Gender Stereotyping in Academia: Evidence from Economics Job Market Rumors Forum

Alice H. Wu*

August 2017

Abstract

This paper examines whether people in academia portray and judge women and men differently in everyday "conversations" that take place online. I combine methods from text mining, machine learning and econometrics to study the existence and extent of gender stereotyping on Economics Job Market Rumors forum. Through a topic analysis, I find that the discourse tends to become significantly less academic or professional oriented, and more about personal information and physical appearance when women are mentioned. The words with the strongest predictive power on gender, selected by the Lasso-logistic model, provide a direct look into the gender stereotyping language on this forum. Moreover, a panel data analysis reveals the state dependence between the content of posts within a thread. In particular, if women are mentioned previously in a thread, the topic is likely to shift from academic to personal. Finally, I restrict the analysis to discussions on specific economists, and find that high-profile female economists tend to receive more attention on EJMR than their male counterparts.

*I would like to thank my thesis advisor David Card for his invaluable guidance and support. I also thank Janet Currie, David Romer, Ulrike Malmendier, Wei Jiang, Alessandra Fenizia, Sydnee Caldwell and Jonathan Tan for very helpful discussions and suggestions at different stages of the project. All errors are my own. Contact information: alice.hw15@gmail.com.

Despite remarkable gains in educational attainment in recent decades, women are still underrepresented in math-intensive fields like economics, engineering, and computer science (Ceci et al. 2014). Some analysts believe that this continued underrepresentation reflects the impact of subtle day-to-day interactions that convey the message that women do not "belong" in science, technology, engineering, and math (STEM) fields (e.g. Kahn et al. 2017). Recent controversies surrounding the sexist culture in Silicon Valley underscore the ubiquity of such gender stereotyping (e.g. Seetharaman 2017; Wootson 2017). Women as tokens in a mainly male group are likely to face disproportionate attention and distortions to fit presumed stereotypes (Kanter 1997). Gender stereotypes can also be reinforced by confirmation biases that lead people to over-weight signals that fit with their prior beliefs and ignore others that challenges these beliefs (Bordalo et al. 2016a).

Althought there is an emerging literature in economics formally modeling gender stereotypes and testing for them in lab experiments (Bordalo et al. 2016b), the topic remains understudied. One difficulty is that and even in lab settings gender biases may not manifest in easily observable behavior. Another difficulty is that even if researchers were able to observe interpersonal interactions in a controlled setting, subjects who are concerned about political and social correctness would not necessarily reveal their true stereotype beliefs.

In this paper, I aim to examine whether people in academia portray and judge women and men differently in everyday "conversations" that take place online. With a focus on the field of economics, I use text scraped from Economics Job Market Rumors[1] (EJMR), an online forum where the majority of users are current graduate students in Ph.D. economics programs. The forum was established to share information about job applications and results in each year's hiring cycle. On top of this, it is active all year round, and users post anonymously about economics-related or miscellaneous issues. Anonymity presumably eliminates any social pressure participants may feel to edit their speech, and thus creates a natural setting to capture what people believe but would not openly say. There is a fair amount of gender-related discussions on this forum, which can address women and men in general or be in reference to specific economists. Despite the presence of moderators[2] who regularly remove offensive or inappropriate material, the remaining posts still show significant evidence of gender stereotyping. For example, Table 1 shows the top 10 words

[1] More information about EJMR on https://www.econjobrumors.com/topic/about-ejmr

[2] Moderation policy: https://www.econjobrumors.com/topic/request-a-thread-to-be-deleted-here

with the strongest predictive power for each gender, selected by the Lasso-Logistic model discussed in Section 2.2.

Table 1: Words with the strongest predictive power for gender

Most "female"		Most "male"	
Word	Marginal Effect	Word	Marginal Effect
hotter	0.388	homosexual	-0.237
hot	0.285	homo	-0.228
attractive	0.260	philosopher	-0.204
pregnant	0.252	keen	-0.182
gorgeous	0.251	motivated	-0.171
beautiful	0.249	fieckers	-0.164
tits	0.247	slides	-0.160
lesbian	0.242	nordic	-0.156
bang	0.229	filling	-0.152
horny	0.224	textbook	-0.148

Notes: the marginal effect of word w is the change in probability of a post being classified as *female*, i.e. 1 if it is discussing women, when it contains one more word w.

Text as data is relatively new to economics research (Gentzkow et al. 2017). The advances in text mining and machine learning enable researchers to extract meaningful features unique to the textual data. This study provides an illustration of the use of these tools in combination with econometric methods to get at a phenomenon that is otherwise difficult to quantify. First, I use a list of gender classifiers (e.g. "she", "he") to identify posts related to gender. At the post level, I approach the main research question through two channels: *Topic Analysis* and *Word Selection*. The first channel provides an overview of the content of posts, whereas the second captures the effects of gender on language at finer granularity. In topic analysis, I classifiy the most frequent $10,000$ words into 15 categories and mainly consider two topics of interest: (i) *Academic/Professional*; (ii) *Personal/Physical*. Results show that a $Female = 1$ post on average contains 43% significantly less academic or professional terms, and 192% more terms about personal information or physicial attributes[3]. As for the word selection, I use a logistic regression with ℓ_1-norm penalty, a method often applied to the analysis of high-dimensional text data (e.g. Gentzkow et al. 2016). The model I train with 5-fold cross validation identifies about $3,600$ words with meaninful predicitive power for gender. The top "female" and "male" words sorted by marginal

[3]The results here are based on the sample of gender-related posts identified by all possible gender classifiers (Level 1). See Figure 1 for details.

effect as shown in Table 1 further illustrates the pervasiveness of gender stereotyping in this online forum.

I extend the analysis to gender-related thread environments, where I preserve a thread if its title or at least one of its post is related to gender. A user is prompted to click on a thread by its title. Titles related to female, i.e. containing a *female* word, attract 2 more posts on average than those related to male. Topic analysis at the thread level yields consistent findings that the the discussion becomes significantly less academic or work oriented, and more about personal information or physicial appearance when women are being mentioned more than men. In addition, tests of state dependence between adjacent posts reveal that a post tends to deviate from being academic/professional when its prior one is related to female. These results demonstrate the pervasiveness of gender stereotyping in daily online discussions among people in the field of economics. This unwelcoming atmosphere online may be discouraging current female graduate students and early career academics.

Finally, I present a difference-in-difference analysis on the attention recevied by a comparable set of 190 female and 190 male high-profile economists who rank among the Top 5% of Authors on the RePEc ranking[4], and a second analysis on a cohort of 204 assistant professors (45 women, 159 men) from Top 20 economics departments[5] in the U.S.. I estimate the amount of attention one receives by the number of results returned via a name search on EJMR. Among high-profile economists, women get more attention than their male counterparts, and the difference is widening for relatively less prominent economists. Among junior faculty, women working at the top 5 departments are discussed more than men on the forum, but this trend is reversed for those at lower-ranked departments.

The rest of the paper is structured as follows. Section 1 describes the raw data available on the EJMR forum, and the construction of my samples using different levels of gender classifiers. Section 2 presents the topic analysis on individually gender-related posts and the word selection using Lasso-logistic model. Section 3 discusses the popularity of threads in relation to gender, replicates topic analysis at the thread level and tests for state dependence between posts in a panel

[4]RePEc ranking of Top 5% Authors (Last 10 Years Publications), as of September 2016: `https://ideas.repec.org/top/top.person.all10.html`

[5]based on U.S. News ranking of best graduate programs in Economics as of 2013 and 2017, and RePEc ranking of top Economics Departments.

data setting. Section 4 uses an alternative design to analyze the attention comparable economists of different genders receive.

1 Data

By July 2nd, 2017, there are about 289,205 threads[6] available on the site of EJMR forum in a span of six years. I scrape the first page of each thread initiated or updated from 2014 to July 2016[7], and build up a dataset of 1,143,416 posts across 131,913 threads. Without preexisting dictionaries, I use the open-vocabulary strategy (Schwartz et al. 2011) to consider the most frequent 10,000 words that emerge from the raw text. I record the word counts in a N-by-10,000 sparse matrix, where $N = 1,143,416$, the number of posts.

In order to subset the gender-related posts, a list of gender classifiers was extracted from the top 10,000 words, which contain 44 words associated with females, and 134 words associated with males. On each post, I define $Female = 1$ if it includes any word indicating female, $Female = 0$ if it includes any word indicating male, and NA otherwise. Under this classification criteria, 26,002 posts that contain both female and male words become duplicate observations. To resolve this issue, I design a Lasso-Logistic propensity score model, and use 5-fold cross validaton to train the model on a subset of posts that contain only "female" or only "male" words. As a result, 9,044 of the duplicate posts are reclassified as "female" posts, and 16,958 are reclassified as "male" posts. In Section 2.2, I discuss this p-score model in detail and show a list of words with the strongest predictive power for gender.

Table 2: Summary of the text data

	All	Gender Related	Female (Level 1)	Male (Level 1)
Number of Posts	1,143,416	237,863	56,171	181,692
Number of Threads	131,913	74,679	22,708	66,796

Notes: "Level 1" uses the most inclusive set of gender classifiers to identify gender related posts. Duplicate observations have been resolved by the Lasso-Logistic model in Section 2.2.

Table 2 provides an overview of the entire dataset and of gender-related posts, using all pos-

[6] Estimated by 35 *threads/page* $\times$ 8,263 *number of pages* on 07/02/2017.

[7] The exact time stamp of a post is not available. On the website, threads are arranged in a reverse chronological order, according to "Freshness". i.e. the time the last post was added.

sible gender classifiers. Around 20.8% of all individual posts are gender-related, and they spread across 56.6% of the threads. I construct two samples, summarized in Table 3: (i) Sample 1 includes 237,863 individual posts that are gender-related; (ii) Sample 2 is a panel data set that preserves the entire page of a given thread if it includes at least one gender related post. Sample 2 includes 810,998 posts (including Sample 1) across 74,679 threads.

Table 3: Summary of Sample 1 and Sample 2

	Number of Posts				Number of Threads
	All	Gender Related	Female (Level 1)	Male (Level 1)	
Sample 1	237,863	237,863	56,171	181,692	74,679
Sample 2	810,998	237,863	56,171	181,692	74,679

Notes: Sample 2 look at the same set of threads as Sample 1, but preserves all posts on a given page rather than gender-related ones only. "Level 1" uses the most inclusive set of gender classifiers to identify gender related posts. Duplicate observations have been resolved by the Lasso-Logistic model in Section 2.2.

Gender Classifiers

Among the gender classifiers, the most straightforward ones are "she", "he" etc., while others can refer to a group or identity such as "women", "men","wife", "husband". In addition, there are first and economists' last names emerging from the top 10,000 words, most of which are male names and therefore result in a big gap in the total number of "female" and "male" words. Based on the characteristics of these classifiers, I divide them into four groups, and define four levels cumulatively as illustrtaed in Figure 1. Level 1 uses all classifiers, whereas Level 4 is the most restrictive using only "she", "he", "her", "him", "his", "herself", "himself". Figure 2 provides an overview of the number of female and male posts identified based on each level. Such specifications are particularly useful for robustness checks in later sections. The more restrictive specifications also address cases where posters refer themselves as "bros" or "guys" rather than the specific people they are discussing about.

Figure 1: Levels of Gender Classifiers

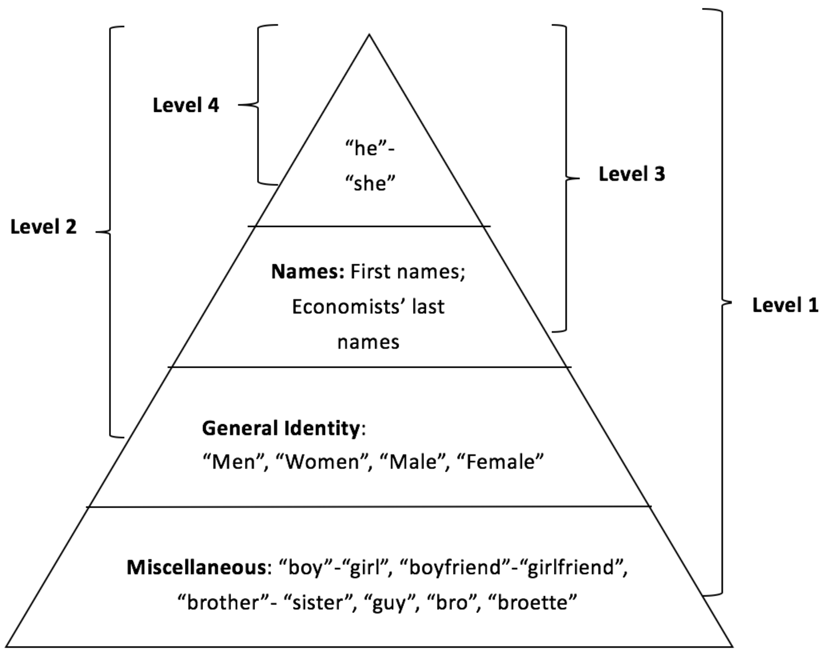
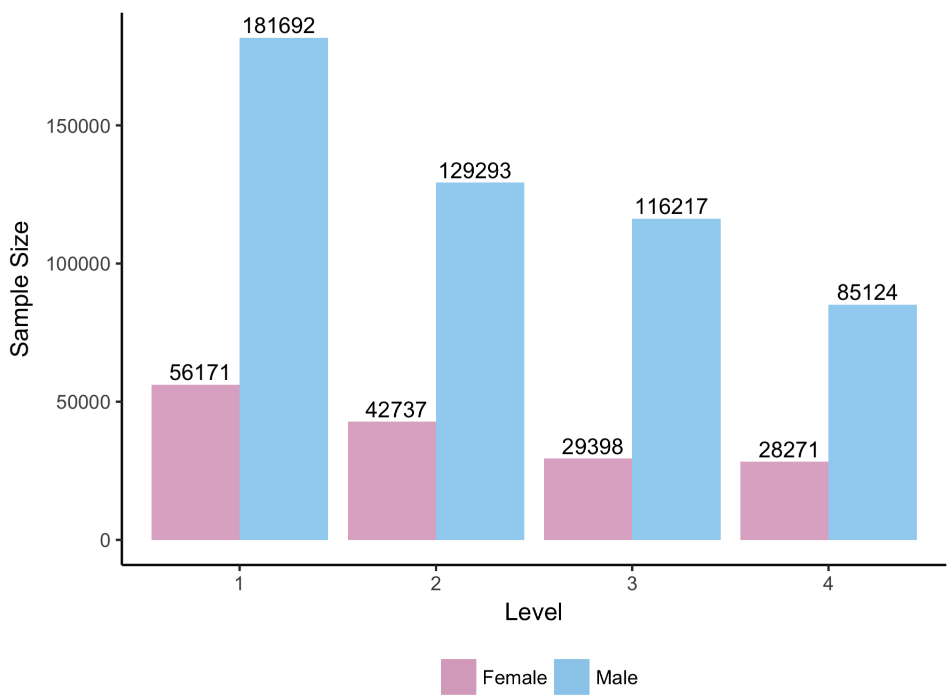


Figure 2: Number of Female and Male Posts by Level



2 Analysis on Gender-related Posts

In this section I restrict my analysis to gender-related posts (Sample 1), and the sample size varies by the level of gender classifiers defined above. First, I present the topic analysis, which shows that $Female = 1$ posts are much less academic/professional oriented and focus more on physical appearance than $Female = 0$ ones, and the results are consistent across different identifications. Second, I design a Lasso-Logistic propensity score model to sort out the words with the strongest predictive power on gender. The contrast between the female and male predictors is in line with the main conclusion from the topic analysis, and further reveals the patterns of gender portrayals in social media.

2.1 Topic Analysis

To capture the content of the text, I manually classify the top $10,000$ words into 15 categories. Table 4 explains how I group certain categories to consider two topics of interest: (i) *Academic/Professional*; (ii) *Personal/Physical*.

Table 4: Categories of Words for Topic Analysis

Category	No. Words	Examples
i. Academic/Professional		
Economics	177	"economics", "macro", "empirical","QJE", "Keynesian"
Academic-General	$1,515$	"research", "papers","tenure", "teaching", "professor"
Professional	138	"career", "interview", "payrolls", "placement", "recruit"
ii. Personal/Physical		
Personal Information	118	"family","married", "kids", "relationship","lifestyle"
Physical Attributes	125	"beautiful", "handsome", "attractive", "body","fat"
Gender related	86	"gender", "femine", "masculine", "sexist", "sexual"

Notes: "Gender related" category under *Personal/Physical* are not used as gender classifiers. The complete list of 15 categories can be found at Appendix I.

For each post, I count the number of occurrences of words from each category, which provides an explicit representation of a post's association with a given topic. For example, a post that includes 8 economics terms is considered more academic than a post with only 3 such terms. I use two benchmark models to estimate the gender differences in topics. The first model looks at the effects of gender on the sum of word frequencies in each topic, while the second uses an indicator $1[Topic_i > 0]$ of whether any word from a given topic occurs:

$$(i): Topic_i = \gamma_0 + \gamma_1 Female_i + e_i$$

$$(ii): 1[Topic_i > 0] = \theta_0 + \theta_1 Female_i + u_i$$

$$Topic \in \{Academic/Professional,\ Personal/Physical\}$$

Table 5 presents the results on the *Academic/Professional* topic. At Level 1 where all gender classifiers are used to identify gender-related posts, it shows on average there are 4.07 academic or job related words in each post associated with male, but 1.76 less (a significant 43.2% decrease) when it is asscoiated with female. In terms of probability, 70.6% of the "male" posts include at least one academic/work term, while 57.4% of "female" posts do.

Table 5: Academic/Professional

	Level 1		Level 2		Level 3		Level 4	
	counts	$1[counts > 0]$	counts	$1[counts > 0]$	counts	$1[counts > 0]$	counts	$1[counts > 0]$
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
$Female_i$	−1.756	−0.132	−1.893	−0.143	−2.084	−0.164	−2.241	−0.165
	(0.035)	(0.002)	(0.043)	(0.003)	(0.053)	(0.003)	(0.055)	(0.003)
Constant	4.067	0.706	4.444	0.742	4.676	0.758	4.918	0.764
	(0.017)	(0.001)	(0.022)	(0.001)	(0.024)	(0.001)	(0.027)	(0.002)
N	233,433	233,433	168,293	168,293	142,131	142,131	111,535	111,535
R^2	0.011	0.014	0.011	0.018	0.011	0.022	0.015	0.026
Adj. R^2	0.011	0.014	0.011	0.018	0.011	0.022	0.015	0.026
F Stat.	2,575.068	3,373.472	1,931.830	3,145.749	1,548.165	3,153.588	1,657.292	2,923.917

Notes: Standard errors are in parentheses. Restrict to posts with ≥ 3 and ≤ 252 words, roughly 98% of each sample. "Level 1" to "Level 4" refer to increasingly restrictive levels of gender classifiers to identify gender-related posts (see Figure 1). The odd numbered regressions estimate model (i), while the even numbered ones estimate model (ii).

One potential issue of using Level 1 gender classifiers is that it picks up a large amount of posts talking about "girlfriend" or "boyfriend" etc. that are necessarily not academic/work oriented. The higher the level of classifiers, the more likely it focuses on posts about people within the Economics community, including professors, colleagues and candidates. For instance, Level 4 only uses "he" or "she" and the like to identify gender-related posts. The sample restriction through gender classifiers is not a perfect filter, but Level 4 does successfully reduce 50% of the sample identified by Level 1, and the comparison across levels provides an opportunity for robustness check. Therefore, I test the models on the gender sample identified by each level, and find that the null hypothesis $E[Academic_i|Female_i = 0] = E[Academic_i|Female_i = 1]$ is rejected consistently across all four levels. The gap in the number of *Academic/Professional* terms is estimated to fall between 42% to 45%, with Level 3 and Level 4 showing the biggest difference. As the sample becomes more restrictive, the average number of *Academic/Professional* terms does increase for both genders, which helps illustrate the validity of such restrictions, i.e. the posts becoming more centered at the Economics community.

Table 6: Personal/Physical

	Level 1		Level 2		Level 3		Level 4	
	counts	$1[counts > 0]$	counts	$1[counts > 0]$	counts	$1[counts > 0]$	counts	$1[counts > 0]$
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
$Female_i$	0.883	0.282	0.846	0.248	0.776	0.234	0.724	0.208
	(0.007)	(0.002)	(0.009)	(0.003)	(0.011)	(0.003)	(0.012)	(0.003)
Constant	0.458	0.256	0.518	0.274	0.505	0.262	0.598	0.301
	(0.004)	(0.001)	(0.005)	(0.001)	(0.005)	(0.001)	(0.006)	(0.002)
N	233,433	233,433	168,293	168,293	142,131	142,131	111,535	111,535
R^2	0.059	0.066	0.047	0.052	0.036	0.041	0.031	0.036
Adj. R^2	0.059	0.066	0.047	0.052	0.036	0.041	0.031	0.036
F Stat.	14,596.460	16,461.220	8,309.106	9,192.315	5,310.392	6,136.646	3,512.172	4,115.958

Notes: Standard errors are in parentheses. Restrict to posts with ≥ 3 and ≤ 252 words, roughly 98% of each sample. "Level 1" to "Level 4" refer to increasingly restrictive levels of gender classifiers to identify gender-related posts (see Figure 1). The odd numbered regressions estimate model (i), while the even numbered ones estimate model (ii).

For the other topic - *Personal/Physical*, I also estimate the benchmark models on posts identified by each level of gender classifiers. As shown in Table 6, at Level 1, a "female" post on average include 1.341 terms related to personal info or physical attributes, almost three times of what occurs in an average "male" post. Even though the overall number of *Personal/Physical* terms seems smaller than *Academic* ones, it is worth noting that this category includes a significant portion of words related to physical appearance or sexual content, which are inappropriate in a forum for economists. In terms of probability, 53.8% of "female" posts at Level 1 includes at least one term associated with this topic, more than double of the portion of "male" posts. The gender difference shrinks as the sample becomes more restrictive, but it is mainly driven by a small increase of the number of such terms in "male" posts, and on average a "female" post consistently has about 1.3 terms under this topic.

A more sophisticated approach is to consider a topic as a distribution over a given vocabulary (Blei et al. 2003; Blei 2012) instead of a single focus. In this setting, *Academic/Professional* terms and *Personal/Physical* ones can occur in a post simultaneously and interact with each other. Therefore, I define the topic difference as

$$\textit{Topic Diff} = \frac{\textit{No. academic words} - \textit{No. personal words}}{\textit{Total No. words}}$$

, which is a simple proxy for the tendency of a post being academic oriented relative to being personal[8]. Table 7 shows similar estimates of the gender difference in *Topic Diff* across increasingly restricitive levels of sample restriction: a $Female_i = 0$ ("male") post on average contains around 6.7%-7.6% more *Academic/Professinal* terms than *Personal/Physical* ones, while it is merely a 1.1%-1.9% difference for a typical $Female_i = 1$ ("female") post. The results are significant and robust across different levels.

Table 7: Topic Difference

	Topic Diff (Academic-Personal)			
	Level 1	Level 2	Level 3	Level 4
$Female_i$	-0.056	-0.054	-0.057	-0.049
	(0.0004)	(0.001)	(0.001)	(0.001)
Constant	0.067	0.072	0.076	0.066
	(0.0002)	(0.0003)	(0.0003)	(0.0003)
N	233,433	168,293	142,131	111,535
R^2	0.064	0.062	0.059	0.058
Adjusted R^2	0.064	0.062	0.059	0.058
F Statistic	15,889.900	11,067.110	8,850.571	6,918.905

Notes: Standard errors are in parentheses. Restrict to posts with ≥ 3 and ≤ 252 words, roughly 98% of each sample. "Level 1" to "Level 4" refer to increasingly restrictive levels of gender classifiers to identify gender-related posts (see Figure 1). The model I estimate is $(\textit{Topic Diff})_i = \beta_0 + \beta_1 Female_i + \epsilon_i$.

[8] *Total No. words* in the definition refers to the total number of words in a post, including but not limited to *Academic/Professional* or *Personal/Physical* ones.

In summary, a post related to female is much less academic-oriented, and more likely to discuss one's physical appearance, which deviates from the main purpose of this academic forum. This finding is robust to different model specifications and identification strategies.

2.2 Word Selection using Lasso-Logistic Model

I design a propensity score model to predict the gender a post is related to by the number of occurrences of the $10,000$ most frequent words, excluding the gender classifiers and the last names of celebrities (not economists)[9]. The goal of this model is twofold: first, to resolve the duplicate posts that include both "female" and "male" classifiers at Level 1; second, to figure out the words with the strongest predictive power for gender. Finally, I also apply this model to the gender-related posts identified by Level 4 classifiers for a robustness check.

Let W_i denotes the frequencies of words, and assume the posterior probability is:

$$\begin{aligned} P(Female_i = 1|W_i) &= \frac{exp(\theta_0 + W_i'\theta)}{1 + exp(\theta_0 + W_i'\theta)} \\ P(Female_i = 0|W_i) &= \frac{1}{1 + exp(\theta_0 + W_i'\theta)} \end{aligned}$$

Write the likelihood of each observation as:

$$P(Female_i|W_i) = P(Female_i = 1|W_i)^{Female_i} \times P(Female_i = 0|W_i)^{(1-Female_i)}$$

Assume the observations are independent, the log likelihood of N observations is

$$\begin{aligned} l_N(\theta) &= log(\Pi_{i=1}^{N} P(Female_i|W_i)) \\ &= \sum_{i=1}^{N} Female_i(\theta_0 + W_i'\theta) - log(1 + exp(\theta_0 + W_i'\theta)) \end{aligned}$$

[9] The names of celebrities (not economists) are not used as gender classifiers.

I estimate θ on words through the following objective function:

$$\hat{\theta}_\lambda = argmin_\theta \left(-l_N(\theta)\right) + \lambda||\theta||_1$$
$$\text{where } \|\theta\|_1 = \sum_{j\geq 1} |\theta^j|.$$

Lasso regularization, i.e. the ℓ_1-norm penalty, promotes sparsity in the estimator $\hat{\theta}_\lambda$. It shrinks coefficients on variables with little explanatory power to zero, and thus is particularly useful for variable selection in high dimensional data. Lasso has become a popular approach in computational linguistics (e.g. Eisenstein et al. 2011). Gentzkow et al. (2016) also use this strategy to identify the most partisan phrases in Congressional speech. In this case, each W_i is a $9,674$-by-1 vector of word counts[10]. The logistic-Lasso model sorts out words with the strongest predictive power on gender. The estimator $\hat{\theta}_\lambda$ is biased, but the variance of the model is reduced, and tends to yield more accurative predictions.

There are $211,861$ posts that include only "female" words or only "male" words at Level 1. I use 75% of them, i.e. $158,843$ posts, to train the model and select an optimal tuning parameter λ through 5-fold cross validation. The remaining $53,018$ posts are assigned to the test set to select the best cutoff on p-scores, which turns out to be 0.40. Finally, I apply the model to the $26,002$ duplicates, and classify $9,044$ of them to $Female = 1$ and the rest to $Female = 0$. As for the variable selection, the coefficients of $6,088$ words are shrinked to zero; that is, they are considered irrelevant to the gender identification of a post. The average marginal effect of word k is estimated by:

$$\begin{aligned}\textit{Word k's marginal effect} &= P(Female_i = 1|W_{i,(-k)}, W_{ik} + 1) - P(Female_i = 1|W_{i,(-k)}, W_{ik}) \\ &= \frac{1}{N}\sum_i P(Female_i = 1|Wi) \times (1 - P(Female_i = 1|Wi))\, \widehat{\theta^k_\lambda}\end{aligned}$$

where W_{ik} is the frequency of word k in post i, and $W_{i,(-k)}$ is the vector of frequencies of words other than k in post i.

The left half of Table 8 displays the top 30 words with the strongest predictive power for gender at Level 1. A significant proportion of the words considered most "female" are under the

[10] I exclude the gender classifiers and the last names of celebrities (not economists) from the $10,000$ words.

second topic - *Personal/Physical*, whereas there are more academic/job related ones on the list of words suggesting “male”. For example, the word “attractive” increases the probability a post is discussing female by 26.0%, while each additional occurrence of “motivated ” decreases the probability by 17.1%.

To check the robustness of the words selected by Lasso, I apply the p-score model to posts identified by Level 4 gender classifiers, and the results are shown in the right half of Table 8. Level 4 uses the most restrictive set of classifiers - “he”, “she” etc. (see Figure 1). The most “male” words at Level 4 do turn out to be quite different from those at Level 1, with a 83.3% turnover rate among the top 30. Academic words like “mathematician” and “wharton” emerge, with 13.7% and 16.8% marginal effects on a post being related to male respectively. On one hand, the restriction on gender classifiers does help identify “male” posts that are more academic or professional oriented. On the other hand, the comparision between the top “female” words shows that the discussions related to women consistently tend to deviate from academic topics, no matter how restrictive the sample selection is. This finding is in line with the results from the topic analysis that $Female = 1$ posts include significantly less *Academic/Professional* terms, and significantly more *Personal/Physical* terms. Moreover, the robustness of the top “female” words reveals the intensity of gender stereotyping involved in the discussions on EJMR.

Table 8: Top 30 Words with the strongest predictive power for $Female_i = 1$

Level 1				Level 4			
Most "female"		Most "male"		Most "female"		Most "male"	
Word	Marginal Effect	Word	Marginal Effect	Word	Marginal Effect	Word	Marginal Effect
hotter	0.388	homosexual	-0.237	hotter	0.336	juicy	-0.218
hot	0.285	homo	-0.228	lesbian	0.301	keys	-0.196
attractive	0.260	philosopher	-0.204	bb	0.276	adviser	-0.181
pregnant	0.252	keen	-0.182	sexism	0.245	bully	-0.174
gorgeous	0.251	motivated	-0.171	tits	0.233	prepare	-0.173
beautiful	0.249	fieckers	-0.164	anal	0.223	fought	-0.170
tits	0.247	slides	-0.160	marrying	0.215	wharton	-0.168
lesbian	0.242	nordic	-0.156	feminazi	0.213	austrian	-0.164
bang	0.229	filling	-0.152	slut	0.194	fieckers	-0.163
horny	0.224	textbook	-0.148	hot	0.193	homo	-0.153
slept	0.224	adviser	-0.140	vagina	0.187	genes	-0.151
marry	0.221	fenance	-0.138	boobs	0.176	e7ee	-0.142
attracted	0.216	fiekers	-0.137	pregnant	0.176	mathematician	-0.137
0,0	0.213	bowl	-0.136	pregnancy	0.174	advisor	-0.134
cute	0.209	gay	-0.132	cute	0.167	burning	-0.132
breasts	0.205	bench	-0.129	marry	0.167	pricing	-0.131
sexy	0.202	mountain	-0.128	levy	0.159	philly	-0.126
pregnancy	0.195	humble	-0.124	gorgeous	0.157	band	-0.124
dumped	0.194	iraq	-0.122	horny	0.153	kfc	-0.124
feminazi	0.189	rust	-0.121	crush	0.153	nobel	-0.123
feminist	0.187	amusing	-0.120	beautiful	0.151	cmt	-0.121
raped	0.185	speeches	-0.119	secretary	0.150	amusing	-0.120
dated	0.184	affected	-0.119	dump	0.148	greatest	-0.119
cheerful	0.182	mere	-0.119	shopping	0.148	textbook	-0.118
ugly	0.179	chill	-0.118	date	0.144	goals	-0.117
marrying	0.178	bugs	-0.118	nonprofit	0.141	irate	-0.116
blonde	0.175	rip	-0.113	intentions	0.140	roof	-0.116
crush	0.173	recession	-0.111	sexy	0.140	pointing	-0.116
date	0.172	brilliant	-0.108	dated	0.138	episode	-0.115
naked	0.172	salmon	-0.108	prostitute	0.138	tries	-0.114

Notes: the marginal effect of word w is the change in probability of a post being classified as *female*, i.e. 1 if it is discussing women, when it contains one more word w.

3 Analysis on Gender-related Threads

To capture a more complete picture of the gender-related discussons, I extend the analysis to threads that contain at least one $Female = 1$ or $Female = 0$ post. Using Level 1 gender classifiers, I construct Sample 2, a panel data that contains $810,998$ individual posts under $74,679$ threads. From a user's perspective, he or she first reads the title of a thread, and then decides whether to continue reading the posts under it and contribute to the discussion. Based on this decision making process, I first look at how gender can affect the popularity of a thread. In addition, I replicate the topic analysis to the aggregate data at the thread level, and finally test for the state dependence between adjacent posts within a discussion. Given a thread, I define the following variables:

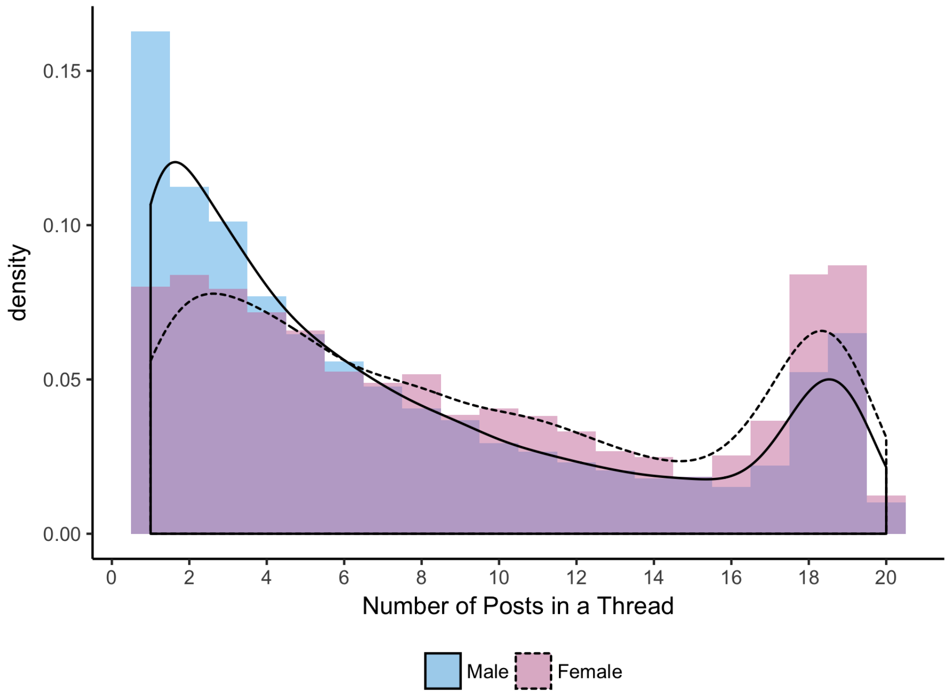
$$\begin{aligned}
&nPosts = \textit{Number of posts within a thread}\\
&nFemale = \textit{Number of (Female=1) Posts}\\
&nMale = \textit{Number of (Female=0) posts}\\
&nGender = nFemale + nMale\\
&(\%Female - \%Male) = \frac{nFemale - nMale}{nPosts}
\end{aligned}$$

3.1 Popularity of Threads in relation to Gender

The popularity of a thread can be measured by $nPosts$, the numer of posts within the discussion. The title of each thread is particularly important in determining whether the thread will attract a lot of attention. I define $Female_{t,0} = 1$ for thread t if its title includes a female classifier at Level 1, $Female_{t,0} = 0$ if including a male classifier, and NA if neither case above is true. About 20.9% of the titles in Sample 2 are directly related to gender, i.e. including at least one gender classifier. Figure 3 breaks down the distribution of $nPosts$ for threads of titles related to female vs. male.

For threads with "male" titles, i.e. $Female_{t,0} = 0$, the mass of the distribution is more highly concentrated on the left than those with "female" titles. It suggests that titles related to female are more likely to attract readers' attention and induce them to partcipate in the discussions. About 60% of the threads with gender-related titles have fewer than 8 posts. However, there are spikes in the distribution of $nPosts$ on the right for both types of threads. In particular, about 18% of

Figure 3: Popularity by Gender in Titles



threads under "female" titles have ≥ 18 posts, and about 12% for the male counterparts. The main reason for the spikes is that I scrape up to one page for each thread, resulting in the maximum of $nPosts$ at 20. The higher concentration at $nPosts \geq 18$ for threads under "female" titles implies they tend to be more popular and induce an intensive discussion (over 1 page).

To address the censoring from above, I define an indicator for a thread t being likely to go over one page: $D_t = 1$ if $nPosts_t \geq 18$ and 0 otherwise[11]. In this setting, I have the latent variable $nPosts_t^* = nPosts_t$ if $D_t = 0$.

Table 9 reports the regression outputs of $nPosts$ on gender in a title (Female, Male or Neutral), using both OLS and the censored model. Threads with gender-related titles are on average shorter than those with neutral titles that do no include any gender classifier. However, the comparison between threads with "female" and "male" titles show that a thread with a "female" title on average has 1.85 more posts than its "male" counterpart according to OLS, and 2.07 more

[11] The censored model above does not have a single threshold at 20, because some of the posts with special characters or completely repeating posts above are eliminated in the data cleaning process, contributing to the spike at 18 and 19. In addition, since I did not add in new posts after July 2016, threads with over 18 posts may be updated and go over one page. The variable D_t is considered as a indicator for whether the true $nPosts_t^*$ is unobserved due to censoring.

accoridng to the censorsed regresison.

Table 9: Popularity of Threads on Gender

	Number of Posts ($nPosts$)	
	(1) OLS	(2) Censored
Title:		
$Female_{t,0} = 1$	−1.242	−1.414
	(0.097)	(0.117)
$Female_{t,0} = 0$	−3.093	−3.481
	(0.062)	(0.074)
Constant	10.401	11.112
	(0.025)	(0.030)
No. Threads	74,679	74,679
No. Threads Censored	0	$15,591$
R^2	0.033	
Adjusted R^2	0.033	
Log Likelihood		−217,095.600
Akaike Inf. Crit.		434,199.200

Notes: Standard Errors in parentheses. In regression (2), a thread t is censored if $D_t = 1$, i.e. given the observed no. posts it contains, it is projected to exceed one page.

3.2 Topic Analysis on Thread Environment

I replicate the topic analysis on this panel data. First, I aggregate the data to the thread level, where I look at $\overline{Academic}_t$ and $\overline{Personal}_t$, the mean occurrences of *Academic/Professional* terms and *Personal/Physical* terms across all posts within a thread. Second, I test for the state dependence between adjacent posts in a thread, in particular whether a gender-related post will affect the topic of the following one.

I. Topics at the thread level

For each thread, I use $\%Female - \%Male = \frac{nFemale - nMale}{nPosts}$ as an aggregate measure of the representation of "female" posts relative to "male" ones, which are indentified through Level 1 gender classifiers. I divide this measure into quartiles, where the first quarter $[-1, -0.364)$ corresponds to threads that most heavily center on men while the last quarter $[0, 1]$ refers to threads that include more posts related to female than to male.

Table 10 shows the OLS outputs for both the unweighted and the weighted versions. Threads that are mostly centered on men (Quartile 1) on average has 3.52 *Academic/Professional* terms per post. The more "female" posts a thread contains, the lower the mean number of *Academic/Professional* terms, and the gap is as wide as 46.2% - 48.0% for Quartile 4, where the number of "female" posts exceed that of "male" posts. This result is consistent with the topic analysis on gender-related posts (Table 5), which shows a "female" post on average contains 43.2% (1.76 words) less than a "male" post.

Table 10: Mean Frequencies of Words by Topic

	$\overline{Academic_t}$		$\overline{Personal_t}$	
	(1)	(2)	(3)	(4)
Posts: $(\%Female - \%Male)$				
Quartile 1: $[-1, -0.364)$ (base)				
Quartile 2: $[-0.364, -0.176)$	−0.662	−0.314	−0.108	−0.013
	(0.045)	(0.035)	(0.009)	(0.007)
Quartile 3: $[-0.176, 0)$	−0.539	−0.450	−0.124	0.035
	(0.045)	(0.038)	(0.009)	(0.008)
Quartile 4: $[0, 1]$	−1.688	−1.566	0.263	0.513
	(0.045)	(0.034)	(0.009)	(0.007)
Constant	3.518	3.391	0.393	0.347
	(0.032)	(0.022)	(0.006)	(0.005)
Weighted		X		X
N	74,679	74,679	74,679	74,679
R^2	0.019	0.030	0.030	0.086
Adjusted R^2	0.019	0.030	0.030	0.086
F Statistic	493.961	748.272	782.810	2,229.114

Notes: Standard errors are in parentheses. Each title can be classfied as $Female_{t,0} = 1$, $Female_{t,0} = 1$ or not related to gender. Columns (2) and (4) use #gender-related posts in each thread as the weight.

As for the *Personal/Physical* topic, the unweighted OLS model shows that threads in Quartile 4 contain about 66.9% significangtly more terms about personal information and physical attributes. By contrast, the increase becomes even more drastic to 148.0% when I use *nGender*, the number of gender-related posts as weights. The weights seem to have larger influence on results for this topic, which is likely because the words under *Personal/Physical* are more directly associated with gender discussions than *Academic/Professional* ones. Also, note the decreases in $\overline{Personal_t}$ for threads in Quartile 2 and 3 relative to 1 are in line with the observation that when $\%Female - \%Male$ is close to 0, it is either because a thread is very balanced in posts related to "female" or "male" or because the discussion overall is not really related to gender. Therefore, it is important to put more weight on threads that contain more gender-related posts.

II. State Dependence in Topics

I regard a thread as *state-dependent* if the revealed topic of a post depends on that of the prior posts. Within each thread, a post can either be an immediate reaction to the title exclusively, or a response to some of the previous posts. Here I explore the relationship between each post (p) and its prior one ($p-1$) through the effects on *Academic/Professional* and *Personal/Physical* terms.

To control for the unobserved heterogeneity between threads, I consider the posts within a thread to be equally informative, and then assume the linear combination of $Topic_{t,0}$ of the title and the mean $\overline{Topic_t}$ across posts within thread t is able to absorb the omitted information. I use the following control function:

$$
\begin{aligned}
&Topic_{t,p} = \omega_0 + \omega_1 \overline{Topic_t} + \omega_2\, Topic_{t,0} + \omega_3\, Topic_{t,p-1} + \nu_{t,p} \\
&Topic \in \{Academic/Professional,\ Personal/Physical\} \\
&t - thread;\ p - post
\end{aligned}
$$

In addition, I want to explore whether there is a stronger or weaker state dependence when the prior post is related to gender,i.e. $Female_{t,p-1} = 0\, or\, 1$. Therefore, I add an interaction term between the lagged gender variable and the lagged characteristic of interest:

$$
\begin{aligned}
Topic_{t,p} = \omega_0 &+ \omega_1 \overline{Topic_t} + \omega_2\, Topic_{t,0} \\
&+ \omega_3\, Topic_{t,p-1} + \omega_4 Female_{t,p-1} + \omega_5\, Topic_{t,p-1} \times Female_{t,p-1} + \nu_{t,p}
\end{aligned}
$$

In Table 11, all models display a mean reversion pattern on average, as the coefficients on the lagged variable $Y_{t,p-1}$ are significantly negative. For the *Academic/Professional* topic, the reversion effect is 29% stronger for each academic word the prior post contains when it is "female", i.e.$Female_{t,p-1} = 1$. That is, even though each post tends to use less *Academic/Professional* terms relative to its prior one, it is significantly more likely to deviate from an academic-oriented discussion if women are mentioned previously. Finally, column (4) reveals that the impact of following a "female" post on the use of Personal/Physical words is significantly positive and it counteracts about 76.8% of the mean reversion effect when the previous post includes one $Personal/Physical$ term.

Table 11: State Dependence in Topics

	Academic/Professional		Personal/Physical	
	(1)	(2)	(3)	(4)
$\overline{Topic}_t$	0.986	0.985	1.008	1.003
	(0.004)	(0.004)	(0.004)	(0.004)
$Topic_{t,0}$ (titles)	0.043	0.042	0.052	0.045
	(0.010)	(0.010)	(0.006)	(0.006)
State Dependence				
$Topic_{t,p-1}$	−0.071	−0.069	−0.080	−0.073
	(0.002)	(0.002)	(0.002)	(0.002)
$Female_{t,p-1}$		−0.002		0.059
		(0.028)		(0.007)
$Female_{t,p-1} \times Topic_{t,p-1}$		−0.020		−0.014
		(0.004)		(0.003)
Constant	0.106	0.114	0.027	0.016
	(0.016)	(0.018)	(0.003)	(0.003)
N	210,406	210,406	210,406	210,406
R^2	0.244	0.244	0.243	0.244
Adjusted R^2	0.244	0.244	0.243	0.244

Notes: Standard errors are in parentheses. Restrict to posts where the prior one of each is gender-related, i.e. $Female_{t,p-1} \in \{0, 1\}$. Regressors $\overline{Topic}_t$ and $Topic_{t,p-1}$ refer to the same topic as the dependent variable in each model.

In summary, the topic analysis at the aggregate thread level reaches the same conclusion as the analysis at the post level that a discussion on EJMR become significantly less academic or professional oriented, and more about personal information or physical appearances when it is related to women rather than men. And the test of state dependence shows that a post is also more likely to deviate from an academic focus when the prior post is "female".

4 Alternative Design: Analysis at the Economist level

While the previous sections study the patterns in all gender-related discussions, this final part of the paper examines whether gender plays a role in determining how much attention an economist receives on EJMR. In this alternative design, I select two cohorts of economists: (1) 380 high-profile economists who ranked among Top 5% Authors on RePEc; (2) 204 assistant professors in Top 20 U.S. programs in Economics by U.S. News Ranking. Using a difference-in-difference approach, I find that high-profile female economists tend to receive more attention than their male counterparts, and the gap is widening for relatively lower-ranked economists. The junior cohort shows different patterns when I group economists by the ranking of their current institutions.

4.1 Selection of Economists and Sample Construction

Economists most likely being discussed on EJMR are either prominent senior faculty, or tenure-track junior economists who had been through the job market recently. Based on this observation, I build up my samples of both senior and junior economists.

For the senior cohort, I generate a balanced set of female and male economists, who are comparable according to the RePEc ranking of the Top 5% Authors [12] I find 190 female economists among the top $2,422$ authors. For each of them, a coin is tossed to decide whether the male economist who ranks 1 above or below will be included in the control group. I use each economist's rank as a proxy for his or her prominence in the field of economics. Hence I have a sample of 190 female and 190 male high-profile economists. For the junior cohort, I select all assistant professors in Top 20 U.S. programs in Economics, according to U.S. News Graduate School ranking [13] I find 45 female and 159 male junior faculty among these schools.

Given the 584 economists in total, I search by each person's full name within EJMR forum and then preserve as many threads in which he or she is mentioned as possible. Then I keep all the posts on a given page of a thread [14]. As a result, I construct a data set of $3,299$ unique

[12] RePEc ranking of Top 5% Authors (Last 10 Years Publications), as of September 2016: https://ideas.repec.org/top/top.person.all10.html. To identify the gender of each economist, I match the overall ranking with a separate RePEc ranking on female economists: https://ideas.repec.org/top/top.women.html.

[13] based on U.S. News ranking of best graduate programs in Economics as of 2013 and 2017, and RePEc ranking of top Economics Departments.

[14] In each query, I maximize the number of results Google display, but if there are over 20 results, the amount of URLs I can successfully scrape is shrinked by 25% on average.

threads. There is no restriction on the years of the discussions in this data set. Among 380 senior economists, there are 278 economists (145 women, 133 men) mentioned at least once in EJMR. Among 204 junior faculty, 187 economists (38 women, 149 men) were mentioned at least once. Seniority increases the attention one receives significantly. On average, a high-profile economist is discussed in 20.5 threads, whereas an assistant professor occurs in 14.8 threads.

4.2 Difference-in-Difference Analysis of Gender on Attention

Given the number of search results- N_i on each economist[15], I define A_i, a metric that represents the amount of attention person i receives as

$$A_i = asinh(N_i) = log(N_i + \sqrt{1 + N_i^2})$$

I estimate the following difference-in-difference specification:

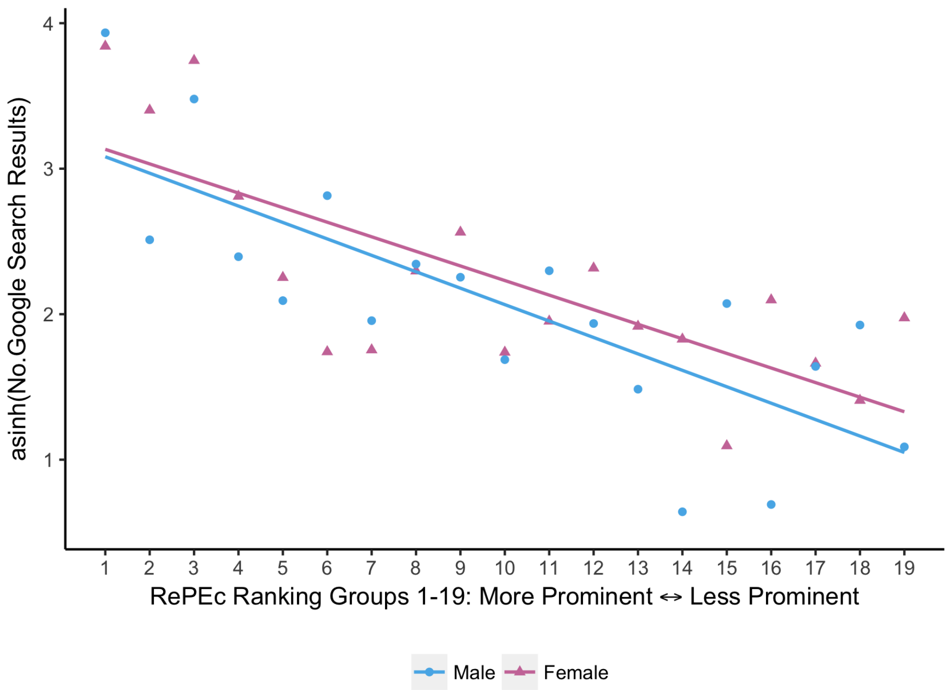
$$A_i = \gamma_0 + \gamma_1 Female_i + \Gamma' Group_i + \Lambda'(Female_i \times Group_i) + \varepsilon_i$$

where Λ are the coefficients of interest. For the high-profile cohort, each "Group" contains 10 female economists and 10 male economists, based on their RePEc ranking. Figure 4 shows that the higher ranked an economist is, the more attention one receives on EJMR. Female economists tend to receive more attention than their male counterparts, and this gap, though insignificant, is widening as the economist ranking goes down. This finding is in line with the hypothesis that women as the minority group are more "visible" (Kanter 1997).

For the junior cohort, since I do not have a measure of prominence at invidual level, I split them into 6 groups by the ranking of their current departments. Figure 5 reveals that junior faculty in higher ranked institutions receive significantly more attention. Female assistant professors receive more attention than their male counterparts in the first two groups (top 5 economics departments), but this trend is reversed for people in relatively lower ranked departments. In other words, for women the amount of attention one gets is more sensitive to the prestigiousness of the institutions. However, note the junior cohort is imbalanced in gender: 45 women and 159 men. The gender

[15] The number of threads in the final dataset is considered as an alternative measure, and it gives consistent results.

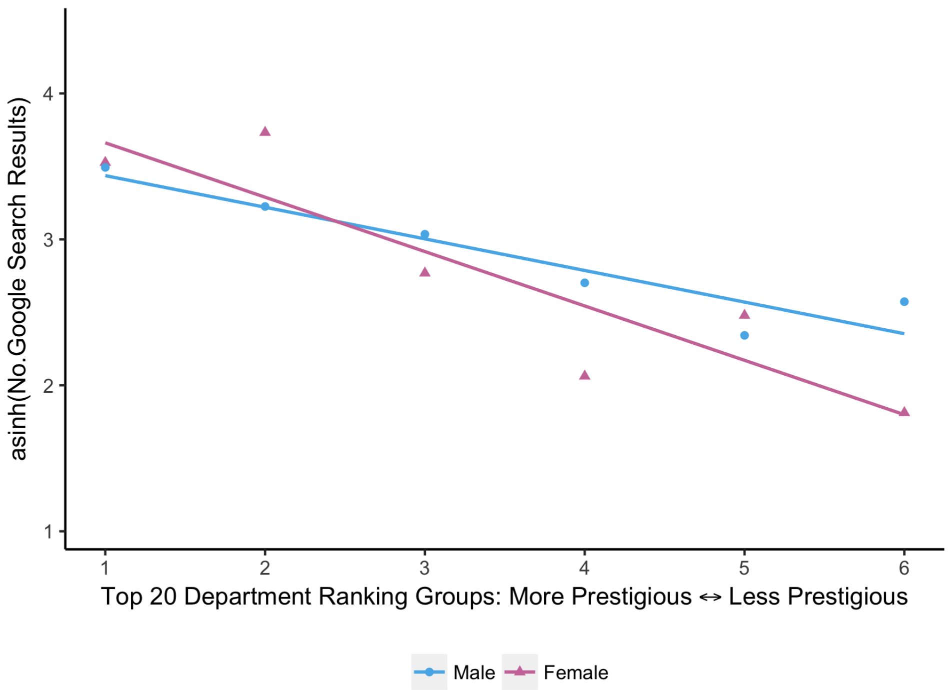
Figure 4: 380 High-profile Economists (190 female,190 male)



differences can be exaggerated if there are outliers among men who receive much more attention than their peers. For a more careful analysis on the junior cohort, it would be helpful to use the publication information of each economist as a measure of individual achievement in lieu of the institutional ranking.

For both the high-profile and the junior cohorts, the selection is limited as I focus on the best people in the field in terms of their academic and professional achievements. A more informative analysis would require expanding the sample of economists to be more representative of the overall academic community. It is also worth mentioning that there is no clear relationship between the prestigiousness of a department one works at and his or her own prominence. In particular, junior faculty within the same department ranking group are not as comparable as the high-profile economists within the same RePEc ranking group based on individual performance. Therefore, the results for high-profile and junior cohorts should be viewed separately.

Figure 5: 204 Assistant Professors (45 female, 159 male)



5 Conclusions

Gender stereotyping can take a subtle or implicit form that makes it difficult to measure and analyze in economics. In addition, people tend not to reveal their true beliefs about gender if they care about political and social correctness in public. The anonymity on the Economics Job Market Rumors forum, however, removes such barriers, and thus provides a natural setting to study the existence and extent of gender stereotyping in this academic community online.

I approach the question of gender stereotyping through a combination of natural language processing, machine learning and econometric methods. With more than one million posts on the Economics Job Market Rumors forum over two years, I identify gender-related posts by four increasingly restrictive levels of gender classifiers. The first research question I focus on is whether there are significant differences in topics when a discussion is related to women rather than men. I count the number of *Academic/Professional* words and *Personal/Physicial* words in each post. A $Female = 1$ post on average contains 43% significantly less academic or professional terms, and 192% more terms about personal information or physical attributes[16], and the conclusions are

[16]the results here come from the sample using Level 1 gender classifier.

consistent across all levels of gender classifers I specify. Topic analysis at the thread level also shows similar patterns when I compare threads focusing on women with those focusing on men. Besides, the tests for state dependence between adjacent posts show that a post tends to deviate more drastically from being academically or professionally oriented if its prior post is related to women. Meanwhile, such a reversion pattern almost vanishes when it comes to the use of *Personal/Physical* words.

Word selection by the Lasso-Logistic model on gender-related posts yields a more straightforward demonstration of the pervasiveness of gender stereotyping at finer granularity. Words with the strongest association with *female* are mostly inappropriate, and the occurrence of these words in a forum that was meant to be academic and professional exposes the issues of explicit biases in social media.

Finally, results from the difference-in-difference analysis of gender on the attention received by economists supports the hypothesis that women as a minority in this community receive disproportionate attention (Kanter 1997). In the high-profile cohort where I have 190 pairs of female and male economist of similar RePEc ranking, I find that female economists on average receive more attention than their male counterparts, and the gap widens for relatively less prominent groups. In the junior cohort of 45 female and 159 male assistant professors, female faculty working at Top 5 U.S. economics departments get more attention on EJMR than their male counterparts, but the trend is reversed for those at lower-ranked institutions.

In conclusion, my results suggest the need for changes to maintain an inclusive online environment for everyone in the academic community. The casual setting of this online forum cannot be an excuse for gender stereotyping conversations, and the freedom to express one's opinions anonymously should not be abused to create a sense of isolation, which can be discouraging and harmful to the academic and professional development of all genders.

References

[1] Bordalo, Pedro, Katherine Coffman, Nicola Gennaioli, and Andrei Shleifer. 2016. "Stereotypes." Quarterly Journal of Economics 131 (4): 1753-1794.

[2] Bordalo, Pedro, Katie Coffman, Nicola Gennaioli, and Andrei Shleifer. 2016. Working Paper. "Beliefs about Gender".

[3] Ceci, S.J., Ginther, D.K., Kahn, S. and Williams, W.M., 2014. "Women in academic science A changing landscape". Psychological Science in the Public Interest, 15(3), pp.75-141.

[4] Eisenstein J, Smith NA, Xing EP. 2011. "Discovering sociolinguistic associations with structured sparsity". Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies-Volume 1. Association for Computational Linguistics, pp. 1365–1374.

[5] Ginther, D.K. and Kahn, S., 2006. "Does science promote women? Evidence from academia 1973-2001" (No. w12691). National Bureau of Economic Research.

[6] Gentzkow, M., Shapiro, J.M. and Taddy, M., 2016. "Measuring polarization in high-dimensional data: Method and application to congressional speech" (No. w22423). National Bureau of Economic Research.

[7] Gentzkow, Matthew and Kelly, Bryan T. and Taddy, Matt. "Text As Data". February 15, 2017.

[8] Hutto, C.J. and Gilbert, E.E. 2014. "VADER: A Parsimonious Rule-based Model for Sentiment Analysis of Social Media Text". Eighth International Conference on Weblogs and Social Media (ICWSM-14). Ann Arbor, MI, June 2014.

[9] Kahn, S. and Ginther, D.K. 2017. "Women and STEM." Prepared for The Oxford Handbook on the Economics of Women. Averett, S., Argys, A. and Hoffman, S.D. eds.

[10] Kanter, Rosabeth Moss. 1997. "Some Effects of Proportions on Group Life: Skewed Sex Ratios and Responses to Token Women." American Journal of Sociology, vol. 82, no. 5, 1977, pp. 965–990. www.jstor.org/stable/2777808.

[11] Pennebaker, J.W., Booth, R.J., Boyd, R.L., and Francis, M.E. 2015. "Linguistic Inquiry and Word Count: LIWC2015". Austin, TX: Pennebaker Conglomerates (www.LIWC.net).

[12] Schwartz HA, Eichstaedt JC, Kern ML, Dziurzynski L, Ramones SM, Agrawal M, et al., 2013. "Personality, Gender, and Age in the Language of Social Media: The Open-Vocabulary Approach". PLoS ONE 8(9): e73791. doi:10.1371/journal.pone.0073791.

[13] Seetharaman, Deepa. 2017. "Facebook's Female Engineers Claim Gender Bias". Wall Street Journal. 2 May.

[14] Wootson Jr., Cleve R.. 2017. "A Google engineer wrote that women may be unsuited for tech jobs. Women wrote back". Washington Post. 6 August.

Appendix I. Categories of Words (Topic Analysis)

Table 12: Categories of Words

Category	No. Words	Examples
Gender Classifiers (All - Level 1)		
Female	44	"she", "female"
Male	134	"he", "male"
Academic/Professional		
Economics	177	"economics", "macro", "empirical","QJE", "Keynesian"
Academic-General	1,515	"research", "papers","tenure", "teaching", "professor"
Professional	138	"career", "interview", "payrolls", "placement", "recruit"
Personal/Physical		
Personal Information	118	"family","married", "kids", "relationship","lifestyle"
Physical Attributes	125	"beautiful", "handsome", "attractive", "body","fat"
Gender related	86	"gender", "femine", "masculine", "sexist", "sexual"
Swear Words		
Swear	78	"shit", "wtf", "asshole"
Intellectual		
Intellectual-Positive	115	"intelligent", "creative","competent"
Intellectual-Neutral	29	"brain", "iq", "ability"
Intellectual-Negative	134	"dumb","ignorant","incompetent"
Miscellaneous		
Emotion/Feelings	74	"happy","depressing"
Emojis	11	":)", ";)", ":p"
Others	7,222	"years", "places", "everything"
Total	10,000	

Notes: "Gender related" category under *Personal/Physical* are not used as gender classifiers.