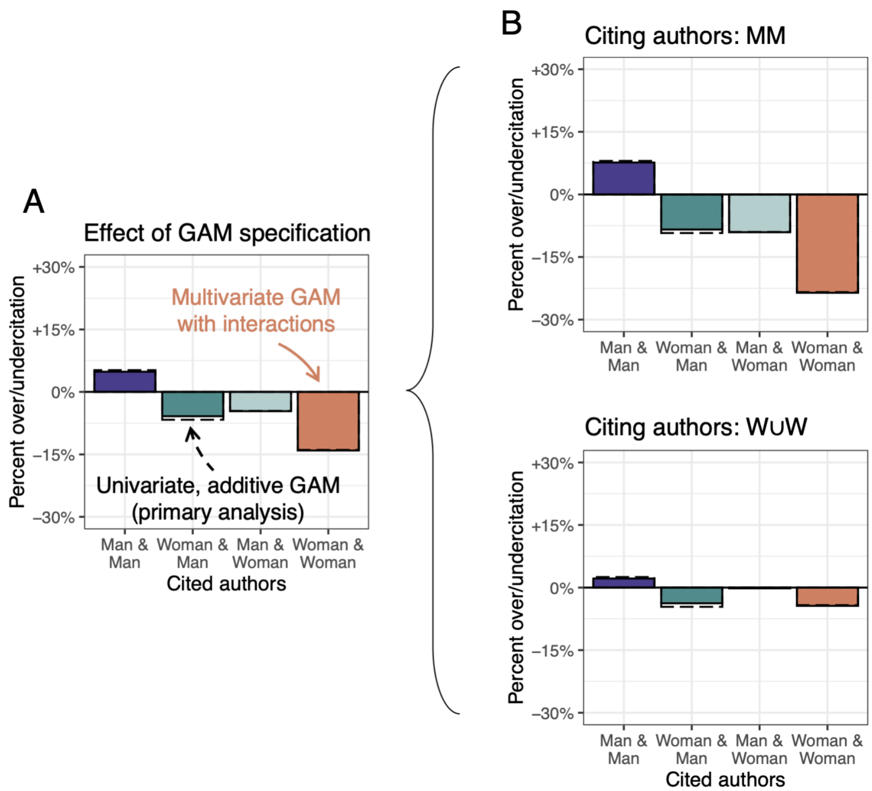


**Fig. S2. Impact of specifying a more complex GAM model on author gender.** *(A)* Overall over/undercitation of author gender categories under the primary GAM specification (dotted) and under a more complex specification (colored). Values are calculated from 303,886 total citations. *(B)* Over/undercitation of author gender categories under alternate GAM specifications within MM reference lists (top) and W∪W reference lists (bottom). Values are calculated from 277,030 total citations (146,077 within MM reference lists, and 130,953 within W∪W reference lists).

***Construction of graph-preserving null model.*** To determine the extent to which observed effects could be explained by the structure of the citation graph, a graph null model was constructed to assess significance of primary analyses. This null model was derived from the randomization of *cited* papers' author gender categories. To incorporate other characteristics of papers that may be relevant to citation behavior, randomization was carried out by probabilistically drawing new gender categories for each paper according to their GAM-estimated gender probabilities. Randomly sampling gender categories for each paper therefore produces a null model in which cited author gender is conditionally independent of citation rates and citing author gender (conditional on the characteristics included in the GAM model), while the structure of the citation graph and the long-tailed nature of the citation distribution are both preserved. 10,000 randomizations were carried out, and $p$-values were calculated by taking the proportion of randomizations for which the absolute value of a given null estimate was greater than the absolute value of its respective observed estimate. To account for multiple comparisons, reported $p$-values were corrected according to the Holm-Bonferroni procedure. Figure S3 shows the null distributions for the primary citation imbalance measures overlaid on the observed values. Table S10 shows the means, standard errors, and 95% confidence intervals of the null distributions across 10,000 randomizations, for all primary estimates reported in the paper.

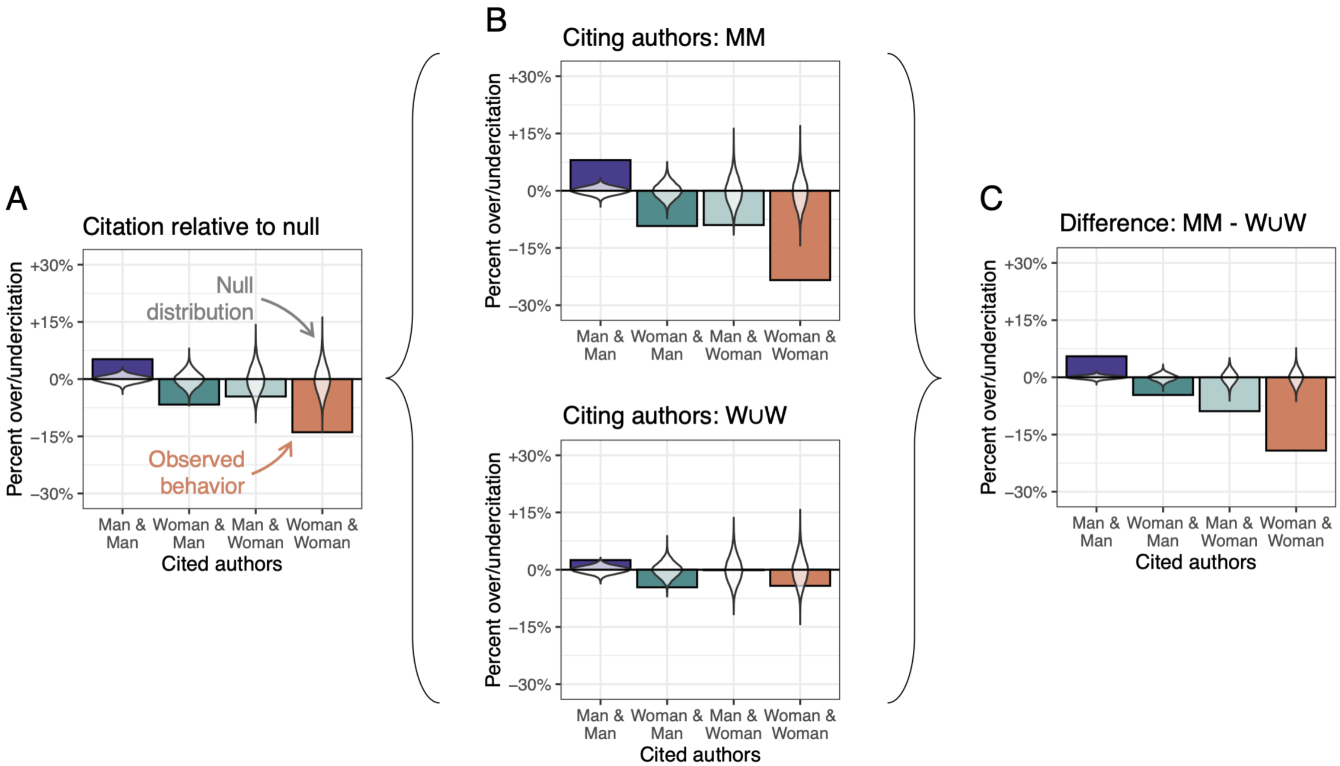


**Fig. S3. Visualization of the distribution of citation behavior under the null hypothesis that citation is independent of cited author gender conditional on specific paper characteristics.** *(A)* Overall over/undercitation of author gender categories (colored bars) with the graph-preserving null distributions overlaid (translucent violins). Values are calculated from 303,886 total citations. *(B)* Over/undercitation of author gender categories and overlaid null distributions within MM reference lists (top) and W∪W reference lists (bottom). Values are calculated from 277,030 total citations (146,077 within MM reference lists, and 130,953 within W∪W reference lists). *(C)* Difference between citation rates within MM and W∪W reference lists and overlaid null distributions. Across all panels, violin plots represent the null distribution of each over/undercitation measure. Distributions are created using 10,000 randomizations of the graph-preserving null model. The end of each violin's tails represents the minimum and maximum of that null distribution; quantiles are not plotted, but detailed summary statistics for each null distribution are given in the table below.

| Effect | Est. | Null mean | Null SE | Null 2.5/97.5%-interval |
|---:|---:|---:|---:|---:|
| Overall - MM citation rate | 5.21 | 0.00 | 0.93 | [-1.84, 1.80] |
| Overall - WM citation rate | -6.69 | 0.00 | 1.93 | [-3.67, 3.85] |
| Overall - MW citation rate | -4.57 | 0.04 | 3.38 | [-6.25, 7.05] |
| Overall - WW citation rate | -13.92 | -0.05 | 3.73 | [-7.06, 7.65] |
| MM ref. list - MM citation rate | 8.03 | 0.00 | 0.97 | [-1.89, 1.85] |
| MM ref. list - WM citation rate | -9.26 | 0.00 | 2.01 | [-3.80, 4.02] |
| MM ref. list - MW citation rate | -9.00 | 0.04 | 3.55 | [-6.63, 7.39] |
| MM ref. list - WW citation rate | -23.43 | -0.04 | 3.96 | [-7.50, 8.05] |
| W∪W ref. list - MM citation rate | 2.53 | 0.00 | 0.95 | [-1.88, 1.85] |
| W∪W ref. list - WM citation rate | -4.63 | 0.01 | 1.93 | [-3.66, 3.83] |
| W∪W ref. list - MW citation rate | -0.12 | 0.04 | 3.37 | [-6.24, 7.07] |
| W∪W ref. list - WW citation rate | -4.23 | -0.06 | 3.73 | [-7.06, 7.53] |
| MM/W∪W diff - MM citation rate | 5.50 | 0.00 | 0.43 | [-0.83, 0.83] |
| MM/W∪W diff - WM citation rate | -4.63 | -0.01 | 0.87 | [-1.71, 1.69] |
| MM/W∪W diff - MW citation rate | -8.88 | 0.00 | 1.58 | [-3.13, 3.07] |
| MM/W∪W diff - WW citation rate | -19.19 | 0.02 | 1.86 | [-3.56, 3.71] |
| Overall - MM citation trend | 0.41 | 0.00 | 0.09 | [-0.17, 0.17] |
| MM ref. list - MM citation trend | 0.54 | 0.00 | 0.10 | [-0.19, 0.19] |
| W∪W ref. list - MM citation trend | 0.29 | 0.00 | 0.09 | [-0.17, 0.17] |
| MM/W∪W diff - MM citation trend | 0.25 | 0.00 | 0.07 | [-0.15, 0.15] |
| MM - MM uncond. overcitation | 5.46 | 0.18 | 0.61 | [-1.04, 1.38] |
| WM - MM uncond. overcitation | 3.04 | 0.19 | 0.64 | [-1.08, 1.45] |
| MW - MM uncond. overcitation | 2.39 | 0.21 | 0.70 | [-1.14, 1.56] |
| WW - MM uncond. overcitation | -0.71 | 0.11 | 0.68 | [-1.15, 1.50] |
| MM - MM network overcitation | 3.50 | 0.19 | 0.61 | [-1.02, 1.39] |
| WM - MM network overcitation | 1.92 | 0.21 | 0.64 | [-1.06, 1.43] |
| MW - MM network overcitation | 1.60 | 0.22 | 0.71 | [-1.18, 1.58] |
| WW - MM network overcitation | -0.38 | 0.12 | 0.69 | [-1.16, 1.53] |

**Table S10. Means, standard errors, and 2.5/97.5-percentiles of the null distribution for each primary estimate.** Null distributions for each estimate were calculated using 10,000 randomizations of the graph-preserving null model.

***Potential confounding impact of research subfields.*** Though the role of neuroscience subfields is not directly accounted for within the primary analyses, it is important to understand whether and to what extent relationships between gender and subfield confound our results. To assess this possibility, we conducted a sensitivity analysis on a subset of *Journal of Neuroscience* papers. *Journal of Neuroscience* was chosen because it contains the most articles within our dataset, and because it uses a consistent sub-disciplinary classification scheme for its papers. Specifically, almost all of its papers are classified as either *behavioral/systems*, *systems/circuits*, *neurobiology of disease*, *development/plasticity/repair*, *behavioral/systems/cognitive*, or *cellular/molecular*. Two separate author gender-predicting generalized additive models (GAMs) were fit to the subset of 21,338 articles with one of these classifications. The first did not include the subfield classification, and the second did. Estimates of the over/undercitation of author genders within these 21,338 articles were then calculated as per the primary analyses using each estimated model. Figure S4 shows the results, with dashed lines indicating the results without subfield classifications (i.e., the same model as was used in the primary manuscript), and colored bars indicating results after accounting for subfields. The results suggest that subfields likely have little impact on either the extent of citation imbalance or the discrepancy in citation behavior across citing author genders.

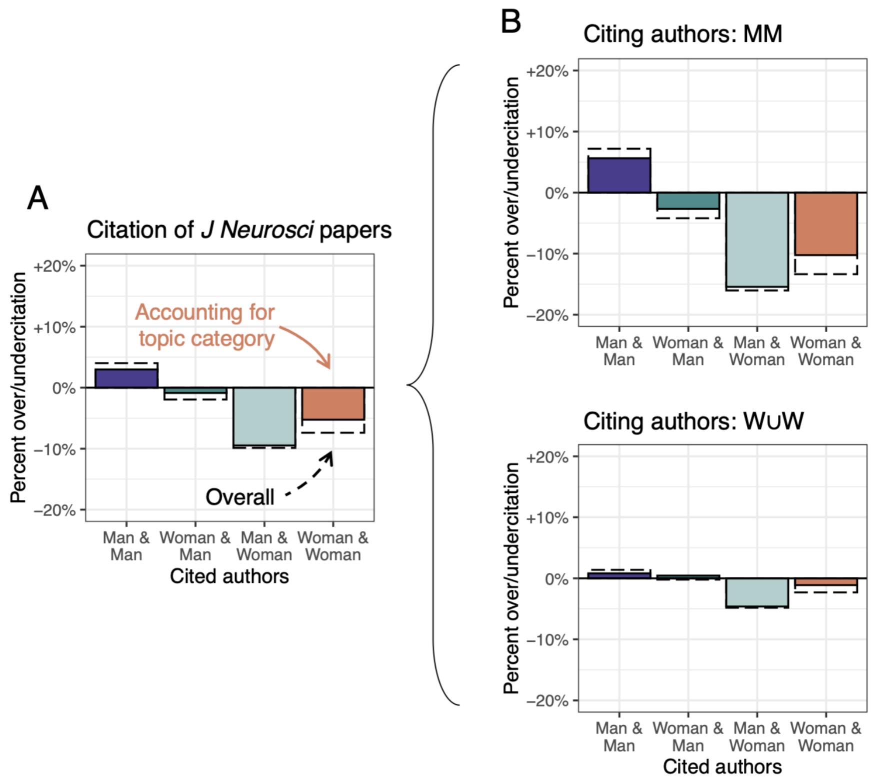


**Fig. S4. Impact of research subfields on citation behavior, using a subset of *Journal of Neuroscience* papers with sub-disciplinary classifications.** *(A)* Overall over/undercitation of author gender categories before accounting for subfield (dotted) and after accounting for subfield (colored). Values calculated from 79,032 total citations. *(B)* Over/undercitation of author gender categories before and after accounting for subfield within MM reference lists (top) and W∪W reference lists (bottom). Values are calculated from 72,476 total citations (37,708 within MM reference lists, and 34,768 within W∪W reference lists).

***Potential differential effects among high- and low-citation papers.*** Because the citation distribution is long-tailed, it is of particular interest to determine whether the observed overcitation of men is driven primarily by a subset of highly cited papers, over whether over/undercitation occurs relatively consistently for high- and low-citation papers. To address this question, we separately quantified over/undercitation of author gender categories within the top 50% and bottom 50% of the citation distribution. In other words, papers in the data were separated into two groups according to a median split (the median number of citations at time of data retrieval was 50), and citation rates relative to expectation were assessed within each group according to the same procedure used for the primary analyses. While the papers in the top half of the distribution accounted for three-quarters of all citations in the data, the patterns of over/undercitation were similar for both sets (Figure S5). Specifically, below-median MM/WM/MW/WW papers were cited +4.8%, -2.2%, -6.6%, and -14.6% relative to expectation, while above-median papers were cited +5.3%, -7.6%, -4.1%, and -13.8% relative to expectation. This suggests that the observed imbalance is not driven by a 'rich club' of highly cited papers by men.

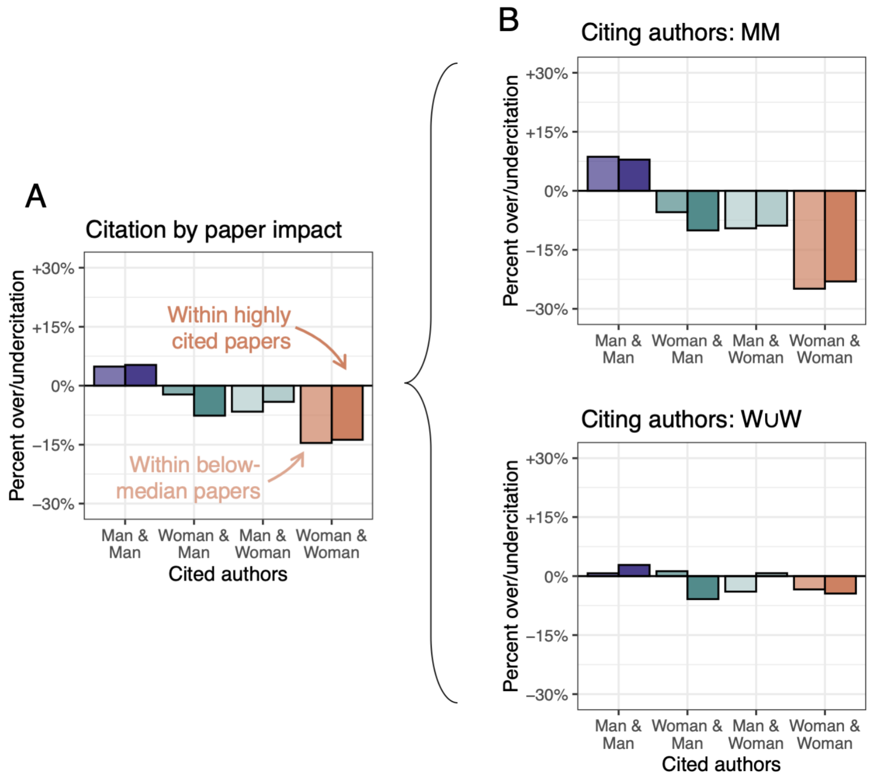


**Fig. S5. Citation behavior after stratifying cited papers by their status at the top or bottom of the citation distribution (by median split).** *(A)* Overall over/undercitation of author gender categories among low-citation (left, transparent) papers and high-citation (right, opaque) papers. Values are calculated from 48,744 citations to low-citation papers, and 255,142 citations to high-citation papers. *(B)* Over/undercitation of author gender categories among low-citation papers and high-citation papers within MM reference lists (top) and W∪W reference lists (bottom). Values are calculated from 23,651 citations from MM papers to low-citation papers, 122,426 citations from MM papers to high-citation papers, 20,932 citations from W∪W papers to low-citation papers, and 110,021 citations from W∪W papers to high-citation papers.

***Role of seniority/productivity in network structure and MM overcitation.*** Scholars' seniority and productivity have a potentially important role in constructing the collaboration network and shaping citation behavior. Therefore, it was of interest to determine the extent to which the most senior (i.e., in our data, most highly productive) scientists drive the network homophily and citation patterns. To assess this possibility, we performed a median split of articles based on the total number of papers the first/last-author team had published within the context of our dataset (i.e., top 5 journals, 1995-2018). For the purpose of the analysis, teams who published $> 8$ papers in this timespan are referred to as *more productive* and teams who published $\leq 8$ papers are referred to as *less productive*.

The network structure showed subtle differences between the two groups. More productive teams tended to have very slightly less man-author overrepresentation within their networks (Figure S6A). Additionally, more productive teams tended to have higher MM-paper overrepresentation in their networks when men were the last-author, and lower MM-paper overrepresentation when women were last-author (Figure S6B). In terms of citation patterns, higher and lower productivity teams tended to show similar citation behavior across author genders both before (Figure S6C) and after accounting for network effects (Figure S6D). Yet interestingly, the network structure of more productive teams appeared to be more strongly associated with citation behavior than that of less productive teams (Figure S6D). Specifically, across all papers, a one percentage point increase in MA overrepresentation corresponded to a 0.09 percentage point increase in MM overcitation. When broken down by less/more productive teams, this value was 0.05 and 0.15, respectively. Similarly, across all papers, a one percentage point increase in MM-paper overrepresentation corresponded to a 0.24 percentage point increase in MM overcitation. When broken down by less/more productive teams, this value was 0.21 and 0.28, respectively. This finding may suggest that more senior authors' citation behavior is more closely tied to the structure of their social networks.

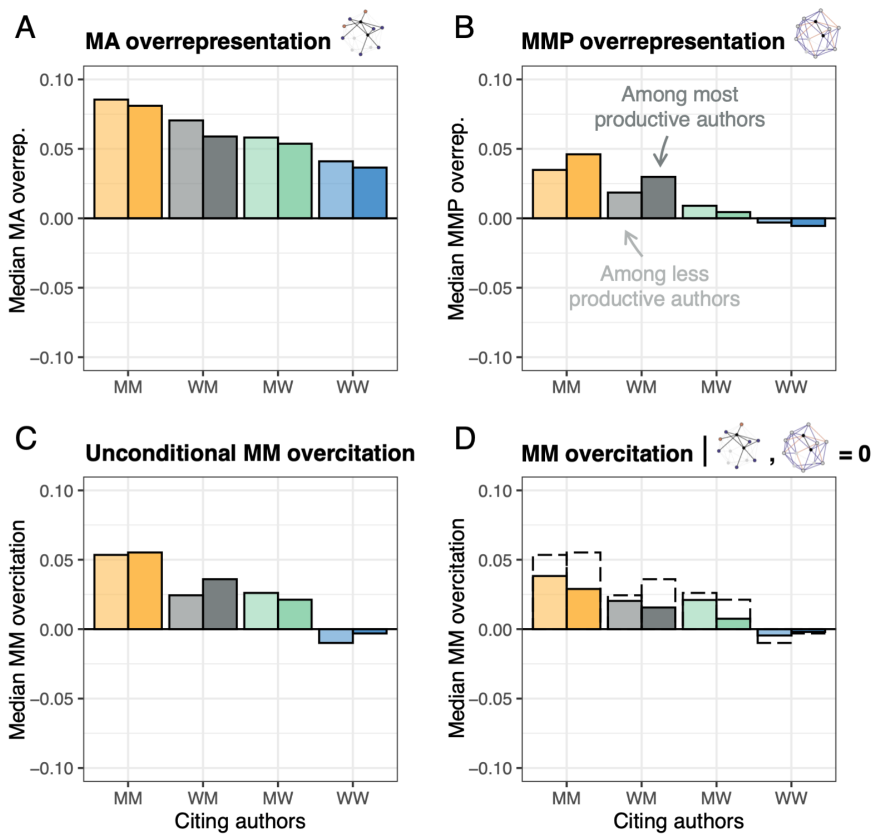


**Fig. S6. Relationship between seniority/productivity, network structure, and citation patterns.** *(A)* Man-author overrepresentation by author gender among low productivity (left, transparent) and high productivity (right, opaque) teams. *(B)* MM-paper overrepresentation by author gender among low productivity and high productivity teams. *(C)* Overcitation of MM papers (not accounting for network structure) by author gender among low productivity and high productivity teams. Shows similar patterns of citation between low and high seniority/productivity teams. *(D)* Overcitation of MM papers by author gender, after accounting for network structure. Shows that network structure explains more of the citation imbalance within high seniority teams than low seniority teams. Values are calculated across 27,636 author teams (7,720 high-seniority MM teams, 6,431 low-senority MM teams, 3,789 high-seniority WM teams, 3,922 low-seniority WM teams, 1,280 high-seniority MW teams, 1,826 low-seniority MW teams, 913 high-seniority WW teams, and 1,755 low-seniority WW teams).

***Quantile regression vs. linear regression for estimating MM overcitation.*** The analyses in the “The relationship between social networks and citation behavior" section were conducted using quantile regression to fit the conditional median of the outcome. Quantile regression was chosen, as opposed to linear regression, because of the bounded and skewed nature of the outcome measure. However, to ensure that the results are robust to this modeling choice, here we compare the results of that section using quantile regression to the results that would be obtained using linear regression. Table S11 shows that the results, while somewhat attenuated using linear regression, are consistent with the results obtained with quantile regression.

| | MM teams | | WM teams | | MW teams | | WW teams | |
|---|---|---|---|---|---|---|---|---|
| Effect | Est. | 95% CI | Est. | 95% CI | Est. | 95% CI | Est. | 95% CI |
| Unconditional *median* MM overcitation (perc. points) | 5.5 | [5.1, 5.9] | 3.1 | [2.6, 3.6] | 2.3 | [1.6, 3.0] | -0.7 | [-1.6, 0.3] |
| *Median* MM overcitation given network (perc. points) | 3.5 | [3.1, 3.9] | 2.0 | [1.5, 2.6] | 1.6 | [0.8, 2.3] | -0.3 | [-1.0, 0.6] |
| Unconditional *mean* MM overcitation (perc. points) | 4.7 | [4.4, 5.0] | 2.4 | [2.0, 2.8] | 1.6 | [0.9, 2.2] | -1.2 | [-2.0, -0.4] |
| *Mean* MM overcitation given network (perc. points) | 3.1 | [2.8, 3.4] | 1.3 | [0.9, 1.8] | 1.2 | [0.6, 1.9] | -1.2 | [-2.0, -0.4] |

**Table S11. Comparison between overcitation results obtained using quantile regression and linear regression**. Values are calculated across 27,636 author teams (14,151 MM teams, 7,711 WM teams, 3,106 MW teams, and 2,668 WW teams)

## Reference

1. Molly M. King, Carl T. Bergstrom, Shelley J. Correll, Jennifer Jacquet, and Jevin D. West. Men Set Their Own Cites High: Gender and Self-citation across Fields and over Time. *Socius: Sociological Research for a Dynamic World*, 3:237802311773890, December 2017. ISSN 2378-0231, 2378-0231. doi: 10.1177/2378023117738903.