AEA Papers and Proceedings 2018, 108: 175–179
https://doi.org/10.1257/pandp.20181101

GENDER ISSUES IN ECONOMICS

Gendered Language on the Economics Job Market Rumors Forum†

By Alice H. Wu*

Women are underrepresented in math-intensive fields (Ceci et al. 2014; Kahn and Ginther 2017), and analysts have noted that the representation gap is as large or larger in economics than in STEM (science, technology, engineering, and math) fields on average (e.g., Bayer and Rouse 2016). Among various mechanisms that have been proposed to explain this gap,[1] one that seems particularly relevant but that has not yet been evaluated systematically, is the role of an unwelcoming culture that reinforces stereotypical beliefs of men as an in-group in the field and women as an out-group (e.g., Tajfel and Turner 1986; Tonso 1996).

This paper attempts to assess the existence of an unwelcoming or stereotypical culture using evidence on how women and men are portrayed in anonymous discussions on the Economics Job Market Rumors forum (EJMR). As its name suggests, EJMR was established to share information about job interviews and outcomes in each year's hiring cycle, though it is active year-round. EJMR users post anonymously about economics-related or miscellaneous issues. Anonymity presumably eliminates social pressures that constrain participants' speech in other public settings, leading to a record of postings that reveal what participants believe but would not otherwise openly express.

I use a Lasso logistic model to measure gendered language in EJMR postings, identifying the words that are most strongly associated with discussions about one gender or the other. I find that the words most predictive of a post about a woman (*female* words) are generally about physical appearance or personal information, whereas those most predictive of a post about a man (*male* words) tend to focus on academic or professional characteristics. Despite some intervention by EJMR moderators, the top *female* words include several explicitly sexual terms. Gendered language is also shown to be widespread: about one in five posts about women (*Female* posts) contains at least one of the top 50 *female* words selected by Lasso, many of which are arguably inappropriate for a professional forum. Finally, I evaluate the robustness of the word-selection process through a subsampling exercise, which provides more confidence in the conclusion of differential portrayal of women and men on the forum.

I. Data

I scraped 2,217,046 posts on the first and last page of 223,475 threads on EJMR initiated or updated between October 2013 and October 2017. In the absence of a pre-existing dictionary, I identified the most frequent 10,000 words from the raw text and recorded the word counts for each word in each post. To determine the gender of the subject of each post, I extracted a list of 57 female classifiers (e.g., "she"/"woman") and a list of 236 male classifiers (e.g., "he"/"man")

* University of California, Berkeley (email: haowen.wu@berkeley.edu). I am extremely grateful to David Card and Janet Currie for their guidance and support. I also thank David Romer, Patrick Kline, Amanda Pallais, Jesse Rothstein, Michael Reich, Will Dobbie, Justin Wolfers, Alessandra Fenizia, Sydnee Caldwell, Ulrike Malmendier, Wei Jiang, Dick Startz, Patrick Button, Jonathan Tan, seminar participants at UC Berkeley, Harvard Business School, and Bowdoin College for helpful comments.

† Go to https://doi.org/10.1257/pandp.20181101 to visit the article page for additional materials and author disclosure statement.

[1] For example, recent studies examine course-taking patterns and comparative advantage (Card and Payne 2017), the impacts of role models (Carrell, Page, and West 2010), and stereotype beliefs (Reuben, Sapienza, and Zingales 2014; Bordalo et al. 2016).

from the top 10,000 words. The gap between the numbers of female and male classifiers is driven by the different numbers of female and male names among the top 10,000 words.

I consider a post to be *Female* if it contains any female classifier and *Male* if it contains any male classifier. Using the comprehensive list of gender classifiers, I identify 444,810 gendered (*Female* or *Male*) posts, comprising over 20 percent of all posts over the past four years. These gendered posts are from 138,477 threads, representing about 62 percent of all threads in the past four years.

II. Lasso Logistic Model

I fit a Lasso logistic model to predict whether a gendered post is *Female* or *Male* using the types of words in the post. My hypothesis is that an unwelcoming or stereotypical culture will lead EJMR participants to use terms to describe men that emphasize their fit and position within the field and terms to describe women that de-emphasize their professional accomplishments. Specifically, letting $\mathbf{w}_i$ denote a vector of counts for each of the most common words (excluding all gender classifiers) that are present in gendered post i, I estimate a Lasso-regularized logistic model for the probability that the post is *Female*, as follows:

$$\hat{\theta}_\lambda = \arg\min_\theta - \log(\Pi_{i=1}^{N} P(Female_i \,|\, \mathbf{w}_i)) + \lambda||\theta||_1,$$

where $\|\theta\|_1 = \sum_{j\geq 1} |\theta^j|$. The Lasso regularization helps identify words with the strongest predictive power while avoiding over-fitting. I estimate the model using only gendered posts that refer uniquely to one gender or the other, excluding posts that contain classifiers for both genders (which account for about 10 percent of gendered posts).

A. *Model Training Process*

There are 401,734 posts that include only female or only male classifiers from the comprehensive list. I use a 75 percent random sample to train the model and select an optimal tuning parameter λ^* through 5-fold cross validation. I then select the *p*-score threshold that minimizes the mean squared error for predicting gender on the remaining 25 percent as the test set.[2] This leads to the selection of a threshold of $p^* = 0.40$ for assigning a post to be *Female*. I use the same threshold to assign genders for the posts that include both female and male classifiers: 31.8 percent of the posts that contain classifiers for both genders are re-classified to *Female*, and the rest to *Male*.

B. *Gendered Words*

The estimated model identifies about 4,500 words with nonzero predictive power for determining gender. I sorted these words by their marginal effect, i.e., the increase in the probability that the subject of a post is *Female* given an additional occurrence of each word. Table 1 displays the top ten words selected by Lasso.

The table reveals that the words that are most predictive of a *Female* post are typically about: (i) physical appearance; (ii) personal or family information; or (iii) gender issues/sexism. The words "hot" and "attractive" increase the predicted probability that a post is discussing a female by approximately 27.1 percent and 24.5 percent, respectively. While such terms could be viewed as complimentary in other settings, in this setting they arguably reflect the treatment of women as an out-group who are to be judged by nonprofessional standards (e.g., physical appearance). For example, there is a thread titled "Cute, unmarried HRM AP is doing a seminar at my school. Can I ask her out?",[3] which judges a female economist based on her appearance, with no reference to professional-related attributes.

In contrast the words that predict a *Male* post include more academically and professionally oriented terms. For example, "adviser," "supervisor," and "Nobel" are in the 30 most predictive *male* words, and each increases the probability that a post is discussing a male by about 13 percent–15 percent. Nevertheless, the Lasso model also picks up a few offensive (and potentially out-group-defining) terms such as "homo," suggesting an unwelcoming online environment for some subgroups of males.

[2] See online Appendix Figure 1 for a plot of MSE at each *p*-score cutoff.

[3] This thread was initiated and last updated 2 years ago. It contains 20 posts and has 1,238 views.

TABLE 1—TOP 10 WORDS MOST PREDICTIVE OF FEMALE/MALE

Most *female*		Most *male*	
Word	ME	Word	ME
Hotter	0.422	Homo	−0.303
Pregnant	0.323	Testosterone	−0.195
Plow	0.277	Chapters	−0.189
Marry	0.275	Satisfaction	−0.187
Hot	0.271	Fieckers	−0.181
Marrying	0.260	Macroeconomics	−0.180
Pregnancy	0.254	Cuny	−0.180
Attractive	0.245	Thrust	−0.169
Beautiful	0.240	Nk	−0.165
Breast	0.227	Macro	−0.163

Notes: The model was trained on a 75 percent sample of gendered posts that contain only female or only male classifiers from the comprehensive list. ME—the marginal effect of word *w* is the change in probability that a post is discussing a female, when it contains an additional word *w*. The words that predict *Female* (*Male*) are sorted in descending (ascending) order of the ME.

The moderation policy on the EJMR forum is based both on an automatic censorship of words and on reports by users. The evidence of stereotyped and offensive language captured here suggests that either the moderators did not remove the threads reported by users, or that the users themselves tolerated such content and did not complain.

To make inferences about the pervasiveness of gendered language, I consider the frequency of the words selected by Lasso.[4] Some of the most *female* words also turn out to be relatively common. For example, the word "hot" shows up in about 3.5 percent of the *Female* posts, and ranks as the third most common term in *Female* posts, whereas the third most common word in *Male* posts is "job." Overall, about 19.4 percent of all *Female* posts include at least one of the top 50 *female* terms, most of which highlight physical attributes or personal information.

III. Robustness Check

One concern about assigning gender to posts based on the comprehensive list of gender classifiers is that it may over-identify gendered words that occur in personal discussions about "girlfriends" or "boyfriends," which are included as classifiers. To address this concern, I conduct a robustness check by replicating the analysis using gendered posts identified only by gender pronouns (e.g., "he" or "she"). Relative to gendered posts identified using the comprehensive list, more of the posts identified using pronouns, referred to as the pronoun sample, pertain to specific individuals. As a result, the model trained on the pronoun sample should pick up more academic or professional terms for both genders.

Following the same procedure, I train a Lasso logistic model[5] on 35,850 *Female* posts and 103,449 *Male* posts in the pronoun sample. As expected, the estimated model based on the pronoun sample identifies a few more academic terms. For example, "AEJ" (ME: 13.6 percent) and "RCT" (ME: 13.3 percent) appear among the top *female* words.[6] The marginal effects of terms such as "advisor," "Nobel," and "promoted" among the top *male* words become stronger. Nevertheless, an overwhelming majority of the *female* words continue to focus on non-academic aspects. For example, six out of the ten most *female* words selected when gender is determined by the comprehensive list of classifiers also appear when it is determined by pronouns only (see Table 2).

Finally, to evaluate the pervasiveness of gendered words identified using the two alternative sets of classifiers, Figure 1 plots the fraction of *Female* (*Male*) posts that contain at least one of the 50 words most strongly associated with *Female* (*Male*) under the two alternatives.

This trend plot for data in the most recent year reveals several interesting patterns of gendered language. First, there is a large gap between the pervasiveness of the top *female* versus top *male* words selected by Lasso, particularly when gendered posts are identified using the comprehensive list of gender classifiers. Across all months, about 17.2 percent to 19.6 percent of all *Female* posts identified by the comprehensive list include at least one of the top 50 *female* words, but for *male* words the equivalent measures

[4] See online Appendix Table 1, 2, and 3 for the top 50 *female* and *male* words selected by Lasso, the number of gendered posts each of the words occurs in, and the most frequent 50 words in gendered posts, respectively.

[5] For an additional check, I train a Lasso-regularized linear probability model on the pronoun sample, and the top 50 *female* or *male* words selected by the linear Lasso are shown in online Appendix Figure 2 and online Appendix Figure 3.

[6] For word selection by Lasso logistic on the pronoun sample, see online Appendix Tables 4, 5, and 6.

Table 2—Top 10 Words Most Predictive of Female/Male (*Pronoun sample*)

Most *female*		Most *male*	
Word	ME	Word	ME
Pregnancy	0.292	Knocking	−0.329
Hotter	0.289	Testosterone	−0.204
Pregnant	0.258	Blog	−0.183
Hp	0.238	Hateukbro	−0.176
Vagina	0.228	Adviser	−0.175
Breast	0.220	Hero	−0.174
Plow	0.219	Cuny	−0.173
Shopping	0.207	Handsome	−0.166
Marry	0.207	Mod	−0.166
Gorgeous	0.201	Homo	−0.160

Note: The model was trained on a 75 percent sample of gendered posts that contain only feminine pronouns or only masculine pronouns.

range from 7.3 percent to 9.3 percent. In the pronoun sample, the gap in pervasiveness shrinks: the top *female* words selected when gender is identified only by pronouns become less common, whereas the top *male* words become more common.

Second, there is larger month-to-month variation in the pervasiveness of the top *female* words selected through the pronoun sample than through the complete sample, especially during the job market season. It is disturbing to see that within the pronoun sample, the fraction of *Female* posts that include at least one of the top 50 *female* terms can be 3 to 4 percentage points higher in particularly active months of the job market (December 2016, February and March 2017) than other months. Such variation suggests that the competitive environment of the job market may lead to more gendered discussions about female and male candidates.

Third, there is evidence of some effect of media discussions about the content of EJMR postings in August 2017. A *New York Times* article by Justin Wolfers,[7] citing results from Wu (2017), raised some concerns about the gendered discussions on EJMR. This treatment appears to have led to a decline in the occurrences of the top *female* words in the pronoun sample in the following two months, which may

[7] Wolfers, Justin. 2017. "Evidence of a Toxic Environment for Women in Economics." *New York Times*, August 18. https://www.nytimes.com/2017/08/18/upshot/evidence-of-a-toxic-environment-for-women-in-economics.html.

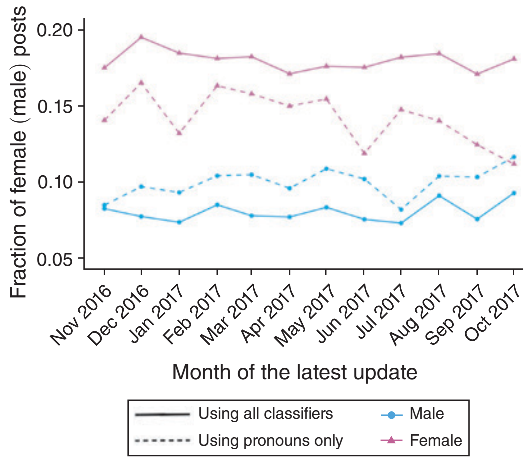

Figure 1. Fraction of *Female* (*Male*) Posts that Include any Top 50 *Female* (*Male*) Words, under Two Alternatives

Notes: The solid lines plot the fraction of *Female* (*Male*) posts identified by the comprehensive list of gender classifiers that include at least one of the top 50 *female* (*male*) words selected by the Lasso-Logistic model. The dashed lines plot the equivalent measures for the word selection based on gendered posts identified by pronouns only. For threads initiated or updated from November 2016 to October 2017, I identified the month of its most recent post from the rough time stamps on the main pages of EJMR.

reflect either a decrease in the usage of gendered words or an increase in censoring by EJMR moderators. If the censoring is playing a more important role, however, then this trend should not be interpreted as a change in the underlying beliefs or attitudes of the posters.

To summarize, despite the differences in the pervasiveness of gendered words selected under the two alternatives, this robustness check confirms that the postings about women tend to highlight physical appearance, personal information, and sexism, whereas those about men are more academically or professionally oriented.

IV. Discussion

This paper illustrates the use of text analytic techniques to measure gendered language between posts pertaining to women and men. The gendered posts may not necessarily talk about specific female or male academics, but they play a large role in shaping the overall atmosphere on this forum for economists, which may consolidate the perception of men as an in-group versus women as an out-group.

However, an analysis at the word level provides an incomplete picture of the stereotyping behavior on EJMR. Wu (2017) designs a topic analysis and provides an econometric framework for quantifying stereotyping in the dynamics of conversation. Wu (2017) also shows that high-profile female economists tend to receive more attention than their male counterparts, which may suggest that the work by women is more heavily scrutinized.

REFERENCES

Bayer, Amanda, and Cecilia Elena Rouse. 2016. "Diversity in the Economics Profession: A New Attack on an Old Problem." *Journal of Economic Perspectives* 30 (4): 221–42.

Bordalo, Pedro, Katherine Coffman, Nicola Gennaioli, and Andrei Shleifer. 2016. "Stereotypes." *Quarterly Journal of Economics* 131 (4): 1753–94.

Card, David, and A. Abigail Payne. 2017. "High School Choices and the Gender Gap in STEM." National Bureau of Economic Research Working Paper 23769.

Carrell, Scott E., Marianne E. Page, and James E. West. 2010. "Sex and Science: How Professor Gender Perpetuates the Gender Gap." *Quarterly Journal of Economics* 125 (3): 1101–44.

Ceci, Stephen J., Donna K. Ginther, Shulamit Kahn, and Wendy M. Williams. 2014. "Women in Academic Science: A Changing Landscape." *Psychological Science in the Public Interest* 15 (3): 75–141.

Kahn, Shulamit, and Donna Ginther. 2017. "Women and STEM." National Bureau of Economic Research Working Paper 23525.

Reuben, Ernesto, Paola Sapienza, and Luigi Zingales. 2014. "How Stereotypes Impair Women's Careers in Science." *Proceedings of the National Academy of Sciences of the United States of America* 111 (12): 4403–08.

Tajfel, Henri, and John C. Turner. 1986. "The Social Identity Theory of Intergroup Behavior." In *Psychology of Intergroup Relations*, edited by William G. Austin and Stephen Worchel, 7–24. Chicago: Nelson-Hall.

Tonso, Karen L. 1996. "The Impact of Cultural Norms on Women." *Journal of Engineering Education* 85 (3): 217–25.

Wu, Alice H. 2017. "Gender Stereotyping in Academia: Evidence from Economics Job Market Rumors Forum." Unpublished.