Studies in Communication Sciences 15 (2015) 5–11

Contents lists available at ScienceDirect

Studies in Communication Sciences

journal homepage: www.elsevier.com/locate/scoms

Old glass ceilings are hard to break: Gender usage trends in annual reports☆

Tim Loughran*, Bill McDonald[1]

Mendoza College of Business, University of Notre Dame, Notre Dame, IN 46556-5646, United States

ARTICLE INFO

Article history:
Received 17 November 2014
Accepted 11 March 2015

Keywords:
EDGAR
Gender
Firm age
Annual reports
Textual analysis

ABSTRACT

We examine gender usage in a sample of 89,195 annual reports filed with the SEC during 1996–2013. We find that, after adjusting for other effects, annual reports by younger firms use proportionally more female-linked words than documents created by older, more mature companies. This finding likely reflects gender-related cultural differences between young and old firms. We also report that gender usage differs dramatically across both industry and market values of equity. Historically male dominated industries and industries that do not sell directly to retail customers have lower ratios of female/male word usage while industries characterized as business-to-consumer have substantially higher relative female counts. Larger companies have higher public accountability and thus, as expected, have annual report language that more frequently uses female titles and personal pronouns.

1. Introduction

In this paper, we examine gender usage trends in annual reports filed with the Securities and Exchange Commission (SEC) by publicly-traded firms. The annual report (also known as Form 10-K) is the central periodic filing through which managers communicate with shareholders. Specifically, we tabulate the number of times the titles *Mr.*, *Ms.*, *Miss*, and *Mrs.*, and personal pronouns *his*, *her*, *he*, and *she* appear in a sample of 89,195 annual reports of publicly-traded companies during the 1996–2013 time period. Starting with the pioneering papers by Tetlock (2007) and Tetlock, Saar-Tsechansky and Macskassy (2008), researchers have tabulated word counts in newspaper articles, earnings conference calls, and annual reports to gauge sentiment or to look for patterns in word usage. Our paper extends the literature's understanding of the changing pattern of gender specific terms appearing in business communications.

Much of the prior literature on gender in the workplace focuses on the wage gap and factors that might explain such differences, beyond simple discrimination.[2] In this empirical paper, we provide a different perspective on the issue of gender at work by examining how gender-related issues might be implicitly revealed in management's communication with shareholders.

We first empirically document trends in usage of gender-related titles and pronouns across annual report filings. As women play an increasingly important role in companies, business documents should show more frequent usage of female titles and pronouns. Most of the documented gender word usage appearing in annual reports relates to descriptions of top managers and directors. Extant literature focused on wages, documents a stagnant period for the gender gap until 1970, substantial convergence through the 1990s, a plateau from 1995 to 2000, and an uptick in the first few years of this decade (e.g., see Mulligan & Rubinstein, 2008 or Borghans, ter Weel, & Weinberg, 2006).

Anecdotally, and consistent with this literature, women are increasingly becoming Chief Executive Officers (CEOs) of the largest and most powerful U.S. publicly-traded companies. For example, currently the CEOs of General Motors (Mary Barra), Hewlett-Packard (Meg Whitman), IBM (Virginia Rometty), Yahoo! (Marissa Mayer), and Pepsi (Indra Nooyi) are all females. However, as noted

☆ We thank Robert Battalio and an anonymous referee for helpful comments on the paper. This paper was produced in conjunction with the "Discourse approaches to financial communication" conference held in Ascona, Switzerland in February 2014.

* Corresponding author. Tel.: +1 574 631 8432.
E-mail addresses: Loughran.9@nd.edu (T. Loughran), mcdonald.1@nd.edu (B. McDonald).
[1] Tel.: +1 574 631 5137.

[2] See, for example, Goldin (2014). We will detail the relevant literature in a subsequent section.

http://dx.doi.org/10.1016/j.scoms.2015.03.004
1424-4896/

by Fortune magazine (June 3, 2014), only 4.8% (24 out of 500) of the Fortune 500 firms are headed by women. So although the number of women CEOs is higher than ever before, the relative percentage of top female managers is significantly lower than the frequency of women in the workforce. We examine how artifacts of these changes are reflected in the communications of managers.

As might be expected, our results document a dramatic increase in relative frequency of female gender terms (*Ms.*, *Mrs.*, *Miss*, *her*, and *she*) in annual reports since 1996.[3] This directly reflects the rise of women managers/directors in publicly-traded firms. However, like the fraction of female CEOs, female terms like *Ms.* and *she* are still significantly less likely to occur in annual reports than words like *Mr.* and *he*. For our sample, we find that in 1996, the average annual report contained only 0.152 counts of the female-related titles and pronouns compared to 0.335 for the average firm in 2013, i.e., there has been a 120% increase in female gender words in annual reports since 1996. Yet, the count of male gender words (e.g., *Mr.* and *he*) is 1.734 in 1996 compared with 2.390 in 2013. Thus, although we expect both counts to increase as an artifact of the increasing size of annual reports documented in Loughran and McDonald (2014), the relative changes are notably different. As with the gender of top managers, publicly-traded companies listed on U.S. stock exchanges disproportionally use male personal pronouns and titles in their business communication documents.

We also consider the ratio of female/male titles and pronouns across industries and relative to the market capitalization of firms. We find pronounced industry differences that closely align with a firm's proximity to the ultimate individual consumer. Industries that directly sell products and services to consumers, like Publishing, Banking, Personal Services, Apparel, Healthcare, Pharmaceutical, and Retail have significantly higher usage of female gender terms than industries that do not directly interact with retail customers. Industries with low female-to-male ratios include Agriculture, Electrical Equipment, Textiles, Aircraft, Fabricated Products, Coal, and Oil. Since women historically are not the primary purchasers of Boeing aircraft or products from Arch Coal, it is expected that these industries would have relatively fewer female gender terms. Lower historical top management employment of women by aircraft production and rural coal companies is reflected in the low value female/male ratio for these particular industries.

Our paper also finds that firms with larger market values of equity have significantly higher values of the female/male ratio. Larger firms obviously have higher public profiles. That is, more attention is placed by investors and the media on large, well-known firms like Bank of America, Microsoft, Johnson & Johnson, Genentech, and lululemon athletica (a yoga clothing retailer). Hence, one would expect larger firms to more frequently use female titles and personal pronouns in their primary written communication with investors and Wall Street analysts.

After adjusting for industry effects, time effects, and firm size, there is little reason to expect gender usage to vary as a function of a firm's age, unless older firms are slower to respond to cultural changes. In managers' communications, we observe that older firms have significantly lower usage of female titles and pronouns. These rhetorical artifacts likely represent a corporate culture established in a period of gender inequality, and underscore that culture is slow to change. Viewing the firm through the lens of management's writing provides a different and complementary perspective to those studies whose focus has been on wage differentials.

[3] Note that, as documented in Loughran and McDonald (2014), the size of annual report filings has increased substantially over the 1996–2013 period. Thus we expect the absolute counts to increase over time. Our discussion and subsequent tests focus on the relative comparisons of the gender-related words.

2. Literature review

Most studies on gender in the workplace focus on participation rates and wage differences. Goldin (2006, p. 1) labels female involvement in the economy as "the most significant change in the labor markets during the past century." Goldin divides the changes into phases, with the final "revolutionary" phase beginning in the mid-1970s. The revolutionary period is broadly defined by women now participating in the workplace because work is part of their fundamental identity, with their focus shifting from jobs to careers. Cultural shifts are difficult to precisely measure, but much research in the area has focused on gender differences in performance and even more so on what Goldin (2014, p. 1093) labels as a "summary statistic for gender differences"—wages. When focusing on the gender gap in wages, the crux of the issue is to separate out those effects that—although possibly an artifact of discrimination by society at a broader level—logically map to differences in wages.

For example, in earlier periods when females were just beginning to broadly enter the work force, their experience was less than their male counterparts. Thus it was not a condemnation of business that wages might be accordingly differentiated. Similarly, in earlier periods, men were more likely to have higher levels of education, once again justifying their higher wages. Goldin (2014) argues that a grand convergence has occurred in the wage gap as these structural artifacts have all but disappeared. However, she goes on to emphasize that the ultimate assimilation of gender differences cannot occur as long as there are temporal inflexibilities—long hours and particular hours—within the workplace.

The wage gap is a strong reflection of gender issues in employment and has been studied extensively in Goldin's body of work along with many others (see, for example, Blau & Kahn, 2000, 2006, 2013, or Niederle & Vesterlund, 2007). Newton and Simutin (2014) consider another perspective on the gender pay gap. They examine the role of age and gender of the CEO. Using a large sample of publicly-traded firms with available COMPUSTAT Execucomp data, Newton and Simutin (2014) report that older, male CEOs significantly undercompensate women executives relative to male workers at the same firm. They believe that older top managers were raised in a culture where paying females less than males (for the same work) was considered socially acceptable.

Another view on females in the labor force is provided by Ahern and Dittmar (2012) who consider the role of females on corporate boards by examining the impact of a 2003 Norwegian law mandating female board representation. The law required that publicly traded firms have women account for at least 40% of their board of directors. At the time of law's passage, only 9% of Norwegian directors were female. As might be expected given the dramatic quota requirement, Ahern and Dittmar (2012) document a significantly negative stock market reaction to the initial announcement of the law. Interestingly, Norwegian companies with female directors did not experience much of a decline at the announcement. However, Norwegian companies with zero women directors had a −3.5% decline in their market values as of the quota's announcement.

The authors found that the quota directly caused Norwegian boards to become younger and less experienced. Given the limited pool of Norwegian women with high levels of managerial experience at the time of the law's passage, it is not completely surprising that companies were forced to appoint directors with significantly less business experience than their existing male directors. What is more controversial in their findings is the subsequent decline in both Tobin's *Q* ((Total assets − common equity + market value)/Total assets) and operating performance of Norwegian companies following the quota. The firm shareholders had to directly endure the costs of the quota law. The paper notes that their results

Table 1
Annual report sample creation.

	Dropped	Sample size
Annual reports filed with the SEC during 1996–2013		197,082
Eliminate amended annual report filings	8,382	188,700
Market value of equity requirement	99,505	89,195

This table reports the impact of various data filters on the initial annual report sample. We remove all annual reports with whose net file size is less than 5000 characters.

are "consistent with boards of directors that are less effective monitors and advisors" (p.188). Burke (1999), Huse and Solberg (2006), and Peterson and Philpot (2007) provide additional examples of studies considering the impact of females in upper management.

Nalikka (2009) documents the impact of women managers on the level of a firm's disclosure policy. Their sample is for only 108 non-financial companies traded on the Helsinki Stock Exchange in 2008. Nalikka (2009) finds that women Chief Financial Officers (CFOs) are significantly linked with higher voluntary disclosures in their firm's annual reports. Interestingly, neither female CEOs nor the proportion of female directors on the company's board have any impact on a firm's voluntary disclosure policy. CFOs clearly play a more important role in the production of business documents than the CEOs, who apparently spend more time on strategic planning than with annual report creation.

As previously discussed, our approach takes a very different perspective on the broader topic of gender in the workplace and focuses on differences revealed by language choices made by firm managers when communicating with investors. Such differences should reflect the composition of managements and boards, along with cultural norms in the firms.

3. Data and variable definitions

Our starting point is downloading all annual reports available on the SEC's Electronic Data Gathering, Analysis, and Retrieval (EDGAR) database during the 1996–2013 time period. In 1993, the SEC began to mandate electronic filings on EDGAR by companies, a policy which was slowly phased in over a three-year period. Thus, our sample begins in 1996 when all firms were required to electronically file. To analyze the annual reports, we follow the parsing procedure recommended by Loughran and McDonald (2014). We exclude exhibits from the analysis, as many of these are legal documents that might have their own particular gender bias. We require all annual reports to have a net file size of at least 5000 characters. This screen removes many annual reports that simply provide explanations for deferred filings or alternate locations.

Table 1 reports the impact of the two data requirements in creating the final sample. Initially, there are 197,082 annual reports filed with the SEC on EDGAR. We download all annual reports (i.e., 10-K, 10-K405, 10KSB, 10-KSB, and 10KSB40 filings) from the EDGAR website. The "KSB" post-scripted filings are ones associated with small businesses, a distinction that was eliminated in 2002. The "405" post-scripted filings were variants of the annual reports that were also eliminated in 2002.[4] None of these distinctions have an impact on our study and we will refer to the collection of forms simply as annual reports.

Often, firms file amended documents with the SEC. As an example, General Electric filed its fiscal year 2004 annual report on March 1, 2005. After discovering that it had incorrectly accounted for certain derivative transactions, General Electric filed an amended annual report (Form 10-K/A) on May 6, 2005 reporting the firm's restated financial statements for the fiscal years of 2004, 2003 and 2002. Following the procedures of other papers in the textual analysis literature, we exclude all amended annual reports from the analysis. This screen removes 8,382 firm-year observations.

The second requirement is the availability of the company market value of equity one day prior to the annual report filing date. Non-publicly traded companies are required to file annual reports with the SEC if the firm has more than $10 million in assets and more than 500 shareholders. Requiring an available market value of equity (shares outstanding multiplied by stock price) on the Center for Research in Security Prices (CRSP) database eliminates 99,505 observations. Our final sample contains 89,195 annual reports filed on EDGAR by publicly-traded companies during 1996–2013.

Our key variables are a count of the number of times female and male titles and pronouns appear in a firm's annual report. The variable, *Female*, is the tabulation of how often the words *Ms*, *Ms.*, *Mrs*, *Mrs.*, *Miss*, *her*, and *she* occur in a firm's annual report.[5] The variable, *Male*, is the tabulation of how often the words *Mr*, *Mr.*, *his*, and *he* occur in a firm's annual report. We also create the ratio, *Female/Male*, which is the occurrence of female titles and pronouns divided by the count of male titles and pronouns in the annual report if the variable *Male* has a non-zero value. To lessen the impact of outliers, the *Female/Male* ratio is winsorized at the 5% level.

As expected, there are some firms with a large number of female or male word counts. For example, retailers J Crew, Children's Place, Paul Harris Stores (a Midwest clothing retailer), and Dollar General all had one annual report with more than 100 references to female titles or pronouns. The annual report filed on March 15, 2004 by Martha Stewart Living Omnimedia had a count of 111 occurrences of the words *Mrs.*, *Ms.*, *Miss*, *she*, and *her*. In that annual report, many of the female-linked terms relate to a discussion of CEO and founder Martha Stewart's conviction of conspiracy and making false statements to federal investigators related to the sale of her shares in another company. Her felony conviction occurred only days prior to the annual report filing.

Since our sample is restricted to publicly-traded firms, all companies in the final sample have available market value of equity. We calculate the market value of equity (share price multiplied by the number of shares outstanding) as of the day prior to the annual report filing. Market values are in millions of US dollars and are from the CRSP database.

CRSP also reports the date when firms start trading on the three major U.S. stock markets (New York Stock Exchange, American Stock Exchange, and Nasdaq). For some companies, the first starting day was their initial public offering (IPO) date while for other firms the date represents the starting point of the CRSP database. For example, Facebook has a start date of May 18, 2012 (its historic IPO date) while General Electric has a beginning date of December 31, 1925 (i.e., the start date for CRSP). We use the CRSP start date as a proxy for age. We calculate the age of each firm, in terms of days, as of their annual report filing date. *Age* is defined as the number of days between the annual report filing date and the first day the firm traded on a stock exchange. As an example, the first CRSP day listed for Google is August 19, 2004. After its initial public offering,

[4] The "405" variant is an annual report with an indication on the first page that a "disclosure of delinquent filers pursuant to Item 405" was not included in the current filing. Until this distinction was eliminated in 2002, a substantial portion of annual reports were categorized as 10-K405. Because there was confusion and inconsistency in its usage, the SEC eliminated the 405 classification.

[5] Note that *Ms* and *Mrs.* are predominant in the counts of female titles. In our entire sample, there are only 50 occurrences of the token *Miss*. Also note that the parse is done in such a way that only capitalized tokens are counted, thus it is very unlikely that *Miss* would be miscounted as the homonym *miss* to the extent it is hard to imagine the term *miss* being used at the beginning of a sentence.

Table 2
Time trend in female and male usage for publicly-traded firms filings annual reports, 1996–2013.

Year	Female	Male	Winsorized Female/Male
1996	0.152	1.734	0.118
1997	0.171	1.864	0.135
1998	0.177	1.885	0.141
1999	0.183	1.932	0.147
2000	0.206	1.997	0.156
2001	0.215	1.955	0.161
2002	0.213	1.957	0.166
2003	0.231	2.065	0.171
2004	0.244	2.085	0.180
2005	0.256	2.188	0.182
2006	0.265	2.262	0.188
2007	0.290	2.459	0.197
2008	0.291	2.366	0.201
2009	0.282	2.295	0.201
2010	0.282	2.268	0.209
2011	0.311	2.402	0.215
2012	0.307	2.322	0.227
2013	0.335	2.390	0.232

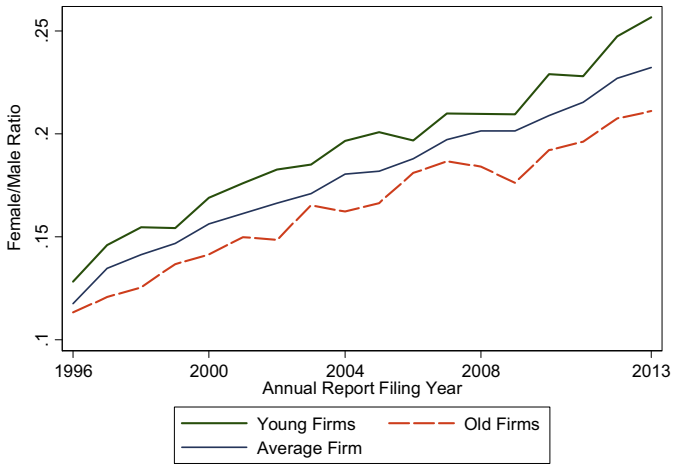

Fig. 1. Time trend of Female/Male counts in annual reports across young and old firms, 1996–2013.

Google's first annual report filing was on March 30, 2005. Since the number of days between Google's annual report filing date and its beginning date on the CRSP data base is 223, Google has an *Age* value of 223 for filing year 2005.

The *Age* variable allows us to categorize the sample into young and old firms. *Young firms* are defined as companies in the lowest *Age* tercile while firms in the highest *Age* tercile are categorized as *old firms*. A number of companies are defined as *young firms* in 1996, however toward the end of our sample period the same firm is classified as an *old firm*. See our Appendix for more detailed variable definitions.

4. Empirical results

We first report the time-series trend of the *Female*, *Male*, and *Female/Male* variables. Table 2 reports the trend in the three variables throughout our sample period. Three patterns are clearly present in the data. First, as expected, female titles and personal pronouns occur much less frequently than their male counterparts. For example, in 1996, the average annual report only mentions female titles or personal pronouns 0.152 times. Male titles or personal pronouns (i.e., *Mr.*, *he*, and *his*) occurs an average of 1.734 times in annual reports. In 1996, 63% of the sample has a *Female* count of zero while by the end of the sample period, only 42% have no occurrences of female titles and personal pronouns in their annual report. The lower frequency of female titles or personal pronouns in annual reports reflects the lower number of women who are senior managers or directors in addition to stylistic choices in writing.

Second, notice the dramatic increase in the values of *Female* and *Male* during the sample period. The average count of female titles and personal pronouns increases more than 100% during the 1996 to 2013 time period (0.152–0.335). The average count of male titles and personal pronouns also rises significantly (1.734–2.390). Part of the higher female and male counts is due to the increasing length of annual reports (see Li, 2008; Loughran & McDonald, 2014). As managers create longer annual reports, all words should exhibit increases in total counts.

Third, the *Female/Male* ratio generally rises each year. Recall that the *Female/Male* variable is winsorized at the 5% level and is only available for firms with at least one male title or personal pronoun in the annual report. The rise in the *Female/Male* ratio is consistent with the increasing proportions of top women managers and board members as companies have become more open to hiring and promoting females to critical positions. The increase in the ratio could also reflect the increasingly important role women play in household purchasing decisions.

4.1. Trend for young and old firms

One of our hypotheses is that young and old firms will differ significantly in their usage of female personal pronouns and titles. Older, more mature companies will be less likely to use female phrases in their annual reports than younger, less established companies. Manager's word choices should proxy for the type of culture in the company. To test this hypothesis, we categorize firms into young, middle, and old groupings.

Fig. 1 plots the *Female/Male* ratio for the young and old firms on a yearly basis. The yearly average is also presented in the figure. Notice the upward trend in all three of the lines. As time has progressed, both young and old firms exhibit increases in the average *Female/Male* ratio. Importantly, as predicted, young firms consistently have higher values for the *Female/Male* ratio. In every year, 1996–2013, young firms more commonly use female words like *Ms.*, *her*, and *she* scaled by the male-linked word count than older, more established companies.

Older companies would likely have corporate cultures that are less responsive to changes in social norms. Similarly, the language usage in annual reports may be slow to change for more established firms. This is reflected in their lower *Female/Male* ratios compared to younger publicly-traded companies. Younger firms might be expected to be more open to employment of women managers and directors. Also, firms that recently went public would have created their initial annual report with a fresh perspective of social norms.

4.2. The Female/Male ratio by Fama–French industries

We would expect traditionally male dominated industries to have a significantly lower ratio for *Female/Male* while industries that cater directly to women customers, are founded by women, and/or have numerous women executives should have higher relatives counts of female-linked words. To examine differences across industries, we categorize all firms into industries defined by Fama and French (1997). The two authors, in a widely-cited finance article, categorize firms into 48 different industries on the basis of their Standard Industrial Classification (SIC) code.

Fig. 2 reports the average *Female/Male* ratio by the 48 Fama–French industries. Also included in the figure are bars reflecting ± two standard errors. The industry with the largest standard error, Smoke (i.e., tobacco), has the largest relative ranking change

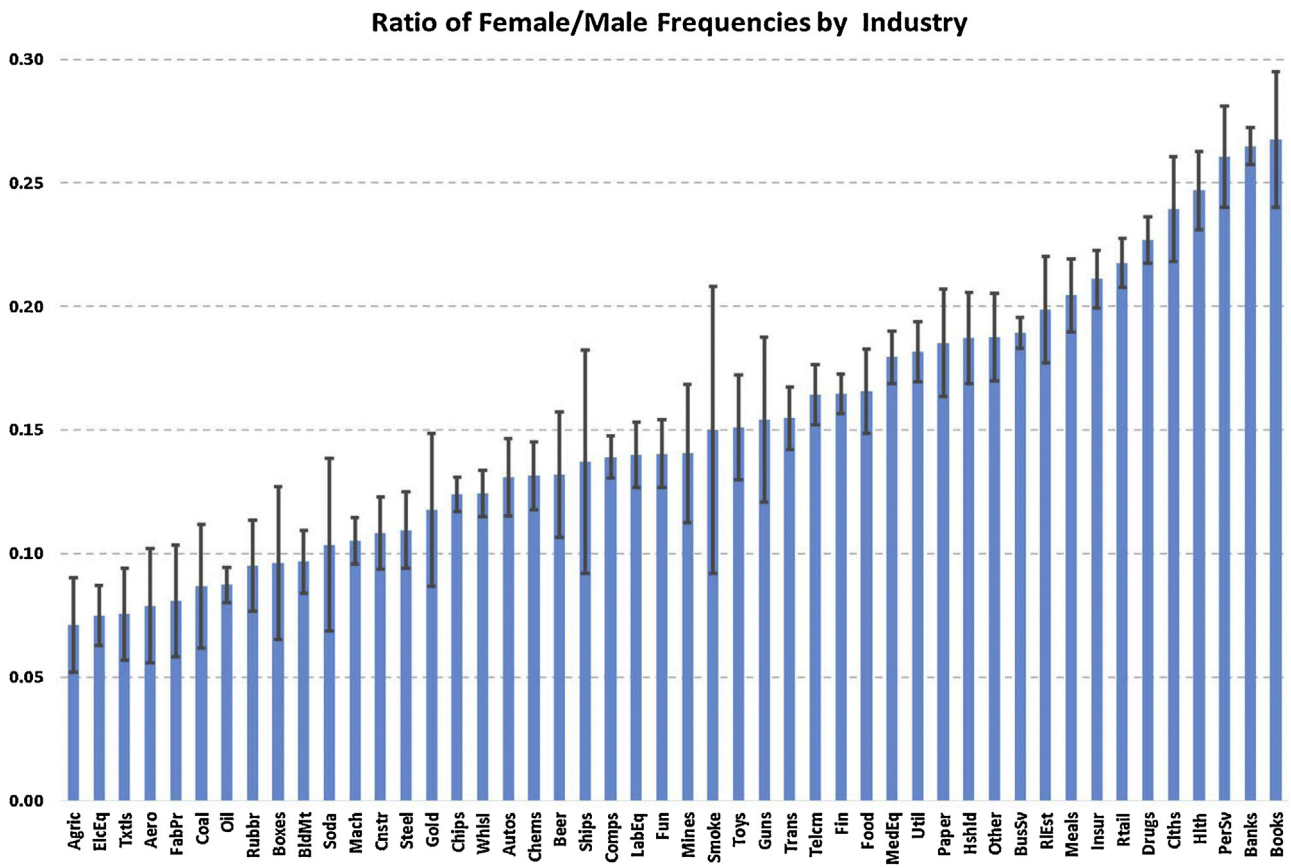

Fig. 2. Graph of the mean Female/Male ratio across Fama–French 48 industries along with bars indicating ± two standard errors.

in the sample. In 1996, the Tobacco industry had the fourth lowest *Female/Male* ratio while the same industry had the highest ratio value by 2013. Two firms, Philip Morris and Lorillard, are solely responsible for the high Tobacco *Female/Male* ratio value in calendar year 2013. As should be expected given the concentrated nature of the sector, Tobacco has the smallest number of firm-year observations of any of the Fama & French 48 industries.

Generally, the pattern in Fig. 2 is one that should be expected. The seven industries with the lowest values for *Female/Male* are Agriculture, Electrical Equipment, Textiles, Aircraft, Fabricated Products, Coal, and Oil. These industries are more male-dominated, often involve extracting minerals from the earth, and tend not to sell products directly to main street customers. These industries tend to have an intermediary between themselves and the ultimate consumers of their product.

Importantly, the industries with the highest *Female/Male* values tend to sell products directly to retail customers. Since women are more likely in the customer mix of the products in these industries, it is not surprising that the firm's annual reports use higher relative frequencies of female-linked words. The top seven industries in terms of relative female titles and personal pronouns are Publishing, Banking, Personal Services, Healthcare, Apparel, Pharmaceutical Products, and Retail. It is important to note that even the industries with the highest *Female/Male* ratios have values that are significantly less than one.[6] The industries of Publishing, Banking, and Personal Services report average *Female/Male* ratios of only slightly more than 0.25. Thus, even the industries with high female personal pronoun and title usage have only about one female-linked word for every four male gender terms.

4.3. Regression results

We have provided descriptive results showing that younger firms have higher *Female/Male* ratios than older firms and that there are dramatic differences across industries for the ratio. In Table 3, we report regressions where the dependent variable is the winsorized *Female/Male* ratio. We include two independent variables: *Log(Market Value)* and *Log(Age)*, using the log transform in both cases to reduce the positive skewness inherent in each measure. As noted before, *Market value* is the CRSP stock price multiplied by the number of shares outstanding on the day prior to the annual report filing date (in $ millions). *Age* is defined as the number of

Table 3
Regressions with winsorized *Female/Male* ratio as the dependent variable.

Variable	(1)	(2)
Log(Market Value)	0.872	1.119
	(4.54)	(4.29)
Log(Age)	−0.834	−1.913
	(−2.74)	(−3.77)
Industry dummies	Yes	Yes
Year dummies	Yes	No
Constant	Yes	Yes
R^2	5.21%	6.24%
Sample size	76,747	3,527

Note: In column (2), the regression is only for annual report filings in calendar year 2012. See the Appendix for variable definations. The *t*-statistics, reported parenthetically, are based on errors clustered by year and industry in column (1) and by industry in column (2).

[6] It is interesting to mention that only 5,083 firm-year observations (out of a potential sample of 76,747) have values of the *Female/Male* ratio of at least one.

days between the annual report filing date and the first day the firm traded on a stock exchange.

In the first column of Table 3, the independent variables of interest are *Log(Market Value)* and *Log(Age)*. The sample size is 76,747 in Table 3 regressions because 12,448 firm year observations have a *Male* value of zero and thus have a missing *Female/Male* ratio. The regressions in column (1) include an intercept, year dummies, and industry dummies. The *t*-statistics, reported parenthetically, are based on errors clustered by year and industry. The coefficients on both of the variables are statistically significant at conventional levels. The coefficient on *Log(Market Value)* is positive with a *t*-statistic of 4.54. This implies that larger firms have significantly higher *Female/Male* ratios. Larger firms have higher public accountability and should be more sensitive to their language usage than smaller firms. Conversely, the coefficient on *Log(Age)* has a negative sign and a *t*-statistic of −2.74.[7] Thus, older firms have significantly lower values of *Female/Male* after controlling for firm size and industry. This is consistent with our time-series evidence in Fig. 1. Annual reports are an important document prepared by firm managers which are meant to inform investors and analysts about current and future operations. It appears that older companies have corporate cultures that are slow to incorporate female-linked gender terms into their business communications with shareholders.

A criticism of the regression in column (1) is that over the time series, the sample includes repeated firms that only incrementally age. That is, do we really expect to see much variation or change in the *Female/Male* ratio in a firm from one year to the next? Column (2) in Table 3 tests the robustness of our results by considering a cross-section of only one year. We select only firms with filings in the year 2012, producing a sample of 3,527 firms. The regressions in the second column of Table 3 include an intercept and Fama and French (1997) 48 industry dummies. The *t*-statistics are calculated using errors clustered by industry. Consistent with the broader sample, both the *Log(Market Value)* and *Log(Age)* variables are statistically significant—*Log(Market Value)* significantly positive, and *Log(Age)* significantly negative. Once again, large firms have higher *Female/Male* values while older companies have lower relative female gender usage terms.

5. Conclusion

The gender landscape in the workplace has undergone revolutionary change in the past generation. The participation in employment, the participation in management and director positions, and the relative wages of females has been studied extensively. We provide another perspective on the elements of this revolution by examining gender usage in annual reports filed with the SEC over the period 1996–2013 by publicly-traded firms. Our descriptive results document a clear gap in the usage of male versus female titles and pronouns, and that this gap has narrowed over the past eighteen years, results wholly consistent with the extant literature on gender in the workplace.

In examining the broad trends more carefully, the extent of heterogeneity across industries is surprising and consistent with the simple notion that firms more directly tied to the ultimate consumer are more likely to have a more female inclusive culture. Not surprisingly, we also find that larger firms tend to be more representative in their use of gender-related words. Presumably, larger firms are more in the public eye and must maintain a higher sensitivity to the dynamics of the culture in which they operate.

More notably we observe a pattern in the data where older firms are less likely to be evolving to a more inclusive culture. We are not aware of this dynamic being identified before. This structural rigidity is consistent with the notion that a male dominant culture is hard to change.

Future research should try to explicitly link gender usage in annual reports to wage differences. Such an analysis would be most appropriate at an industry level. Additionally, given that much of an annual report is written or at least carefully edited by auditors, it could be possible that auditing firms might have distinctly different gender footprints in their writing. Even without a more evolved set of hypotheses, the use of gender in companies' mandated filings provides a reflective look at how gender usage has changed over the past generation along with cultural norms in society as a whole.

Appendix A. Variable definitions

Female	Count of all occurrences of *Ms.*, *Ms*, *Mrs.*, *Mrs*, *Miss*, *her*, and *she* in an annual report (i.e., Form 10-K).
Male	Count of all occurrences of *Mr.*, *Mr*, *his*, and *he* in an annual report (i.e., Form 10-K).
Female/Male	Defined as the ratio of Female/Male for firms with non-zero values of Male. To lessen the impact of outliers, this variable is winsorized at the 5% level.
Log(Market Value)	The natural logarithm of the CRSP stock price multiplied by the number of shares outstanding on the day prior to the annual report filing date (in $ millions).
Log(Age)	The natural logarithm of the number of days between the annual report filing date and the first day the firm traded on a stock exchange. As an example, the first CRSP day listed for Google is August 19, 2004. After its initial public offering, Google's first annual report filing was on March 30, 2005. Since the number of days between Google's annual report filing date and its beginning date on the CRSP data base is 223, Google has an *Age* of 223 for filing year 2005.
Young Firms	Defined as firms with an age, at the time of the annual report filing, in the bottom third of all publicly-traded companies.
Old Firms	Defined as firms with an age, at the time of the annual report filing, in the top third of all publicly-traded companies.

References

Ahern, K. R., & Dittmar, A. K. (2012). The changing of the boards: The impact on firm valuation of mandated female board representation. *The Quarterly Journal of Economics*, *127*, 137–197.

Blau, F. D., & Kahn, L. M. (2000). Gender differences in pay. *Journal of Economic Perspectives*, *14*, 75–99.

Blau, F. D., & Kahn, L. M. (2006). The gender pay gap: Going, going . . . but not gone. In F. D. Blau, M. C. Brinton, & D. B. Grusky (Eds.), *The declining significance of gender?* (pp. 37–66). New York: Russell Sage Foundation.

Blau, F. D., & Kahn, L. M. (2013). Female labor supply: Why is the United States falling behind? *American Economic Review*, *103*, 251–256.

Borghans, L., ter Weel, B., & Weinberg, B. A. (2006). *People people: Social capital and the labor market outcomes of underrepresented groups.* NBER Working Paper No. 11985.

Burke, R. (1999). Women on Canadian corporate boards of directors: Getting the numbers right!. *Corporate Governance: An International Review*, *7*, 374–378.

Fama, E. F., & French, K. R. (1997). Industry costs of equity. *Journal of Financial Economics*, *43*, 153–193.

Goldin, C. (2006). The quiet revolution that transformed women's employment, education, and family. *The American Economic Review*, *96*, 1–21.

Goldin, C. (2014). A grand gender convergence: Its last chapter. *The American Economic Review*, *104*, 1091–1119.

Huse, M., & Solberg, A. G. (2006). Gender-related boardroom dynamics: How Scandinavian women make and can make contributions on corporate boards. *Women in Management Review*, *21*, 113–130.

Li, F. (2008). Annual report readability, current earnings, and earnings persistence. *Journal of Accounting and Economics*, *45*, 221–247.

Loughran, T., & McDonald, B. (2014). Measuring readability in financial disclosures. *Journal of Finance*, *69*, 1643–1671.

Mulligan, C. B., & Rubinstein, Yona. (2008). Selection, investment, and women's relative wages over time. *The Quarterly Journal of Economics*, *123*, 1061–1110.

[7] Note that if we do not include the industry and time dummy variables in column (1) both *Log(Market Value)* and *Log(Age)* remain statistically significant, with *t*-statistics of 4.52 and −3.68, respectively.

Nalikka, A. (2009). Impact of gender diversity on voluntary disclosure in annual reports. *Accounting & Taxation*, *1*, 101–113.
Newton, D., & Simutin, M. (2014). *Of age, sex, and money: Insights from corporate officer compensation on the wage inequality between genders*. Working paper, Concordia University.
Niederle, M., & Vesterlund, L. (2007). Do women shy away from competition? Do men compete too much? *Quarterly Journal of Economics*, *122*, 1067–1101.
Peterson, C. A., & Philpot, J. (2007). Women's roles on US Fortune 500 boards: Director expertise and committee memberships. *Journal of Business Ethics*, *72*, 177–196.
Tetlock, P. C. (2007). Giving content to investor sentiment: The role of media in the stock market. *Journal of Finance*, *62*, 1139–1168.
Tetlock, P. C., Saar-Tsechansky, M., & Macskassy, S. (2008). More than words: Quantifying language to measure firms' fundamentals. *Journal of Finance*, *63*, 1437–1467.