The predictive power of the sentiment of financial reports

Jan-Hendrik Meier[1[0000-0002-3080-2210]], Walid Esmatyar[1], and Rasmus Frost[1]

[1] Kiel University of Applied Sciences, Sokratesplatz 2, 24149 Kiel, Germany
jan-hendrik.meier@fh-kiel.de

Abstract. The present study examines the predictive power of the tone or sentiment of 10-K annual and 10-Q quarterly financial statements for future corporate development. The sentiment indicator was calculated using word lists developed for financial texts by Loughran and McDonalds [23] and Henry [14] and applying a conventional and a tf-idf weighted word count. The results show that the sentiment indicator is of significant incremental prognostic quality both for the next quarter and the quarter following it. Unlike suggested by previous literature, neither the scope and content of the word lists nor the weighting method applied had a significant influence on forecasting quality.

Keywords: Sentiment Analysis, Sentiment, Text Mining, Text Analysis, Prediction, Tone, tf-idf.

1 Introduction

Both professional financial analysts and interested private investors have been able to develop their skills in analyzing financial statements in recent years. Sophisticated computer-aided analysis tools help them to process increasing amounts of data efficiently and free from bias to create decision-relevant information. Of interest in this context is that the technical solutions almost exclusively focus on quantitative financial data and figures, while qualitative financial data, mainly texts, are hardly considered in the analyses. The reason is that figures are considered to be more reliable and less prone to manipulation than textual data because of their better verifiability; in contrast, texts offer more room for interpretation.

Numerous studies have shown that the texts of annual financial statements contain important information that cannot be obtained directly from key figures. While complex verbal statements – despite all advances in the field of natural language processing – cannot yet be adequately interpreted and evaluated by applying a software, implicit information in the tone of printed texts can be detected and evaluated by means of sentiment analysis. The present analysis deals with the question of whether a positive or negative tone in annual and quarterly financial statements can be used to forecast future earnings figures, thus allowing insight into a company’s future development. In order answer this question, a sentiment analysis based on a total of 19,390 annual and quarterly financial statements of the companies listed in the Standard & Poor's 500 Index (S&P 500) for the years 2005 to 2015 was conducted. The results show that the sentiment indicators for annual and quarterly financial statements provide a highly significant incremental contribution to predicting the return on assets (ROA) for the next two quarters. The present study provides empirical evidence on the basis of a very broad random sample and confirms its findings by using two established text-mining methods and comparing their results. The study also assesses the extent to which the scope of the word lists used and the weighting method influence forecasting quality.

Section 2 of this study presents a literature review and provides an overview of the current state of research. Section 3 describes the methodology used and describes the sample. Section 4 presents and discusses the results. This study concludes with a summary.

2 Literature Review and Current State of Research

Two different methods dominate the field of sentiment analysis in empirical accounting and capital market research. Early research used test subjects or the researchers' own subjective assessment to categorize the tone of company publications as positive or negative [11, 16, 19]. Core [5] points out that the work load necessary for this manual approach is too high to evaluate sufficiently large samples, and notes that a computer-assisted evaluation of the texts can help to reduce the workload and to increase the analyses' accuracy and objectivity. Regarding computer-assisted evaluation, two approaches can be identified in the literature, a fully-automated and a semi-automated approach [20]. The statistical approach uses machine learning and is applied among others by Li [21], who uses information from management discussion and analysis sections (MD&A) in 10-K annual and 10-Q quarterly reports. The author evaluates 30,000 randomly selected sentences and subjectively determines if their tone is positive, negative, neutral or uncertain. The words of this training data set are then used to train a naive Bayesian algorithm to determine the largest statistical correlation of the remaining texts with one of the categories. However, research using text analysis prefers a second method, which employs predefined word lists which determine for individual words whether they imply a positive or negative tone. In this method, the number of positive and / or negative words in each text is counted, and it is assumed that the order of the words is irrelevant for the tone [22, 24].

Four different word lists have been established in accounting and capital market research by the following authors: Osgood [30], Hart [13], Henry [14], and Loughran and McDonald [24]. While word lists by Loughran/McDonald and Henry were developed specifically for financial texts, the word lists by Osgood and Hart were originally designed for research in psychology and, respectively, in political communication.

Empirical accounting and capital market research primarily analyzes the tone of 10-K annual financial statements, 10-Q quarterly statements, company press releases, articles in print media, articles on Internet platforms, and other company publications [18]. The sentiment of these publications is used to predict the development of companies, that is, either the price development of their securities or the development of assets, earnings, and financial ratios.

Tetlock [33] based his analysis on the word list by Osgood and showed that a pessimistic tone in press releases results in a negative investor reaction and thus a decrease in stock returns. The press releases were taken from the Wall Street Journal column "Abreast of the Market". A number of other studies have also confirmed Tetlock's results with respect to short-term market responses to tone [1, 3, 8, 9, 26]. In addition, both Solomon [32] and Huang, Teoh and Zhang [17] conclude that this effect is reversed over time: an initial reaction normalizes over time and is therefore to be considered as a market overreaction. Garcia [12] points out that a negative tone in newspaper articles results in a much stronger stock-market reaction when the economy is in recession.

The two profit figures regularly forecasted in empirical studies are the return on assets and the unexpected earnings. The latter figure is calculated as the difference between earnings per share and the average of earnings per share expected by analysts. Tetlock, Saar-Tsechansky and Macskassy [34] come to the conclusion that a negative tone in press releases causes negative earnings in the following quarters. Other studies arrive at similar results based on the tone in 10-K annual financial statements or of social media posts [3, 9]. Loughran and McDonald [22] on the other hand find that a negative tone in 10-K financial statements follows a positive development of unexpected earnings. A distortion of the results due to the applied word list can be ruled out since both results were obtained using the Osgood as well as the Loughran and McDonald word list. Thus, the studies differ mainly due to the analyzed text types. Thus, Loughran and McDonald [22] conclude with reference to the results of Tetlock, Saar-Tsechansky and Macskassy [34] that due to the independency of journalists the tone of newspaper articles is a better indicator of future profits, while managers of companies can use 10-K annual financial statements to pursue impression or expectation management.

Davis, Piger and Sedor [6] show based on Hart's word list that the tone of corporate earnings announcements in the press is a good predictor of the company's return on assets for the next four quarters. Furthermore, they present evidence that the stock market reacts directly to the sentiment of these communications. The authors argue that managers can report their own earnings expectations more subtly by tone choice than by simply presenting figures. Using tone, managers provide the capital market with important signals.

Davis and Tama-Sweet [7] investigate 10-K and 10-Q quarterly statements in addition to press releases and confirm Davis, Piger and Sedor's results based on their expanded sample. In addition, the authors find that managers' tone choice is with significant frequency more optimistic and less pessimistic in earnings announcements in the press than in the MD&A sections of financial statements. The authors conclude that managers use the freedom afforded to them by press releases to strategically influence the stock market.

In their analysis, Ferris, Hao and Liao [10] evaluate securities prospectuses of companies in the technology sector that are about to go public. They show a forecast of the return on assets for a period of three years into the future to be possible with this information. Huang, Teoh and Zhang present opposite findings for the predictive capacity of press sentiment for the return on assets. According to their results, the return on assets shows a negative sign for three years if the press sentiment was previously positive. Davis, Piger and Sedor [6] and Davis and Tama-Sweet [7] use the Hart word list in their investigations, whereas Huang, Teoh and Zhang [17] apply the other three word lists.

Against the background of the inconsistency in the results of previous research, the present analysis aims at investigating the predictive power of the sentiment in 10-K and 10-Q financial statements for the return on assets on the basis of a representative sample. For this purpose, two different word lists were used to assure that the results are not influenced by word-list choice. The lists by Loughran/McDonald and Henry were applied since they were designed for the analysis of financial texts.

3 Sample and Method

In order to analyze the predictivity of the sentiment in financial reports, a sentiment analysis based on English-language 10-K annually and 10-Q quarterly US financial statements was conducted and the predictivity of the sentiment on the company's performance in terms of return on assets (ROA) was examined. The analysis was based on the 10-K and 10-Q filings of all companies listed in the Standard & Poor's 500 Index (S&P 500) as of March 31, 2016, i.e., the 500 largest listed companies in the United States. For these companies, data from 2005 to 2015 were retrieved. The data were obtained from the Securities and Exchange Commission's (SEC) Electronic File Gathering, Analysis and Retrieval System (EDGAR). The initial sample contained a total of 20,435 annually and quarterly financial statements. The size of the sample was reduced by 1,045 to 19,390 documents as not all relevant financial data pertaining to an annual report were available, which was partly due to incomplete financial years or quarters. Associated financial data were retrieved from the Thomson Reuters Eikon database and randomly tested using annual report data.

The preparation of the texts for processing was based on the so-called parsing procedure by Loughran and McDonald [27, 28]. Since the documents were partly HTML documents, HTML tags were removed [24]. Likewise, tables and figures were removed with the exception of text tables.[1] Furthermore, the texts were cleared of frequently recurring words and words not related to the content (stopwords) [27]. In a final step, numbers, special characters, and single letters were removed.

As mentioned above, this study used the word list by Loughran and McDonald and – to countercheck the results – the much shorter list by Henry. The Loughran and McDonald word list contains a total of six word categories for positive, negative, uncertain, litigious, strongly modal, and weakly modal expressions [22]. This study uses only the 355 positive and 2,355 negative words of this list. The word list by Henry contains a total of 105 positive and 85 negative words [14]. Loughran and McDon-

[1] As in Loughran and McDonald [27, 28], tables embedded in HTML Code are analyzed to ascertain if the proportion of numerals contained in them is greater than 15% of the total of all numerals and letters. If this is the case, the entire table, including its content, is deleted since it is assumed that it only presents financial-reporting figures and, therefore, is irrelevant for textual analysis.

ald [24] argue that managers have far more ways of implying positive or negative sentiment than can be captured by Henry's word list. In particular, Henry's list lacks entries such as loss, losses, adverse and impairment, all of which are considered negative in the corporate context.

The relevance of a positive or negative word for a certain text can be determined unweighted, that is, by ordinary counting of the words, or by so-called term frequency – inverse document frequency (tf-idf) weighting. The latter attributes more influence to a positive or negative word if it occurs frequently in the document (term frequency, tf), but lessens its influence if the word occurs in many documents and, as a result, appears "ordinary" (inverse document frequency, idf). The tf-idf weighting logic used in this study follows Chrisholm and Kolda [4]. The word frequency $tf_{w,d}$ is calculated as follows [4, 22, 24]:

$$tf_{w,d} = \frac{\left(1+log_2(tf'_{w,d})\right)}{\left(1+log_2(a_d)\right)} \tag{1}$$

Here, $tf'_{w,d}$ reflects the frequency of word w in document d, while a_d represents the average word count in document d. The inverse document frequency idf_w is calculated based on the following equation:

$$idf_w = log_2\left(\frac{N}{df_w}\right) \tag{2}$$

N represents the total number of documents and df_w the number of documents in which the word w occurs at least once. The tf-idf weight results from the multiplication of both terms, $tf_{w,d} \cdot idf_w$. The resulting weight measure is then used as the count for the respective word, while in the unweighted approach each occurrence of a word is counted equally.

The tf-idf weighting has been criticized by Henry and Leone [15] because its results are contingent upon the totality of documents used in a sample, and, thus, the result for a particular document can vary significantly depending on the other documents included in the sample. The authors showed that the weighting does not lead to improved results, thus refuting the findings of Loughran and McDonald [22], who observed a significant improvement in the results if tf-idf weighting was applied. Due to contradictory results, this study applied both methods, the unweighted count and the tf-idf weighting, in parallel and tests them against each other.

Since for every company i only one 10-K annual or 10-quarter financial statement is available in each quarter t, the document index d can clearly be replaced by the index i,t. To determine the sentiment of each quarter or annual document, the sentiment indicator $SENT_{i,t}$ or, respectively, $SENT_d$ can be calculated according to the following equation:

$$SENT_{i,t} = \frac{tf'_{pos,d} - tf'_{neg,d}}{tf'_{pos,d} + tf'_{neg,d}} = \frac{tf'_{pos,i,t} - tf'_{neg,i,t}}{tf'_{pos,i,t} + tf'_{neg,i,t}} \tag{3}$$

$tf'_{pos,d}$ denotes the number of weighted or, respectively, unweighted positive words in the document d, $tf'_{neg,d}$ denotes the number of weighted or, respectively, unweighted negative words [15]. Accordingly, the sentiment indicator can have a value between 1 for a perfectly positive sentiment and -1 for a perfectly negative sentiment.

This study analyzed the predictivity of the sentiment for the return on assets; therefore, the return on assets $ROA_{i,t+n}$ with a lag of n quarters was considered as a dependent variable in the econometrics. The return on assets $ROA_{i,t+n}$ was calculated based on earnings before interest and taxes (EBIT) and the average assets of the observed quarter.

The control variables used in this study were based on the relevant literature [6, 7, 17]. The natural logarithm of the market capitalization at the time t was used as an indicator for the size of the company $SIZE_{i,t}$. The correction for company size is necessitated due to the small firm effect according to Banz [2]. In order to control for expected growth as well as other key balance sheet figures and company news, the market to book ratio $MBR_{i,t}$ was used as an indicator that incorporates a variety of publicly known information and news. To take the leverage effect and the findings of Modigliani and Miller [29] into account, the leverage ratio $LEV_{i,t}$ is also used as a control variable in the model. Since the return on assets $ROA_{i,t}$ is a good indicator of the return on assets in the following quarters $ROA_{i,t+n}$, it is also used a control variable. Finally, the Boolean variables $DLOSS_{i,t}$ for loss quarters and $D10K_{i,t}$ for annual reports are integrated into the model. They compensate for the effects of differing company responses in loss quarters as well as the potential differences in the predictivity of annual financial statements compared to quarterly financial statements. To avoid outlier-induced results, all control variables were winsorized at the first and 99th percentile. For an overview of the variables used please refer to Table 1.

Table 1. Variables and data sources

	Variable	Expression	Thompson Reuters Codes
Explained Variable	$ROA_{i,t}$	$\frac{EBIT_{i,t}}{Total\ Assets_{i,t} - Total\ Assets_{i,t-1}}$	TR.EBIT TR.TotalAssetsReported
Explanatory Variable	$SENT_{i,t}$	$\frac{tf'_{pos,i,t} - tf'_{neg,i,t}}{tf'_{pos,i,t} + tf'_{neg,i,t}}$	
Control Variables	$SIZE_{i,t}$	$\ln(MarketCap_{i,t})$	TR.CompanyMarketCap
	$MBR_{i,t}$	$\frac{MarketCap_{i,t}}{Total\ Equity_{i,t}}$	TR.CompanyMarketCap TR.TotalEquity

	$LEV_{i,t}$	$\frac{Total\ Liabilities_{i,t}}{Total\ Assets_{i,t}}$	TR.TotalLiabilities TR.TotalAssetsReported
	$DLOSS_{i,t}$	1, if Net $Income_{i,t} < 0$ 0, else	TR.NetIncome
	$D10K_{i,t}$	1, if 10 K-Filing 0, if 10 Q-Filing	

In the analysis, the word lists by Loughran and McDonald and Henry were each used both with and without tf-idf weighting to predict the return on assets for one and two quarters. This results in four explanatory variables. While the list of Loughran and McDonald returns negative sentiment indicators in most cases, the list of Henry returns almost positive sentiments. This effect can be attributed to the strong deviation of the two lists. All other variables show inconspicuous behavior.

As in the analysis, the word lists by Loughran and McDonald and Henry were each used both with and without tf-idf weighting to predict the return on assets for one and two quarters, this results in a total of eight variants to be calculated. The corresponding pooled OLS, fixed-effects, and random-effects panel data models were calculated for all variants. However, the F- and Hausmann specification tests used to compare the models show the superiority of the fixed-effects model with fixed-effects for each company α_i and every quarter γ_t; the model is applied according to the following equation:

$$ROA_{i,t+n} = \alpha_i + \gamma_t + \beta_1 \cdot SENT_{i,t} + \beta_2 \cdot SIZE_{i,t} + \beta_3 \cdot MBR_{i,t} + \beta_4 \cdot LEV_{i,t} + \beta_5 \cdot ROA_{i,t} + \beta_6 \cdot DLOSS_{i,t} + \beta_7 \cdot D10K_{i,t} + \varepsilon_{i,t} \quad (3)$$

In addition to the panel data models, all variants were calculated by means of a robust least absolute deviations (LAD) regression in order to rule out outlier-induced results. However, the panel data models and LAD regression were not significantly different from each other; thus, the validity of the panel data models can be assumed. Multicollinearity problems were excluded by the analysis of the variance inflation factors that all showed numbers below three. All eight variants were tested for heteroscedasticity using the Breusch-Pagan test. All variants show heteroskedasticity problems. Therefore, the results shown in Table 1 include the heteroskedasticity-robust standard errors according to White (HC 0) and the significance values calculated on their basis. Endogeneity problems were precluded due to the time lag between the explained variable and explanatory variables and, therefore, no separate consideration is required.

4 Results and Discussion

The descriptive statistics and the correlation matrix are presented in Table 2 and Table 3 respectively. The results of the analysis are summarized in Table 4 for one quarter (Lag 1) and Table 5 for two quarters (Lag 2). For the purpose of clarity, individual fix effects, which do not show any abnormalities, have been omitted. The same applies to the fixed quarterly effects, which, albeit particularly significant in financial crisis quarters, do not show abnormalities.

All models show significant model statistics, which implies a high prediction character. The models with Lag 1 have an explained variance R^2 of – depending on the model – around 20%, which can be considered as high in this type of models. The Lag 2 models are naturally weaker in terms of the explained variance as forecasts for the distant future are generally associated with weaker prediction qualities. Accordingly, experiments with higher lags only sporadically show significant models, which is why they are not presented here.

The sentiment indicator has a significant positive impact in all tested models. Therefore, it can be used as a predictor for future corporate development forecasts. It is important to underline that in the present analysis the sentiment indicator was corrected for the influence of several significant indicators. In particular, the market-to-book ratio already contains the aggregated information of a large number of balance sheet ratios and company news, which are incorporated in the price by the market. The sentiment indicator thus contains incremental information beyond the information imparted by balance sheet figures and thereby provides additional forecasting power. This implies that companies use texts to impart information about positive or negative developments which cannot yet be expressed in figures.

For sentiment analysis, it does not seem to matter whether the extensive word list by Loughran and McDonald or the much shorter list by Henry is used. The quality measures of both word lists seem to differ only marginally. It also appears to be irrelevant if the sentiment indicator is determined by mere counting or by applying the tf-idf weighting. The advantages of tf-idf weighting in other areas of text analysis, such as the creation of word clouds for faster collection of content, can hardly be transferred to sentiment analysis. Although the method is capable of identifying relevant words in documents, this advantage is mainly limited to nouns that are frequent in individual documents but rare in other documents. Adjectives and adverbs, which often carry a positive or negative connotation in the corporate publications in question, tend to be underweighted by the method because they are commonplace in everyday language.

Table 2. Descriptive Statistics

Explanatory Variables

	Loughran/McDonald (tf-idf)			Loughran/McDonald (unweighted)		
	positive	**negative**	**$SENTI_{i,t}$**	**positive**	**negative**	**$SENTI_{i,t}$**
Min.	-	-	-0.864	-	-	-0.888
1st Qu.	169.400	371.900	-0.511	80.000	172.000	-0.518
Median	236.700	602.800	-0.426	133.000	341.000	-0.405
Mean	269.200	688.900	-0.405	188.100	487.700	-0.390
3rd Qu.	350.300	945.200	-0.319	248.000	647.000	-0.278
Max.	880.200	2,506.300	0.567	2,620.000	5,347.000	0.619

	Henry (tf-idf)			Henry (unweighted)		
	positive	**negative**	**$SENTI_{i,t}$**	**positive**	**negative**	**$SENTI_{i,t}$**
Min.	-	-	-0.745		-	-0.862
1st Qu.	243.300	178.500	0.085	163.000	119.000	0.025
Median	305.100	229.100	0.154	251.000	192.000	0.136
Mean	326.000	237.300	0.154	336.700	266.600	0.133
3rd Qu.	402.500	293.400	0.224	433.000	331.500	0.245
Max.	745.400	522.900	0.688	3,216.000	2,666.000	0.843

Control-Variables

	$SIZE_{i,t}$	**$MBR_{i,t}$**	**$LEV_{i,t}$**	**$ROA_{i,t}$**	**$DLOSS_{i,t}$**	**$D10K_{i,t}$**
Min.	18.920	0.501	0.129	-0.019	-	-
1st Qu.	22.540	1.606	0.471	0.013	-	-
Median	23.190	2.595	0.605	0.024	-	-
Mean	23.300	3.649	0.600	0.028	0.077	0.238
3rd Qu.	23.930	4.215	0.741	0.038	-	-
Max.	27.310	25.220	0.959	0.100	1.000	1.000

Table 3. Correlation Matrix

		Explained Variables		Explanatory Variables ($SENT_{i,t}$)				Control Variables					
		$ROA_{i,t+1}$	**$ROA_{i,t+2}$**	**LMD (tf.idf)**	**LMD**	**Henry (tf-idf)**	**Henry**	**$SIZE_{i,t}$**	**$MBR_{i,t}$**	**$LEV_{i,t}$**	**$ROA_{i,t}$**	**$DLOSS_{i,t}$**	**$D10K_{i,t}$**
Explained Variables	**$ROA_{i,t+1}$**	1.000											
	$ROA_{i,t+2}$	0.706	1.000										
Explanatory Variables ($SENT_{i,t}$)	**LMD (tf.idf)**	0.133	0.126	1.000									
	LMD	0.127	0.122	0.909	1.000								
	Henry (tf-idf)	0.164	0.170	0.523	0.466	1.000							
	Henry	0.260	0.252	0.408	0.377	0.771	1.000						
Control Variables	**$SIZE_{i,t}$**	0.108	0.098	-0.061	-0.058	-0.001	0.066	1.000					
	$MBR_{i,t}$	0.368	0.368	0.115	0.106	0.116	0.203	0.094	1.000				
	$LEV_{i,t}$	-0.337	-0.338	-0.076	-0.061	-0.132	-0.210	0.034	0.104	1.000			
	$ROA_{i,t}$	0.732	0.654	0.160	0.158	0.192	0.297	0.117	0.402	-0.399	1.000		
	$DLOSS_{i,t}$	-0.241	-0.183	-0.095	-0.117	-0.101	-0.135	-0.145	-0.010	0.094	-0.323	1.000	
	$D10K_{i,t}$	-0.038	0.016	0.002	0.060	0.230	0.045	-0.006	-0.006	-0.002	0.006	0.058	1.000

Table 4. Results of fixed-effects panel data models with heteroskedasticity-robust standard errors (White's HC 0) for lag 1

	Loughran/McDonald (tf-idf)					Loughran/McDonald (unweighted)				
	Coeff.	Std.Error	t-Value	p-Value		Coeff.	Std.Error	t-Value	p-Value	
$SENT_{i,t}$	0.005	0.002	2.743	0.006	**	0.006	0.001	3.840	0.000	***
$SIZE_{i,t}$	0.005	0.001	4.374	0.000	***	0.005	0.001	4.339	0.000	***
$MBR_{i,t}$	0.001	0.000	5.936	0.000	***	0.001	0.000	5.924	0.000	***
$LEV_{i,t}$	-0.006	0.005	-1.202	0.229		-0.006	0.005	-1.162	0.245	
$ROA_{i,t}$	0.403	0.049	8.272	0.000	***	0.402	0.049	8.240	0.000	***
$DLOSS_{i,t}$	-0.002	0.001	-2.647	0.008	**	-0.002	0.001	-2.533	0.011	*
$D10K_{i,t}$	-0.004	0.002	-1.895	0.058	.	-0.004	0.002	-1.947	0.052	.
Observations			19,390					19,390		
R^2			0.198					0.199		
Adjusted R^2			0.175					0.176		
F Statistic (df = 51; 18844)			91.2	0.000	***				0.000	***
F-Test OLS vs. FE (F) (df = 494; 18844)			8.5	0.000	***			8.5	0.000	***
Hausmann FE vs. RE (df = 51)			2,655.9	0.000	***			2,735.6	0.000	***
Breusch-Pagan (BP) (df = 51)			19,563.0	0.000	***			19,490.0	0.000	***

	Henry (tf-idf)					Henry (unweighted)				
	Coeff.	Std.Error	t-Value	p-Value		Coeff.	Std.Error	t-Value	p-Value	
$SENT_{i,t}$	0.011	0.002	4.631	0.000	***	0.006	0.002	3.729	0.000	***
$SIZE_{i,t}$	0.004	0.001	4.137	0.000	***	0.004	0.001	4.138	0.000	***
$MBR_{i,t}$	0.001	0.000	6.048	0.000	***	0.001	0.000	5.981	0.000	***
$LEV_{i,t}$	-0.007	0.005	-1.345	0.179		-0.007	0.005	-1.348	0.178	
$ROA_{i,t}$	0.399	0.048	8.246	0.000	***	0.399	0.049	8.219	0.000	***
$DLOSS_{i,t}$	-0.002	0.001	-2.683	0.007	**	-0.002	0.001	-2.699	0.007	**
$D10K_{i,t}$	-0.005	0.002	-2.220	0.026	*	-0.004	0.002	-1.937	0.053	.
Observations			19,390					19,390		
R^2			0.200					0.199		
Adjusted R^2			0.177					0.176		
F Statistic (df = 51; 18844)			92.2	0.000	***			91.7	0.000	***
F-Test OLS vs. FE (F) (df = 494; 18844)			8.5	0.000	***			8.4	0.000	***
Hausmann FE vs. RE (df = 51)			2,706.5	0.000	***			2,694.8	0.000	***
Breusch-Pagan (BP) (df = 51)			19,559.0	0.000	***			19,449.0	0.000	***

Table 5. Results of fixed-effects panel data models with heteroskedasticity-robust standard errors (White's HC 0) for lag 2

	Loughran/McDonald (tf-idf)					Loughran/McDonald (unweighted)				
	Coeff.	**Std.Error**	**t-Value**	**p-Value**		**Coeff.**	**Std.Error**	**t-Value**	**p-Value**	
$SENT_{i,t}$	0.005	0.002	2.630	0.009	**	0.006	0.002	3.821	0.000	***
$SIZE_{i,t}$	0.005	0.001	3.893	0.000	***	0.005	0.001	3.857	0.000	***
$MBR_{i,t}$	0.002	0.000	5.971	0.000	***	0.002	0.000	5.957	0.000	***
$LEV_{i,t}$	-0.020	0.006	-3.451	0.001	***	-0.020	0.006	-3.408	0.001	***
$ROA_{i,t}$	0.166	0.081	2.054	0.040	*	0.165	0.081	2.031	0.042	*
$DLOSS_{i,t}$	-0.000	0.001	-0.063	0.950		0.000	0.001	0.018	0.985	
$D10K_{i,t}$	0.001	0.001	1.444	0.149		0.001	0.001	1.311	0.190	
Observations			19,390					19,390		
R^2			0.107					0.107		
Adjusted R^2			0.081					0.082		
F Statistic (F) (df = 51; 18844)			44.1	0.000	***			44.5	0.000	***
F-Test OLS vs. FE (F) (df = 494; 18844)			11.6	0.000	***			11.6	0.000	***
Hausmann FE vs. RE (df = 51)			5,650.7	0.000	***			5,705.0	0.000	***
Breusch-Pagan (BP) (df = 51)			21,713.0	0.000	***			21,677.0	0.000	***

	Henry (tf-idf)					Henry (unweighted)				
	Coeff.	**Std.Error**	**t-Value**	**p-Value**		**Coeff.**	**Std.Error**	**t-Value**	**p-Value**	
$SENT_{i,t}$	0.015	0.003	5.264	0.000	***	0.009	0.002	4.802	0.000	***
$SIZE_{i,t}$	0.004	0.001	3.583	0.000	***	0.004	0.001	3.538	0.000	***
$MBR_{i,t}$	0.002	0.000	6.066	0.000	***	0.002	0.000	6.017	0.000	***
$LEV_{i,t}$	-0.021	0.006	-3.602	0.000	***	-0.021	0.006	-3.620	0.000	***
$ROA_{i,t}$	0.161	0.081	1.985	0.047	*	0.159	0.081	1.966	0.049	*
$DLOSS_{i,t}$	-0.000	0.001	-0.025	0.980		-0.000	0.001	-0.036	0.971	
$D10K_{i,t}$	0.000	0.001	0.401	0.688		0.001	0.001	1.269	0.204	
Observations			19,390					19,390		
R^2			0.110					0.110		
Adjusted R^2			0.084					0.084		
F Statistic (F) (df = 51; 18844)			45.7	0.000	***				0.000	***
F-Test OLS vs. FE (F) (df = 494; 18844)			11.7	0.000	***			11.6	0.000	***
Hausmann FE vs. RE (df = 51)			5,689.3	0.000	***			5,683.1	0.000	***
Breusch-Pagan (BP) (df = 51)			21,573.0	0.000	***			21,503.0	0.000	***

5 Conclusions

The present study demonstrated the predictive power of the sentiment of 10-K annual and 10-Q quarterly statements for future corporate development. The sentiment indicator has significant incremental prognostic quality both for the following quarter and the subsequent quarter. If the forecast horizon was three quarters or more, only isolated cases of significant predictiveness were detected. As a result, analysts and investors should include the sentiment of corporate publications into their analyses to gather latent information from company and detect subtle signals from management that is of value in their decision-making. Corporate publications contain valuable "between the lines" information which may be relevant to the assessment of potential opportunities and risks. Accordingly, a professionalized evaluation of textual data can provide an information advantage.

The present study drew on two very different word lists, the Loughran and McDonald and the Henry word lists, which are both geared toward financial texts. Both were used in conventional word-counting as well as tf-idf-weighting approaches. However, the quality of the word lists used and the weighting approach taken influenced forecasting quality to a much lesser extent than suggested by the existing literature. The calculated models were almost identical in terms of quality.

References

1. Allee, K. D., DeAngelis, M. D.: The Structure of Voluntary Disclosure Narratives: Evidence from Tone Dispersion. Journal of Accounting Research, 53(2), 241-274 (2015).
2. Banz, R.: The relationship between return and market value of common stocks. Journal of Financial Economics, 9 (1981), 3-18.
3. Chen, H., De, P., Hu, Y., Hwang, B.: Wisdom of Crowds: The Value of Stock Opinions Transmitted Through Social Media. Review of Financial Studies, 27(5), 1367-1403 (2014).
4. Chrisholm, E., Kolda, T. G.: New Term Weighting Formulas for the Vector Space Model in Information Retrieval. 1999. Oak Ridge, Tennessee.
5. Core, J. E.: A review of the empirical literature discussion. Journal of Accounting and Economics, 31, 441-456 (2001).
6. Davis, A. K., Piger, J. M., Sedor, L. M.: Beyond the Numbers: Measuring the Information Content of Earnings Press Release Language. Contemporary Accounting Research, 29(3), 845-868 (2012).
7. Davis, A. K., Tama-Sweet, I.: Managers’ Use of Language Across Alternative Disclosure Outlets: Earnings Press Releases versus MD&A. Contemporary Accounting Research, 29(3), 804–837 (2012).
8. Dougal, C., Engelberg, J., Garcia, D., Parsons, C. A.: Journalists and the stock market. Review of Financial Studies, 25(3), 639-679 (2012).
9. Feldman, R., Govindaraj, S., Livnat, J., Segal, B.: Management’s tone change, post earnings announcement drift and accruals. Review of Accounting Studies, 15(9), 915-953 (2010).

10. Ferris, S. P., Hao, Q., Liao, M.: The Effect of Issuer Conservatism on IPO Pricing and Performance. Review of Finance, 17, 993-1027 (2010).
11. Francis, J., Philbrick, D., Schipper, K.: Shareholder Litigation and Corporate Disclosures. Journal of Accounting Research, 32(2), 137-164 (1994).
12. Garcia, D.: Sentiment during Recessions. The Journal of Finance, 68(3), 1267-1300 (2012).
13. Hart, R.P.: Redeveloping DICTION: Theoretical Considerations. In: West, M. D. (Eds.): Theory, method, and practice in computer content analysis, New York, 43-60.
14. Henry, E.: Are Investors Influenced by How Earnings Press Releases are Written? Journal of Business Communication, 45(4), 363-407 (2008).
15. Henry, E., Leone, A. J.: Measuring Qualitative Information in Capital Markets Research: Comparison of Alternative Methodologies to Measure Disclosure Tone. The Accounting Review, 91(1), 153-178 (2016).
16. Hoskin, R. E., Hughes, J. S., Ricks, W. E.: Evidence on the Incremental Information Content of Additional Firm Disclosures Made Concurrently with Earnings. Journal of Accounting Research, 24, 1-32 (1986).
17. Huang, X., Teoh, S. H., Zhang, Y.: Tone Management. The Accounting Review, 89(3), 1083-1113 (2014).
18. Kearney, C., Liu, S.: Textual sentiment in finance: A survey of methods and models. International Review of Financial Analysis, 33, 171-185 (2014).
19. Lang, M. H., Lundholm, R. J.: Voluntary Disclosure and Equity Offerings: Reducing Infor-mation Asymmetry or Hyping the Stock?. Contemporary Accounting Research, 17(4), 623-662 (2000).
20. Li, F.: Textual Analysis of Corporate Disclosures: A Survey of the Literature. Journal of Accounting Literature, 29, 143-165 (2010).
21. Li, F.: The Information Content of Forward-Looking Statements in Corporate Filings – A Naïve Bayesian Machine Learning Approach. Journal of Accounting Research, 48(5), 1049-1102 (2010).
22. Loughran, T., McDonald, B.: When Is a Liability Not a Liability? Textual Analysis, Diction-aries, and 10-Ks. The Journal of Finance, 66(1), 35-65 (2011).
23. Loughran, T., McDonald, B.: 2014 Master Dictionary. 03.2015. As of: 01.11.2017. URL: https://www3.nd.edu/~mcdonald/Word_Lists.html
24. Loughran, T., McDonald, B.: Textual Analysis in Accounting and Finance: A Survey. Journal of Accounting Research, 54(4), 1187-1230 (2016).
25. Mayew, W. J., Sethuraman, M., Venkatachalam, M.: MD&A Disclosure and the Firm's Ability to Continue as a Going Concern. The Accounting Review, 90(4), 1621-1651 (2015).
26. Mayew, W. J., Venkatachalam, M.: The Power of Voice: Managerial Affective States and Future Firm Performance. The Journal of Finance, 67(1), 1-43 (2012).
27. McDonald, B.: Documentation for Stage One Parse. Last updated: 03.09.2013. As of: 13.08.2017, URL: https://www3.nd.edu/~mcdonald/Word_Lists.html.
28. McDonald, B.: Documentation of Master Dictionary and Document Dictionaries. Last updated: 18.03.2014. As of: 13.08.2017, URL: https://www3.nd.edu/~mcdonald/Word_Lists.html.
29. Modigliani, F., Miller, M.: Corporate Income Taxes and the Cost of Capital: A Correction. The American Economic Review, 53(1), 433-443 (1963).
30. N.N.:Descriptions of Inquirer Categories and Use of Inquirer Dictionaries. As of 01.11.2017. URL: http://www.wjh.harvard.edu/~inquirer/homecat.htm.

31. Price, S. M., Doran, J. S., Peterson, D. R., Bliss, B. A.: Earnings conference calls and stock returns: The incremental informativeness of textual tone. Journal of Banking & Finance, 36, 992-1011 (2012).
32. Solomon, D. H.: Selective Publicity and Stock Prices. The Journal of Finance, 67(2), 599-637 (2012).
33. Tetlock, P. C.: Giving Content to Investor Sentiment: The Role of Media in the Stock Market. The Journal of Finance, 62(3), 1139-1168 (2007).
34. Tetlock, P. C., Saar-Tsechansky, M., Macskassy, S.: More Than Words: Quantifying Language to Measure Firms' Fundamentals. The Journal of Finance, 63(3), 1437-1467 (2008).
35. Twedt, B., Rees, L: Reading between the lines: An empirical examination of qualitative attributes of financial analysts' reports. Journal of Accounting and Public Policy, 31, 1-21 (2012).