

Supplementary

Gender representation

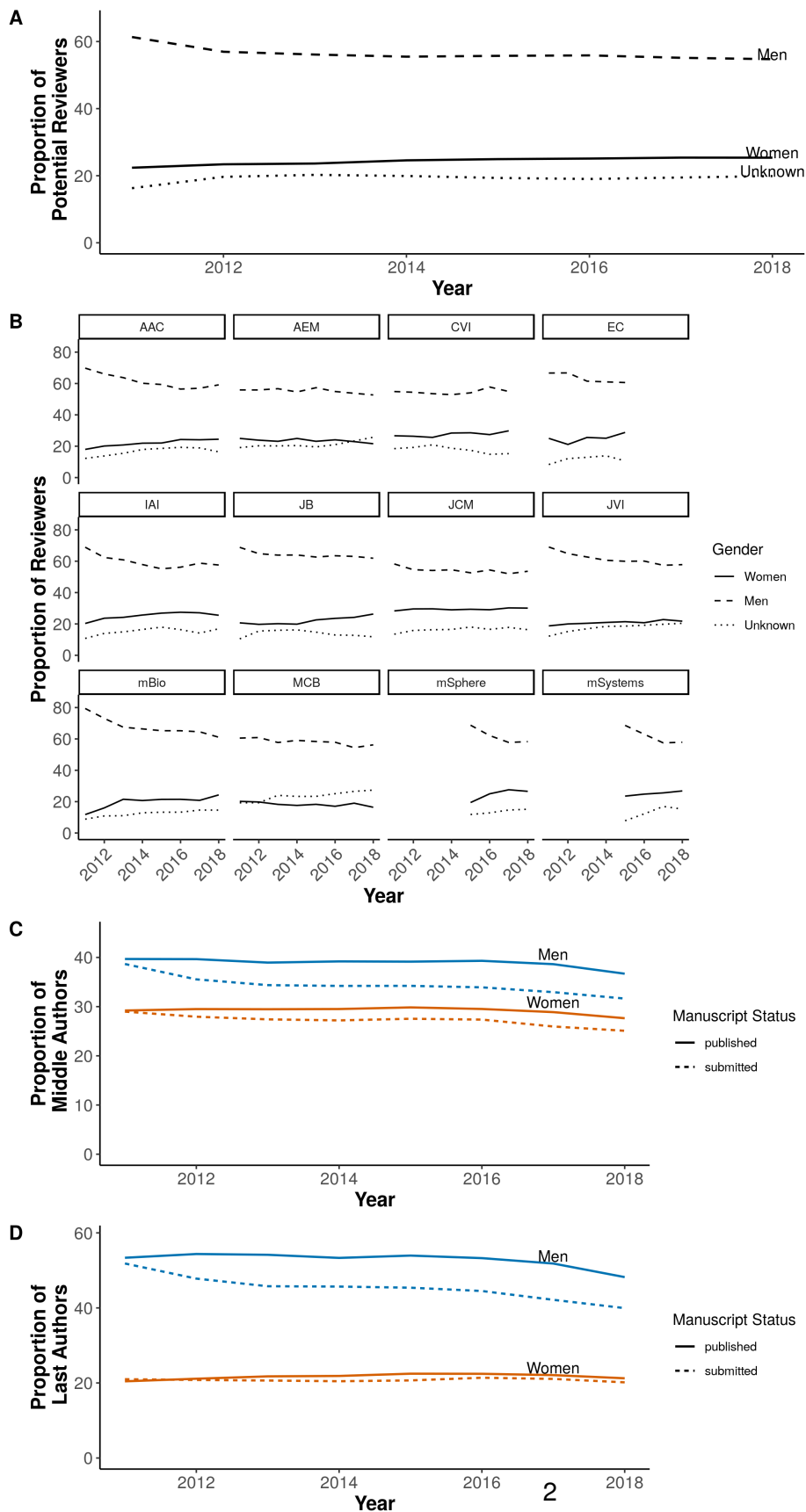


Figure S1. The proportion of (A) potential reviewers at all ASM journals combined, (B) reviewers at each ASM journal, (C) all submitting (dashed line) and publishing (solid line) middle authors, and (D) all submitting and publishing last authors by gender.

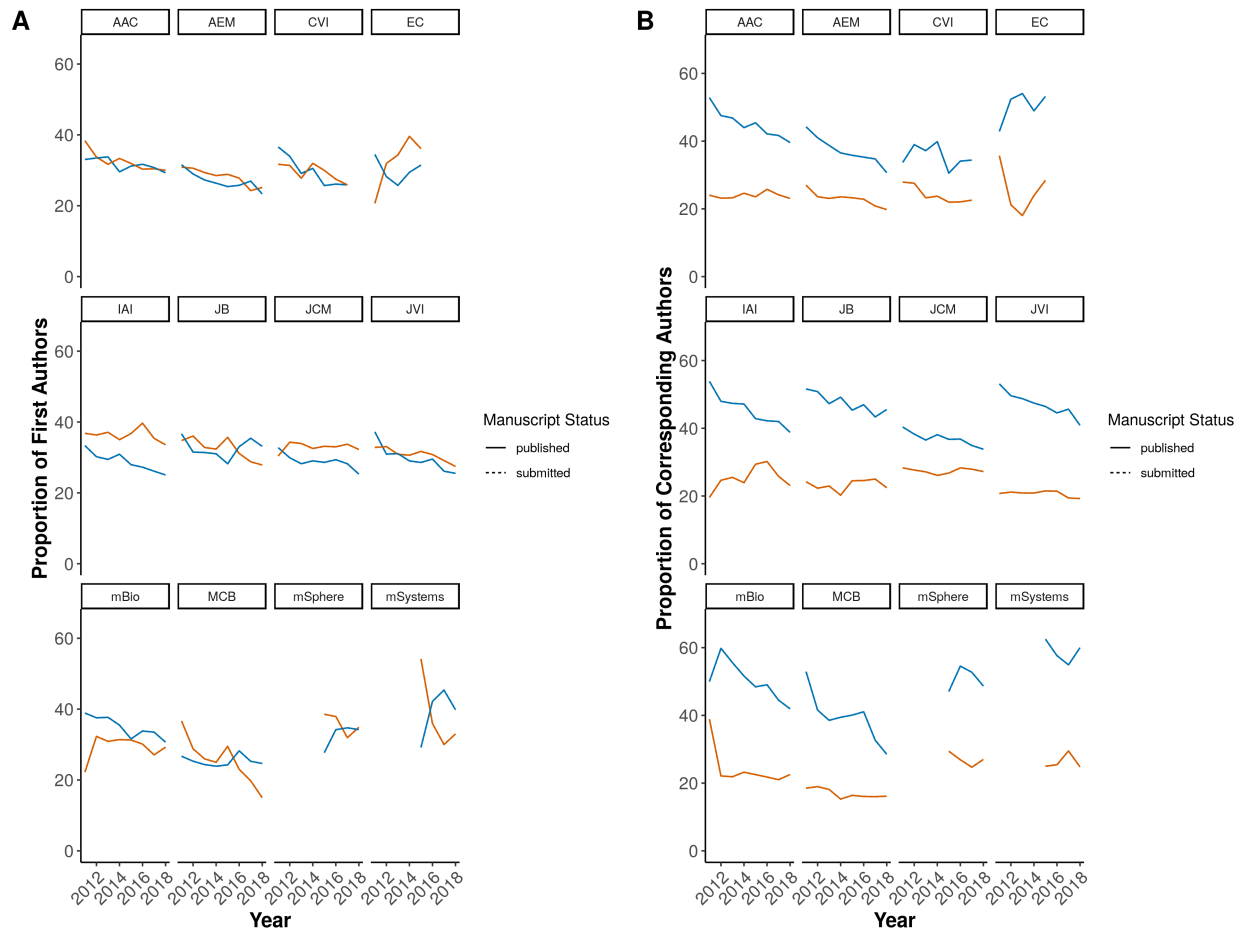


Figure S2. The proportion of all submitting (dashed line) and publishing (solid line) (A) first and (B) corresponding authors by gender at each ASM journal.

Gender bias

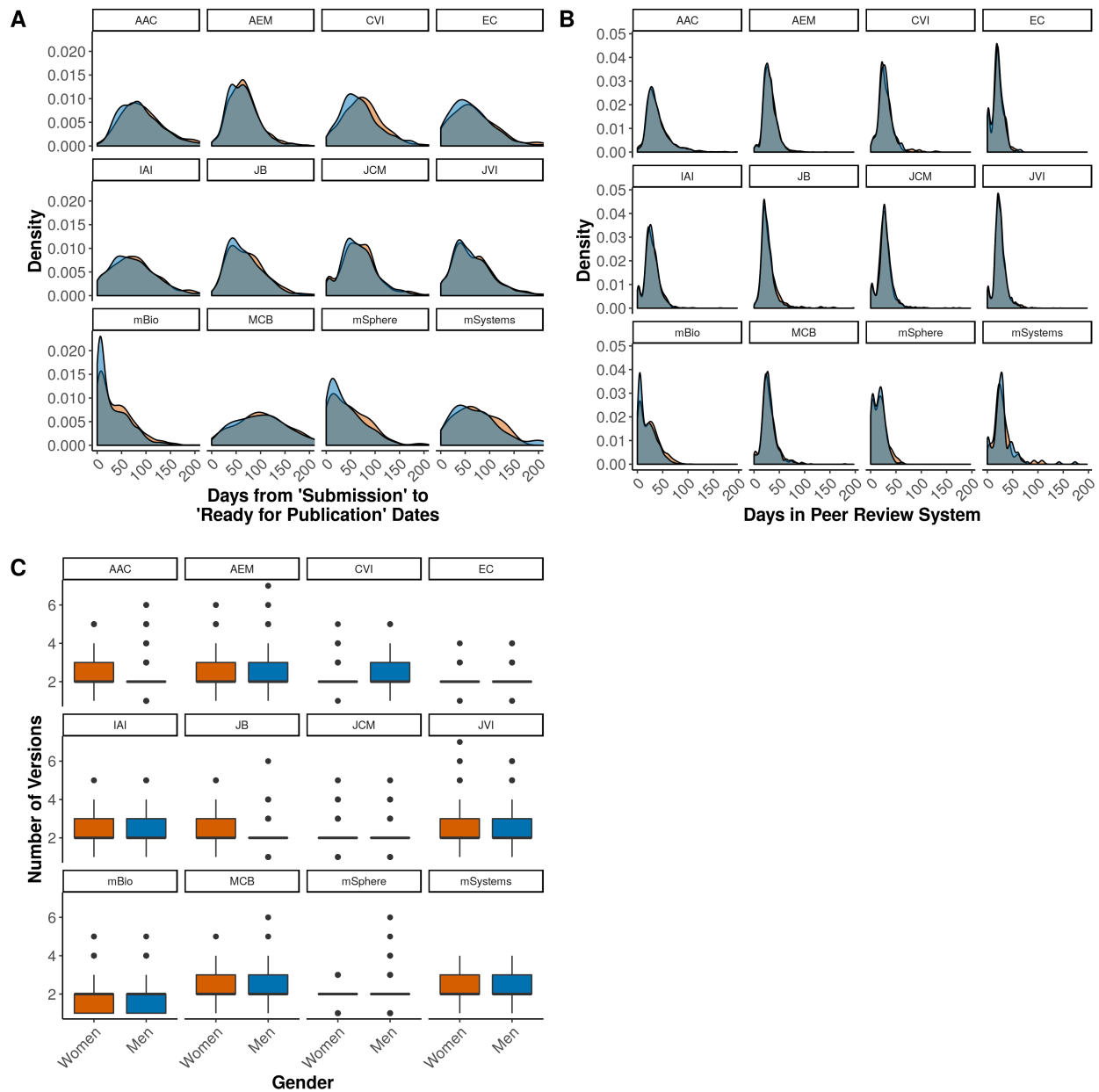


Figure S3. Comparison of time to final decision and impact by gender.** The number days (A) between when a manuscript is initially submitted then finally published and (B) that a manuscript spends in the ASM peer review system. How the impact of papers published by men (blue) versus women (orange) vary 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.

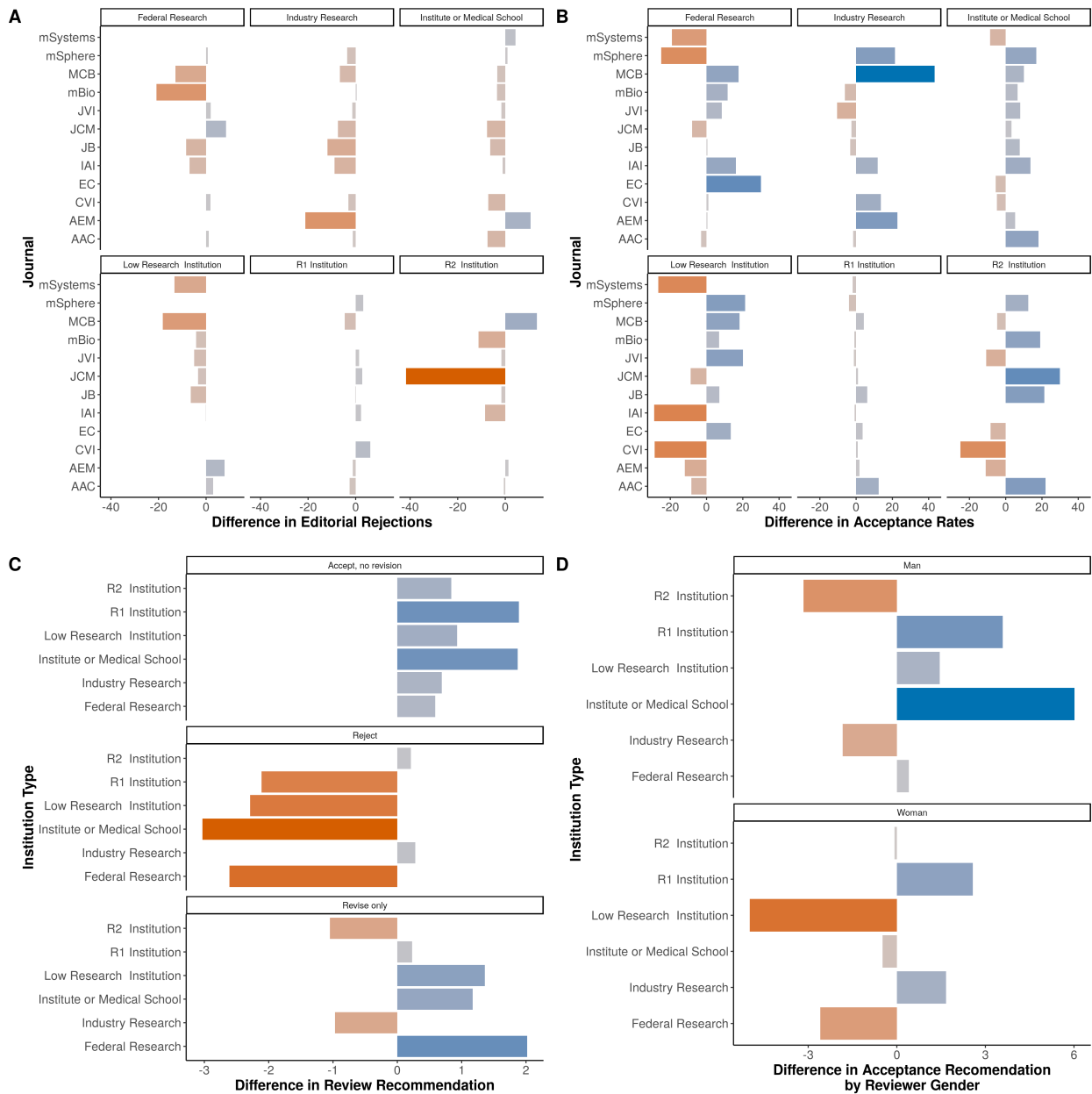


Figure S4. Difference in acceptance and rejections by institution type.

Gender prediction and assignment

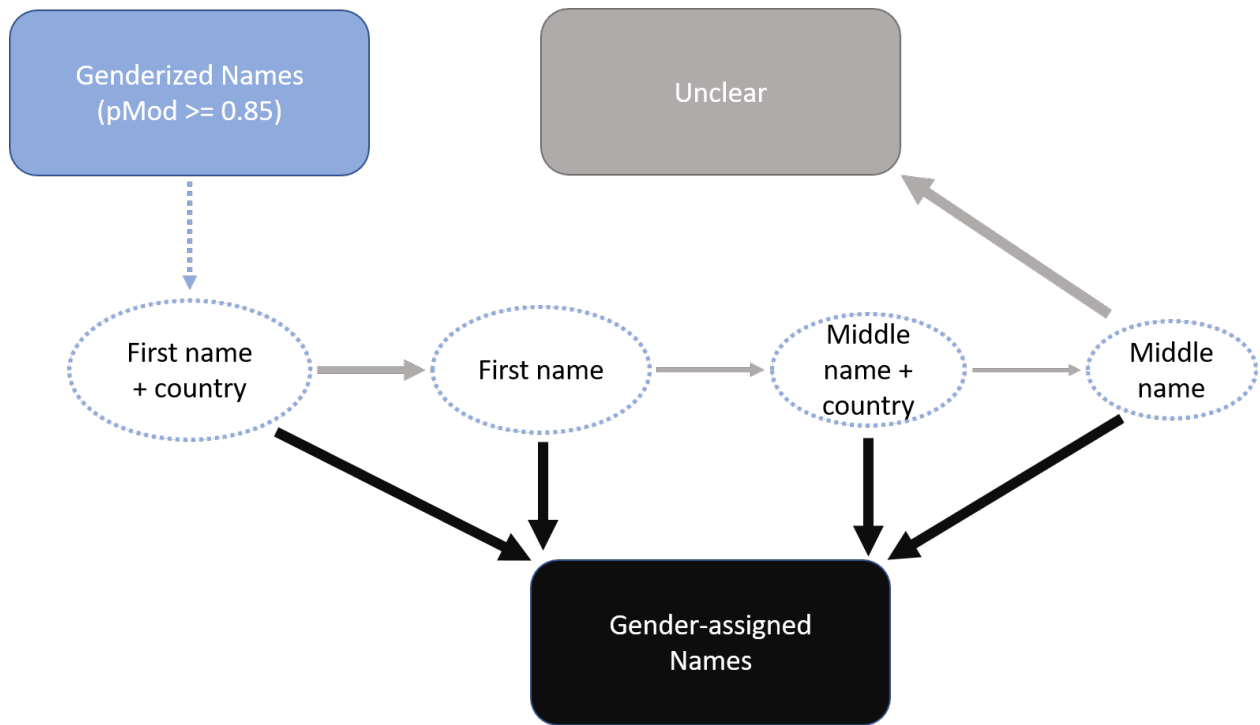


Figure S5. Schematic of gender prediction and assignment.

Validating gender analysis

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

Measure	First Names			Plus Country Data		
	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

$$Impact_c = \left| \frac{(\% \text{ Unpredicted}_c - \% \text{ Unpredicted}_{Total}) \times \left(\frac{Observations_c}{Observations_{Total}} \right)}{\% \text{ Unpredicted}_{Total}} \right|$$

Figure S6. Equation for calculating negative bias by genderize. C indicates an individual country.

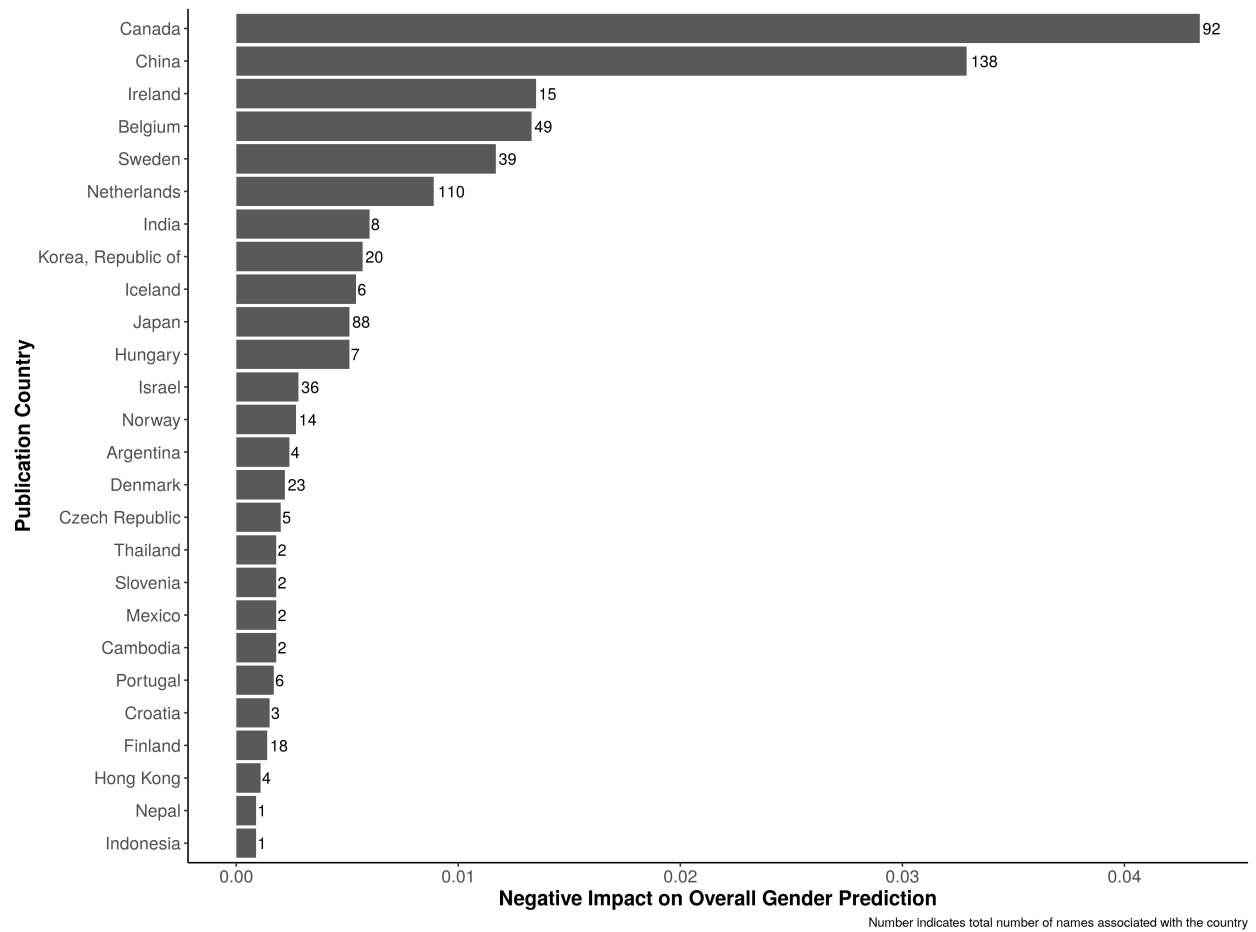


Figure S7. The negative impact of each country on the overall gender prediction of the validation dataset. Number indicates the total number of names associated with each country.

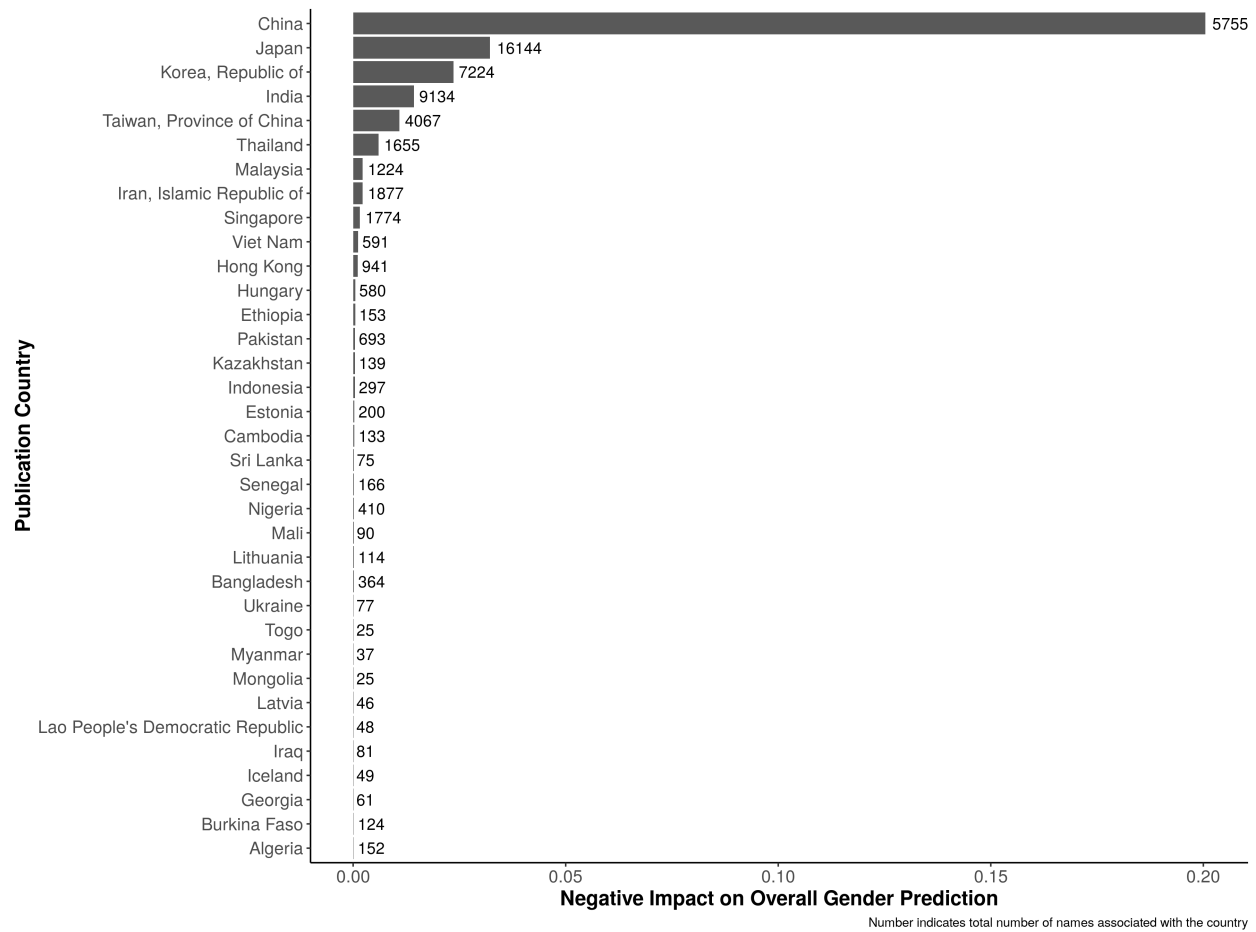


Figure S8. The negative impact of each country on the overall gender prediction of the full dataset. Number indicates the total number of names associated with each country.