Not all words are equal: Sentiment dynamics and information content within CEO letters ☆

Kris Boudt[a,b,1], James Thewissen[a,c,2]

January 15, 2015

ABSTRACT

CEO letters should be perceived as dynamic objects with a common pattern in the intratextual linguistic sentiment and with the position of a word in the text influencing its informativeness for future firm performance. Using a sample of DJIA firms between 2000 and 2011, we identify a U-shaped observed frequency of positive words within a text and a decreasing smirk in the intratextual number of negative words. It follows that the difference (called the net sentiment) shows a smirk with more positive than negative words overall and a sharply increasing prevalence of positive words towards the end of the letter. Based on these stylized facts, we use a novel weighting scheme to aggregate the within-text net sentiment dynamics into a single proxy for the CEO's sentiment and demonstrate that it has more predictive power for the firm's performance over the next year than the equally-weighted measures of net sentiment used in the prior literature. The key findings of this paper are that the location of a word in CEO letters contains information value and that the intratextual analysis of CEO sentiment increases the prediction accuracy of future firm performance.

KEYWORDS: CEO sentiment, Firm profitability, Intratextual analysis, Sentiment dynamics

AUTHORS INFO

[a] *Solvay Business School, Vrije Universiteit Brussel, Belgium*

[b] *Faculty of Economics and Business, V.U.University of Amsterdam, The Netherlands*

[c] *Accounting, Finance and Insurance, Katholieke Universiteit Leuven, Belgium*

[1] Kris.Boudt@vub.ac.be

[2] James.Thewissen@kuleuven.be

PAPER INFO

JEL CLASSIFICATION: *G14, G17, M40*

☆This research was carried out thanks to financial support of the Intercollegiate Center for Management Science, the Junior Mobility Program of the KULeuven and the Dutch science foundation. We thank Özgür Arslan-Ayaydin, Peter de Goeij, Nitish Ranjan Sinha, Richard Taffler, Geert Van Campenhout and Cynthia Van Hulle for their valuable comments, as well as the participants at various seminars and conferences. This paper won the 2014 Centro Stefano Franscini Award for best paper at the 2014 International Conference on Discourse Approaches to Financial Communication, Switzerland, and the Award for best paper at the 2014 R/Finance conference, Chicago.
Corresponding author: James Thewissen, I.C.M. Fellow, Katholieke Universiteit Leuven, Korte Nieuwstraat 33, 2000 Antwerpen, Belgium.

1 Introduction

There seems to exist a kind of periodical oscillation of mental receptivity or attention. (Ebbinghaus, 1885)

The serial position effect is the tendency of a person to recall the first and last items of a series best, and the middle items worst. This cognitive effect, documented first by Ebbinghaus (1885), is hard-coded in human genes and is often cited as one of the reasons a good speech ends with a strong conclusion. It also implies that the sentiment in a corporate disclosure should not be uniformly distributed over the text and that the informativeness of intratextual sentiment may depend on its position within the text. These patterns are particularly relevant for the study of the predictive power of qualitative information in financial disclosures for forecasting future firm performance.

We find that the panel of letters to shareholders written by the Chief Executive Officers (CEOs) of Dow Jones Industrial Average (DJIA) constituents over the period 2000-2011 share a common pattern in the intratextual use of positive and negative words. These dynamics are synthesized by a U-shape in the intratextual dynamic of CEOs' positive sentiment, with a significantly larger peak at the end of the letter. We argue that this pattern is intuitive and in line with theories from the narratology and computational linguistic literature as well as the empirical evidence on impression management by CEOs (Clatworthy and Jones, 2003; Patelli and Pedrini, 2013) and CEO overconfidence (Heaton, 2002).

One of the major consequences of the non-uniform distribution of the intratextual number of positive and negative words is that total sentiment measures that aggregate the intratextual sentiment without considering the position in the text may be suboptimal. In fact, today's workhorse in estimating sentiment consists of a simple spread between the percentage of words that can be classified as positive and those that can be classified as negative. Prior research has shown that this measure has predictive value that is incremental to "hard", quantifiable information (Abrahamson and Amir, 1996; Davis, Piger, and Sedor, 2012; Li, 2010; Patelli and Pedrini, 2013; Price, Doran, Peterson, and Bliss, 2012).

We propose an improved method to measure the sentiment of a text, defined as the author's positivity or negativity (Henry, 2008). The key change is that we consider the words in CEO letters to be unequal and to have information value that depends on their position within the letter. We define a weighted measure of sentiment, where the weights are a function of the position of a word in a text. The weights are optimized to maximize the sentiment measure's predictive power for the firm's future return on assets. To avoid overfitting, the weights are parsimoniously specified as a linear combination of third-order Almon polynomials that are smooth functions of the word's position in the text (Almon, 1965).

We then evaluate the gains of optimizing the intratextual sentiment weights for forecasting future performance of the DJIA constituents between 2000 and 2011. We find an economically and statistically significant increase in the R-square of the prediction model relative to the classical approach used in the prior literature. This result indicates that the structure of the sentiment within CEO letters provides a signal to investors concerning future performance and that an intratextual analysis is required to accurately measure CEO sentiment within the CEO letter. The sentiment measures defined in the literature (Abrahamson and Amir, 1996; Patelli and Pedrini, 2013) are thus potentially inefficient, as they assume all words to be equal, irrespective of their position within the document. This result is robust to controlling for 'hard' information. In fact, we find that the equally-weighted measure becomes insignificant after controlling for hard information, whereas our optimized measure of sentiment remains an important earnings-predicting variable.

To summarize, our main contributions are threefold. First, we show the existence of a common intratextual pattern of sentiment in CEO letters. Second, we develop a more efficient sentiment aggregation method to predict future firm performance. Third, we show that the proposed sentiment is effectively a better predictor of future firm performance than the traditional equally-weighted measure of sentiment.

We proceed as follows. Section 2 first sets our motivation and develops our hypotheses. Section 3 describes our sample as well as the word libraries used. Section 4 shows the presence of a common pattern in CEO sentiment dynamics within annual letters. Section 5 explains our new weighted measure of sentiment. Section 6 contains the main analysis of the forecast performance of the sentiment measure for future firm performance. Section 7 presents the conclusions and sketches directions for further research.

2 Literature Review

The last decade has witnessed a substantial increase in the number of studies on the informational value of corporate textual disclosures. The contributions of our research are at the tangency of the literature on estimating the sentiment of a text, the narratology and computational linguistic literature and the study of the information content of CEO letters. We first review these strands of literature and then formulate our research hypotheses.

2.1 Measuring linguistic sentiment

The building block of a sentiment measure is the qualification of words as positive, negative or neutral. This is usually performed by a content analysis that verifies whether the words belong to a pre-specified list of positive and negative words, called a dictionary. The numbers of positive and negative words are then aggregated into a manageable metric for further analysis. For both the method of sentiment identification at the individual word level and the aggregation at the text level, different approaches exist. However, no universal approach that fits all purposes has arisen to date.

First, regarding the choice of the library used, most of the early research uses general lists of words, such as the Diction software program, that automatically generates a score of a document's optimism. These libraries were built for the study of sociological and psychological text and may not be suitable for the content analysis of corporate disclosures. The current trend in text analysis research is to refer to domain-specific dictionaries. For the analysis of CEO letters, this implies the use of specialized financial dictionaries, such as those developed by Loughran and McDonald (2011) and Henry (2008). Although domain-specific libraries are progressively being defined, a consensus has yet to be reached as to which library to use. This explains why

Table 1: **Overview of sentiment metrics used in prior literature**

Author(s)	Dictionary	Sentiment	Metric	All words equal?
Abrahamson and Amir (1996)	CEO letters library	Negative	$NegSent_{j,t} = \frac{1}{B}\sum_{b=1}^{B} NegSent_{b,j,t}/TW_{b,j,t}$	Yes
Tetlock et al. (2008)	General Inquirer	Negative	$NegSent_{j,t} = (NW_{j,t} - \mu_{NegSent_{j,t-1}})/\sigma_{NegSent_{j,t-1}}$	Yes
Henry (2008)	Diction 5.0	Positive Negative Net	$PosSent_{j,t} = \frac{1}{B}\sum_{b=1}^{B} NW_{b,j,t}$ $NegSent_{j,t} = \frac{1}{B}\sum_{b=1}^{B} PW_{b,j,t}$ $NetSent_{j,t} = \frac{1}{B}\left(\frac{PosSent_{b,j,t}-NegSent_{b,j,t}}{PosSent_{b,j,t}+NegSent_{b,j,t}}\right)$	Yes
Sadique, In, and Veeraraghavan (2008)	Henry list	Positive Negative Net	$PosSent_{j,t} = \frac{1}{B}\sum_{b=1}^{B} PW_{b,j,t}/TW_{b,j,t}$ $NegSent_{j,t} = \frac{1}{B}\sum_{b=1}^{B} NW_{b,j,t}/TW_{b,j,t}$ $NetSent_{j,t} = \frac{1}{B}\sum_{b=1}^{B} (PosSent_{b,j,t} - NegSent_{b,j,t})$	Yes
Demers and Vega (2010)	a) Diction 6.0 b) General Inquirer c) Loughran and McDonald d) Factor analysis	Positive Negative Net	$PosSent_{j,t} = \frac{1}{B}\sum_{b=1}^{B} PW_{b,j,t}/TW_{b,j,t}$ $NegSent_{j,t} = \frac{1}{B}\sum_{b=1}^{B} PW_{b,j,t}/TW_{b,j,t}$ $NetSent_{j,t} = \frac{1}{B}\sum_{b=1}^{B} (PosSent_{b,j,t} - NegSent_{b,j,t})$	Yes
Davis et al. (2012)	Diction 5.0	Positive Negative Net	$PosSent_{j,t} = \frac{1}{B}\sum_{b=1}^{B} PW_{b,j,t}/TW_{b,j,t}$ $NegSent_{j,t} = \frac{1}{B}\sum_{b=1}^{B} NW_{b,j,t}/TW_{b,j,t}$ $NetSent_{j,t} = \frac{1}{B}\sum_{b=1}^{B} (PosSent_{b,j,t} - NegSent_{b,j,t})$	Yes
Patelli and Pedrini (2013)	Diction 5.0	Positive Negative Net	$PosSent_{b,j,t} = \frac{1}{B}\sum_{b=1}^{B} PW_{b,j,t}/TW_{b,j,t}$ $NegSent_{j,t} = \frac{1}{B}\sum_{b=1}^{B} NW_{b,j,t}/TW_{b,j,t}$ $NetSent_{j,t} = \frac{1}{B}\sum_{b=1}^{B} (PosSent_{b,j,t} - NegSent_{b,j,t})$	Yes
This paper	a) Loughran and McDonald b) Diction 7.0 c) Abrahamson and Amir	Net	$NetSent_{j,t}(w) = \sum_{b=1}^{B} w_b NetSent_{b,j,t}$ where $w_b(\theta) = \theta_1 + \sum_{c=1}^{3} \theta_{1+c} P_c(b/B) + \sum_{c=1}^{3} \theta_{3+c} P_c((B-b)/B)$ and $P_c(u) = (1-u^c)u^{3-c}$ is a c^{th} order Almon polynomial of u	No

Henry (2008) uses Diction 5.0 to define a list of positive and negative words found in earnings press releases, which we refer to as the *Henry list*. $PW_{b,j,t}$ is the number of positive words for firm j in bin b for fiscal year t, $NW_{b,j,t}$ is the number of negative words for firm j in bin b for fiscal year t, $TW_{b,j,t}$ is the total number of words for firm j in bin b for fiscal year t. $\mu_{NegSent_{j,t-1}}$ and $\sigma_{NegSent_{j,t-1}}$ is the average and standard deviation of $NegSent_{j,t-1}$. Refer to Subsection 4.1 for details on the notations.

research usually refers to various multiple lists of words to evidence the robustness of their results.

The second and main component of the sentiment measure is how the metric is constructed. As we show in Table 1, sentiment is sometimes measured as the spread in the proportion of positive and negative words in the document (Davis et al., 2012; Davis and Tama-Sweet, 2012; Demers and Vega, 2010; Patelli and Pedrini, 2013) or simply as the proportion of negative words (Abrahamson and Amir, 1996; Tetlock, Saar-Tsechansky, and Macskassy, 2008). The common element in the metrics used in prior research is that they assume all of the words in a text to be equal, irrespective of their position.

2.2 The value of the position of a word within financial disclosures

The value of the position of a word within a text has been thoroughly investigated in the narratology and computational linguistics literature. We will focus on two generally accepted theories: the serial position effect and peak-end-rule theory.

According to the serial position effect, readers recall information better when it is presented first (primacy) or last (recency) in a vector of words, rather than in the middle (Baddeley and Hitch, 1977; Glanzer and Cunitz, 1966; Roediger and Crowder, 1976). Some studies have examined this issue in prose. One finds that recall of propositions in the text was higher for the first propositions, followed by the last propositions and finally the middle propositions in two of eight passages (Freebody and Anderson, 1986), while another study finds only primacy effects (Frase, 1969). Furthermore, Deese and Kaufman (1957) find both primacy and recency effects and Meyer and McConkie (1973) and Kieras (1980) find that information was recalled better if it appeared early in the text and that this information was more important than other information in the text. This evidence thus suggests that position and information value interact in some way.

This pattern in readers' recall is usually referred to as the U-shaped free-recall curve and is consistent with the position method defined by Edmundson (1969) in computational linguistics. Edmundson (1969) develops automated text summarization techniques to aid readers in accessing information at a faster pace and defines a weight-based method that computes the weight of each sentence based on certain features, such as cue phrase, keyword (i.e., term-frequency-based), title and location. He evaluates each of the criteria by comparison against manually created extracts. He finds that the combination of cue phrases/title/location dominates word frequency measures in the creation of better extracts, with keywords alone being the worst performing algorithms and location being the best individual feature. This research suggests that the relationship between the position of a word in the text and its information value should be considered to optimally measure sentiment in corporate disclosures.

The peak-end-rule theory developed in Varey and Kahneman (1992) predicts that the peak and final event of an experience influences the evaluation more than all other events in the experience, which contradicts a simple hedonic calculus in which years of please and pain are summed or averaged. Experiences that end very well or with a large positive moment are rated as more pleasurable than longer, more moderately pleasant experiences despite the total happiness experienced ostensibly being greater in the longer case (Diener, Wirtz, and Oishi, 2001; Do, Rupert, and Wolford, 2008; Fredrickson and Kahneman, 1993). As a consequence, following the peak-end rule theory, investors reading two sentiment-neutral CEO letters (both with the same number of positive and negative words) have a more positive (negative) assessment of the firm's future performance depending on whether positive words are at the end (beginning) and negative words at the beginning (end) of the letter and whether a large positive (negative) peak occurred in the letter.

Based on the serial position effect and peak-end-rule, we argue that the location of a word in a text matters is important when measuring sentiment within financial disclosures and that a metric of sentiment in which all words are equal is potentially suboptimal. In the next subsection, we explain why we focus on CEO letters and develop our hypotheses.

2.3 CEO annual letters to shareholders

Annual reports discharge managers' obligations to report to shareholders and stakeholders on their management of the business and provide key information for assessing past performance and appraising future opportunities. They have become an important popular medium for communicating company strategies (Diffenbach and Higgins, 1987). A company earns and maintains its credibility by convincing others that it is conducting a sound strategy and has an effective planning capability. If others are to know that a company is well run strategically, both internal and external communications must identify with this objective (Kohut and Segars, 1992). The CEO letter to shareholders occupies a key position to accomplish this objective in annual reports.

CEO letters occupy the strategic first pages in annual reports. The main objective of annual letters to shareholders is to express the CEO's views on the firm's current strategy and performance, as well as how they expect the firm to perform in the future (Bowman, 1984). A CEO letter can thus be viewed as a communication tool addressed to the firm's shareholders that outlines past operating results and identifies new areas of potential corporate growth and profitability (Kohut and Segars, 1992). An accurate measure of a CEO's expectations contained within this letter should thus be of interest to investors and other stakeholders as means to assess future performance.

Although some maintain that the prose in annual reports is written by public relations specialists, the typical CEO spends considerable time outlining the contents of the report, writing out much of it and adapting most of it to his taste (Bowman, 1984). While most sections of annual reports are the subject of extensive analysis[1], CEO letters have long been discarded by academic research. We find only two research papers on the informational value of CEO letters, namely Abrahamson and Amir (1996) and Patelli and Pedrini (2013), who find that CEO letters contain incremental information value and are a key documents for predicting future return on assets. This lack of research is surprising, as CEO letters are the most widely read section of the annual report (Courtis, 1982; Lee and Tweedie, 1975). The SEC commissioned a report that examined the information-gathering activities of a large sample of individual investors. The results show that 91% of the respondents acknowledge reading corporate annual reports. Of the narrative components, 77% of these investors report reading the president's letter at least "somewhat thoroughly", while 74% identify it as being at least "moderately useful" for information purposes (U.-S. Congress, House Committee on Interstate and Foreign Commerce, 1977).

The limited research on CEO letters could be explained by the fact that CEO letters are unaudited and, unlike disclosures to the SEC, the message in these letters can to a great extent be shaped as the CEO sees fit. Letters to shareholders are thus perceived as marketing tools highlighting the firm's mission, objectives, strategies and financial performance. According to Clatworthy and Jones (2006), CEO letters offer examples of impression management, as their textual characteristics reflect the self-serving goals of CEOs rather than the actual performance results

[1] Past research finds that the textual components of annual report contain informational value for predicting future firm performance. For instance, Davis et al. (2012) find that the sentiment included in earnings press releases significantly predicts future firm performance (measured as return on assets). Price et al. (2012) study sentiment within conference calls and find that conference call linguistic tone is a significant predictor of abnormal returns and trading volume. Li (2010) looks at 10-K filings and finds that the risk sentiment within those documents predicts future earnings.

being communicated. When engaging in impression management, CEOs control the information released in the CEO letters to influence shareholders' judgment about managerial performance.

Impression management explains why Abrahamson and Amir (1996) focus exclusively on negative words to measure CEO sentiment.[2] They argue that the positive words in CEO letters to shareholders serve as 'sugar coating' and are irrelevant and ritualistic elements for analysis. Reports of negative outcomes are much more infrequent than those of positive outcomes and typically concern much more important matters for the firm. When computing the sentiment of a CEO letter, it is thus important to identify and allocate a lower weight to these positive words.

CEOs' relative freedom to choose the information reported and the lack of restrictions on how this information is disclosed make the CEO letter an interesting avenue to disclose valuable information. CEOs will include in their letters (non-financial) explanations and interpretations, which cannot be included in the audited financial statements (Abrahamson and Amir, 1996). As for the MD&A section, letters to shareholders are a primary medium to provide users with potentially useful information not included in audited financial statements. Their importance as a complementary means of communicating with shareholders could explain why the length of CEO letters has substantially increased over the last 20 years. From an average of 1,230 words per letter between 1987 and 1988 (Abrahamson and Amir, 1996), the average number of words in CEO letters has increased to approximately 1,900 as of 2012. We will show in the next sections that, despite being influenced by managers' self-serving goals, CEO letters do contain incremental information value.

2.4 Hypothesis development

Our main hypothesis is that a weighted measure of CEO sentiment in which the weights are defined as a function of the position of the (positive or negative) word in the text is more accurate in predicting future performance than its equally-weighted counterpart. To test this hypothesis, we proceed in two ways. First, we investigate the intratextual distribution of net sentiment in CEO letters and show that this distribution is far from uniform and has an explainable periodic shape. Second, based on these stylized facts, we investigate whether allocating weights to words as a function of their position in their text increases the prediction accuracy of future firm performance relative to the equally-weighted metrics used in prior literature.

2.4.1 Hypotheses regarding the shape of the intratextual dynamics of sentiment

CEO letters are part of the firm's discourse and share common features that we call stylized facts. The first is that CEO letters are logically organized discourses in which the most salient elements of the text are discussed at the beginning and end of the text, while the more neutral elements are discussed in the middle. This implies that the number of words classified as positive will be higher at the beginning and at the end. A second observation is that letters to shareholders are

[2] Tetlock (2007) finds that the negative words in firm-specific news have a much stronger correlation with stock returns than other words. They argue that this result is consistent with a large body of psychology literature that shows that negative information has more impact and is more thoroughly processed than positive information across a wide range of contexts (Baumeister, Bratslavsky, Finkenauer, and Vohs, 2001; Rozin and Royzman, 2001).

influenced by CEOs' impression management techniques (Clatworthy and Jones, 2003; Patelli and Pedrini, 2013). This means that the presentation of accounting information in CEO letters attempts to control and manipulate the impression conveyed to users and, ultimately, influence stock prices. For that reason, CEO letters are sometimes thought of as being purly communication tools used to strategically manipulate the perceptions and decisions of stakeholders instead of means of incremental information dissemination (Yuthas, Rogers, and Dillard, 2002). Therefore, following the serial position effect and for a given total number of positive words, we expect CEOs to be disproportionately more positive at the beginning and end of the letter than in the middle, where the firm's operations and developments are discussed. This forms our first hypothesis:

H1a: Textual positive sentiment within CEO letters to shareholders is U-shaped on average.

Because the end of the letter is recalled best, we expect the end of the letter to contain a much larger number of positive words than the beginning. The U-shape of CEOs' positive sentiment within their letters can also be understood in the context of the peak-end-rule, which predicts that framing financial performance in positive terms with a peak at the end will cause investors to think about the results in terms of increases relative to the reference point (average) (Kahneman, Fredrickson, Schreiber, and Redelmeier, 1993).

The pattern of CEOs' negative sentiment is more difficult to predict. Since the Sarbanes-Oxley Act of 2002 and the establishment of the Public Company Accounting Oversight Board in the USA (and similar bodies in other countries), CEOs have to be more conscious of the words they choose in discharging their accountability to stakeholders. In the wake of recent accounting and corporate governance scandals, audit committees, regulatory authorities and others involved in the oversight of CEOs are now more alert to their obligations implicit in the narratives signed by CEOs, especially concerning past or prospective negative events.

We expect the use of negative words to be a trade-off for the CEO. On the one hand, it is important for the CEO letter to be in agreement with prior knowledge to assist the reader's comprehension (Pearson, Hansen, and Gordon, 1979). Readers will stop reading the CEO letter if it is unrealistic. Because our sample includes economically volatile periods with two important crises (the dot-com bubble in 2001 and the great recession in 2007-2009), realistic CEO letters cannot avoid the use of negative words. On the other hand, the CEO wants to maximize the firm value and communicate positively to investors. We expect that the CEO optimally achieves these objectives by placing the majority of the negative words at the beginning of the text. Because the introduction is thereby realistic, the CEO will avoid losing the reader, as the text is in agreement with his understanding of the economic situation. Because of the recency theory and peak-end rule, investors will remember these negative words less after having read the entire text.

H1b: Textual negative sentiment within CEO letters to shareholders is characterized by a decreasing smirk on average.

As a result and consistent with the peak-end rule, we expect CEOs' net sentiment, measured as the spread between positive and negative words, to show an increasing smirk.

H1c: Textual net sentiment within CEO letters to shareholders is characterized by an increasing smirk on average.

2.4.2 Hypotheses regarding the weighted sentiment proxy

One of the major consequences of the non-uniform distribution of the intratextual number of positive and negative words is that total sentiment measures that aggregate the intratextual sentiment without considering the position in the text may be suboptimal. If the beginning and end of the letter are dominated by impression management and overconfidence biases, it implies that these parts of the text contain less information value and should be underweighted when measuring sentiment.

To make the point above more formally, denote the true net sentiment underlying the CEO letter of firm j in year t as $NetSent^*_{j,t}$. The $NetSent^*_{j,t}$ is a latent variable and thus can only be observed through proxies. Imagine that we can split the text in B parts of equal text length and that, for the letter of firm j in year t, the observed proxy for net sentiment of text part b is given by $NetSent_{b,j,t}$. To the best of our knowledge, the prior literature has almost always solved the inference problem of estimating $NetSent^*_{j,t}$ based on $NetSent_{b,j,t}$ $(b = 1, \ldots, B)$ by taking the mean value of the intratextual sentiment $NetSent_{b,j,t}$:

$$NetSent^{\text{EW}}_{j,t} = \frac{1}{B}\sum_{b=1}^{B} NetSent_{b,j,t}, \tag{2.1}$$

see e.g. Abrahamson and Amir (1996); Patelli and Pedrini (2013).

A regression representation of $NetSent_{b,j,t}$ with $NetSent^*_{j,t}$ as explanatory variable provides interesting insight on the estimation properties of $NetSent^{\text{EW}}_{j,t}$:

$$NetSent_{b,j,t} = \alpha_b + \beta_b NetSent^*_{j,t} + \varepsilon_{b,j,t}. \tag{2.2}$$

The equally–weighted proxy can be shown to have good estimation properties in the absence of intratextual dynamics, i.e. when, for all b, $\alpha_b = 0$ and $\beta_b = 1$, and the error term $\varepsilon_{b,j,t}$ has zero mean. In that special case, $NetSent^*_{j,t}$ is the intercept parameter of the regression:

$$NetSent_{b,j,t} = NetSent^*_{j,t} + \varepsilon_{b,j,t}, \tag{2.3}$$

and $NetSent^{\text{EW}}_{j,t}$ is the ordinary least squares estimate of $NetSent^*_{j,t}$. If we can additionally assume that the $\varepsilon_{b,j,t}$ satisfy the Gauss-Markov conditions, then the equally–weighting approach is even efficient.

But, when there is heterogeneity in the regression parameter and/or the variance of the error terms, i.e. when some parts of the text are systematically more informative than others, then it may be more efficient to measure sentiment as a weighted average of the intratextual net sentiment such that the words with a higher (lower) information value are overweighted (resp. underweighted). This leads us to formulate the following hypothesis:

H2: When there is intratextual dispersion in sentiment, a weighted measure of CEO sentiment *with weights that are a function of the position of a word in a text is more informative for future firm performance than their equally-weighted counterparts.*

3 Data Collection and Financial Dictionaries

In the next sections, we will analyze the intratextual dynamics of sentiment in CEO letters over the period 2000 and 2011. This section first describes our data set of CEO letters and then introduces the libraries used to extract the sentiment from the observed words.

3.1 Collection of DJIA CEO letters

We hand-collect the CEO letters of the firms included in the Dow Jones Industrial Average Index (DJIA) for the twelve consecutive fiscal years 2000 to 2011. We choose the DJIA constituents for reasons of importance and tractability. The DJIA encompasses 30 of the largest firms in the United States and is considered a leading indicator of the stock market. We obtain the letters from each firm's respective website. If the annual report is not directly available, we contact the firm's public relations department. Because firms typically file their annual reports in the next calendar year, our sample mostly covers fiscal years 2001 to 2012. To avoid double-counting, we only select the text portion and delete any table, graph or figure included in the letter. Each letter is then saved as a text file for compatibility with our content analyzer.

In terms of comparison, Abrahamson and Amir (1996) and Patelli and Pedrini (2013) cover two-year periods between 1987–1988 and 2008–2009, respectively. Although both papers cover a larger cross-section of firms, the 12-year period of this paper stands in clear contrast to the short time-series adopted in their research and covers different market regimes, while the 1987–1988 and 2008–2009 periods adopted by Abrahamson and Amir (1996) and Patelli and Pedrini (2013) correspond to a high-market-uncertainty regime.

Our second hypothesis focuses on the relationship between CEO sentiment and future firm performance, which requires a date on which the letter was made publicly available. Thus, we manually collect each firm's annual report SEC filing date on the Edgar system. The firm is required to have at least one filing date at the SEC over the 2000-2012 period. If we obtain no SEC filing date for some years, we extrapolate the missing date(s) based on the latest date available for that firm.

As described below, our analyses require stock price and accounting data. Market prices and returns data are taken from the Center for Research in Security Prices (CRSP) database, while the COMPUSTAT database is our source for accounting data. Our final sample consists of 342 CEO letters with a total of 1,002,054 words.

3.2 Financial dictionaries

As mentioned in Subsection 2.1, the content analysis approach to quantifying sentiment in CEO letters to shareholders requires the identification of the tone of the individual words in the letter. This is usually achieved by matching the words with so-called dictionaries of positive and negative words. We will use three different libraries of words, each of which has been used to study firms' financial disclosures. The first is obtained from Loughran and McDonald (2011), who provide finance-oriented (positive and negative) lists of words. The lists are publicly available on the authors' website.[3]

The second library is obtained from Diction 7.0, a dictionary-based text analysis program that incorporates general content dictionaries designed to capture the tone within a document. Diction generates a score for the sentiment within a document as the sum of three positive components (praise, satisfaction and inspiration) minus the sum of three negative components (blame, hardship and denial). However, it does not provide the user with the location of the positive and negative words within the text. For that reason, instead of relying on the score provided by Diction, we use the Diction word lists to count both optimism-increasing words and optimism-decreasing words in the full text of each CEO letter.

As a third library, we refer to the negative list of words established by Abrahamson and Amir (1996). To construct their library, they require two "coders" (other than the authors) to examine the list of words independently and note every word connoting bad news that appears more than 30 times across all letters. These coders then meet to compare their list of negative words and to resolve any disagreements. They finally establish a list of words that is specifically designed for the study of CEO letters. The Abrahamson and Amir (1996) library is available in their paper.

We report in Table 2 the number of words found in CEO letters between 2000 and 2012 for each library. We find that 40,034 words out of the total 1,002,054 words in our sample can be matched with words in the Loughran and McDonald (2011) library, 47,429 with words in Diction and 2,835 with words in the Abrahamson and Amir (1996) library. A comparison between the words found by Diction and Loughran and McDonald (2011) demonstrates the broader scope of the Diction library, especially for negative words. The top five words in Diction include 'not', 'needs', 'no' and 'hard', those of the Loughran and McDonald (2011) contain more financially oriented words such as 'crisis', 'critical', 'challenges'.

We expect dictionaries with both positive and negative words to be more powerful in predicting future performance than dictionaries based strictly on negative language. In addition, prior studies suggest that generic linguistic algorithms such as Diction may yield noisy measures of "positive" and "negative" linguistic tone in the context of financially oriented text passages. For instance, Loughran and McDonald (2011) show that each discipline has its own dialect in which words take on specific meanings in specific contexts that may not translate effectively in other disciplines. For these reasons, we use in the following sections the word lists provided by Loughran and McDonald (2011) as our main library and report the results of Diction and Abrahamson and Amir (1996) for comparison purposes.[4]

[3] http://www3.nd.edu/~mcdonald/Word_Lists.html

[4] We verify in Subsection 6.3. the robustness of the Loughran and McDonald worldlists for forecasting future firm performance on our sample of CEO letters. We effectively find that sentiment construction from the Diction 7.0 word lists or the negative word list from Abrahamson and Amir (1996) is less predictive for future ROA than the Loughran and McDonald (2011) word lists.

Table 2: **Frequency of (top 5) words in DJIA CEO letters between 2000-2011**

	Loughran and McDonald (2011)		Diction 7.0		Abrahamson and Amir (1996)	
Words found	Positive	Negative	Positive	Negative	Positive	Negative
No.	29,726	10,308	37,248	10,181	-	2,835
%	2.967	1.029	3.717	1.016	-	0.283
Top 5 words found	Word	No.	Word	No.	Word	No.
Positive	Strong	1,763	Growth	4,470	-	-
	Leadership	1,080	Strong	1,763	-	-
	Opportunities	1,077	Important	1,116	-	-
	Innovation	1,006	Health	1,090	-	-
	Better	1,006	Leadership	1,080	-	-
Negative	Challenges	536	Not	1953	Difficult	328
	Difficult	328	No	673	Crisis	281
	Critical	324	Needs	628	Tough	235
	Crisis	281	Risk	598	Losses	212
	Challenging	275	Hard	305	Loss	175

Note: This table reports the number and frequency of (positive & negative) words in DJIA CEO letters between 2000 and 2011. The Loughran and McDonald (2011), Diction 7.0 and Abrahamson and Amir (1996) list of words are used to identify the positive and negative words. This table also reports the top 5 words found in the CEO letters and their associated frequency.

4 Intratextual Dynamics of CEO Sentiment

In this section, we first introduce the definition of the intratextual sentiment proxies. We then provide strong empirical evidence in favour of the hypotheses H1a, b and c on the shape of the intratextual dynamics of CEO sentiment within the letters to shareholders.

4.1 Notation

The length of each text is standardized to correspond to the $[0, 1]$ interval, which we divide in B bins such that each bin contains the same number of total words. For each bin, we then compute the percentage number of positive (resp. negative) words out of the total number of words in each bin. As such, the positive CEO tone in bin b ($b = 1, \ldots, B$) for firm j for the CEO letter of fiscal year t is

$$PosSent_{b,j,t} = \frac{PW_{b,j,t}}{TW_{b,j,t}}, \tag{4.1}$$

where $PW_{b,j,t}$ and $TW_{b,j,t}$ are the number of positive words and the total number of words for firm j in bin B for fiscal year t, respectively. Similarly, the negative CEO tone for bin b for firm

j for the CEO letter of fiscal year t is given by

$$NegSent_{b,j,t} = \frac{NW_{b,j,t}}{TW_{b,j,t}}, \tag{4.2}$$

where $NW_{b,j,t}$ is the number of negative words for firm j in bin b for fiscal year t.

Finally, for each bin, we also compute the difference between the positive and negative tone, and call this the net sentiment of that bin:

$$NetSent_{b,j,t} = PosSent_{b,j,t} - NegSent_{b,j,t}. \tag{4.3}$$

4.2 Findings

In Figure 1, we report the average optimism ($PosSent_{j,t}$) and pessimism ($NegSent_{j,t}$) in CEO letters to shareholders for the Loughran and McDonald (2011) dictionary. The average number of positive financial words in a letter is 3.200% and is significant at a 99% confidence level. As expected, the average number of negative words is substantially lower than the average positive tone, with a mean of 0.865%.[5] This result suggests that CEOs are acting opportunistically and engage in impression management strategies throughout the letter. From the standardized measure of CEO sentiment, the statistically significant difference between the average positive and negative tone over a letter is an indication that CEOs attempt to manipulate the tone of CEO letters, distracting from past performance and distorting expectations of future performance. This result is consistent with prior research on CEO behavior, which shows that CEOs conceal bad news by not reporting it to the same extent as good news (see, e.g., Clatworthy and Jones, 2003). A look at the prevalence of the positive values of $NetSent_{j,t}$ in Table 6 in Appendix confirms this trend. CEOs are consistently more optimistic than pessimistic, even during financial crises; only 8 out of 342 (2%) CEO letters have a negative value for their net sentiment between 2000 and 2012, among which five are for JPMorgan between 2007 and 2011.

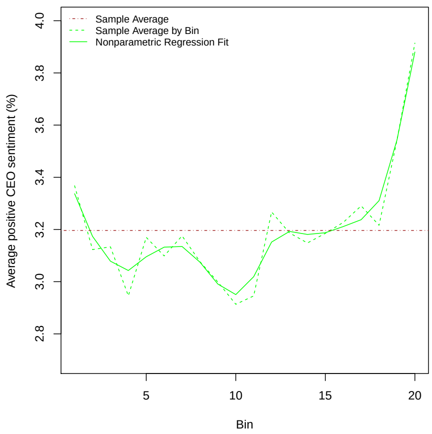
When we divide each letter into 20 bins, the dynamics of the positive and negative sentiment within the letter show a clear contrast to the overall standard measure of CEO sentiment.[6] Consistent with Hypothesis 1a, we find in Figure 1a a U-shaped trend for positive tone, where the average tone decreases from 3.4% in the first bin to approximately 3% between bins 4 and 18. The positive tone then increases again to 4% in the last two bins. When divided into bins, we clearly identify in Figure 1c a decreasing smirk in negative words within the letters. Starting with an average of 1.4% in bin 1, the average number of negative words progressively declines to 0.8% towards the end of the letter. A single-sided t-test shows that the negative tone at the end of the letter is statistically lower than that at beginning at a 99% confidence level. This result is consistent with Hypothesis 1b. As shown in Figure 6 and 7 in Appendix, we find a similar shape in positive and negative words regardless of the library we use to measure CEO sentiment.

[5]In Figure 6 in Appendix, the Diction-library-based measure of $PosSent_{j,t}$ is also significantly higher than its negative counterpart $NegSent_{j,t}$. The average value of $PosSent_{j,t}$ equals 3.864%, while that of $NegSent_{j,t}$ equals 0.813%. As shown in Figure 7, the average CEO sentiment of the Abrahamson and Amir (1996) library is substantially lower than that for the other two list of words, with an average of 0.220%.

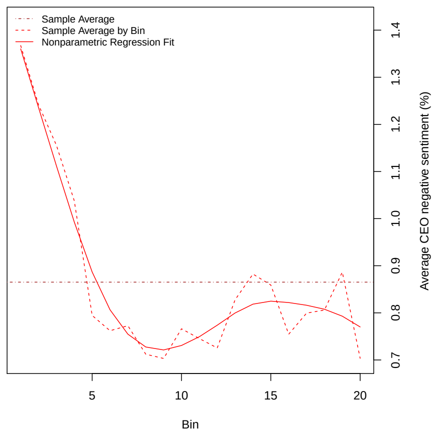
[6]Results are qualitatively similar for 10 bins.

Figure 1: **Intratextual dynamics of CEO sentiment following the financial library of Loughran and McDonald (2011) to identify positive and negative words**

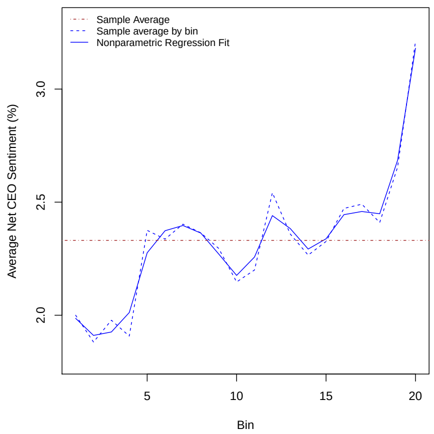
Note: This figure depicts the dynamics of CEO sentiment within letters to shareholders. The length of each text is standardized to correspond to a $[0, 1]$ interval, which is divided in B bins such that each bin contains the same number of total words. For each bin, the percentage number of positive words out of the total number of words in each bin is reported (Figure 1a). Similarly, for each bin, the percentage number of negative words out of the total number of words in each bin is computed (Figure 1b). For each bin, the net sentiment is then measured as the spread between the positive and negative tone (Figure 1c). Positive and negative tones are measured based on the Loughran and McDonald (2011) word lists.



(a) U-shape in average positive sentiment by bin



(b) Decreasing smirk in average negative sentiment by bin



(c) Increasing smirk in average net sentiment by bin

Our intratextual analysis of CEO sentiment identifies a new and subtler form of impression management that occurs within the CEO letter. An intratextual analysis shows that CEOs first refer to negative past events and, whenever they introduce bad events, they swamp them with many positive words. They then progressively talk about the future in positive terms while clearly reducing the negative tone within their letter. CEOs' swamping strategy is compatible with the serial position effect. Because investors will recall the end of the text best (recency effect), CEOs will increase the number of positive words towards the last bins. Similarly, the first bins report an above-average positive sentiment, as these are the bins that are recalled more frequently than the middle items (the primacy effect). These middle items report an average $PosSent_{j,t}$ that is lower than the letter average, with a nadir at the 10^{th} bin. This impression management strategy is in line with the concealment theory of impression management, but at an intratextual level.

The U-shape with a large peak at the end is also consistent with the hypothesis of CEO overconfidence. CEOs refer to the past in the first bins and express an optimism that is slightly higher than the letter average (0.2 percentage point higher than average in the first bin). Towards the end of their letter, CEOs tend to mention their plans for the company's future and become substantially more optimistic than in the first bin (difference of one percentage point relative to the first bin). This pattern is consistent with the hypothesis in Heaton (2002) and Malmendier and Tate (2005) of optimistic and overconfident CEOs who tend to see the future as brighter than the past.

There is an increasing smirk in CEO net sentiment within the letter. Figure 1c shows an increasing value for net sentiment within the letter that peaks at 3.2% at the end. This is consistent with the peak-end-rule and the recency effect, confirming Hypothesis 1c.

Overall, these results show that CEO letters are strategically crafted corporate discourses. Positive and negative words are distributed over the text in such a way that readers are left with a positive impression about the firm.

5 A Position-Weighted Measure of Sentiment

The next question that we address is how to aggregate the intratextual net sentiment measures $NetSent_{b,j,t}$ into a single overall sentiment per text that has predictive power for future firm performance. In this section, we first outline our estimation methodology and then discuss the resulting differences between the standard equally-weighted measure of textual sentiment versus the position-weighted sentiment estimates obtained for our panel of CEO letters over the period 2000-2011.

5.1 Methodology

As mentioned in our hypotheses and literature review, the workhorse aggregation technique in the literature has hitherto been to use simple averaging. The underlying assumptions are:

- linearity: the total sentiment measure of the CEO letter j in year t is given by the linear mapping of the B intratextual sentiment measures $NetSent_{b,j,t}$ on $NetSent_{j,t}$ with weights $w = (w_1, \ldots, w_B)'$:

$$NetSent_{j,t}(w) = \sum_{b=1}^{B} w_b NetSent_{b,j,t}, \tag{5.1}$$

and where all weights sum to unity;

- equal importance of each part of the text: since the bins have the same text length, this implies an equal weighting, i.e. w is set to

$$w^{\text{EW}} = (1/B, 1/B, \ldots, 1/B)'. \tag{5.2}$$

Next, we will keep the linearity assumption, but relax the assumption of constant intratextual weights for firms with strong intratextual dynamics in sentiment. In subsection 5.1.1 we first describe our approach to identify the CEO letters for which intratextual weighting may be more optimal and explain the use of Almon polynomials to model these weights. Then, in Subsection 5.1.1, the least squares estimation approach is developed in order to calibrate the weights in such a way that the resulting position-weighted sentiment measure is a better predictor for future firm performance.

5.1.1 The Almon approach to specifying flexible and smooth intratextual weights

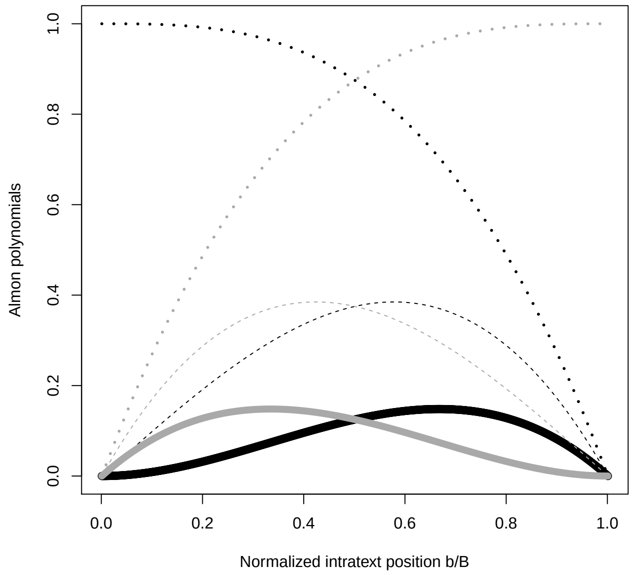
For simplicity and ease of exposition, we will keep the linearity assumption, but expect that, for firms with strong intratextual dynamics in sentiment, a higher predictive power can be obtained by using a sentiment measures that differentiates the weights across the text bins. We measure the intratextual dispersion using the standard deviation of net sentiment:

$$SD_{j,t} = \left[\frac{1}{B-1} \sum_{b=1}^{B} \left(NetSent_{b,j,t} - NetSent_{j,t}^{\text{EW}} \right)^2 \right]^{\frac{1}{2}}.$$

Note that SD reaches its minimal value of zero when intratextual sentiment is perfectly uniformly distributed across all bins. The more intratextual dispersion, the higher SD is. Whenever $SD_{j,t}$ is small (in practice, smaller than a threshold value κ), and thus the intratextual sentiment does not deviate much from the equally-weighted measure, we recommend to stick to the benchmark equal-weighting approach. However, when there is sufficient scope for differentiating across texts parts (i.e. $SD_{j,t}$ is large enough), then we will use optimized weights.

The most straightforward way to estimate these intratextual weights is to use dummy variables. The disadvantage is that, because of estimation error, the weights may be erratic and show jumps, whilst a smooth intratextual weight pattern is expected. We impose smoothness and parsimony on the optimized weights by the specifying the weights as a combination of Almon polynomials of the normalized bin index (b/B): the B weights can be represented as a linear combination of first, second and third–order Almon polynomials of the normalized bin

Figure 2: **Left- (grey) and right-centered (black) Almon polynomials of order one (full), two (dashed) and three (dotted) used to model the intratextual weights**



Note: This graph reports the third-order Almond polynomials denoted as $P_c(u)$ in Equation (5.3), for $u = b/B$ (left-centered) and $u = (B-b)/B$ (right-centered).

index (b/B):

$$w_b^{\text{Almon}}(\theta) \quad = \quad \theta_1 + \sum_{c=1}^{3} \theta_{1+c} P_c(b/B) + \theta_{4+c} P_c((B-b)/B), \tag{5.3}$$

where $P_c(u) = (1-u^c)u^{3-c}$ is a c^{th}-order Almon polynomial of u. These polynomials are positively valued functions such that the positivity of the weights is guaranteed by requiring $\theta \geq 0$. We further require that all weights add up to unity:

$$\sum_{b=1}^{B} w_b^{\text{Almon}}(\theta) = 1. \tag{5.4}$$

The flexibility of the third-order Almon polynomials is illustrated in Figure 2. The linear combination of those six curves can closely fit almost every smooth periodic pattern and has been used to fit periodic patterns by Andersen, Bollerslev, and Cai (2000) and Boudt, Croux, and Laurent (2011), among others.

Based on these assumptions, the two unknowns needed to define the optimized weights are the threshold κ determining the use of equal–weights versus optimized intratextual weights, and the value of the parameter vector θ. These values will be application–specific and we will next describe a data–based procedure to determine the weights' particular functional form.

5.1.2 Panel threshold least squares estimation of intratextual weights

The optimized weights are application-specific. Our focus is to use the textual sentiment for predicting future firm performance. Based on Engelberg (2008) and Davis et al. (2012), we proxy future firm performance using return on assets (ROA).[7] Because of the yearly frequency of our data, we resort to panel estimation in order to estimate the weights on the intratextual sentiment.

This leads us to specify the following fixed effects regression model:

$$
\begin{aligned}
ROA_{j,t+1} \quad &= \quad \alpha_j + \beta_1 \cdot NetSent^{\text{EW}}_{j,t} \cdot (1 - I\left[SD_{j,t} > \kappa\right]) \\
&\quad + \beta_2 \cdot NetSent_{j,t}(w^{\text{Almon}}(\theta)) \cdot I\left[SD_{j,t} > \kappa\right] + \epsilon_{j,t+1},
\end{aligned} \tag{5.5}
$$

where $I\left[SD_{j,t} > \kappa\right]$ is a dummy variable that equals one if the intratextual dispersion in sentiment is higher than a threshold value of κ and is zero otherwise. $NetSent^{\text{EW}}_{j,t}$ refers to the equally-weighted measure of net sentiment.

In the panel, we distinguish between texts for which there is a sufficiently high intratextual dispersion in sentiment ($SD_{j,t} > \kappa$) from those for which the intratextual sentiment dynamics are less important. We estimate the parameters (α_j, β_1, β_2, θ, κ) by non-linear least squares. This can be performed in a computationally convenient way by a loop over κ and by noting that, for a given value of κ, the model is linear in the parameters α, β_1 and the parameter product $\tilde{\theta} = \beta_2 \cdot \theta$.[8] Because of the bound constraint that $\theta \geq 0$, $\tilde{\theta}$ is also bound constrained, but the problem of least squares estimation is still simple to solve by reformulating the estimation problem conditional on κ as a bound constrained quadratic optimization problem that can be easily solved numerically.[9] The constraint that all intratextual weights need to sum up to unity

[7] To avoid any look-ahead bias, we start measuring ROA in the quarter following the quarter in which the annual report has been filed at the SEC. Specifically, firm future performance $ROA_{j,t+1}$ is measured as the sum of quarterly earnings before extraordinary items $Y_{j,q+i,t+1}$ (Compustat data item #18) over the four quarters after the SEC filing quarter q, scaled by total assets (#6) at the end of quarter q. Each item is defined in Table 7.

[8] The linearity in $\tilde{\theta} = \beta_2 \cdot \theta$ becomes obvious by using (5.1) and (5.3) to rewrite the regressor in (5.5) as:

$$
\begin{aligned}
\beta_2 \cdot NetSent_{j,t}(w^{\text{Almon}}(\theta)) &= \beta_2 \sum_{b=1}^{B} w_b^{\text{Almon}}(\theta) NetSent_{b,j,t} \\
= \quad (\beta_2\theta_1) \sum_{b=1}^{B} NetSent_{b,j,t} &+ \sum_{c=1}^{3} (\beta_2\theta_{1+c}) \sum_{b=1}^{B} P_c(b/B) NetSent_{b,j,t} + (\beta_2\theta_{4+c}) P_c((B-b)/B) NetSent_{b,j,t}.
\end{aligned}
$$

[9] Given a value of κ, the least squares regression estimator minimizes a sum of squared residuals that can be rewritten as $(y-Xb)'(y-Xb)$, with y the vector of future ROA, X the matrix of explanatory variables and b the parameter estimates. This is equivalent to minimizing $-2b'X'y + b'X'Xb$, which is a quadratic objective function of b.

in Equation (5.4) implies that the estimated θ is the normalized version of the estimated $\tilde{\theta}$.

5.2 Equally- and Position-weighted sentiment in CEO letters

Let us now investigate the resulting optimized weights for our 2000-2011 sample of CEO letters using the Loughran and McDonald (2011) library to decode the intratextual sentiment. We first estimate sentiment over the full sample and then consider the estimation on rolling samples of three years, which will be important for the prediction analysis in Section 6.

The first step is to estimate the threshold value κ that distinguishes the CEO letters for which the equally–weighted measure is used versus those for which the weights are heterogeneous and depend on the position in the text. Using the full 2000-2011 sample, we estimate κ at 1.9%, resulting in 81.579% of the letters for which position–specific weights are to be used. The corresponding pattern of estimated weights is reported in Figure 3b.

First of all, we see that the optimized weights have a strong inversely bell-shaped intratextual variation: it starts at 1.9% for the words used in the beginning of the text, 7% in the middle of the text and almost 0% at the end of the text. It follows that, compared to the equally-weighted measure of CEO sentiment, the pattern of weights shows that the optimized sentiment measure will underweight the textual sentiment at the beginning and end of the text and overweight the sentiment at the middle of the text. This is consistent with Hypothesis 2 and indicates that an equally-weighted measure of CEO sentiment may not be optimal to predict future performance.

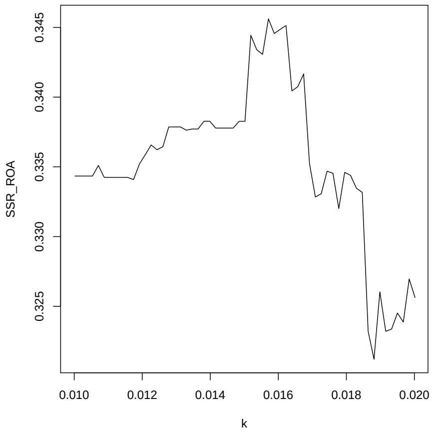
Such a strong intratextual variation does not necessarily imply a large difference between the position– and equally–weighted measures. Indeed, if the measure of sentiment were the same for each bin, then the weighting has no impact on the sentiment measure. Figure 4 investigates the impact of the weighting on the estimated sentiment. It shows the scatter plot of the position–weighted versus the equally–weighted sentiment measure for our panel of DJIA CEO letters over the period 2003–2011.[10] We see that there is a strong agreement across the measures, but that differences exist. To inspect these differences, consider as the reference line, the 45° line corresponding to equality of the two approaches. We find that 14% of the observations are on the line (perfect agreement), 54% below the line (the position-weighted measure is more pessimistic than the equally-weighted measure) and 32% under the line. The outcome that position-weighted leads on average to a less optimistic view on sentiment is as expected, since, as explained in the hypothesis section, managers have a tendency to be overly optimistic and convey this optimism by overloading the beginning and end of the text with positive words.

In the next section, we evaluate the forecasting performance of the position-weighted sentiment measure. To avoid look ahead bias, we use, at each point in time, the sample of the three most recent years, in order to estimate the optimal intratextual weights. Denote these estimates based on years $t-2$, $t-1$ and t as $\hat{\theta}_t$ and $\hat{\kappa}_t$. The estimated threshold varies between 1.068 % and 2.003% of the sample with an average value of 1.572%, corresponding to 124 firms. Figure 5 plots the kappa values for each rolling sample, as well as the percentage of firms with an intratextual standard deviation of sentiment that is higher than kappa. For each rolling sample, the

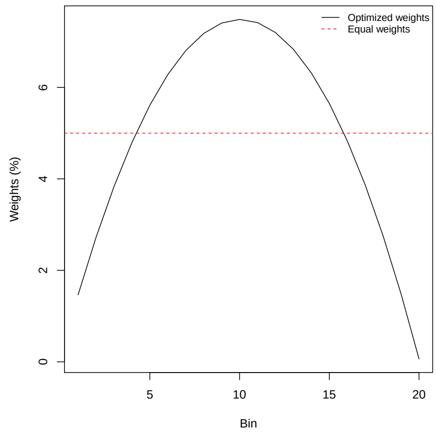
We report in Figure 3a the sum of squared residuals as a function of the threshold parameter κ.

[10] Table 6 and 8 detail the equally- and position-weighted sentiment measures per CEO letter between 2000 and 2011.

Figure 3: **The sum of squared residuals as a function of the threshold parameter κ (left figure) and the bell shape in optimized weights of intratextual net sentiment as a function of position of a word in the text (right figure)**



(a) Sum of squared residuals as a function of the threshold parameter kappa (κ)



(b) Pattern of weights of intratextual net sentiment as a function of position of a word in the text

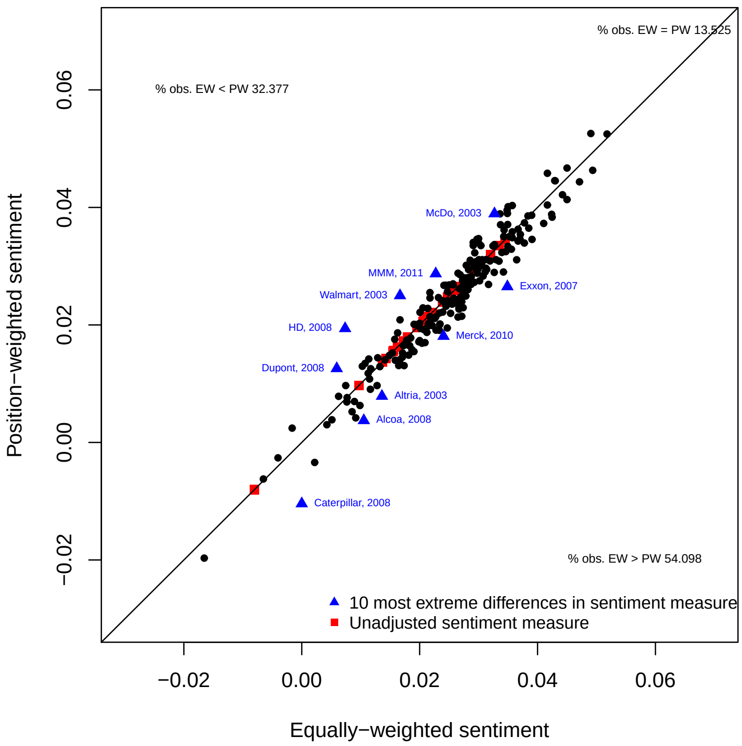
value of kappa and the percentage of firms with high values of intratextual standard deviation of sentiment remain fairly stable.

6 CEO Sentiment and Future Firm Performance

We now turn to the question of the economic relevance of using the position-weighted sentiment measure rather than the equally weighted sentiment measure for predicting future firm performance. We do this analysis for the out-of-sample period 2003-2011, where, as described in the previous section, the weights underlying the position-weighted sentiment measure are estimated on the three preceding years.

In Subsection 6.1, we first do the rat race of comparing models that predict future ROA, using only the information in the intratextual net sentiment. Then in Subsection 6.2 we control for other influences that may have an impact on ROA.

Figure 4: **Out of sample estimation – Scatter plot of equally- and position-weighted sentiment measures**



6.1 Comparison of pure sentiment-based prediction models for future ROA

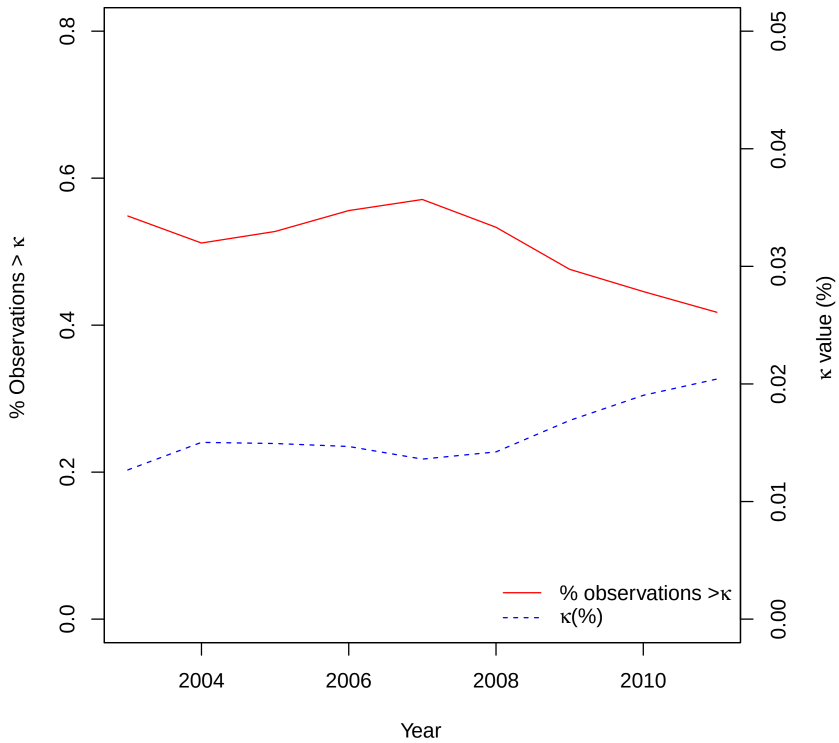
The benchmark model is the traditional approach consisting of a linear prediction model in which the equally-weighted measure of sentiment is used to forecast future ROA:

$$ROA_{j,t+1} \quad = \quad \alpha_j + \delta_{t+1} + \beta \cdot NetSent^{\mathrm{EW}}_{j,t} + \epsilon_{j,t+1}, \tag{6.1}$$

and where α_j and δ_{t+1} correspond to the firm and year fixed effects. Under this "EW model" approach, we thus regress future $ROA_{j,t+1}$ on $NetSent^{\mathrm{EW}}_{j,t}$, as in Abrahamson and Amir (1996) and Patelli and Pedrini (2013).[11]

[11]Our primary interest is in predicting future performance. We are not interested in the behavioral interpretation of the coefficient of the impact of sentiment on future performance, which would require to deal with the endogeneity of the sentiment variable and either require instrumental variable type of estimation or, at least mitigate the endogeneity issue by taking the performance and sentiment variables in first differences, as recommended by Li

Figure 5: **Out of sample estimation – Plots of kappa values and the percentage firms with intratextual standard deviation higher than kappa**



Our leading hypothesis is that, for CEO letters with marked intratextual sentiment dynamics, a more accurate forecast can be obtained by using the proposed position-weighted sentiment measure through the "PW model":

$$ROA_{j,t+1} = \alpha_j + \delta_{t+1} + \beta_1 \cdot NetSent^{\text{EW}}_{j,t} \cdot (1 - I\left[SD_{j,t} > \hat{\kappa}_t\right]) \qquad (6.2)$$

$$+\beta_2 \cdot NetSent_{j,t}(w^{\text{Almon}}(\hat{\theta}_t)) \cdot I\left[SD_{j,t} > \hat{\kappa}_t\right] + \epsilon_{j,t+1}.$$

We will compare the two test regression based on their goodness of fit as measured by the adjusted R^2, but also through an F-test comparing their fit with the one of the generalized unre-

(2010) and Kravet and Muslu (2013).

Table 3: **Equally-weighted versus position-weighted CEO sentiment and future firm performance**

	Panel A: Without fixed effects			Panel B: With fixed effects		
	EW model	PW model	GUM	EW model	PW model	GUM
Parameter estimates						
(Intercept)	$\mathbf{0.024}^{***}$	$\mathbf{0.031}^{***}$	$\mathbf{0.029}^{***}$			
	(0.009)	(0.009)	(0.009)			
$NetSent^{\text{EW}}_{j,t}$	$\mathbf{2.400}^{***}$		1.016	$\mathbf{0.724}^{**}$		0.012
	(0.325)		(1.051)	(0.294)		(0.744)
$NetSent_{j,t}(w^{\text{Almon}}(\hat{\theta}_t)) \cdot I\left[SD_{j,t} > \hat{\kappa}_t\right]$		$\mathbf{2.320}^{***}$	1.368		$\mathbf{0.775}^{***}$	0.764
		(0.312)	(1.035)		(0.285)	(0.732)
$NetSent^{\text{EW}}_{j,t} \cdot (1 - I\left[SD_{j,t} > \hat{\kappa}_t\right])$		$\mathbf{1.003}^{**}$	0.075		0.117	0.106
		(0.457)	(1.080)		(0.366)	(0.790)
Goodness of fit statistics – F-test of equal fit between GUM and its restrictions (EW model, PW model)						
R^2	0.167	0.193	0.195	0.930	0.932	0.932
Adj. R^2	0.164	0.186	0.184	0.915	0.916	0.916
RSS	0.780	0.756	0.754	0.184	0.179	0.179
F-test EW/PW vs GUM	4.109	0.572	-	2.669	0.001	-
$pvalue$ EW/PW vs GUM	0.018	0.450	-	0.071	0.988	-

Note: This table presents the estimation results for the EW, PW and GUM models. Panel A reports the results for the EW (Equation (6.1)), PW (Equation (6.2)) and GUM (Equation (6.3)) models without firm and year fixed effects. In Panel B, the EW, PW and GUM models control for firm and year fixed effects. The equally- and position-weighted measures of CEO sentiment are defined by Equation (5.2) and Equation (5.3), respectively. The word lists used to estimate sentiment is from the Loughran and McDonald (2011) library. The significance of coefficients is tested using Newey-West standard errors. *, **, and *** denote statistical significance at the 10 percent, 5 percent, and 1 percent levels, respectively, based on a two-sided t-test.

stricted model (GUM):

$$\begin{aligned} ROA_{j,t+1} &= \alpha_j + \delta_{t+1} + \beta \cdot NetSent^{\text{EW}}_{j,t} + \beta_1 \cdot NetSent^{\text{EW}}_{j,t} \cdot (1 - I\left[SD_{j,t} > \hat{\kappa}_t\right]) \\ &\quad + \beta_2 \cdot NetSent_{j,t}(w^{\text{Almon}}(\hat{\theta}_t)) \cdot I\left[SD_{j,t} > \hat{\kappa}_t\right] + \epsilon_{j,t+1}. \end{aligned} \tag{6.3}$$

The results of these regressions are reported in Panel A (without firm and year fixed effects) and Panel B (with fixed effects) of Table 3. We test for the significance of the coefficients using Newey-West standard errors. The first result shown in Table 3 is that the sentiment in CEO letters contain information to predict future firm performance. This can be seen through the positive and significant coefficient of sentiment in the EW model of Panel A and B and its large explanatory power in predicting future firm performance (as measured by the $Adj.R^2$). This result implies that, despite their strategic approach to communicate with shareholders, managers use language in their annual letters to communicate information about the firm's future performance.

The second finding in Table 3 is that the proposed position-weighted measure has a significantly higher power to predict future firm performance than the traditional position-weighted measure. The $Adj.R^2$ increases from 16.4% (EW Model, Panel A) to 18.6% (PW Model, Panel A) once the sentiment measure considers the position of a word in the document. The increase in $Adj.R^2$ is statistically significant at a 95% confidence level. From the F-tests comparing the GUM with its restricted versions, we observe that, ignoring the position of a word in a document, decreases significantly the fit of the model, while omitting the equally–weighted sentiment variable has no significant effect on the fit of the model. We further see in Panel B of Table 3 that the results are qualitatively similar when correcting for firm and year fixed effects. In the PW model, it is interesting to observe that for firms with strong intratextual dynamics of sentiment, the effect of sentiment on future firm performance is significantly higher in the model without fixed effects than in the model with fixed effects (2.3 vs. 0.8).

The main conclusion is that, not all words in a CEO letter are equal and their informational value depends on their position within the letter. For firms with strong intratextual dynamics in sentiment, the total sentiment of the text is more accurately measured by a position-weighted average of the intratextual sentiment. The traditional sentiment measures defined in the literature are thus potentially inefficient, as they assume all words to be equal, regardless of their position within the document.

6.2 Multivariate prediction model: Controlling for hard information

In this subsection, we expand the number of regressors in the EW Model (Equation (6.1)) and the PW Model (Equation (6.2)) in order to test whether the explanatory power of CEOs' net sentiment in predicting future firm performance survives after controlling for "hard information". We define hard information as quantitative information easily processed from annual reports. We select hard information variables from the set of earnings-predicting covariates identified by Fama and French (2006), each of which is defined in Table 7 and explained in the Appendix.

The estimates are reported in Table 4. Panel A presents the results for the EW Model, Panel B to the PW Model, and Panel C for the GUM model. Consistent with Hypothesis 2, the main evidence of this section is that our conclusions persist after we control for the earnings-predicting variables defined by Fama and French (2006).

Indeed, the coefficient of $NetSent_{j,t}(w^{\text{Almon}}(\hat{\theta}_t)) \cdot I\left[SD_{j,t} > \hat{\kappa}_t\right]$ in the multivariate PW Model (1) is positive and significant at a 95% confidence level, which indicates that there is at least some forward-looking information in the sentiment of CEO letters that is incremental to more quantitative financial and accounting information. After including past stock market performance, the coefficient remains positive and significant at a 90% confidence level. On the other hand, the coefficient of the equally-weighted net sentiment measure $NetSent^{\text{EW}}_{j,t} \cdot (1 - I\left[SD_{j,t} > \hat{\kappa}_t\right])$ is negative and insignificant at traditional confidence levels, with a coefficient of -0.089. In addition, from the F-tests comparing the GUM with its restricted versions,

Table 4: **Equally-weighted versus position-weighted CEO sentiment and future firm performance**

	Panel A: EW model		Panel B: PW model		Panel C: GUM	
	Model 1	Model 2	Model 1	Model 2	Model 1	Model 2
Parameter estimates						
Sentiment measures						
$NetSent^{\text{EW}}_{j,t}$	$\mathbf{0.445}^{*}$	0.356			-0.339	-0.531
	(0.264)	(0.276)			(0.640)	(0.639)
$NetSent_{j,t}(w^{\text{Almon}}(\hat{\theta}_t)) \cdot I\left[SD_{j,t} > \hat{\kappa}_t\right]$			$\mathbf{0.533}^{**}$	$\mathbf{0.467}^{*}$	0.850	0.963
			(0.258)	(0.275)	(0.660)	(0.669)
$NetSent^{\text{EW}}_{j,t} \cdot (1 - I\left[SD_{j,t} > \hat{\kappa}_t\right])$			-0.089	-0.152	0.214	0.323
			(0.335)	(0.352)	(0.648)	(0.650)
Control variables						
$ROA_{j,t}$	0.049	0.056	0.067	0.074	0.069	0.077
	(0.078)	(0.090)	(0.072)	(0.084)	(0.073)	(0.085)
$\sigma_{ROA,j,t}$	-0.242	-0.201	-0.266	-0.225	-0.268	-0.226
	(0.229)	(0.225)	(0.221)	(0.218)	(0.221)	(0.218)
$+AC_{j,t}$	-0.009	-0.003	-0.011	-0.007	-0.011	-0.006
	(0.010)	(0.011)	(0.010)	(0.011)	(0.010)	(0.011)
$-AC_{j,t}$	-0.008	-0.008	-0.005	-0.004	-0.005	-0.005
	(0.011)	(0.011)	(0.010)	(0.011)	(0.010)	(0.012)
$NegY_{j,t}$	-0.012	0.000	-0.001	0.010	0.000	0.010
	(0.017)	(0.015)	(0.016)	(0.016)	(0.016)	(0.016)
$BTM_{j,t}$	$\mathbf{-0.057}^{***}$	$\mathbf{-0.053}^{***}$	$\mathbf{-0.054}^{***}$	$\mathbf{-0.050}^{***}$	$\mathbf{-0.055}^{***}$	$\mathbf{-0.050}^{***}$
	(0.010)	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)
$MC_{j,t}$	$\mathbf{-0.021}^{**}$	$\mathbf{-0.026}^{**}$	$\mathbf{-0.019}^{*}$	$\mathbf{-0.025}^{**}$	$\mathbf{-0.019}^{*}$	$\mathbf{-0.025}^{**}$
	(0.010)	(0.011)	(0.010)	(0.011)	(0.010)	(0.011)
$Div_{j,t}$	$\mathbf{-0.067}^{*}$	-0.044	$\mathbf{-0.057}^{*}$	-0.034	$\mathbf{-0.057}^{*}$	-0.034
	(0.035)	(0.035)	(0.034)	(0.035)	(0.034)	(0.035)
$Ret_{j,t}$		0.019		0.020		$\mathbf{0.021}^{*}$
		(0.013)		(0.013)		(0.013)
$Ret_{j,t-2}$		0.013		0.013		0.013
		(0.010)		(0.010)		(0.010)
$Ret^{*}_{j,t}$		0.030		0.027		0.027
		(0.021)		(0.021)		(0.021)
Goodness of fit statistics – F-test of equal fit between GUM and its restrictions (EW model, PW model)						
R^2	94.6	94.8	94.8	94.9	94.8	95.0
Adj. R^2	93.1	93.3	93.3	93.4	93.3	93.4
RSS	0.142	0.137	0.138	0.133	0.138	0.133
F-test EW/PW vs. GUM	2.998	3.232	0.217	0.536	-	-
$pvalue$ EW/PW vs GUM	0.052	0.042	0.642	0.465	-	-

Note: This table presents estimation results for Equation (6.1) in Panel A, for Equation (6.2) in Panel B and for Equation (6.3) in Panel C, including control variables and firm and year fixed effects. The equally- and position-weighted measures of CEO sentiment are defined by Equation (5.2) and Equation (5.3), respectively. The word lists used to estimate sentiment is from the Loughran and McDonald (2011) library. The significance of coefficients is tested using Newey-West standard errors. *, **, and *** denote statistical significance at the 10 percent, 5 percent, and 1 percent levels, respectively, based on a two-sided t-test.

Table 5: **Robustness to the choice of library and sentiment type – A comparison of $Adj.R^2$ (in %)**

	Sentiment:	Net		Positive		Negative	
	Hard Information:	No	Yes	No	Yes	No	Yes
Fixed effects	Dictionary						
No	Loughran and McDonald (2011)	18.59	89.32	1.28	89.01	9.04	88.72
	Diction 7.0	3.96	88.76	1.98	88.72	5.20	88.70
	Abrahamson and Amir (1996)	-	-	-	-	13.60	88.75
Yes	Loughran and McDonald (2011)	91.65	93.44	91.41	93.31	91.25	93.16
	Diction 7.0	91.12	93.20	91.13	93.22	91.21	93.17
	Abrahamson and Amir (1996)	-	-	-	-	91.42	93.22

Note: This table reports the robustness of our results to the choice of library, by comparing the $Adj.R^2$ of the PW model (Equation (6.2)) using the Diction 7.0, Loughran and McDonald (2011) and Abrahamson and Amir (1996) word lists. It also tests the inclusion of firm and year fixed effects and of hard financial information. Positive, negative and net sentiment measures are also distinguished.

we see that, ignoring the position–weighting of a word in a text, decreases significantly the fit of the model, while omitting the equally–weighted sentiment variable has no significant effect on the fit of the model. The F-test that compares the $Adj.R^2$ of the EW Model (2) and the GUM Model (2) has a value of 3.232, which is significant at a 95% confidence level.

The bottom line of the regression results in Table 3 and 4 is that, when there is intratextual dispersion in sentiment, a weighted measure of CEO sentiment with weights that are a function of the position of a word in the text is more informative to predict future firm performance than the equally-weighted metrics used in prior literature and that this result holds after we control for hard, financial information.

6.3 Robustness to the choice of library and sentiment type

Section 4 showed that for all three word lists considered (Loughran and McDonald (2011); Diction 7.0; Abrahamson and Amir (1996)), a similar intratextual pattern of sentiment can be found in the sample of DJIA firms over the period 2000–2011. Since Loughran and McDonald (2011) are designed to capture the sentiment in terms of firm performance based on financial corporate disclosures, we focused so far on their word lists in Section 5 and 6. In this Subsection, we verify the superiority of the Loughran and McDonald (2011) word lists for measuring sentiment in order to predict future performance. We run Equation (6.2) with the net sentiment constructed based on the Diction 7.0 and Abrahamson and Amir (1996) word lists and report in Table 5 the corresponding $Adj.R^2$. As expected, the highest $Adj.R^2$ is obtained for our baseline approach: net sentiment using the word lists of Loughran and McDonald (2011). Using the Diction word lists reduce the $Adj.R^2$ by 12%, to 7.0%. In addition, considering only positive or negative sentiment instead of their spread reduced substantially the explanatory power. E.g exclusively using the positive (negative) words in CEO letters reduces the $Adj.R^2$ to 1.28% (9.04%).

7 Conclusions

Investors routinely use summary statistics to avoid being overwhelmed by the massive amount of information available to them. The linguistic sentiment of a corporate disclosure is one such statistic that recently has become popular. Most studies find that the sentiment of financial disclosures is related to both current and future firm profitability, which is consistent with the notion that the linguistic sentiment of a corporate disclosure conveys the managers' private information about the expected future performance of the firm.

A common feature of current summary statistics for linguistic sentiment is that they assign equal weights to the intratextual sentiment. Such an approach is suboptimal under the joint hypothesis that linguistic sentiment is not uniformly distributed within the text and that the informativeness of linguistic sentiment depends on the position within the text. We test these two hypotheses on the CEO letters of the Dow Jones Industrial Average constituents over the period 2000–2011.

Regarding the intratextual dynamics in sentiment, we are the first to show empirical evidence of the presence of a U-shape (resp. decreasing smirk) in the intratextual frequency of positive (resp. negative) sentiment. The intratextual net sentiment shows an increasing smirk. This result is as expected, since according to the serial position effect, readers recall information better when it is presented first (primacy) or last (recency) in a vector of words, rather than in the middle. Consistent with the hypothesis of impression management and optimism of corporate managemnet, CEO letters are thus carefully crafted documents in order to give a positive impression to the reader.

Since intratextual dynamics in sentiment are non-uniform, it is relevant to investigate the accuracy gains that can be obtained by weighting intratextual sentiment when aggregating it into a total sentiment measure used to predict future firm performance. We test this second hypothesis by first proposing a methodology for weighted sentiment measurement and then applying it to the panel of 2000-2011 DJIA CEO letters.

The proposed position-sentiment measurement framework consists of replacing the equal weight design in linguistic sentiment with a flexible weighting scheme that is optimized to predict future firm performance. A convenient threshold least squares estimator is proposed to optimize the weights linking the intratextual sentiment to future firm performance. The threshold is needed to distinguish the letters with negligible sentiment dynamics from the majority of the texts showing significant intratextual dynamics in sentiment. The optimization is done on rolling estimation samples to avoid look-ahead bias.

When modeling the weights as a function of the position in the text, we find that for our sample of 342 CEO letters the optimized sentiment measure significantly outperforms the standard equally-weighted measure in terms of explaining future firm performance. This result is robust to the inclusion of all types of control variables.

We have applied our framework and based our conclusions on the sample of CEO letters of the Dow Jones Industrial Average constituents over the period of 2000 to 2011. An important direction for future research is to test our two hypotheses and apply the proposed optimized sentiment measurement framework on other types of corporate communication tools, such as earnings press releases or forward-looking statements in corporate filings.

References

Abrahamson, E. and E. Amir (1996). The information content of the president's letter to shareholders. *Journal of Business Finance & Accounting 23*, 1157–1182.

Almon, S. (1965). The distributed lag between capital appropriations and expenditures. *Econometrica 33*, 178–196.

Andersen, T., T. Bollerslev, and J. Cai (2000). Intraday and interday volatility in the japanese stock market. *Journal of International Financial Markets, Institutions and Money 10*, 107–130.

Baddeley, A. and G. Hitch (1977). *Recency reexamined.* In S. Dornic (Eds.) Attention and performance vol. VI, pp. 647–667. Hillsdale, NJ: Lawrence Erlbaum and Associates Inc.

Barber, B. and J. Lyon (1997). Detecting long-run abnormal stock reutnrs: The empirical power and specification of test statistics. *Journal of Financial Economics 43*, 341–371.

Baumeister, R., E. Bratslavsky, C. Finkenauer, and K. Vohs (2001). Bad is stronger than good. *Review of General Psychology 5*, 323–370.

Boudt, K., C. Croux, and S. Laurent (2011). Robust estimation of intraweek periodicity in volatility and jump detection. *Journal of Empirical Finance 18*, 353–369.

Bowman, E. (1984). Content analysis of annual reports and alchemy. *Interfaces 14*, 61–71.

Clatworthy, M. and M. Jones (2003). Financial reporting of good news and bad news: Evidence from accounting narratives. *Accounting and Business Research 33*, 171–185.

Clatworthy, M. and M. Jones (2006). Differential reporting patterns of textual characteristics and company performance in the chairman's statement. *Accounting and Business Research 19*, 493–511.

Core, J., R. Holthausen, and D. Larcker (1999). Corporate governance, chief executive officer compensation and firm performance. *Journal of Financial Economics 51*, 371–406.

Courtis, J. (1982). Private shareholder response to corporate annual reports. *Accounting and Finance 22*, 53–72.

Davis, A., J. Piger, and L. Sedor (2012). Beyond the Numbers: Measuring the Information Content of Earnings Press Release Language. *Contemporary Accounting Research 29*, 845–868.

Davis, A. and I. Tama-Sweet (2012). Managers' use of language across alternative disclosure outlets: Earnings press releases versus MD&A. *Contemporary Accounting Research 29*, 804–837.

Deese, J. and R. Kaufman (1957). Serial effects in recall of unorganized and sequentially organized verbal material. *Journal of experimental psychology 54*, 180–187.

Demers, E. and C. Vega (2010). Soft information in earnings announcements: News or noise? *Working paper, INSEAD*.

Diener, E., D. Wirtz, and S. Oishi (2001). End effects of rated life quality: The James Dean effect. *Psychological Science 12*, 124–128.

Diffenbach, J. and R. Higgins (1987). Strategic credibility can make a difference. *Business Horizons 30*, 13–18.

Do, A., A. Rupert, and G. Wolford (2008). Evaluations of pleasurable experiences: The peak-end rule. *Psychonomic Bulletin & Review 15*, 96–98.

Ebbinghaus, H. (1885). *Über das gedächtnis: untersuchungen zur experimentellen psychologie*. Duncker & Humblot.

Edmundson, H. (1969). New methods in automatic extracting. *Journal of the ACM 16*, 264–285.

Engelberg, J. (2008). Costly Information Processing: Evidence from Earnings Announcements. *Working paper, Northwestern University*.

Fairfield, P., S. Whisenant, and T. Yohn (2003a). Accrued earnings and growth: Implications for future profitability and market mispricing. *Accounting Review 78*, 353–371.

Fairfield, P., S. Whisenant, and T. Yohn (2003b). The differential persistence of accruals and cash flows for future operating income versus future profitability. *Review of Accounting Studies 8*, 221–243.

Fama, E. and K. French (1995). Size and book-to-market factors in earnings and returns. *Journal of Finance 50*, 131–155.

Fama, E. and K. French (2001). Disappearing dividends: Changing firm characteristics or lower propensity to pay. *Journal of Financial Economics 60*, 3–43.

Fama, E. and K. French (2006). Profitability, investment and average returns. *Journal of Financial Economics 82*, 491–518.

Frase, L. (1969). Paragraph organization of written materials: The influence of conceptual clustering upon the level and organization of recall. *Journal of Educational Psychology 60*, 394–401.

Fredrickson, B. and D. Kahneman (1993). Duration neglect in retrospective evaluations of affective episodes. *Journal of Personality and Social Psychology 65*, 45–55.

Freebody, P. and R. Anderson (1986). Serial position and rated importance in the recall of text. *Discourse Processes 9*, 31–36.

Glanzer, M. and A. Cunitz (1966). Two storage mechanisms in free recall. *Journal of Verbal Learning And Verbal Behavior 5*, 351–360.

Heaton, J. (2002). Managerial optimism and corporate finance. *Managerial Optimism and Corporate Finance 31*, 33–45.

Henry, E. (2008). Are investors influenced by how earninigs press releases are written? *Journal of Business Communications 45*, 363–407.

Kahneman, D., D. Fredrickson, C. Schreiber, and D. Redelmeier (1993). When more pain is preferred to less: Adding a better end. *Psychological Science 4*, 401–405.

Kieras, D. (1980). Initial mention as a signal to thematic content in technical passages. *Memory And Cognition 8*, 345–353.

Kohut, G. and A. Segars (1992). The president's letter to stockholders: An examination of corporate communication strategy. *Journal of Business Communication 29*, 7–21.

Kravet, T. and V. Muslu (2013). Textual risk disclosures and investors' risk perceptions. *Review of Accounting Studies 18*, 1088–1122.

Lee, T. and D. Tweedie (1975). Accounting information: An investigation of private shareholder usage. *Accounting and Business Research 5*, 280–291.

Li, F. (2010). The Information content of forward looking statements in corporate filings – A naive bayesian machine learning approach. *Journal of Accounting Research 48*, 1049–1102.

Loughran, T. and B. McDonald (2011). When is a liability not a liability? Textual analysis, dictionaries, and 10-ks. *Journal of Finance 66*, 35–65.

Malmendier, U. and G. Tate (2005). CEO overconfidence and corporate investment. *Journal of Finance 60*, 2661–2700.

Meyer, B. and G. McConkie (1973). What is recalled after hearing a passage? *Journal of Educational Psychology 65*, 106–117.

Patelli, L. and M. Pedrini (2013). Is the optimism in CEO's letters to shareholders sincere? Impression management versus communicative action during the economic crisis. *Journal of Business Ethics*, 1–16.

Pearson, P., J. Hansen, and C. Gordon (1979). The effect of background knowledge on young children's comprehension and implicit information. *Journal of Reading Behaviour 11*, 201–209.

Price, S., J. Doran, D. Peterson, and B. Bliss (2012). Earnings conference calls and stock returns: The incremental informativeness of textual tone. *Journal of Banking & Finance 36*, 992 – 1011.

Roediger, H. and R. Crowder (1976). A serial position effect in recall of United States presidents. *Bulletin of The Psychonomic Society 8*, 275–278.

Rozin, P. and E. Royzman (2001). Negativity bias, negativity dominance and contagion. *Personality and Social Psychology Review 5*, 296–320.

Sadique, S., F. In, and M. Veeraraghavan (2008). The impact of spin and tone on stock returns and volatility: Evidence from firm-issued earnings announcements and the related press coverage. *Working paper, Monash University*.

Sloan, R. (1996). Do stock prices fully reflect information in accruals and cash flows about future earnings? *Accounting Review 71*, 289–315.

Tetlock, P. (2007). Giving content to investor sentiment: The role of media in the stock market. *Journal of Finance 62*, 1139–1168.

Tetlock, P., M. Saar-Tsechansky, and S. Macskassy (2008). More than words: Quantifying language to measure fims' fundamentals. *Journal of Finance 63*, 1437–1467.

Varey, C. and D. Kahneman (1992). Experiences extended across time: Evaluation of moments and episodes. *Journal of Behavioral Decision Making 5*, 169–195.

Yuthas, K., R. Rogers, and J. Dillard (2002). Communicative action and corporate annual reports. *Journal of Business Ethics 41*, 141–157.

8 Appendix

Table 6: **Equally-weighted CEO net sentiment by DJIA firm between 2000 and 2011**

Firm	Industry	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	Mean
AE	4	0.016	0.012	0.017	0.031	0.029	0.022	0.021	0.012	0.009	0.021	0.018	0.022	0.019
AIG	4					0.018	0.011	0.023	0.011					0.016
Alcoa	1	0.012	0.004	0.022	0.015	0.014	0.012	0.033	0.008	0.011	0.010	0.023	0.016	0.015
Altria	3	0.011	0.015	0.010	0.014	0.009	-0.002	0.016	0.019					0.011
ATT	8	0.005	0.012	0.013	0.021	0.032	0.027	0.017	0.023	0.028	0.028	0.020	0.016	0.020
BoA	4									0.007	0.008	0.018	0.012	0.011
Boeing	6	0.022	0.009	0.013	0.010	0.006	0.030	0.023	0.028		0.022	0.021	0.030	0.019
Caterpillar	6	0.046	0.026	0.011	0.024	0.035	0.039	0.027	0.020	-0.000	0.026	0.017	0.020	0.024
Chevron	1									0.013	0.028	0.026	0.029	0.024
Cisco	8										0.015	0.040	0.027	0.027
Citi	4	0.021	0.018	0.029	0.029	0.013	0.020	0.021	0.026	0.002				0.020
Coca	3	0.019	0.029	0.039	0.030	0.026	0.029	0.023	0.042	0.025	0.029	0.027	0.021	0.028
Disney	7	0.016	0.016	0.020	0.021	0.020	0.035	0.034	0.039	0.027	0.021	0.030	0.034	0.026
Dupont	4	0.013	0.019	0.023	0.030	0.027	0.029	0.029	0.027	0.006	0.018	0.036	0.043	0.025
Exxon	1	0.047	0.036	0.047	0.042	0.045	0.031	0.034	0.035	0.030	0.038	0.042	0.038	0.039
GE	6	0.005	0.010	0.019	0.019	0.024	0.024	0.026	0.022	0.013	0.016	0.021	0.022	0.018
HD	3	0.030	0.040	0.030	0.035	0.036	0.027	0.022	0.030	0.007	0.030	0.027	0.033	0.029
Honeywell	6	0.018	0.017	0.020	0.020	0.025	0.033	0.033	0.025					0.024
HP	1	0.028	0.013	0.021	0.021	0.034	0.025	0.037	0.018	0.029	0.024	0.031	0.014	0.025
IBM	8	0.005	0.006	0.009	0.017	0.016	0.030	0.022	0.020	0.010	0.016	0.022	0.028	0.017
Intel	8	0.025	0.006	0.024	0.029	0.026	0.023	0.022	0.029	0.028	0.004	0.020	0.020	0.021
IPC	3	0.025	0.023	0.043	0.027									0.030
J&J	5	0.021	0.031	0.030	0.031	0.030	0.023	0.026	0.020	0.027	0.016	0.011	0.024	0.024
JPM	4	0.031	0.022	0.028	0.043	0.020	0.018	0.017	-0.007	-0.017	-0.004	-0.008	-0.008	0.011
Kodak	8	0.016	0.018	0.025	0.028									0.022
McDo	7	0.027	0.026	0.031	0.033	0.036	0.031	0.047	0.052	0.045	0.049	0.035	0.044	0.038
Merck	5	0.034	0.027	0.022	0.030	0.019	0.027	0.017	0.020	0.024	0.027	0.024	0.027	0.025
MMM	2	0.041	0.031	0.049	0.035	0.038	0.034	0.021	0.028	0.005	0.008	0.028	0.023	0.028
MSFT	8	0.028	0.004	0.023	0.042	0.035	0.032	0.037	0.036	0.039	0.028	0.034	0.025	0.030
Pfizer	5					-0.014	0.007	0.030	0.027	0.011	0.014	0.024	0.018	0.015
P&G	3	0.034	0.032	0.027	0.035	0.028	0.029	0.037	0.027	0.049	0.035	0.037	0.032	0.034
Travelers	4										0.020	0.000	0.013	0.011
UTC	6	0.009	0.009	0.014	0.015	0.017	0.017	0.009	0.030	0.028	0.031	0.033	0.041	0.021
Verizon	8					0.029	0.018	0.023	0.027	0.020	0.007	0.024	0.025	0.022
Walmart	7	0.020	0.024	0.013	0.017	0.028	0.028	0.030	0.032	0.036	0.030	0.028	0.034	0.027
Mean		0.022	0.019	0.024	0.026	0.024	0.024	0.026	0.025	0.019	0.021	0.024	0.025	0.023

Industry numbers – 1: Basic Materials, 2: Conglomerates, 3: Consumer Goods, 4: Financial, 5: Healthcare, 6: Industrial Goods, 7: Services, 8: Technology, 9: Utilities,

Note: This table presents the equally-weighted CEO net sentiment by firm and by year. The sample tracks the DJIA index between 2000 and 2011. A blank indicates that either the firm was not included in the DJIA that year or that the CEO letter is missing. Firms are included only if at least one annual report is filed at the SEC between 2000 and 2012. In the case of GM and Kraft, no annual report SEC filing date is available between 2000 and 2012.

Table 7: **Variable definitions**

Panel A – Compustat/CRSP items used to compute the dependent and control variables in Panel B

Item	**Item name**	**Compustat/CRSP item #**
$CSTI_{j,t}$	Cash and short term investments	#1
$CA_{j,t}$	Current assets	#4
$CL_{j,t}$	Current liabilities	#5
$A_{j,t}$	Total assets	#6
$A_{j,q,t}$	Total assets at the end of quarter q	#6
$Y_{j,t}$	Net income before extraordinary items	#18
$Y_{j,q,t}$	Net income before extraordinary items at the end of quarter q	#18
$CSHO_{j,t}$	Common shares outstanding	#25
$Div_{j,t}$	Total dividends per share by ex date	#26
$DCL_{j,t}$	Debt in current liabilities	#34
$L_{j,t}$	Total liabilities	#181
$P_{j,t}$	Closing price fiscal year t	#199
$B_{j,t}$	Book value of equity	$A_t - L_t$
$Prc_{d,j,t+1}$	Price on day d, year t	CRSP

Panel B – Dependent and control variables

Variable	Expected Sign	**Variable name**	**Definition**
$ROA_{j,t+1}$		The sum of quarterly Y over the four quarters after the SEC filing quarter q, divided by $A_{j,q,t+1}$ at the end of quarter q	$\sum_{i=1}^{4} Y_{j,q+i,t+1}/A_{j,q,t+1}$
$ROA_{j,t}$	+	Return on assets over fiscal year t	Y_t/A_{t-1}
$\widehat{ROA}_{j,t-i}$		Average return on assets over the five preceding years	$\frac{1}{5} \cdot \sum_{i=0}^{4} ROA_{j,t-i}$
$\sigma_{ROA,j,t}$	+	Standard deviation of return on assets over the five preceding years	$\sqrt{\frac{1}{5} \cdot \sum_{i=0}^{4} \left(ROA_{j,t-i} - \widehat{ROA_{j,t-i}}\right)}$
$NegY_{j,t}$	-	Negative earnings dummy	$I\left[Y_{j,t} \leq 0\right]$, else 0
$MC_{j,t}$	+	Market capitalization (in bil.$)	$P_{j,t} \cdot CSHO_{j,t}$
$BTM_{j,t}$	-	Book to market ratio	$B_{j,t}/MC_{j,t}$
$AC_{j,t}$		Accruals	$\Delta CA_{j,t} - \Delta CSTI_{j,t} - \Delta CL_{j,t} + \Delta DCL_{j,t}$
$+AC_{j,t}$	-	Positive accruals	$I\left[AC_{j,t} > 0\right] \cdot AC_{j,t}$, else 0
$-AC_{j,t}$	-	Negative accruals	$I\left[AC_{j,t} \leq 0\right] \cdot AC_{j,t}$, else 0
$D_{j,t}$	+	Total dividends	$Div_{j,t} \cdot CSHO_{j,t}$
$Ret_{j,t}$	+	Return over fiscal year t	$(P_{j,t} - P_{j,t-1})/P_{j,t-1}$
$Ret_{j,t-2}$	+	Two-year return for the years up to the end of fiscal year $t-1$	$(P_{j,t-1} - P_{j,t-3})/P_{j,t-3}$
$Ret_{j,t*}$	+	Return between the end of fiscal year t and the earnings release on day d of fiscal year $t+1$	$(Prc_{d,j,t+1} - P_{j,t})/P_{j,t}$

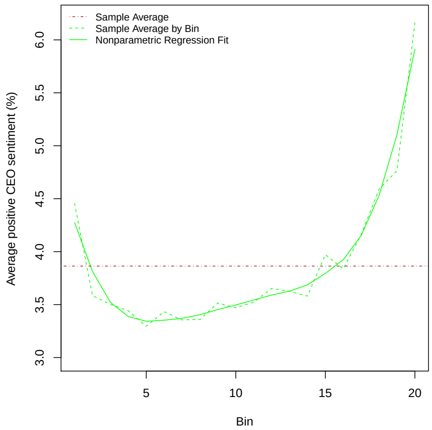
Description of control variables We include firm size MC, which is measured as the natural logarithm of market value of equity (Compustat item #25 $\cdot$ #199) at the end of the fiscal year. We expect smaller firms to be less profitable (Fama and French, 1995). The book-to-market ratio $BTM_{j,t}$ is known to be negatively related to profitability (firms with lower $BTM_{j,t}$ tend to be more profitable). We define book-to-market as the book value of equity (#6-#18), divided by $MC_{j,t}$. There is also evidence that accruals forecast profitability (Fairfield, Whisenant, and Yohn, 2003a,b; Sloan, 1996). We distinguish between positive accruals ($+AC$) and negative accruals ($-AC_{j,t}$), each scaled by the book value of equity. Previous work also shows that dividend-paying firms tend to be more profitable (Fama and French, 2001). We include the ratio of dividends to book equity ($D_{j,t}$). Dividend is defined as the number of shares outstanding (#25), times the total dividend per share (#26).

We also include past firm and stock market profitability as control variables. The return on assets ($ROA_{j,t}$) is measured as the earnings before extraordinary items at the end of fiscal year t, scaled by the total assets at the beginning of the year. The $ROA_{j,t}$ coefficient is predicted to be positive and lower than 1, consistent with prior research documenting mean reversion in performance metrics Barber and Lyon (1997). We also include a dummy variable $NegY_{j,t}$ for negative earnings in fiscal year t. Based on Fama and French (2006), we predict that current stock market performance is positively related to future firm performance. We define $Ret_{j,t}$ as the firm's stock return for fiscal year t and $Ret_{j,t-2}$ as its combined return for years $t-1$ and $t-2$.

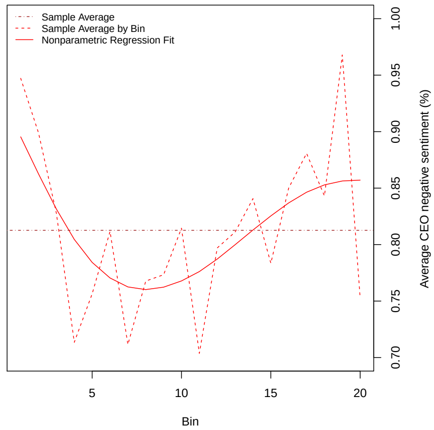
Finally, to complete our model, we add two variables to those introduced by Fama and French (2006). $Ret_{j,t}$ measures a firm's return up to the end of fiscal year t. There is, however, a gap between the end of the fiscal year t and the filing of the annual report at the SEC. For this reason, we compute the return between the end of the fiscal year and the filing of the annual report at the SEC. This return is typically measured in fiscal year $t+1$ and is denoted $Ret_{j,t+1}$. Based on Core, Holthausen, and Larcker (1999), we also introduce a variable $\sigma_{ROA,j,t}$ to capture firm risk as a control in our model. It is defined as the standard deviation of $ROA_{j,t}$ over the preceding five years. We expect a positive relation between future firm performance and past volatility.

Figure 6: **Intratextual dynamics of CEO sentiment following the general purpose library Diction 7.0 to identify positive and negative words**

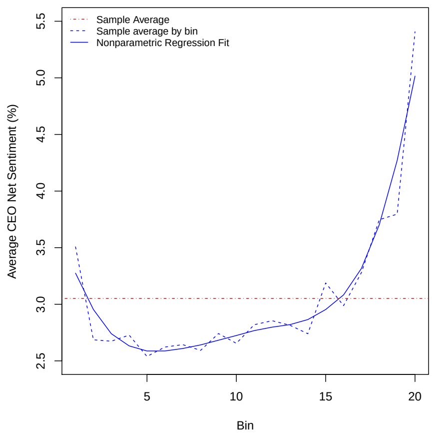
Note: This figure depicts the dynamics of CEO sentiment within letters to shareholders. The length of each text is standardized to correspond to a $[0, 1]$ interval, which is divided in B bins such that each bin contains the same number of total words. For each bin, the percentage number of positive words out of the total number of words in each bin is reported (Figure 6a). Similarly, for each bin, the percentage number of negative words out of the total number of words in each bin is computed (Figure 6b). For each bin, the net sentiment is then measured as the spread between the positive and negative tone (Figure 6c). Positive and negative tones are measured based on the Diction 7.0 word lists.



(a) U-shape in average positive sentiment by bin



(b) Decreasing smirk in average negative sentiment by bin



(c) Increasing smirk in average net sentiment by bin

Figure 7: **Intratextual dynamics of CEO sentiment following the financial library of Abrahamson and Amir (1996) to identify positive and negative words**

Note: This figure depicts the dynamics of negative CEO sentiment within letters to shareholders. The length of each text is standardized to correspond to a $[0, 1]$ interval, which is divided in B bins such that each bin contains the same number of total words. For each bin, the percentage number of negative words out of the total number of words in each bin is reported. Negative sentiment is measured based on the Abrahamson and Amir (1996) negative word list.

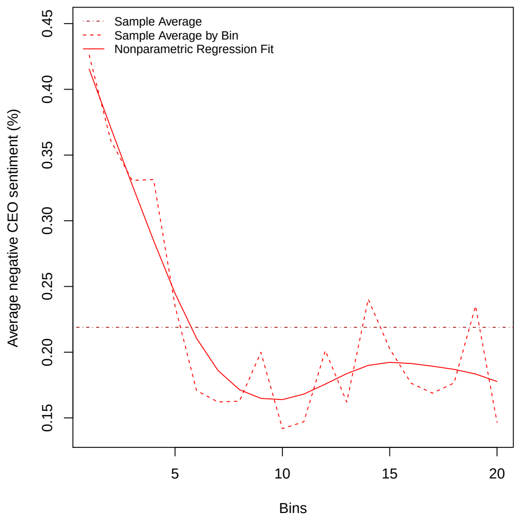


Table 8: **Position-weighted net sentiment by DJIA firm between 2000 and 2011**

Firm	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	Mean
AE				**0.028**	0.029	**0.021**	0.021	**0.013**	**0.007**	**0.020**	**0.017**	0.022	0.020
AIG							**0.021**	**0.012**					0.016
Alcoa				**0.015**	**0.014**	**0.009**	**0.031**	**0.008**	**0.004**	0.010	**0.020**	**0.018**	0.014
Altria				**0.008**	**0.004**	**0.002**	**0.014**	**0.016**					0.009
ATT				**0.023**	**0.027**	**0.025**	0.017	**0.025**	**0.026**	0.028	0.020	0.016	0.023
BoA											0.018	**0.011**	0.014
Boeing				**0.006**	**0.008**	**0.031**	**0.020**	**0.026**		0.022	**0.017**	**0.030**	0.020
Caterpillar				**0.022**	0.035	**0.036**	**0.021**	**0.019**	**-0.010**	**0.025**	**0.015**	0.020	0.020
Chevron											**0.024**	**0.032**	0.028
Cisco												**0.024**	0.024
Citi				**0.028**	**0.010**	**0.020**	**0.019**	**0.027**	**-0.003**				0.017
Coca				**0.034**	**0.027**	**0.031**	**0.019**	**0.038**	**0.027**	**0.031**	**0.029**	0.021	0.029
Disney				0.021	0.020	**0.033**	**0.032**	**0.039**	**0.029**	**0.023**	**0.03**	0.034	0.029
Dupont				**0.034**	**0.028**	**0.030**	**0.027**	**0.023**	**0.013**	**0.017**	**0.031**	**0.045**	0.027
Exxon				**0.046**	**0.041**	**0.029**	**0.037**	**0.027**	**0.029**	**0.037**	**0.039**	**0.039**	0.036
GE				**0.02**	0.024	**0.023**	0.026	**0.025**	**0.014**	**0.019**	0.021	**0.020**	0.021
HD				**0.037**	**0.036**	**0.026**	**0.026**	**0.035**	**0.019**	**0.035**	**0.024**	**0.029**	0.03
Honeywell				**0.020**	**0.024**	**0.031**	**0.033**	**0.022**					0.026
HP				**0.019**	**0.029**	**0.026**	**0.035**	**0.018**	**0.034**	**0.027**	**0.03**	0.014	0.026
IBM				**0.015**	0.016	**0.031**	0.022	**0.022**	**0.013**	**0.013**	0.022	**0.026**	0.020
Intel				**0.027**	**0.024**	**0.021**	**0.020**	**0.027**	**0.03**	**0.003**	**0.020**	0.020	0.021
IPC				**0.023**									0.023
J&J				**0.031**	**0.030**	**0.019**	0.026	**0.017**	**0.027**	0.016	**0.013**	**0.024**	0.023
JPM				**0.045**	0.020	**0.016**	**0.016**	**-0.006**	**-0.020**	**-0.003**	-0.008	-0.008	0.006
Kodak				**0.026**									0.026
McDo				**0.039**	**0.040**	**0.029**	**0.044**	**0.052**	**0.047**	**0.053**	**0.039**	**0.042**	0.043
Merck				**0.030**	**0.015**	**0.021**	**0.013**	**0.017**	**0.024**	0.027	**0.018**	**0.024**	0.021
MMM				**0.032**	**0.034**	**0.035**	**0.020**	**0.025**	**0.004**	**0.007**	**0.026**	**0.029**	0.023
MSFT				**0.040**	**0.040**	**0.033**	**0.036**	**0.033**	**0.035**	**0.026**	**0.036**	**0.019**	0.033
Pfizer							**0.030**	**0.025**	**0.014**	0.014	**0.024**	**0.015**	0.020
P&G				**0.040**	**0.028**	**0.034**	**0.034**	**0.028**	**0.046**	**0.035**	**0.035**	0.032	0.035
Travelers												**0.013**	0.013
UTC				**0.015**	**0.014**	**0.021**	**0.005**	**0.030**	**0.027**	**0.031**	**0.034**	**0.037**	0.024
Verizon							**0.022**	**0.023**	**0.017**	**0.010**	**0.027**	0.025	0.021
Walmart				**0.025**	**0.027**	**0.028**	**0.03**	**0.031**	**0.035**	**0.028**	**0.029**	**0.039**	0.030
Mean				0.027	0.025	0.026	0.025	0.024	0.019	0.022	0.024	0.024	0.024

Note: This table presents the position-weighted net sentiment in CEO letters to shareholders by firm and by year. Values in bold are firm-year observations for which the intratextual dispersion of sentiment is higher than κ. The sample tracks the DJIA index between 2000 and 2011. A blank indicates that either the firm was not included in the DJIA that year or that the CEO letter is missing. The first three years of the sample are used to estimate the weights. Firms are included only if at least one annual report is filed at the SEC between 2000 and 2012. In the case of GM and Kraft, no annual report SEC filing date is available between 2000 and 2012.