Supplementary

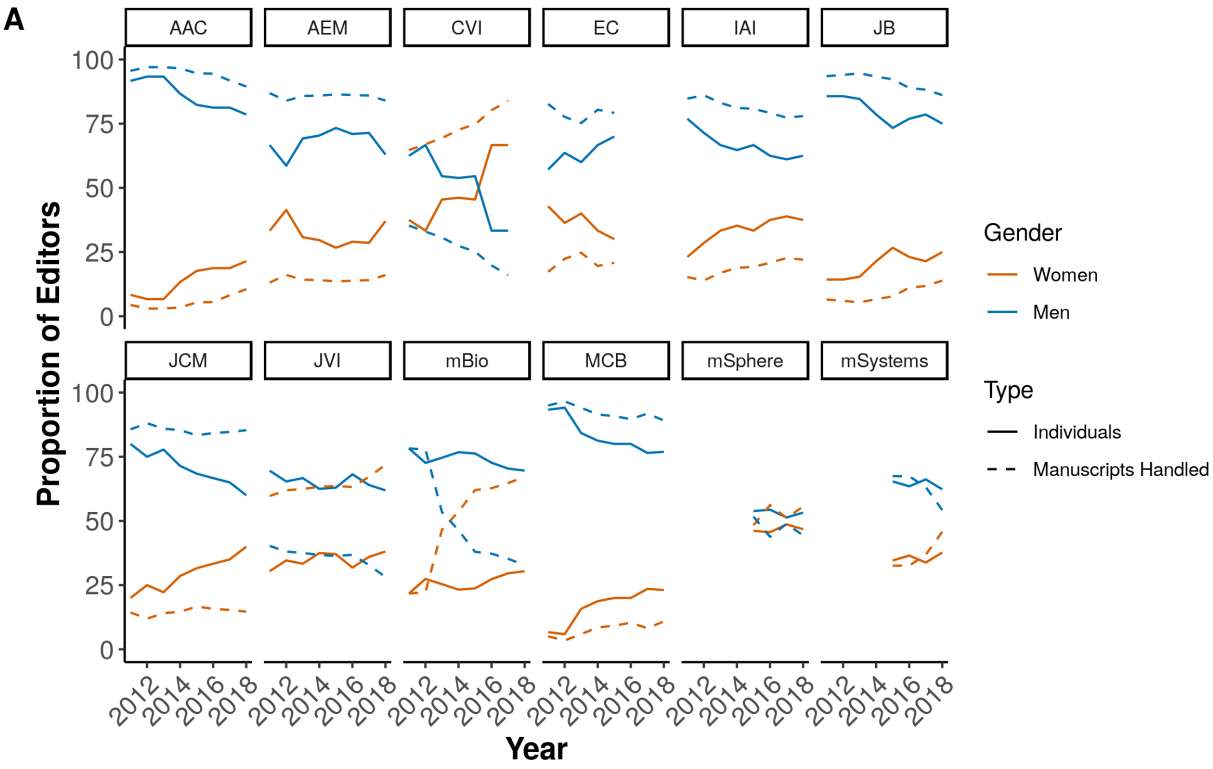

Figure S1. The proportion of editors (solid line) and their workloads (dashed line) at each ASM journal from 2012 to 2018.

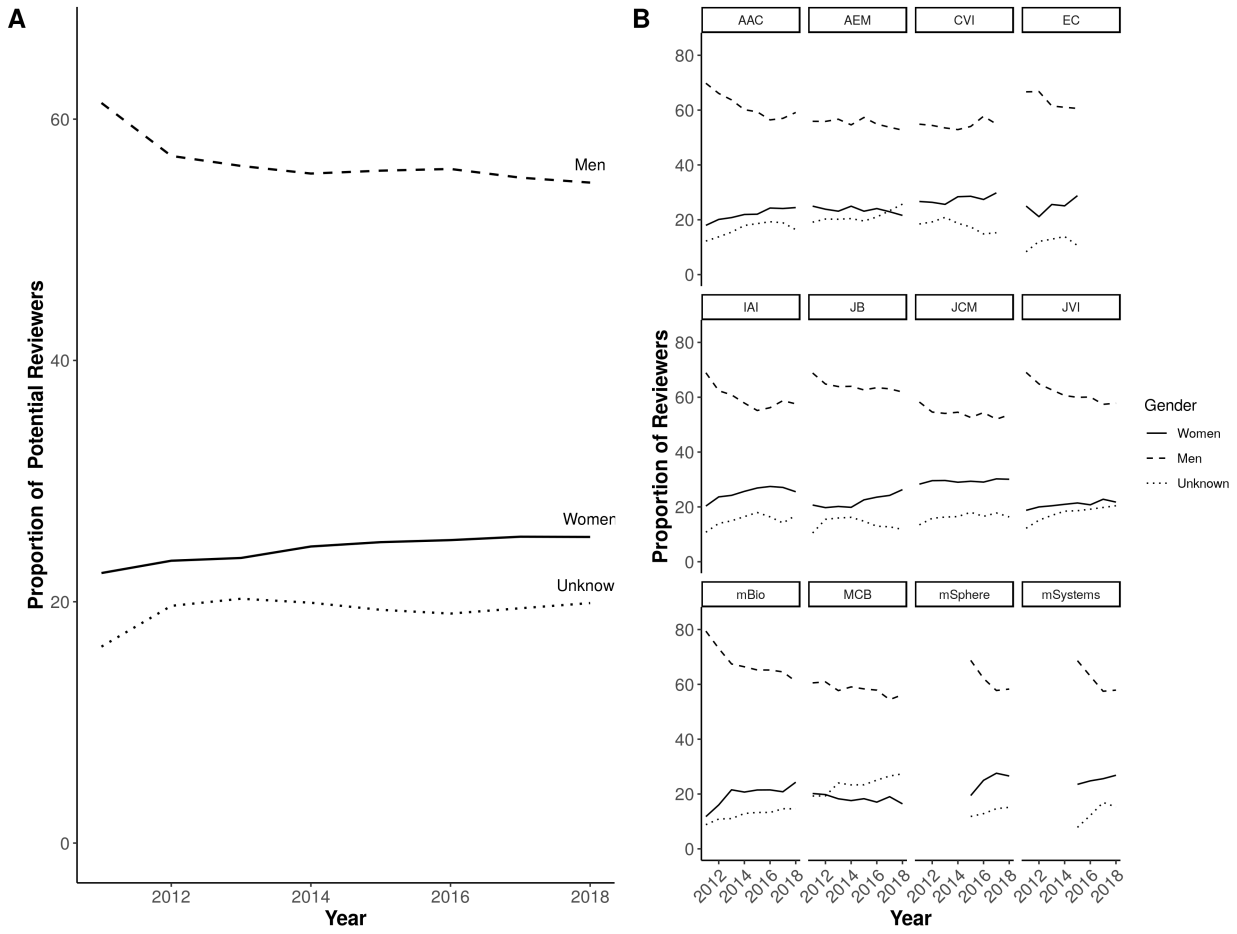

Figure S2. The proportion of (A) potential reviewers at all ASM journals combined, (B) reviewers at each ASM journal.

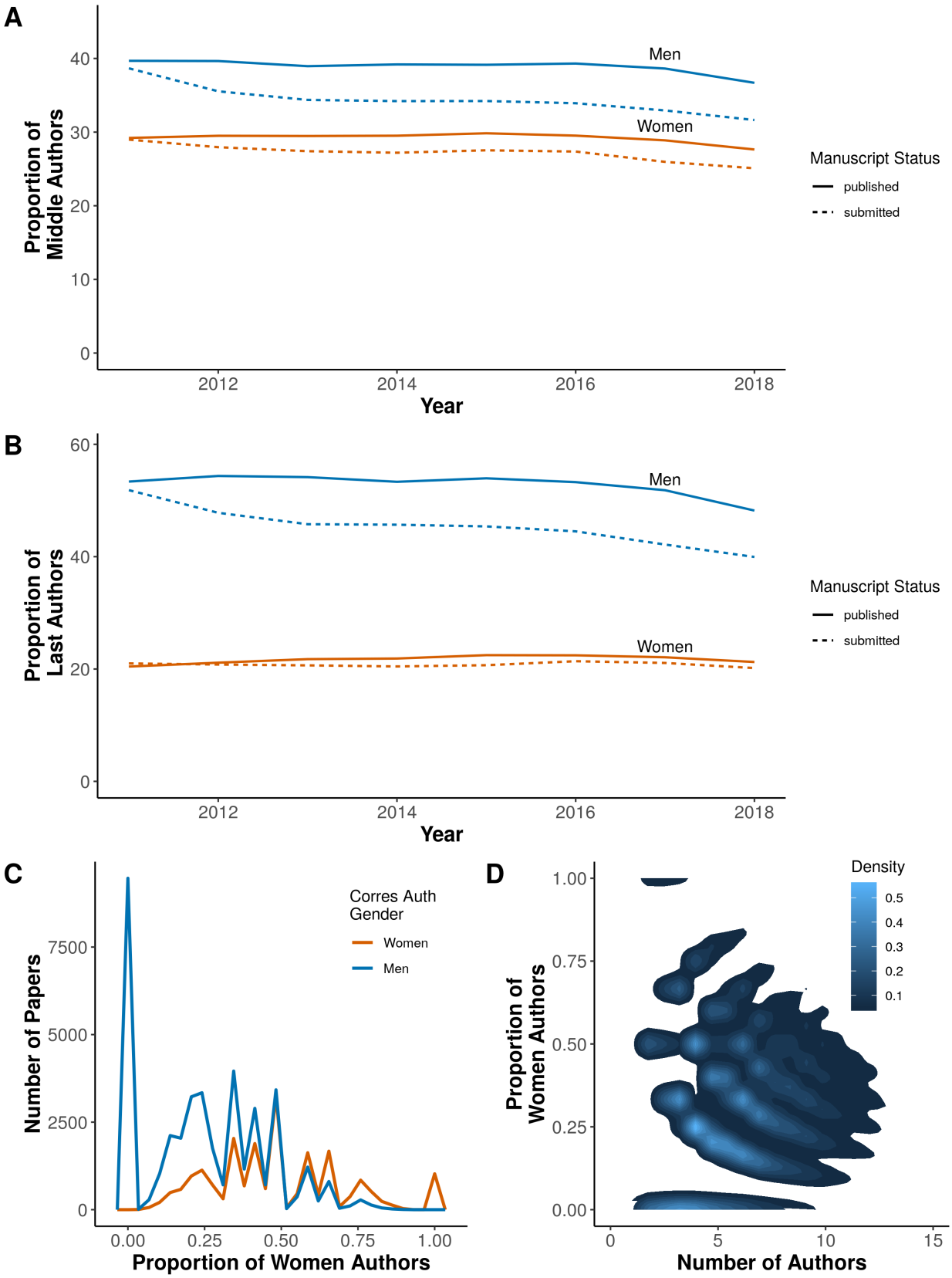
A
Proportion of Middle Authors
40
30
20
10
0
Men
Women
2012
2014
2016
2018
Year
Manuscript Status
published
submitted
B
Proportion of Last Authors
60
40
20
0
Men
Women
2012
2014
2016
2018
Year
Manuscript Status
published
submitted
C
Number of Papers
7500
5000
2500
0
Corres Auth Gender
Women
Men
0.00
0.25
0.50
0.75
1.00
Proportion of Women Authors
D
Proportion of Women Authors
1.00
0.75
0.50
0.25
0.00
Density
0.5
0.4
0.3
0.2
0.1
0
5
10
15
Number of Authors

Figure S3. The proportion of all submitting (dashed line) and publishing (solid line) (A) middle and (B) last authors by gender at each ASM journal. The proportion of women authors on submitted papers according to (D) the gender of the corresponding author or (E) the number of authors.

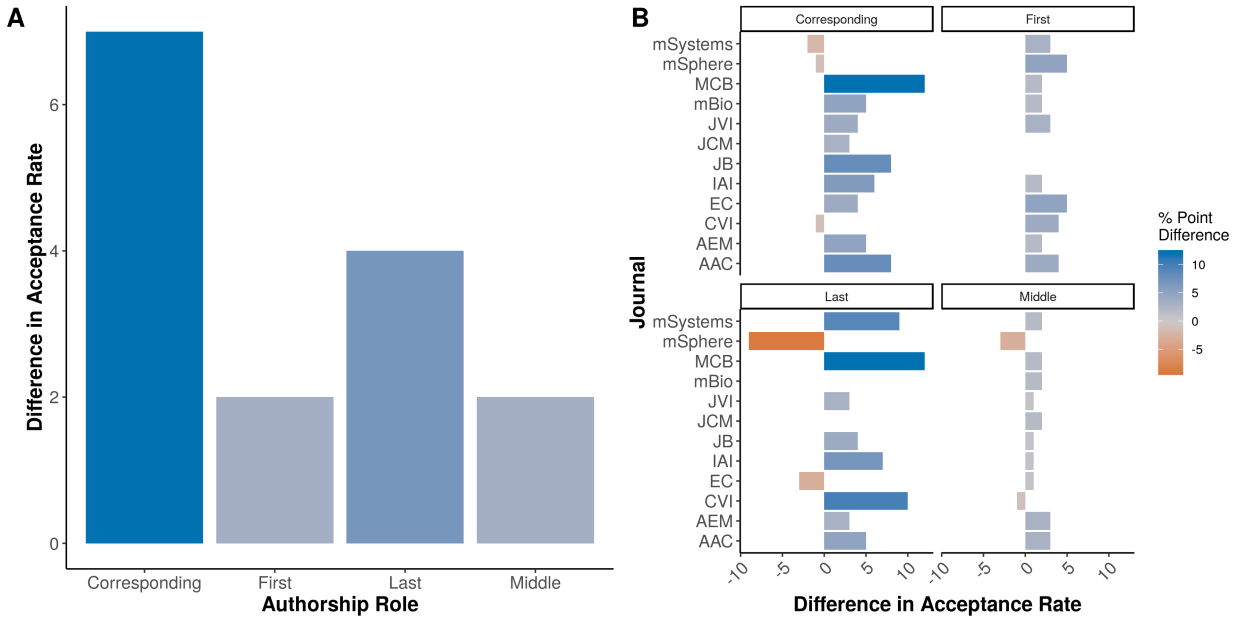

Figure S4. The difference in percentage points of papers accepted. The difference in percentage points of papers accepted at (A) all journals or (B) for each journal.

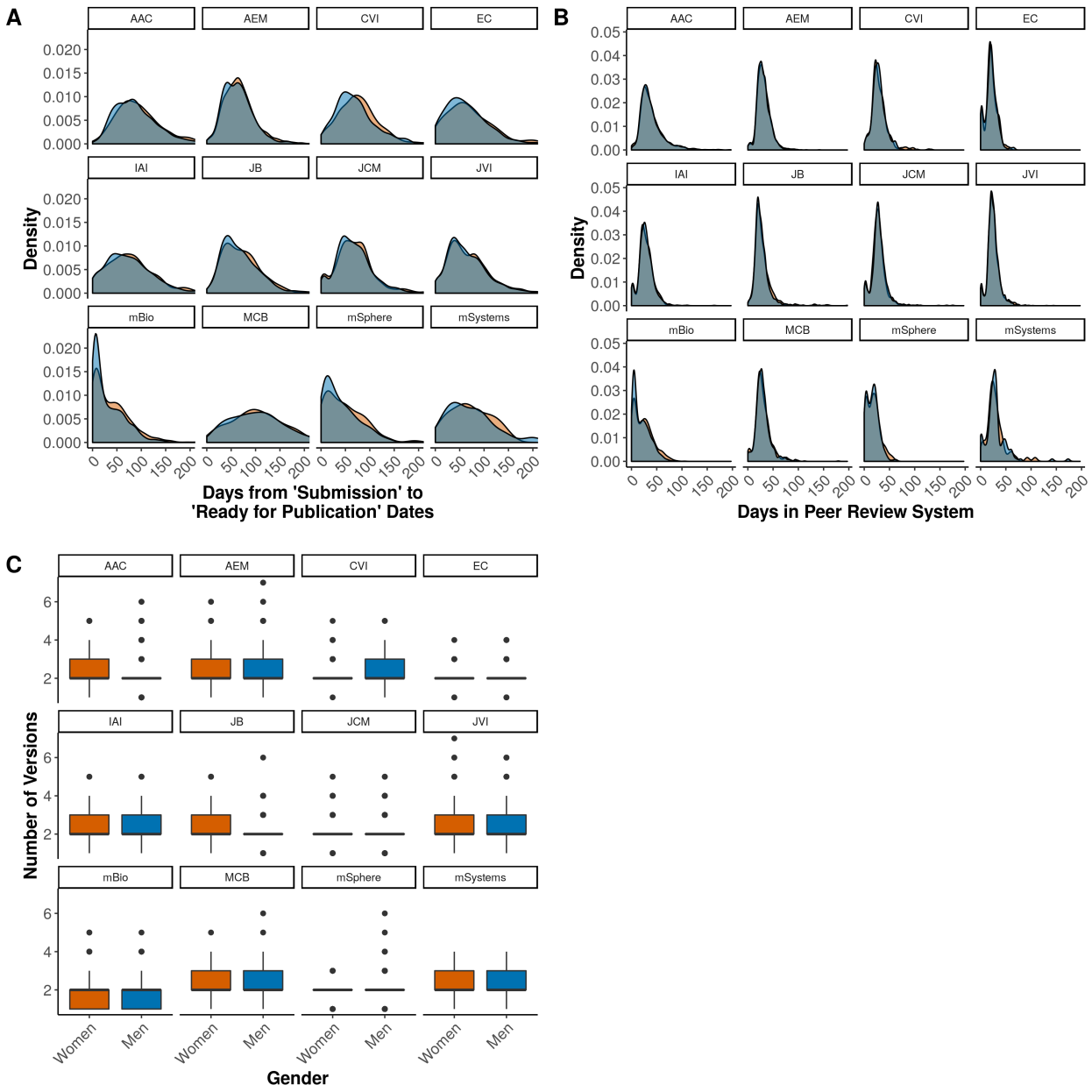

Figure S5. Comparison of time to final decision and impact by gender. The number days (A) between when a manuscript is initally submitted then finally published and (B) that a manuscript spends in the ASM peer review system. How the impact of papers published by men (blue) versus women (orange) vary according to (C) cites and (D) total reads. Citation data includes articles published between 36 and 48 months prior to August 2018. Total reads includes both HTML and PDF online views for articles published between 12 and 24 months prior to August 2018. Impact data are divided by the number of months published.

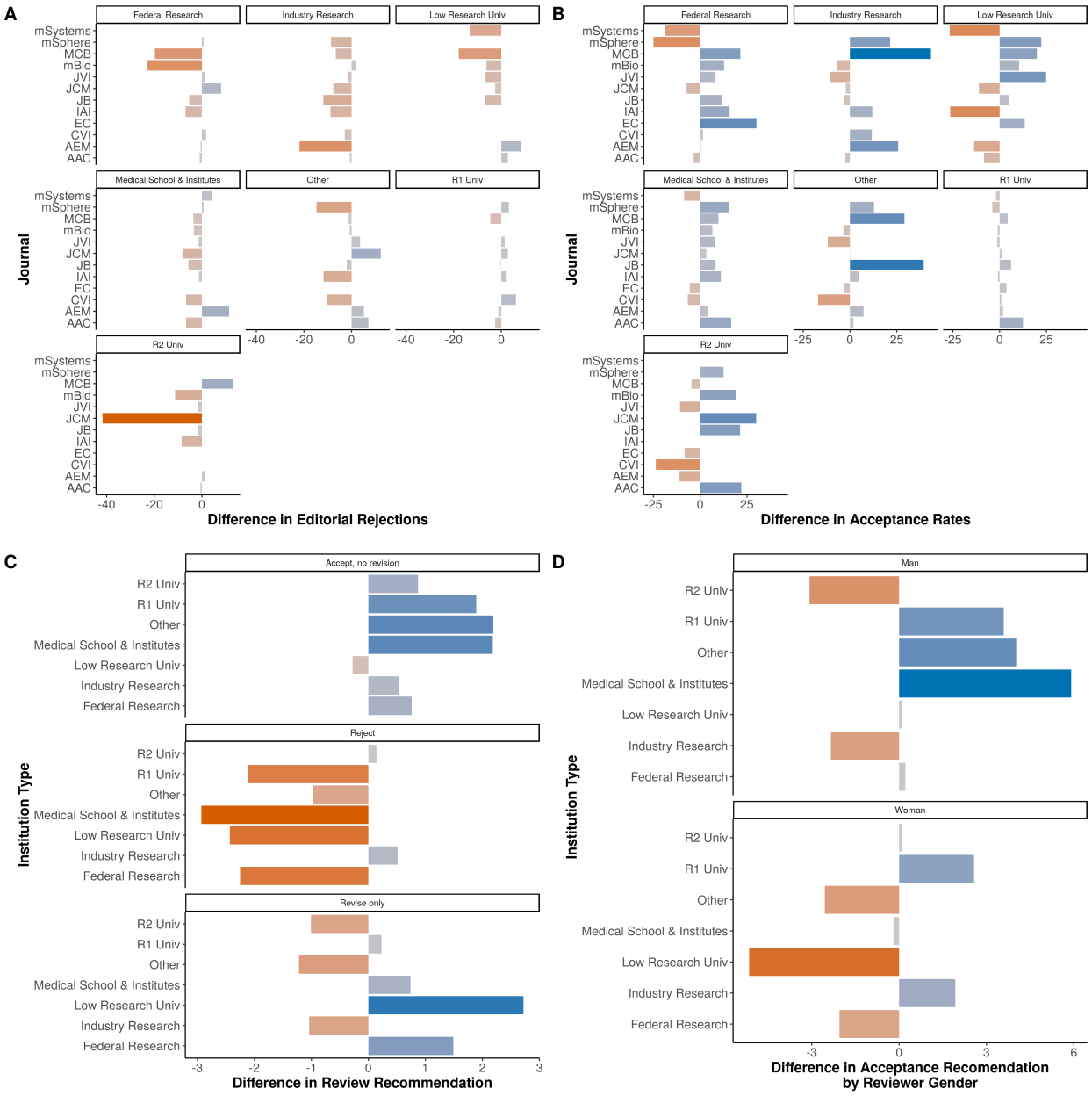

Figure S6. Difference in acceptance and rejections by institution type.

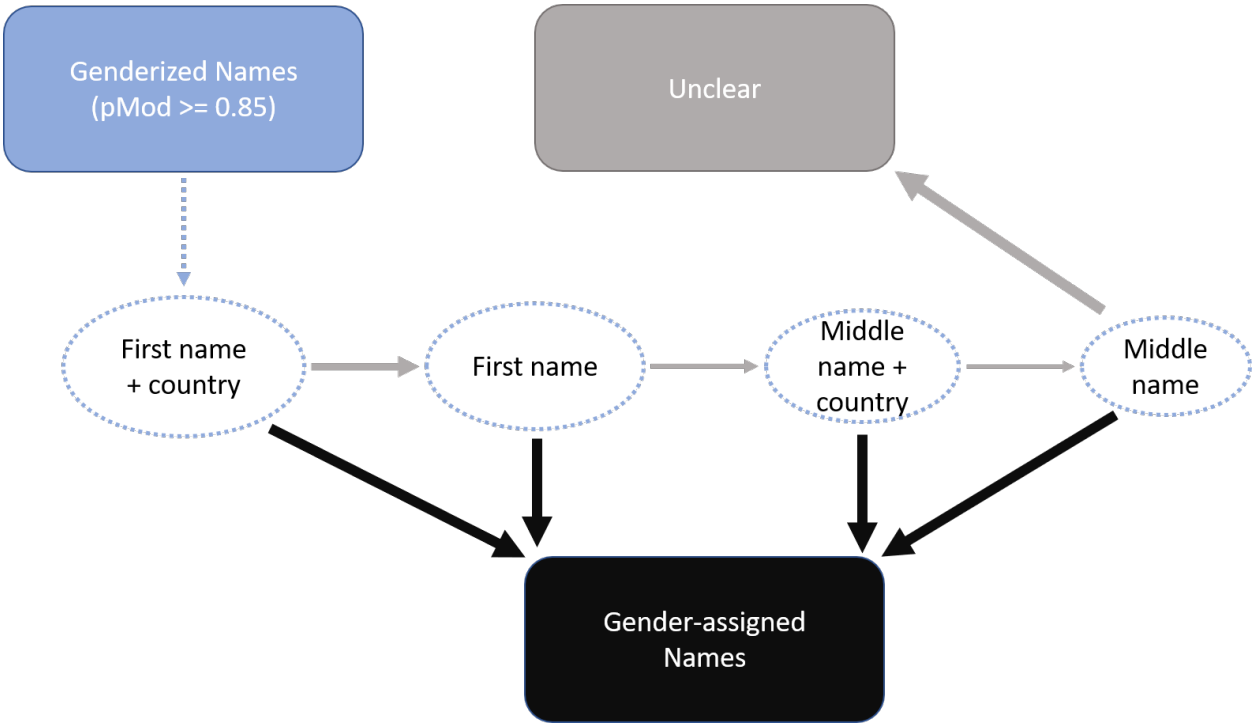

Figure S7. Schematic of gender prediction and assignment.

Table S1. sensitivity/specificity/accuracy of genderize thresholds. Bolded text denotes the accuracy of the threshold used in all further analyses.

	First Names			Plus Country Data		
Measure	p0.5	p0.85	pmod0.85	p0.5	p0.85	pmod0.85
Sensitivity	0.8943	0.9516	0.971	0.9055	0.9471	0.9669
Specificity	0.9339	0.9593	0.972	0.9265	0.9553	0.9727
Accuracy	0.9110	0.9549	**0.9714**	0.9146	0.9507	**0.9695**

$$\boldsymbol{Impact_C} = \left| \frac{(\%\,\boldsymbol{Unpredicted_C} - \%\,\boldsymbol{Unpredicted_{Total}})\; \boldsymbol{x}\; \left(\frac{\boldsymbol{Observations_C}}{\boldsymbol{Observations_{Total}}}\right)}{\%\,\boldsymbol{Unpredicted_{Total}}} \right|$$

Figure S8. Equation for calculating negative bias by genderize. C indicates an individual country.

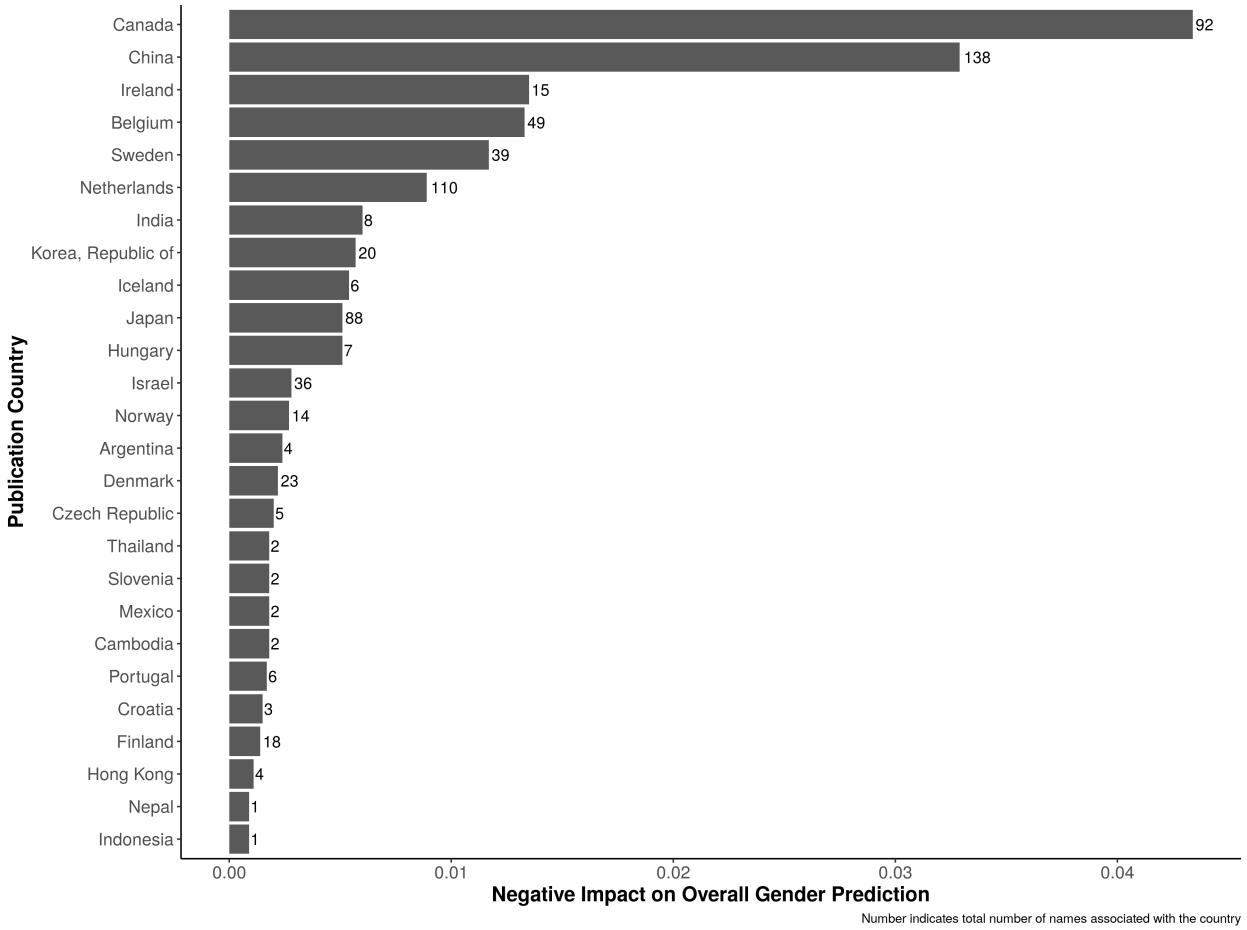

Figure S9. The negative impact of each country on the overall gender prediction of the validation dataset. Number indicates the total number of names associated with each country.

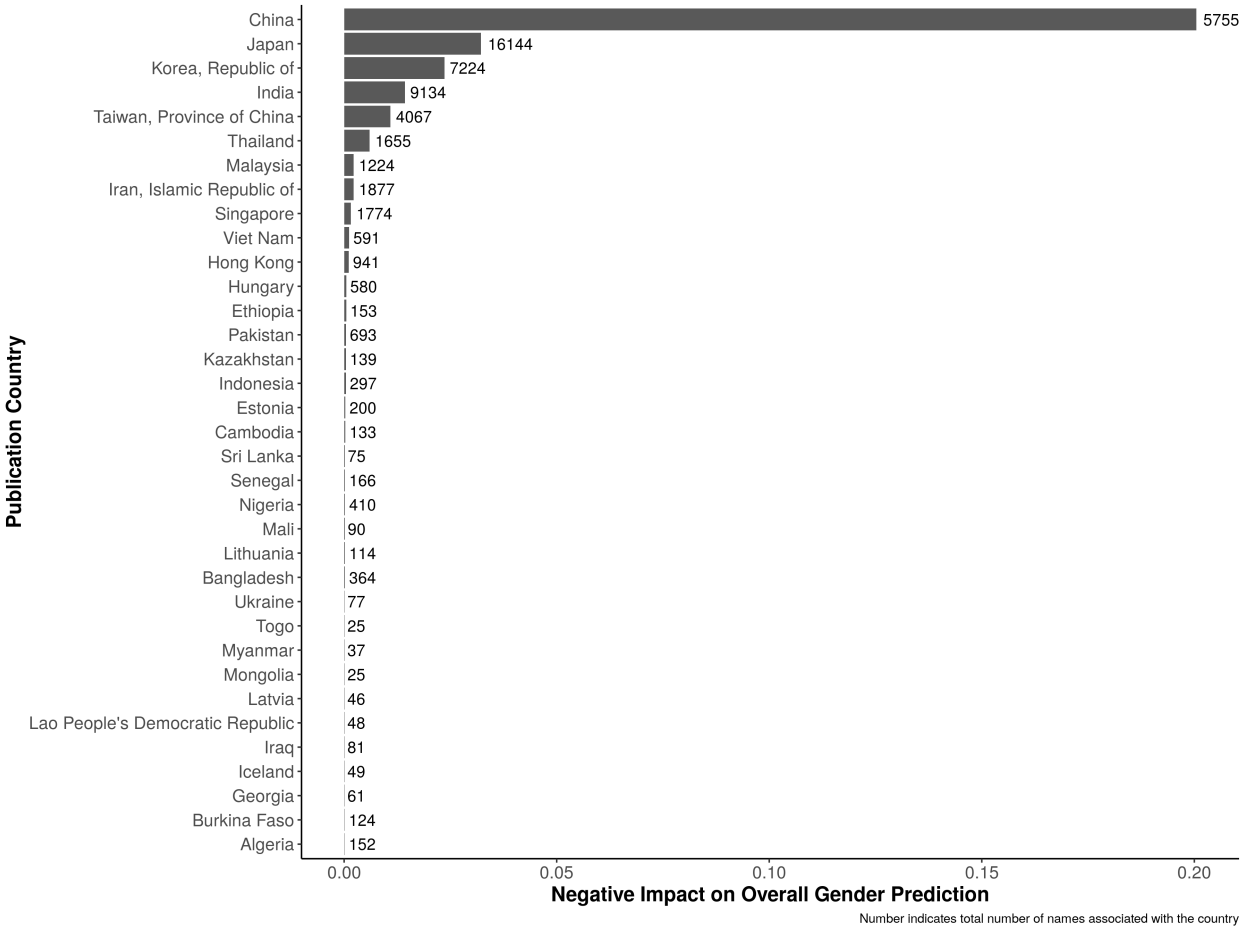

Figure S10. The negative impact of each country on the overall gender prediction of the full dataset. Number indicates the total number of names associated with each country.