Job Satisfaction, Management Sentiment, and Financial Performance: Text Analysis with Job Reviews from Indeed.com

Sijie (Fiona) Feng[1]

[1]Affiliation not available

July 28, 2020

Abstract

Employee job reviews contain information on many firm characteristics that are unobservable to the public. Job review sites such as Glassdoor and Indeed are used by job hunters and current employees to inform each other of prevailing workplace culture. I collect over 1.2 million job reviews from Indeed.com, matched to firm level data was procured from Compustat. I use machine learning and natural language processing methods to obtain qualitative data about firms across the dimensions of job satisfaction and management sentiment. In the cross-section, I find that a 1% increase in ratings increases market value by approximately 0.68-0.73%, revenue by 0.62-1.01%. However, increases in management sentiment may have a non-monotonic effect on financial performance. These non-monotonic relationships are also present in first difference results, which show that there may be trade-offs involved in improving financial results and workplace culture.

Introduction

Job reviews have become a rich source of "inside" information on qualitative factors relevant for both employee satisfaction and the firm's overall financial performance. While available quantitative data has been used extensively to determine what are relevant determinants of positive financial outcomes, the use of text data is a novel approach. I mine text data from publicly listed companies, resulting in approximately 1.2 million job reviews from 2012-2017, retrieved from the careers site Indeed.com. I then gather data on earnings per share, market value, and revenue from Compustat. A key variable that is vital to firm outcomes is the quality of management, which is generally unobservable by parties external to the firm, such as researchers and investors. However, it is observable by employees of the firm. Management is also frequently mentioned in job reviews. This would allow me to construct a measure of the quality of management from the perspective of employees. I use the portion of review text that mention "management" as an indication of employee sentiment towards management. To derive a specific value for management sentiment, I use three different methods of sentiment analysis. I use supervised classifiers Deep Neural Networks and XGBoost, trained on review text and ratings, to predict sentiment for the management portion of the review text. In addition, I construct a measure of how "non-generic" a review is, to counteract the presence of uninformative reviews. This is based on the term frequency-inverse term frequency (tf-idf) of a particular document, with the rationale that reviews with a higher average tf-idf contain more terms that are more uncommon, and thus more non-generic. I discuss my methodology in . The main contribution of this paper is the measurement of previously unobservable firm characteristics using company review text data.

I then conduct both pooled cross-sectional and first difference OLS regressions of financial performance as explained by the prior year's ratings, management sentiment, and review non-generic index. I find strong evidence of there being a positive relationship between ratings and financial performance in the cross

section. A 1% increase in ratings increases market value by approximately 0.68-0.73%, revenue by 0.62-1.01%. There is a significant relationship also between management sentiment, the non-generic index, and financial performance, but their relationships are non-monotonic: positive relationship for certain range of values, and negative relationship in others. This non-monotonic relationship is also present in the first difference regression results, which indicates the strong influence of firm-level fixed effects. One possible explanation for this is the truncation in ratings, although other papers in the literature also indicate "double-edge sword" effects of focusing on improving short-term financial performance at the expense of a positive long-term corporate culture. In this case, a decline in job satisfaction might lead to an increase in financial performance. The estimation strategy and results are discussed in .

Prior Literature A nascent literature has used job review data from prominent jobs site Glassdoor. (Popadak, 2013) uses job review data to assess the effect of active shareholder governance on both financial performance and corporate culture. Corporate culture has been defined by (Schein, 1990) as shared assumptions, values, and beliefs that help employees understand which behaviors are and are not appropriate. Popadak finds that when active shareholders change corporate culture towards being more results driven, there are increases in sales, profitability, and payout. However, in a longer-term horizon of five years, there are impairments to intangible "goodwill" assets and customer satisfaction. These findings are well-aligned with my results that there is likely to be non-monotonicity between cultural or qualitative factors and firm performance. (Luo et al., 2016) finds positive correlation in the cross-section between employee satisfaction and corporate performance. (Ji et al., 2017) find that firms with lower levels of job satisfaction and lower levels of "culture and values" are more likely to be subjected to SEC fraud enforcement actions and securities class action lawsuits, although they are also more likely to beat market earning expectations.

Data and Methodology

I collect rating and review data from the job site Indeed, the largest job site in the US. The time frame available is 2012-2017. Each review contains the following information: the job title of the position held by the reviewer, the date of the review, a rating from 1 to 5 of the company, the location of the job, as well as the text of the review alongside pros and cons. (See 1 for example of a review) In order to match to financial data from Compustat, I select only companies that are available in Compustat. The reviews in my sample are thus exclusively from large publicly traded companies. Overall, there are 1.24 million reviews for 2,738 companies. The average rating across all reviews is 3.73. Most reviewers give their jobs a positive rating of 4 (31%) or 5 (21%), with only 14% of reviewers giving their jobs a negative rating of 1 or 2. (See 8).

As for industries, only a third of the reviews (519,254) have companies that can be mapped to an industry at the NAICS 3-digit level. The top industries by number of reviews are mostly in food, retail, and transportation, as these industries employ larger workforces.

Compustat Data Compustat compiles financial data from the accounting records for publicly listed companies. I download data on Earnings Per Share (EPS), Market Value, and Revenue for all companies 2012-2017, which covers the same timeframe as the reviews data. These variables were selected to be the standard indicators of financial performance.

Text Analysis of Positive and Negative Reviews

The texts of reviews are "stemmed" to their base word (e.g. "competitive" to "compet") and stop words (e.g. "and", "the") are removed. Other words that occur in less than 10 reviews are also removed to constrain the dimensionality of the resulting matrix that is fed into the algorithm. I use both one and two-grams (i.e. sequences of two words) in the analysis. In order to better understand the difference in words used to describe positive experiences of jobs (ratings 4 or 5) compared to negative ratings (1 or 2), I count the frequency of

	2012	2013	2014	2015	2016	2017	All
Mean Rating	3.75	3.77	3.77	3.77	3.74	3.65	3.73
Number of Reviews	115736	171864	184896	203985	196718	365851	1240541
Number of Companies	1687	1892	1985	2109	2128	2390	2738
Mean Earnings Per Share	0.4	0.35	0.38	0.2	0.31	0.42	0.23
Mean Market Value	7346.26	8222.08	8937.79	8970.79	9567.04	11579.22	8946.93
Mean Revenue	1909.65	1843.68	1787.14	1632.13	1701.9	1950.19	1851.98

Table 1: Summary statistics of companies with reviews available from Indeed and Compustat

Industry	NAICS3	Num of Reviews
Food Manufacturing	311	48702
General Merchandise Stores	452	44171
Food Services and Drinking Places	722	42806
Credit Intermediation and Related Activities	522	33845
Building Material and Garden Equipment	444	17288
Administrative and Support Services	561	17194
Computer and Electronic Product Manufacturing	334	16182
Transportation Equipment Manufacturing	336	15310
Clothing and Clothing Accessories Stores	448	14532
Broadcasting (except Internet)	515	14077

Table 2: Top industries at NAICS3 level by number of reviews

Term	Importance	Term	Importance
horribl	0.0060	like famili	0.0017
poor	0.0058	dull	0.0017
terribl	0.0045	yell	0.0016
worst	0.0042	hostil	0.0016
unprofession	0.0029	unless	0.0015
told	0.0028	micro manag	0.0015
fire	0.0026	unorgan	0.0015
dont	0.0022	ensur	0.0015
turnov	0.0022	joke	0.0015
lie	0.0021	bottom line	0.0015
suck	0.0019	excel	0.0015
let go	0.0018	various	0.0015
upper	0.0018	challeng	0.0015
favorit part	0.0017	promis	0.0014
rude		lack	0.0014

Table 3: Most important features for discerning between positive and negative reviews.

terms that appear in positive and negative reviews. 9 show the 20 most common words that appear across all reviews; 10 show the common words for positive reviews; 11 for the negative reviews. While all reviews feature words that describe aspects of the job (*customers, day, company, people, time*), the positive reviews contain words that reflect constructive aspects of work (*learn, enjoy*), while negative reviews mention *pay* , which potentially reflects their dissatisfaction with their compensation.

Another method to finding words that are important for discerning between positive and negative reviews is to train a supervised machine learning algorithm on the labelled reviews, and allow the algorithm to specify which features are most important. Since not all algorithms provide such metrics, I use an implementation of a Gradient Boosting Method called XGBoost. Gradient boosting trains an ensemble of simple models to make predictions. For XGBoost, the component simple models are decision trees. A feature, in this case a term, is important when it is used to make an important decision or "split" in the tree. For a single decision tree, the importance of a term is the amount that the split at the term (i.e. whether the document contains the term or not) improves on the performance measure (e.g. accuracy, whether the predicted label matches the actual training label). 3 provides the results for the top 30 most important terms.

Using supervised machine learning to find important terms provides a vivid description of mostly what makes certain positions unpopular. Besides emotive language such as *horrible, poor, terrible, worst,* there are mentions of *firings* from employees leaving negative reviews. This aligns with a priori expectations which shows that the algorithm is reasonable in finding discerning terms. More interesting is the inclusion of turnover, which clearly represents continuing challenges at the firm leading to many employees leaving. Descriptions of *unprofessional* and *unorganised* work environments are also important determinants of negative job experiences, as is *micromanagement* . On the positive side, reviews that mentioned that they had *favourite parts* of their work were far more likely to leave high ratings, as well as those that described a work environment *like family* .

Predicting Management Sentiment

One important variable that cannot be ascertained from financial data is the quality of management at a firm. Job review data may provide an important signal into the qualitative attributes of a firm. Many reviews contain some mention of management, which encapsulates the employee's subjective opinion about management or managers at the firm. The limitation of this data is that the metrics that the employee uses to form a sentiment of management may be very different from what other interested parties, such as investors, would assess as good or bad management. However, averaged across many employees, the signal is likely to be more informative.

I ascertain management sentiment in three ways. The first two methods use supervised machine learning classification to predict whether management sentiment is likely to be positive or negative. The final method uses Google Cloud SDK to perform sentiment analysis on management review text.

Procuring management review text   I perform Word2Vec on a sample of reviews. Word2Vec is a neural network algorithm that transforms words into vector embeddings, and situates words which are closer to one another in meaning close to each other in the vector space. I gather a list of terms that are most similar to *manag* , the stemmed form of management, which are listed in . For each review that contains a management term, I gather a management phrase consisting of the five words prior and post to the management term. For example, a entire job review from 8 is: "I didn't like the work environment at Walmart at all. Some of the employees where rude and very unfriendly. It wasn't hard work if you could keep up. I didn't feel that I could go to the management with issues that I had. I felt like they would judge me and think less of me." The management portion of the review would constitute a management phrase: "I didn't feel that I could go to the management with issues that I had."

Supervised machine learning classification of management sentiment   I train machine learning algorithms based on the method described in . The classifier would predict the sentiment of the management phrase based on whether it is more likely to be a positive or negative "review". This extracts the particular sentiment about management, separate from the rating of the overall review. While the two are likely to be highly correlated, it is possible that positive reviews contain negative management sentiment, and vice versa. Thus, it is reasonable to look at management sentiment separately from the rest of the job rating. I use two supervised learning algorithms: `XGBoost` and a Deep Neural Network implemented in the Python package `tensorflow`. Both predict positive/negative reviews with an accuracy rate of 0.88 on the test data set. The management sentiment algorithms are trained on review text then predicted on management phrases, resulting in a probability prediction in the range $[0, 1]$.

Google Cloud SDK Sentiment Analysis   Management sentiment does not need to be exclusively trained on job reviews data, since other terms unrelated to sentiment (such as industry related terms, for example *sales, transport* , etc.) may affect the predictions. Google Sentiment Analysis is a state of the art API which can analyse the sentiment of any text, and gives a score between $[-1, 1]$.

5

	$rating$	$sent$	$manag\,sent_{NN}$	$manag\,sent_{XG}$	$manag\,sent_{Goog}$
$rating$	1.00	0.93	0.13	0.53	0.41
$sent$	0.93	1.00	0.13	0.53	0.40
$manag\,sent_{NN}$	0.13	0.13	1.00	0.22	0.22
$manag\,sent_{XG}$	0.53	0.53	0.22	1.00	0.65
$manag\,sent_{Goog}$	0.41	0.40	0.22	0.65	1.00

Table 4: Correlation between review ratings, sentiment, and management sentiments. $NN =$ Deep Neural Network, $XG =$ XGBoost, $Goog =$ Google.

Correlations of review rating, sentiment, and management sentiments for each of the three measures are are reported in 4. While the management sentiment predicted by the Google API and XGBoost are similar, the management sentiment predicted by the Deep Neural Network is less correlated with the other measures.

Non-Generic Index

Not all reviews are equally informative. There are also reports that employees have been encouraged to leave positive reviews on Glassdoor by their employers, which may skew the ratings data. It is likely that such reviews are less descriptive and thus more "generic" than other reviews. In order to account for how generic a review is, I introduce a measure called the Non-Generic Index. This is based on the term frequency-inverse doccument frequency (tf-idf) of a particular document. Roughly, tf-idf measures how important terms (one word or pairs of words or other $n-$grams) are to a particular document. This is measured by the frequency that the term appears in the document, then offset by how frequently the term appears across all documents. To measure how generic a document is, I calculate the mean tf-idf of all unique terms in a document. As discussed in , the more frequent a term across all documents, the lower its inverse document frequency (idf) will be. Thus, if a document contains many words that are commonly used across all reviews, the value of its average tf-idf will be low. For example, a management phrase consisting of "*manag staff great work great*" is more generic and thus has a low non-generic index of 0.0004, while "*leadership skill includ profession busi ethic*" which includes more uncommon words, has a higher non-generic index of 0.001. Each review and management phrase is given its own non-generic index weighting, so that when ratings are averaged by company, the more generic reviews are weighted less than the less generic reviews.

Estimation

Variable Construction

Data aggregation The text data available for each review is the ratings, non-generic weighted ratings, management sentiments, and non-generic weighted management sentiments. Prior to regression estimation, I aggregate each variable in the dataset by company and year. For example, if for Company A there are three reviews of ratings 1,4,5 in 2016, then the rating for Company A would be 3.33. I then remove outliers that are more than 3 S.D.s away from the mean, and rescale the remaining observations to lie between an index of 1-100. These values would roughly represent their overall percentile. In this case, we are seeking to

	rating	sent	$manag\,sent_{NN}$	$manag\,sent_{XG}$	$manag\,sent_{Goog}$	NGI_r	NGI_{ms}	rev	mktv
count	12002.00	11792.00	9427.00	9412.00	6293.00	11966.00	9448.00	8539.00	8565.0
mean	4.10	4.27	4.07	4.14	4.02	3.85	3.87	1.28	1.25
std	0.40	0.32	0.38	0.39	0.40	0.33	0.39	1.09	1.08
min	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
25%	4.00	4.15	4.00	4.05	3.92	3.75	3.80	0.35	0.33
50%	4.18	4.32	4.12	4.22	4.08	3.90	3.94	1.00	0.97
75%	4.27	4.49	4.26	4.35	4.20	4.01	4.06	1.99	1.92
max	4.61	4.61	4.61	4.61	4.61	4.61	4.61	4.61	4.61

Table 5: Summary statistics of data.

explain firm financial performance as measured by annual earnings per share (eps), market value ($mktval$), and revenue (rev) using the independent text variables.

Pooled OLS Regression

I estimate a pooled OLS regression across all companies and all years using the following regression equation:

$$y_{i,t} = \beta_0 + \beta_1 rating_{i,t-1} + \beta_2 manag\,sent_{NN,i,t-1} + \beta_3 manag\,sent_{Goog,i,t-1}$$
$$+ \beta_4 NGI_{r,i,t-1} + \beta_5 NGI_{ms,i,t-1} + year\,FE + \epsilon_{i,t}$$
(1)

Where i is the company and t is the year of observation. The dependent variable y will be either $eps, mktval$ or rev for the current year. The explanatory variables are all taken from the prior year: $rating$ indicates the average rating, $manag\,sent_{NN}$ and $manag\,sent_{Goog}$ represents the average management sentiment using the Deep Neural Network algorithm and the Google API, respectively. The XGBoost predicted management sentiment is excluded as it is never signficant in any of the analyses. NGI_r represents the average non-generic index for review texts, while NGI_{ms} represents the non-generic index for management phrases. A high NGI would represent less generic reviews left for the company i.

I also estimate a regression that includes the product of the non-generic indices with their measures:

$$
\begin{aligned}
y_{i,t} =& \beta_0 + \beta_1 rating_{i,t-1} + \beta_2 manag\,sent_{NN,i,t-1} + \beta_3 manag\,sent_{Goog,i,t-1} \\
& + \beta_4 NGI_{r,i,t-1} + \beta_5 NGI_{ms,i,t-1} + \beta_6 NGI_r \cdot rating_{i,t-1} \\
& + \beta_7 NGI_{ms} \cdot manag\,sent_{NN,i,t-1} + \beta_8 NGI_{ms} \cdot manag\,sent_{Goog,i,t-1} \\
& + year\,FE + \epsilon_{i,t}
\end{aligned}
$$
(2)

Here, $NGI_r \cdot rating_{i,t}$ would represent the average rating for company i in year $t-1$, weighted by the genericity of each review as given by NGI_r. Similarly, $NGI_{ms} \cdot manag\,sent_{NN,i,t-1}$ represents the average deep neural network management sentiment, weighted by the genericness of the management phrase. Thus, the ratings, sentiments, and NGIs will enter non-linearly in the regression. All variables are then logged to allow for clear interpretation of coefficients.

Without directly accounting for fixed company effects, I first estimate whether there are cross-sectional relationships between job rating, management sentiment, and financial performance.

First Difference Regression

It is highly likely that company fixed effects may account for much of the variation in both text-based explanatory variables and the financial variables. A successful company policy or management style may account for both high employee satisfaction and financial performance. To account for within-company fixed effects which may confound variation in both dependent and independent variables, I utilize first differencing within companies for both eqeq (1) (1) and eq (2). Thus, (2) becomes:

$$
\begin{aligned}
\Delta y_{i,t} =& \beta_0 + \beta_1 \Delta rating_{i,t-1} + \beta_2 \Delta manag\,sent_{NN,i,t-1} + \beta_3 \Delta manag\,sent_{Goog,i,t-1} \\
& + \beta_4 \Delta NGI_{r,i,t-1} + \beta_5 \Delta NGI_{ms,i,t-1} + \beta_6 \Delta NGI_r \cdot rating_{i,t-1} \\
& + \beta_7 \Delta NGI_{ms} \cdot manag\,sent_{NN,i,t-1} + \beta_8 \Delta NGI_{ms} \cdot manag\,sent_{Goog,i,t-1} \\
& + year\,FE + \epsilon_{i,t}
\end{aligned}
$$
(3)

Results

The main finding is that job ratings have a significant positive correlation with financial performance measured using all three variables of interest. In the cross section, a 1% increase in the ratings index increases market value by approximately 0.68-0.73%, revenue by 0.62-1.01%, and earnings per share by insig-0.06%. Management sentiment, although correlated with ratings, also plays a significant role in its own righ. However, the relationship between management sentiment and financial performance appears to be non-linear and non-monotonic: while the effect of Neural Network management sentiment ($manag\,sent_{NN}$) on market value is significant and positive, the non-generic index weighted sentiment is significant and *negative* , for both market value and revenue. A 1% increase in $manag\,sent_{NN}$ last period increases market value this period by $(1.04 - 1.27 NGI_{ms})\%$ and revenue by $(1.07 - 1.29 NGI_{ms})\%$. Thus, when $NGI_{ms} \leq 0.8$, an increase in $manag\,sent_{NN}$ would likely increase financial performance; when $NGI_{ms} > 0.8$ an increase in

8

y	eps		mktval		rev	
	(1)	(2)	(3)	(4)	(5)	(6)
$const$	3.717	3.6004***	-1.77***	-3.7223***	-1.1609**	-4.086***
	0.0716	0.0996	0.5219	0.7462	0.5339	0.7962
$NGI_r \cdot rating$		-0.0196		0.1443		-0.6499
		0.0482		0.36		0.4037
$rating$	0.0484	0.0579**	0.7306***	0.6764***	0.6192***	1.0132***
	0.0106	0.03	0.0795	0.2313	0.0682	0.264
$manag\,sent_{NN}$	-0.0052	0.0122	0.0343	1.0382***	0.0227	1.0714***
	0.0072	0.0267	0.0515	0.1794	0.0569	0.1969
$NGI_{ms} \cdot manag\,sent_{NN}$		-0.0198		-1.2682***		-1.289***
		0.0348		0.2472		0.2629
$manag\,sent_{Goog}$	0.0039	0.0796	0.0037	0.948***	-0.0323	0.8788***
	0.0087	0.0519	0.0517	0.2871	0.0531	0.2574
$NGI_{ms} \cdot manag\,sent_{Goog}$		-0.0832		-1.0573***		-0.9966***
		0.0587		0.3168		0.2956
NGI_r	-0.0088	0.0071	-0.1046	-0.1749	-0.1173	0.3621
	0.0126	0.0353	0.0924	0.2711	0.0928	0.2925
NGI_{ms}	0.0132	0.0435**	0.161***	0.9503***	0.2022***	0.9611***
	0.0081	0.0177	0.0587	0.1322	0.0548	0.1244

Table 6: Regression results explaining financial performance, pooled OLS. Standard errors are reported below the coefficient estimates.

9

y	Δeps		$\Delta mktval$		Δrev	
	(1)	(2)	(3)	(4)	(5)	(6)
$const$	3.9312***	3.9308***	1.7558***	1.7557***	1.8222***	1.8255***
	0.0074	0.0074	0.0479	0.0479	0.0479	0.0479
$\Delta NGI_r \cdot rating$		0.0076		0.3261*		0.4769***
		0.0372		0.2034		0.1795
$\Delta rating$	0.0092	0.006	-0.065	-0.259**	-0.1476**	-0.4314***
	0.0104	0.0258	0.0706	0.132	0.067	0.1085
$\Delta manag\,sent_{NN}$	0.0021	0.0174	-0.0596	-0.0909	-0.0581	-0.2724
	0.005	0.0198	0.0371	0.1749	0.0435	0.1903
$\Delta NGI_{ms} \cdot manag\,sent_{NN}$		-0.0195		0.0623		0.2966
		0.0278		0.2232		0.2358
$\Delta manag\,sent_{Goog}$	0.004	-0.0043	0.0735*	0.1776	0.0687*	0.2549
	0.0066	0.0332	0.0441	0.1776	0.0414	0.2395
$\Delta NGI_{ms} \cdot manag\,sent_{Goog}$		0.0118		-0.1145		-0.2217
		0.0367		0.2501		0.2626
ΔNGI_r	-0.0074	-0.0105	-0.0177	-0.2279	-0.0498	-0.366***
	0.0107	0.0305	0.0733	0.1643	0.0738	0.1472
ΔNGI_{ms}	0.0079	0.013	0.023	0.0439	0.0672*	0.0098
	0.0067	0.0139	0.0416	0.1104	0.0383	0.1036

Table 7: Regression results explaining financial performance, first difference. Standard errors are reported below the coefficient estimates.

10

$manag\,sent_{NN}$ would likely decrease financial performance. The same holds for the Google API management sentiment ($manag\,sent_{Goog}$). A 1% increase in $manag\,sent_{Goog}$ last period increases market value this period by $(0.95 - 1.06NGI_{ms})\%$ and revenue by $(0.88 - 1NGI_{ms})\%$.

While the non-generic index measure of the full review text (NGI_r) is insignificant across all measures, the non-generic index of management phrases (NGI_{ms}) is significant and *positive* . However, there is also non-linearity in this variable as measured by the weighted management sentiments.

Overall, at the cross section level, there is likely a positive relationship between last period's job ratings and this period's financial performance. There is likely to be a significant, but non-linear, relationship between last period's management sentiment and non-genericity of management sentiment, and this period's financial performance.

With first differences, the effect of job ratings and management sentiment is less pronounced. The effect of a 1% increase in ratings in year from $t-2$ *relative* to $t-1$ will the relative market value from $t-1$ to t by $(0.33\Delta NGI_r - 0.26)\%$, which will be negative if $\Delta NGI_r < 0.79$. The effect on a relative change in revenue this year will be $(0.48\Delta NGI_r - 0.43)\%$. A 1% relative increase in the non-generic index of a review may also decrease relative revenue by 0.37%.

Conclusion

I introduce measures of previously unobservable firm characteristics using company review text data. I find that there is a positive relationship between ratings and financial performance in the cross section, although the evidence is mixed when examining relative within-firm effects. These findings indicate that there may be trade-offs involved between creating a better workplace culture and short term financial performance.

Appendix

Figure 1: Example of review

1. Ratings: number of stars out of 5

2. Job title

3. Location of review

4. Date of review

5. Text of review

The blue boxes indicate terms that may be important in determining the rating given by the reviewer.
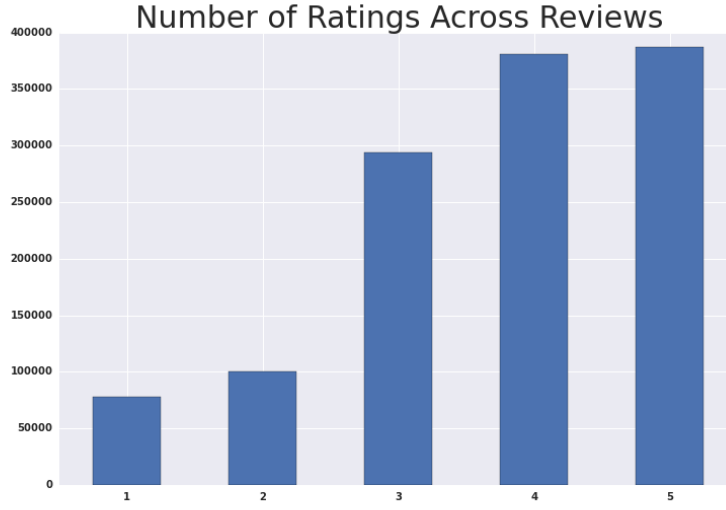
Figure 2: Number of ratings across different reviews.

Management words

manag, manger, supervisor, mgmt, manager, uppermanag, mgt, leader, leadership, mgrs

tf-idf

Term frequency (tf) of a term t in document d is simply given by the raw count of the number of times a word appears in a document. For the document "the cat in the hat," the word "cat" has a term frequency of 1. The inverse document frequency (idf) is a measure of how frequently a term appears in the entire vocabulary of the corpus. For example, if a second document was "the cat ate the rat," then then idf of the term "cat" would be:

$$log \frac{\text{Num Documents}}{1 + \text{Num Documents that term appears in}} = \log \frac{2}{1+2} = 0.10$$

(4)

$\text{tf} \cdot idf = 2 \cdot 0.1 = 0.2$

So the tf-idf of the term "cat" in document d is:

The tf-idf grows with (i) the number of times a term appears in the document; (ii) the rarity of the term across all documents.

13

Rating	Proportion	Example Review
1	6.24%	just a very bad job overall ... way too much for employees getting paid minimum wage and not getting a substantial amount of hours.'
2	8.07%	The building is a dungeon. The pay is not competitive. They thing if they give you free tv services that makes up for not paying better hourly.'
3	23.73%	I didn't like the work environment at Walmart at all. Some of the employees where rude and very unfriendly. It wasn't hard work if you could keep up. I didn't feel that I could go to the management with issues that I had. I felt like they would judge me and think less of me.
4	30.70%	Its an hustle and bustle kind of job,many duties and responsibilities most of the time its fun to meet and help people with their purchase.Help the customers find exactly what they are looking for,helping them get the best sale advantage,helping them with the on line pick up as fast as possible ans cashing them out as quickly as possible.'
5	21.20%	Wish I could have this job forever! This was such a great experience with great potential for advancement and learning. I appreciate every day what I gained from this company.'

Table 8: Example of reviews for each rating

Term	Count
great	422478
custom	405251
day	348203
compani	343606
part	337977
learn	298540
good	294929
time	290837
peopl	290505
enjoy	274470
get	255953
place	233976
help	224524
employe	216242
would	213462
part job	193904
make	187665
cowork	169914
lot	163810
hour	158269

Table 9: Most common terms used, all reviews

Term	Count
custom	263927
part	221609
day	213454
compani	211669
learn	209877
enjoy	194285
good	183258
peopl	177892
time	166793
help	157438
place	151836
get	136694
part job	128350
would	126528
employe	113478
cowork	111674
make	111577
team	111059
alway	105383
lot	102723

Table 10: Most common terms used, positive reviews

Term	Count
compani	57310
get	52532
employe	52051
day	48343
time	47999
custom	45446
peopl	43750
hour	37639
would	36561
good	35335
part	33383
pay	32074
place	32047
like	31512
make	30199
one	26857
dont	26160
great	25584
store	24105
go	23751

Table 11: Most common terms used, negative reviews

References

Yuan Ji, Oded Rozenbaum, and Kyle T Welch. Corporate culture and financial reporting risk: Looking through the Glassdoor. 2017.

Ning Luo, Yilu Zhou, and John Shon. Employee satisfaction and corporate performance: Mining employee reviews on glassdoor. com. 2016.

Jillian Popadak. A corporate culture channel: How increased shareholder governance reduces firm value. *Available at SSRN*, 2345384, 2013.

Edgar H Schein. *Organizational culture.*, volume 45. American Psychological Association, 1990.