UNIVERSIDADE CATÓLICA PORTUGUESA

How tone is used across CEO Letters: the impact of Financial Crisis

Evidence from the UK

Filipa Freire Ferreira

Católica Porto Business School
July 2021

UNIVERSIDADE CATÓLICA PORTUGUESA

How tone is used across CEO Letters: the impact of Financial Crisis

Evidence from the UK

Final Assignment in the modality of Dissertation presented to

Universidade Católica Portuguesa to obtain the master's degree in Finance

by

Filipa Freire Ferreira

under the guidance of Paulo Alexandre Pimenta Alves

Católica Porto Business School
July 2021

Acknowledgments

The conclusion of this thesis is a huge mark in my life. I am beyond happy for having gotten this far because not only did I learn a lot as it also gave me the chance to develop my critical thinking. During the process there were some days that I doubted myself and others were tiredness would come up and tell me that maybe I had to do what I had planned for the day, in the following day, but especially on those days I had my luck charms to motivate me and to lift me when I needed encouraging words.

To my brother, thank you so much for being who you are and for all the comfort that you have given me. I could not do this thesis without you and your helpful and wise insights.

To my amazing boyfriend, for being so supportive and for cooking me dinner when I did not even had time to sit. These gestures seem so simple but are so precious.

To my mom and dad, thank you so much for being my strength, for believing in me and for supporting me in everything that I propose myself to.

To my family, thank you so much for all your kindness and for always rooting for me.

This is such a fulfilling moment in my life, and I am extremely grateful for having had the opportunity to study at Católica Porto Business School and for all the amazing docents that taught me and that made me deepen my passion and interest for Finance.

I want to personally thank all my professors, especially, Professor Paulo Alves, the supervisor of my thesis for the availability, for the advice, and above all, for awakening in me the willingness to be a good and motivated learner and an outstanding professional. I have also to thank Professor Ricardo Ribeiro for all the availability and help that he has given me whenever I needed.

Lastly, I want to personally thank Carolina Leite, who at the beginning of the master's was an acquaintance and rapidly turned out to be one of the best and most supportive friends I could ever find. Thank you, these years by your side have been joyful.

To everyone who stood by me: thank you so much. I am genuinely grateful.

….

Resumo

Estudamos o tom presente nas cartas dos CEOs de empresas cotadas na *London Stock Exchange* (*LSE*), mais especificamente, definindo dois modelos alternativos para medir o tom: o tom da carta como um todo e o tom da carta dividida em secções, sendo que a percentagem escolhida para as secções 1, 2 e 3 da carta são, respetivamente, 10%, 80% e 10%. Com esta diferenciação de modelos, pretendemos analisar se existem diferenças entre o tom do texto como um todo e o tom presente em cada uma das secções. Numa segunda instância, pretendemos analisar se existem diferenças nas médias do tom entre as três secções da carta.

Adicionalmente, estudamos o impacto da crise financeira de 2008 através da análise de mudanças de certas caraterísticas textuais das cartas dos CEOs, nomeadamente, o tom, a incerteza e a causalidade. Para além disto, analisamos ainda se existe uma diferença na relação entre a performance atual de uma empresa e o tom reportado pela mesma em anos de crise face a anos de não crise que possa indiciar qual o tipo de contexto económico mais propício à adoção de técnicas de *impression management*.

Usando vários modelos de regressão linear, realizamos uma análise das caraterísticas textuais na carta como um todo e em cada uma das suas secções.

Existem diferenças entre as secções entre si e entre o tom das secções e o tom da carta como um todo. O tom é explicado pela crise. Embora as palavras de incerteza aumentem durante a crise, o mesmo não acontece para as palavras causais. Quanto à relação entre a performance atual da empresa e o tom reportado pela mesma, não existe nenhuma alteração significante entre anos de crise e não crise.

Palavras-chave: Crise Financeira, *Natural Language Processing*, tom, incerteza, causalidade, Impression Management

Abstract

We study the tone present in CEO Letters of firms listed in *London Stock Exchange* (*LSE*), more specifically, by defining two alternative approaches to assess tone: the CEO Letter as a whole and the weighted approach, where the percentages chosen for Sections 1, 2 and 3 are, 10%, 80% and 10%, respectively. By differentiating these two models, we intend to analyse if differences exist between the whole text tone and the tone present in each section. Secondly, we aim to understand if there are differences in the tone means within sections for the same CEO Letter.

Additionally, we study the impact of financial crisis by analysing the changes in some textual features of CEO Letters, more specifically, tone, uncertainty, and causality. Furthermore, we also analyse if there is a significant change in the relationship between firm's current performance and its reported tone in crisis versus non crisis settings that might help us conclude which scenario is more likely for firms to engage in impression management.

By using several linear regression models, we conduct an analysis of the textual features in CEO Letters as a whole and in its sections, individually.

There are indeed differences between the whole text tone and the tone of each section and, differences in the tone means within sections. Tone is explained by the financial crisis. While uncertainty words increase with Financial Crisis, we do not find any relevant link in the use of causal words. Regarding current firm's performance and tone relationship, we do not observe any significant change in this relationship from crisis to non-crisis years.

Keywords: Financial Crisis, *Natural Language Processing*, tone, uncertainty, causality, Impression Management

Table of contents

List of tables

1. Introduction

Natural Language Processing (NLP) is part of a field of artificial intelligence that focus on communications between humans and computers (Fisher, Garnsey and Hughes, 2016). Human communication can be imprecise and ambiguous, which creates an additional layer of complexity. This is particularly relevant when we are dealing with a massive amount of data and we need to automate the reading. The application of automated methods to process information makes it possible to analyse the content of the text, but also to extract discourse characteristics, such as sentiment, readability, and complexity.

Recent research in Accounting and Finance has suggested that tone delivers signals about the wealth of a given firm and its prospects. Bliss, Doran, Peterson and Price (2012) showed that the tone in conference calls is a predictor of abnormal returns and trading volume. Smith and Taffler (2000) show that tone reflects current firm performance and is also predictive of firm future performance (Davis, Piger and Sedor, 2012). Boudt, Thewissen and Torsin's (2018) results indicate that tone in earnings press releases is more informative and signals future performance for firms with higher information asymmetry, which explains the high market reaction of investors to qualitative information in such an opaque environment.

There has been an increasingly volume of sentiment-based research in the last few years, focusing on many different types of corporate disclosure vehicles, from Management and Discussing Analysis, earnings press releases to annual reports. Despite many supporting the view that managers disclose truthful information, others do recognize that tone in financial narratives is often biased and managed according to managers' incentives and goals. It has been shown in previous literature that firms with poorer results tend to use more complex

language and on the contrary, firms with persistent earnings usually recur to lower complexity (e.g., Li, 2008).

In this thesis, we compute textual metrics using a pre-constructed script that extracts sentiment based on finance-specific dictionaries. Our sample comprises 4,222 CEO Letters from 2005 to 2014. Chief Executive Officer's letter to shareholders (hereafter CEO letters) are a particular type of accounting narratives that are widely read as it contains key information for financial analysts and any other stakeholders (Aerts, Thewissen and Yan, 2019), thus CEO letters are important routine disclosure vehicles of communication as they help readers to understand how powerful corporate leaders make sense of the world. In addition, these types of financial narratives define the vision that management desire to share – or sometimes, impose on – those who read it. Thus, accordingly to Bournois and Point (2006) these documents can be a ''… a subtly revealing medium for understanding how a company works…'' as well as the image that it wants to portray to others.

To assess narratives' characteristics, we make a distinction between two measurement approaches: the whole text and the sections approach. While the former considers the CEO Letter as a whole and develops metrics for the total CEO Letter, the latter treats each section independently, giving rise to the same metrics as for the whole CEO Letter but for each section individually. This split is conducted to effectively capture the differences between sections of a given CEO Letter. Such a detailed and thorough analysis regarding narrative's features came to contribute to accounting and finance research by complementing Financial Statements' analysis and help investors in the decision-making process as quantitative information can often be hard to interpret.

In addition to the CEO Letter's tone, we also study the uncertainty and causality language. Uncertainty words aim to capture imprecision whereas causal words are used by firms to justify or explain performance or any relevant

event that might have affected it. While many recognize the importance of studying causal language, there is still a lack of research conducted regarding this topic and causal language usage.

Thus, we conduct this study with the hope to contribute to sentiment literature by analysing narrative's characteristics particularly during a very impactful event worldwide – the 2008 financial crisis.

The financial crisis had a huge impact in all over the world. It brought a lot of uncertainties regarding the firm's future performance and making market outlook turned out to be a quite challenging task. Babu, Imam and Tan (2019) studied which information is more relevant during economic crisis. Their results suggest that although all information plays an important role in analysts' decision-making process, the authors do recognize that financial information is more verifiable than non-financial information because the latter can be manipulated by investors' and identifying the potential bias is much more complicated.

Having in mind that non-financial information can be distorted and acknowledging that tone can be impacted depending on firm's reporting style, it becomes essential to test whether firm's discourse change as a response to the Financial Crisis. We expect tone to become more negative, or to decrease, and uncertainty and causal words to be more recurrent in CEO Letters published during the Financial Crisis. Furthermore, we also expect a change in the relationship between firm's performance and tone in crisis versus non-crisis periods, which is consistent with prior studies such as the one conducted by Patelli and Pedrini (2014) who state that normally firms have more incentives to distort information during non-crisis periods and therefore, it is expectable to observe a switch in the relationship mentioned earlier.

Our results show that the financial crisis had a significant impact in tone. Uncertainty words also become more frequent in 2008, 2009 and 2010, whereas

for the Causality words we could not find any significant change in the use of these words. As of the relationship between firm's performance and tone, we also could not find any strong evidence to conclude that Financial Crisis years are an important factor in shaping this relationship.

Moreover, with this study we contribute to the Accounting and Financial literature by developing an analysis across CEO Letters and by studying different sentiments. Additionally, we contribute to the work on automated content analysis, a study applied to the U.K., contrary to many studies that focus on the U.S. case specific.

The remainder of the thesis is organised as follows Chapter 2 refers to the literature review presenting a theoretical framework on financial narratives (more specifically, on CEO Letters), on the intratextual dynamics of CEO Letters and the Financial Crisis. Chapter 3 regards methodology, where we describe the process of sample selection and all the variables included in the models. Chapter 4 presents and discusses the results. Chapter 5 concludes the thesis.

2. Literature review

2.1. The insufficiencies of quantitative information

In the last few years, the nature of business has suffered enormous changes which consequently led to the evolution of the business reporting model to fit to the ever-changing information needs of the market and its participants. In 1973, AICPA stated that Financial Statements (FS) should not be limited to quantified information and that amplification in narrative form of data included in the statements should be required.

The reliability, transparency and uniformity of FS allows investors to make accurate decisions. However, corporate scandals (such as Enron and others) have deteriorated investors' confidence in the quality and veracity of the financial system (Rezaee, 2005).

Often, FS are subject to manipulation, earnings management, and fraud. According to Healy and Whalen's (1999) definition, earnings management occur when ''managers use judgement in financial reporting and structure transactions to change Financial Statements to either mislead stakeholders about the underlying economic performance of the company or to influence contractual outcomes that rely on reported accounting numbers''. Fraud and earnings management have the same objective, but the former is outside the Generally Accepted Accounting Principles (GAAP), while earning management is within GAAP. Fraud, using once again, Healy and Whalen's (1999) definition, occurs when managers use accounting practices that are not supported by GAAP to ''alter financial reports to either mislead some stakeholders about the underlying economic performance of the company or to influence contractual outcomes that rely on reported accounting numbers''.

Several studies have focused on the methods used by firms to manipulate their FS. As the name suggests, False Financial Statements (FFS) consists of overstating or understating the number presented in the FS. Prior literature sets a relationship between managerial behaviour and contractual rearrangements – such as compensation schemes, debt covenants as well as asset pricing, information asymmetry, agency and political costs (Francis, 2001; Lambert, 2001). Thus, managers may use discretionary accounting policies to change the company's performance aiming at, for example, affecting their remuneration (Young, 1998; Bushee, 2001), to transfer earnings from good-performance years to bad years (see DeFond and Park, 1997; Guidry, Leone and Rock, 1999) or to defer revenue recognition into future accounting periods to lower the current period's tax charge (Scholes, Wilson and Wolfson, 1992).

The need for additional information and the increasing volume of research conducted around financial narratives came to suggest that quantitative information presented in the FS have its insufficiencies and corporate scandals came to amplify concerns regarding the credibility of numerical financial information. Moreover, as stated by Huang, Zhang and Zheng (2014), quantitative information by itself provides an incomplete image to investors of a firm's economic condition. Understanding quantitative information is not a straightforward process, investors first need to encode the information and then process it (Fiske and Taylor, 1991). The sentiment employed in the qualitative sections of annual reports facilitates encoding and processing of the quantitative disclosures and generally, informs the reader.

2.2. Manual versus automated content analysis

Companies produce a vast variety of financial information that is included in their annual reports, containing both numerical and textual components.

Although qualitative research has attracted many attentions, the need to hand collect and then, manually scoring content constrained work in this field (El-Haj et al. 2019).

Text analysis has become quite popular in accounting and finance research due to textual analysis approaches – such as Information Extraction (IE), Natural Language Processing (NLP) and Corpus Linguistics which came to complement high-quality manual analysis and both methods represent symbiotic approaches (El–Haj, Rayson, Alves and Young, 2018; Alves et al. 2016).

NLP is a subdiscipline of Artificial Intelligence (AI) and has several machine learning methods such as term weighting, Naïve Bayes classification and cosine similarity. NLP methods can be complex, time-consuming, and hard to replicate, furthermore, it has a low level of sophistication in mainstream accounting and finance research. According to El-Haj et al. (2019), the well-known *Bag of Words* approach is the predominant model in the accounting and finance research. The *Bag of Words m*ethod relies on the critical assumption of independence to reduce the extraordinary dimensionality of a document and by independence, it means that the order, and consequently, the direct context of a word is not important (Loughran and McDonald, 2016; El–Haj et al., 2019).

As previously mentioned, Text Analysis which in most accounting and finance applications is based on measuring the tone of narrative an annual report or narrative disclosures (earnings press releases, chairman's letter, CEO letter, etc.) has been widely researched area. Before assessing the measurement of sentiment, it is necessary to decide which dictionary to use to evaluate the intratextual dynamics.

The dictionary helps to allocate words into positive, negative, or other categories (e.g.: uncertainty, forward-looking, causal, etc.) and once the total count of words positive (or negative) is done, it is normally scaled by the total number of words in the document. As stated by Loughran and McDonald (2015),

documents with a relatively high frequency of positive words are considered optimistic and likewise, those with a high frequency of negative words are labelled pessimistic.

Regarding the dictionaries, there are four different lists of words commonly used for accounting and finance research: Henry (2008), Harvard's GI (General Inquirer), Diction and Loughran and McDonald (2011).

There is empirical evidence that suggests that domain-specific dictionaries outperform general dictionaries (Li, 2010; Loughran and McDonald, 2011). A relevant aspect relies on the fact that Diction's optimistic and pessimistic word lists were not tailor-made to analyze financial documents. In fact, many of the users of the Diction's word list focused on political and not on business discourse. One major issue regarding Diction is its pessimistic words' list. More than 45% of all pessimistic words used in a large 10-K sample are the words *not* and *no* (these words are generally considered stop words in the context of business text). Thus, evidence suggests that a large fraction of Diction's optimistic and negative words are misclassified (Loughran and McDonald, 2015).

Prior literature has used Harvard's GI to assess the tone present in newspapers columns/articles or 10Ks, however, Loughran and McDonald (2011) documented that around 75% of the Harvard Dictionary negative words include *taxes, board, capital, liabilities,* and *mine.* In what concerns the Henry's Dictionary, it was the first business-specific dictionary to be created. Henry (2008) created a limited list of both positive and negative words to assess the tone of earnings press releases. While Henry's list has only 85 negative words, Diction has 920 and Loughran and McDonald has 2329 negative words. This suggests that relying exclusively on Henry's Negative word's list might lead to an incomplete assessment for managers to describe current or future operations (Loughran and McDonald, 2015).

2.3. Financial narratives

Gibbins, Richardson and Waterhouse (1990) define financial disclosures as ''…any deliberate release of financial information, whether numerical or qualitative, required or voluntary, or via formal or informal channel'' and defend that there are underlying economic incentives to disclose and not to disclose information. On the one hand, management wants to keep information for themselves as to exploit their inside-firm knowledge. On the other hand, managers might want to share information to enhance the firm's value and its accomplishments.

The importance of financial disclosures is widely consensual. Buchholz, Jaeschke and Lopatta (2018) consider them key elements in lowering information asymmetries between stakeholders (including shareholders) and the management team.

The annual reports constitute a key mandatory disclosure and are composed of two components: the mandatory financial statements and the financial narratives. U.S. registrants might make available, alongside its 10-K fillings, a brochure-type annual report in PDF format that infographics. The characteristics of 10Ks fillings' such as allowing for batch retrieval, plain text formatting and standardized structure facilitate research. However, outside the U.S. annual reports are glossy and unstructured and are almost exclusively presented in digital PDF. This lack of consistency among glossy annual reports gives managers enough discretion regarding the content disclosed, the labels of each section and the sequence in which the information is presented (Alves et al. 2016).

Regarding financial narratives, it normally contains commentary by senior management on the performance obtained during the reporting period, alongside additional information such as the letter to the shareholders and reviews of strategy, risk management, corporate governance, and executive

remuneration policy. Aerts, Thewissen and Yan (2019) recognize the crucial role that financial narratives play in today's business world in justifying and explaining FS. However, according to Alves et al. (2016), the informativeness of Narratives are expected to vary for different reasons. Firstly, there are different teams responsible for crafting each section of the annual reports, thus it is expected that linguistic variations occur and, consequently, it can affect the informativeness of the commentaries. Secondly, author's incentives likely vary with resulting implications for the content and predictiveness of the information (e.g., management-prepared content is much more likely to be positively biased when compared to sections that are written by outsiders). Lastly, dividing each section into different segments helps deliver clarity for identifying key aspects of the reporting entity (strategy, business operation, governance, etc.).

According to Alves et al. (2016), there is an association between exaggerated optimism and obfuscation in management commentary. If there is an overly optimistic tone present in this kind of narratives, it might signal that managers are obfuscating and omitting inside information.

Finally, the tone in narratives can be positive (reflecting good performance) or it might be signalling positive expectations regarding future performance that accounting numbers in quantitative disclosures cannot reflect mostly due to regulations. In both scenarios, the tone is a source of incremental information. According to Huang, Teoh and Zhang (2011) it can also represent an attempt for CEOs to hide poor results, often referred to as tone management.

2.4. CEO Letters

In this study, we focus on CEO Letters to contribute to financial narratives' literature in the accounting and finance literature.

CEO letters do not obey to a particular structure therefore it is expectable that when comparing these letters within different firms, differences will exist. However, we can expect them to cover some of the following themes (Bournois and Point 2006): the market, growth, strategic plans, product mix, imminent losses, future profits, confidence, the embellishment of the year's results and the expectations about the future performance.

Despite prior literature arguing that CEO letters may not be written by the CEO, his signature makes the letter itself the CEO's legal responsibility. Thus, even if other agents help to craft these letters, there is no doubt that CEOs have the legal responsibility for the content of their letters, and due to this, the letters are a useful medium for assessing the intentions of CEOs as they struggle to set a tone at the top (Amernic, Craig and Tourish, 2010).

CEO letters are very rich in terms of their vocabulary diversity and lexical variety (Wang et al., 2012) and are largely unregulated. This lexical variety should not be ignored as it can have a significant impact on how corporate performance is to be defined, measured and, finally how the firm's business model and direction is to be perceived (Aerts, Thewissen and Yan, 2019).

Prior literature has often neglected the ethical dimension of CEO letters which according to Patelli and Pedrini (2014) is quite concerning as most investors rely on these documents to make investment decisions. In addition, the fact that CEO letters are non-audited documents, increase the levels of curiosity and scrutiny around these narrative corporate disclosures and gives managers enough discretionary power to shape the message to be delivered (Boudt and Thewissen, 2019).

A very interesting question has been raised regarding what type of firms usually tend to publish a CEO letter. According to Meek, Robert and Gray (1995) one major motive for firms publishing CEO letters is the constant growth of the company and the need to suppress information asymmetries which will decrease

the cost of capital. From Costa et al.'s (2013) point of view, when managers decide to disclose information, they are acting in accordance with the elements of agency theory. Agency theory assumes that providing (voluntarily) information reduces the agency costs between parties thus managers have more than enough motives for disclosing information.

Costa et al. (2013) analyzed the incentives behind the publication of CEO letters and explained why companies in code-law oriented Latin Countries (i.e., Portugal) published a CEO letter. They found that companies with high levels of profitability and with a high number of foreign subsidiaries are more likely to include a CEO letter in their annual report. Additionally, they state that the likelihood of a firm publishing a CEO letter is much larger if it is audited by a Big 4 or if it is a financial firm whereas non-financial firms are less likely to publish.

2.5. Impression management

Despite prior literature arguing that CEO letters contain valuable information to predict future performance (Davis, Piger and Sedor, 2012), there is increasing empirical evidence that these communication channels are sometimes misused by managers, who take advantage of their inside-information firm knowledge to influence stakeholders through many impression management techniques (Boudt and Thewissen, 2019).

Impression Management can be seen as the tendency for organizations and their management to use and to some extent, to even manipulate information to present themselves in a favourable light (Clatworthy and Jones, 2006). However, the implementation of these practices delivers a signal that financial disclosures are biased which reduces the confidence and credibility of the information behind it (Heaton, 2002; Clatworthy and Jones, 2003).

There are many forms of engaging in Impression Management techniques and therefore, change the expectations of third parties whether it is by selecting only positive information, by choosing which financial measures and ratios to highlight or simply by withholding negative news information (Boudt and Thewissen, 2019). The authors focus their research on the strategic positioning of positive and negative words within a CEO letter and consider it as a subtle form of Impression Management. Characteristics such as poor readability and manipulation of linguistic features are also a source of Impression Management (Merkl-Davies and Brennan, 2007), however, prior literature states that from a behavioural perspective, language usage in financial disclosures can be subject to unintentional biases because of managers' specific traits - e.g.: gender, academic background, experience (Liu and Nguyen, 2019).

Yuthas, Rogers and Dillard (2002) challenge the corporate disclosure literature based on impression management and proposes a different approach based on communicative action. According to Habermas (1984, 1987) there are two components of social action, namely strategic action, and communicative action. The former states that one in society engages in communication driven by self-serving behaviour whereas the latter seeks mutual understanding. While through the lens of impression management, voluntary narrative disclosures are seen as strategic action that is undertaken to influence third parties' perspectives; through the lens of Theory of Communicative Action, these are seen as Communicative Action undertaken to achieve mutual understanding.

Despite manager's desire to deliver a wealthy image of the firm, past literature has also documented that CEOs have incentives to provide high-quality information to maintain organizational legitimacy. In addition, CEO letters are routine disclosure vehicles and if their content were distorted frequently this would jeopardize the firm's credibility thus, it is more likely that CEOs engage

in impression management with non-routine disclosure vehicles due to their extraordinary nature (Liu and Nguyen, 2019).

Clatworthy and Jones (2003) study whether firms with improving and declining performance report good news and bad news in different ways. For conducting this study, they based themselves on the chairman's statement, which is one of the most widely read accounting narratives. They concluded that companies with improving performance concentrate on good news rather than bad news whereas declining performers do not discuss either analyze the motives behind their poor performance.

Some research around narratives has focused on the relationship between these narrative disclosures and future performance. Patelli and Pedrini (2014) study the association between firm performance and the rhetorical features of CEO letters in a large sample of Fortune 500 firms in a context of the global economic crisis and found that optimistic tone is higher in firms with better performance than in declining performance firms. Due to this result, the authors conclude that the tone of the CEO letters is congruent with performance, indicating that there is an adherence to the sincerity principle of discourse ethics and that managers try to preserve organizational legitimacy.

Che, Zhu and Li (2020) study the sentiment present in CEO letters regarding one main theme, corporate social responsibility (CSR) and explain the effectiveness of CEO letters on forecasting a firm's financial performance. Among various machine learning approaches, through a logistic regression it is possible to predict with an accuracy of 70.46% the financial performance. Results suggest that the sentiment within these letters is a vital factor for depicting future performance. Moreover, this result suggests that there is a linear relationship between sentiment and economic performance.

2.6. Tone at the top

CEOs are the active agents of setting the tone at the top, more particularly in placing a strong value on the creation of unified corporate cultures.

Tone at the top can be defined as the set of shared values that emanates from the most senior executives of an organization (Cunningham, 2005).

The concerns around tone at the top increased exponentially with major company crisis and more particularly, with the 2008 financial crisis. These events draw attention to dysfunctional practices among organizations which led to multiple collapses that destroyed prosperity in the global economy and ultimately led to a whole new relationship between the State and the Financial system in many countries. (Amernic, Craig and Tourish, 2010).

According to accounting and finance literature, tone is not a constant metric within a CEO letter. Several studies were conducted to understand the intratextuality dynamics of CEO letters and the tone captured in it. Boudt and Thewissen (2015), studied the intratextual dynamics of CEO Letters and showed through their results that using an approach that considers the position in which words appear in a text has a greater fit in measuring the sentiment within CEO Letter than an approach that does not take this into consideration.

In CEO letters, not all sections give the same information value. Some provide more valuable information and for that, existing literature defends that the classical method of total textual sentiment obtained by equally weighting the intratextual sentiment becomes inaccurate. Boudt and Thewissen (2019) adopt a flexible weighting scheme that is optimized to predict future performance. The approached used by the authors significantly underweights the sentiment at the beginning and end of the text, compared to the sentiment in the middle of the text, as the first and latest parts normally contain a higher level of bias.

Boudt and Thewissen (2019) highlight the importance of the serial position effect of words. That is, readers tend to 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, 1993; Cunitz and Glanzer, 1966). Therefore, it is expectable that firms will inflate the perceived sentiment of CEO letters in which the most salient elements of the text are placed at the beginning and end of the text, while the more neutral elements are discussed in the middle.

However, since the end of the letter is recalled best, Boudt and Thewissen (2019) infer that the end of the letter will contain a larger number of positive words at the end than at the beginning, which leads them to conclude that textual positive sentiment within CEO letters is U-Shaped on average, with a peak in positive sentiment at the end of the text.

The pattern of intratextual frequency of CEO's negative sentiment is not as straightforward as the use of negative words is somehow a trade-off for the CEO. On the one hand, it is important for the CEO letter to be in line with information already disclosed to assist the reader's comprehension (Pearson, Hansen and Gordon, 1979), on the other hand, firms' managers want to maximize value and communicate positively to investors. Thus, it is expectable that the use of negative words will be higher at the beginning of the text due to the recency theory and peak-end rule, which states that investors will remember these words less after having read the entire text.

2.7. Financial crisis

The 2008 financial crisis led to enormous and catastrophic consequences globally. A massive decline in housing prices was the trigger for a full-blown liquidity crisis that emerged in 2007 and dragged for the following years.

The financial crisis led to the bailouts of many large uninsured financial institutions caused by their national governments and to sharp declines in stock prices, which were followed by smaller and more expensive loans for corporate borrowers (Brunnermeier, 2009; Thakor, 2012).

There are different views regarding the causes behind this crisis, according to Thakor (2012) it was a combination of both global macroeconomic factors and U.S. monetary policy that helped to create an environment where banks could enjoy a longer period of sustained profitability and growth. This growth was motivated by advances in information technology that boosted and turned a vast variety of securities marketable. However, all these innovative securities brought higher risks and, at some point in time, led to defaults that were not expectable, which caused distrust among investors and certainly, calling for a crisis (see Gennaioli, Shleifer and Vishny, 2012).

Despite prior literature mentioning Fair Value Accounting (FVA) as the contributor to the Financial Crisis (Magnan, 2009), there are distinctive views that do not see FVA as a significant factor that led to the 2008 Financial Crisis. Laux and Leuz (2010) do not believe that fair-value accounting contributed to U.S. banks' problems in the financial crisis in a relevant manner. According to them, fair values play only a limited role in banks' income statements and regulatory capital ratios except for a few banks with large positions. Finally, supporting the same line of thought, Shaffer (2010) defends that the capital destruction was mainly driven by deterioration in loan portfolios and was then depleted by proprietary trading losses and common stock dividends. Thus, according to the author, these are a result of bank management and its lending practices and not due to accounting rules as proposed by existing literature.

It has been discussed how frequent management explanations of corporate performance in Narratives are subject to a self-serving bias. Despite the large volume of research conducted regarding the truthfulness and sincerity behind

Narrative Disclosures, a few of them have focused on this issue during an economic downturn. There are divergent views regarding the incentives to distort information during a crisis. Patelli and Pedrini (2014) explored the sincerity of rhetorical tone in CEO letters in the wake of the global economic crisis. According to them, the costs of engaging in impression management during economic downturns are much higher and could put at stake the organizational legitimacy and CEO's reputation. These inherent higher costs are due to the social pressure that is imposed during economic downturns to obtain understanding and transparency. Additionally, the fact that negative news are highly expected during a global crisis reduces by far the incentives for distorting information provided in Financial Narratives.

A different view is shared by Bollen, Hassink and Keusch (2012) who give several contributions in the field of Financial Narratives, moreover, they developed a study that compares the narratives of the letters addressed to shareholders of Europe's most highly capitalized companies in crisis and non-crisis settings.

They find that crisis leads to a higher presence of self-serving bias as tough economic conditions are used by managers to present themselves in the best possible scenario. Furthermore, they found that the amount of positive news disclosed did not vary significantly from 2006 to 2008 which leads to conclude that actual performance does not have a particularly strong impact on management's decision to disclose positive news. Given the importance that CEO letters have in capital allocation, investors need to be aware of the potential misleading explanations of performance, particularly during a crisis.

Bollen, Hassink and Keusch (2012) also focused on the usage of linguistic features during financial crisis such as the use of defensive and acclaiming vocabulary and causal wording. They showed that acclaiming and defensive attributions are more frequently used in periods of crisis than in non-crisis

periods. Additionally, and given the significant use of defensive attributions, the authors concluded that companies in 2008 give significantly more explanations relating to the external environment than they did before the Financial Crisis. This last point leads us to conclude that external factors were used by managers as an opportunity to dissociate them from unfavourable outcomes and to highlight their own responsibility for positive results achieved without external influences.

2.8. Research question: the impact of distinct tone approaches and financial crisis on firms' discourse

Prior literature has proven that annual reports and their different individual sections, such as management commentary, are incrementally informative about a firm's performance (Abrahamson and Amir, 1996; Davis and Tama-Sweet, 2012; Patelli and Pedrini, 2014). However, each section has a different informativeness power and consequently, performing a simultaneous analysis of different tone measurement approaches can deliver more interesting results. Alves et al. (2016) identified three arguments in favour of the differences in the predictive ability of different sections of a given annual report. First, each section comprises different content and so, different purposes. Second, Financial Narratives are not written by the same person, which certainly affects the linguistic style reported due to the author's different characteristics, personality traits, background, experience, etc. (see e.g., Bertrand and Schoar, 2003; Argamon et al (2009); Liu and Nguyen, 2019). Lastly, the incentives of preparers also play a fundamental role in financial report's quality.

The objective of this research is to understand if a company's reported tone and discourse characteristics present in CEO letters vary significantly depending on the tone evaluation method used in both Financial Crisis and non-Financial

Crisis settings. Moreover, we pretend to assess whether there are differences in narratives' characteristics between the whole text approach and the sections approach. This research also aims to analyze if a firm's current performance relationship with tone changes significantly from financial crisis years to years of economic stability. Not only this study will assess tone in CEO letters as other speech characteristics such as uncertainty and causality.

As explained by prior literature, financial narratives are often misused by managers and other preparers of Financial Information, this means that on some occasions they might alter the structure of these financial disclosures to fit their goals. They do so in a very strategic manner, however research has identified some of these practices. Managers typically choose the position of words by setting a higher level of positive sentiment in the beginning and at the end – this is what we refer to as the Recency Effect. It has been shown in existing literature that readers recall best the beginning and end parts of the text.

To assess if CEO letters sections deliver different tone measures, we make the following hypothesis:

H1: The reported tone given by different sections of CEO letters is not equal.

According to Boudt and Thewissen (2019), the introduction part and conclusion parts portray a more positive tone when compared to the middle part of the text that is often neutral and displays lower readability, as this latter section normally discusses firm's performance.

Regarding, the different tone evaluation approaches, we assess if different methods lead to significant different tone metrics. More specifically, we compare two evaluation approaches: the whole text versus the sections method to assess tone. Typically, the total textual sentiment, also known as tone, is measured as the spread in the proportion of positive and negative words in a document (Demers and Vega, 2010; Davis et al., 2012, Davis and Tama-Sweet, 2012; Patelli

and Pedrini, 2014; Huang, Zhang and Zheng, 2014) or solely as the proportion of negative words (Abrahamson and Amir, 1996). However, one of the major consequences of nonuniform distribution of tone is that total sentiment measures that aggregate the intratextual sentiment without considering the position in the text may be suboptimal. Therefore, whenever firms engage in impression management with the objective of managing the narrative structure, the equal weighted approach will deliver biased tone measures. Particularly, whenever 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 tone. This position weighted needs to sum to unity and will lead to tone measures that are, on average, more pessimistic than the approaches based on equal weighting of intratextual sentiment (Boudt and Thewissen, 2019). Thus, we formulate the hypothesis regarding tone evaluation approaches:

H2: Whole text approach and sections approach deliver tone different results.

Existing literature focused on intratextual dynamics defend that equal weighted approach (also known as whole text approach) displays a more optimistic tone versus the latter approach. Similarly, equal weighted approach contains on average a more positive (optimistic) tone measure than the position weighted approach, which considers that not all parts of the text have the same relevance and thus, some parts must be underweighted in order not to bias the textual metrics.

Finally, regarding the impact of financial crisis on the reported tone, we develop the following hypothesis:

H3: The Financial Crisis has a negative impact on the tone in the CEO's Letter.

This third assumes that there is a positive relationship between a firm's current performance and its reported tone, which accordingly to prior literature (see e.g., Patelli and Pedrini, 2014) is a clue that firms adhere to the sincerity principle and thus, disclose accurate information, especially in periods of crisis.

A large volume of research has focused on the potential bias that CEO letters contain. Some state that managers act differently because each has his own characteristics, which in fact, can impact financial disclosure's quality. On the one hand, some question the sincerity and truthfulness behind Financial Narratives information, on the other hand, others defend that managers do not want to put at stake their own legitimacy and firm's legitimacy at risk, especially in periods of economic turbulence. As mentioned previously, the inherent costs of engaging in impression management are much higher during crisis periods, thus it is expectable that managers will choose to disclose truthful information in order not to lose investor's trust. Another reason that motivates managers to be sincere is that negative news are much more expected during economic downturns, thus their poor results (or poor management practices) are sort of diluted with the external factors. In line with this last point, we believe that managers on average will deliver more biased information during expansion economic cycles than during recession cycles. Thus, we formulate the following hypotheses:

H4a): The relationship between CEO's letter tone with firm's current performance (ROA) will change from financial crisis versus non-financial crisis periods.

H4b) The relationship between CEO's letter tone with changes in firm's current performance (Ch_ROA) will change from financial crisis versus non-financial crisis periods.

Moreover, these relationships are expected to be negative for non-crisis periods whereas it is expectable to be positive for financial crisis years, which is

consistent with prior literature that defends that managers tend to disclose more sincere information under tough economic conditions.

These last hypotheses assume that firms' reported tone depends also on external factors such as the wealth of the economy. Bollen, Hassink and Keusch (2012) results showed that from 2006 to 2008 there was not a significant change in the amount of positive news disclose, which indicates that actual performance does not have a strong influence on firm's management to disclose positive news. This, as cited by the authors, is clearly an indication of impression management.

As mentioned by existing literature, financial crisis brought a lot of insecurities and uncertainties globally regarding the functioning of the financial system. According to Loughran and McDonald (2011), uncertainty words are related to the notion of imprecision. Consequently, we expect the level of uncertainty words to increase with Financial Crisis on CEO's Letter, thus we formulate the following hypothesis:

H5: The Financial Crisis has a positive impact on the use of uncertainty words in the CEO's Letter.

Causal words are used when one aims at explaining or justifying a given event or outcome. According to Zhang and Aerts (2015), firms use this type of speech when explaining the reason why behind their performance and how it happened. Bollen, Hassink and Keusch (2012) showed that firms use more attribution (causal) speeches in periods of crisis. Moreover, results from Zhang and Aerts (2015) show that there is a significant positive relationship between the use of causal words and not meeting the earnings threshold. Thus, accordingly to past literature, we expect that Financial Crisis influences firm's to use more causal words to explain their performance, as such we make the following hypothesis:

H6: The Financial Crisis has a positive impact on the use of causal words in the CEO's Letter.

3. Methodology

3.1. Sample selection

The narratives data for the CEO letters were obtained from the CFIE Project (CFIE 2019) as described in (El-Haj, Young and Rayson, 2015), that generates a dataset of textual analysis. We collected financial data for UK listed companies between 2004 and 2018, however, we dropped observations prior to 2005 and after 2014 and excluded all financial firms.

Regarding the textual variables, we run a script to count the number of words, number of positive words and number of negative words, which enabled us to compute the tone variable for all CEO letters. In addition, we also counted the number of causality and uncertainty words present in each letter. After, we developed another approach to get tone metrics for each section of each text. We did so by dividing each text into 3 parts: section 1 (introduction), section 2 (development) and section 3 (conclusion) and chose a percentage to determine each part, more specifically, the first and last parts were attributed the same percentage whereas the middle part was given the reminiscent percentage. These percentages were determined by us and are open to debate.

Our final sample comprises 4,222 CEO letters for the periods from 2005 to 2014 published by non-financial firms, corresponding to 1,162 companies.

Table 1 presents all the criteria that we followed to get to our final sample and presents the distribution of our sample by firms and by years.

Table 1 – Sample selection method and the distribution of the final sample

Panel A: Sample selection criteria

		Reports
Population of CEO Letters between 2005 and 2014		11,712
Less		
Unprocessed CEO Letters	6,269	
Multiple CEO Letters during fiscal year	310	
Outliers	101	
Cleaning to equal observations for regressions	810	7,490
Final Sample		4,222

Panel B: Distribution of final sample by years

Count of obs. by year	Years	N	% of total
	2005	382	9.048
	2006	396	9.379
	2007	439	10.398
	2008	456	10.801
	2009	456	10.801
	2010	424	10.043
	2011	424	10.043
	2012	428	10.137
	2013	409	9.687
	2014	408	9.664
Total		4,222	100

As Table 1 shows, over 6,269 CEO Letters were not correctly processed and there were some years where we observed multiple reports for the same year, probably caused by a change in the fiscal year ending date. Moreover, due to the presence of outliers, we dropped the extreme bottom and top one percentiles for the following variables: number of words, ROA, change in ROA, Return, Earnings Yield, Market Value and Book-to-Market Ratio. After subtracting all the CEO Letters that did not meet our criteria, we ended up with a final sample of 4,222 CEO Letters for 1,162 companies.

3.2. Research design

To test our hypotheses, we use different OLS regression models. More specifically, this study comprises 4 models, which will be applied separately to the whole text metrics and for each section individually, aiming to capture better the effects that tone determinants have in each part of the CEO Letter.

We divided each CEO Letters into 3 parts, and we defined weights for each of them. More specifically, sections 1 (for the introduction part), 2 (for the development part) and 3 (for the conclusions part) were given the following weights: 10%, 80% and 10%, respectively. The lower percentages for the introduction and conclusions parts are a response to the fact that these often contain bias. To divide each section, we counted the number of the total sentence of each CEO Letter and multiplied each percentage by the number of the total sentence, as a result, we got to the number of sentences that must be attributed to each section.

First, we conducted a very simple regression model on tone using a set of three dummy variables, each representing section j of the CEO Letter i in year t. Moreover, dS1 represents the dummy variable for section 1, dS2 represents the dummy variable for section 2 and finally, dS3 represents the dummy variable for section 3. The purpose of this regression is to confirm if our expectation that CEO Letters' tone sections are different between themselves and between the tone extracted from the whole text approach. We expect that the latter is more optimistic when compared to the sections because the former approach considers that all sections contain the same relevance. However, we know according to past literature that the initial part is more likely to be subject of bias and the last part normally contains the acknowledgements, thus not underweighting these parts can bias our results.

Then, we developed several models. Model (1) is our base model and includes a set of control variables and firm-year fixed effects. Model (2) incorporates a dummy variable that represents the financial crisis, with the sole objective of confirming if our hypothesis that financial crisis impacts negatively the reported tone is met. Moreover, we expect that tone will decrease in the presence of economic downturns. Model (3) adds to the previous model an interaction term between an accounting measure of firm's performance – ROA – and Financial Crisis, to estimate whether ROA's impact on tone changes depending on whether we are referring to a period of financial crisis or not. Finally, Model (4) is practically identical to model (3), the only difference is that in Model (4) we use the interaction term between the change in ROA and financial crisis that according to our expectations, should have a higher significance than the *roacrisis* interaction term, as changes in firm's performance explain better firm's performance than a static measure of performance (ROA).

Afterwards, we will also run repeat Models (1) and (2) described above for 2 different textual metrics. Moreover, we will regress uncertainty and causality on tone determinants and on firm and year fixed effects, followed by a regression that adds the financial crisis dummy variable to capture the effect of crisis in the use of both uncertain and causal words.

Regarding the differences between crisis versus non-crisis periods in the relationship between firm's current performance (ROA) and its change (Ch_ROA) and tone, we focus on the estimate of β_{12}, whereas for understanding the impact that financial crisis itself has on tone, we focus on the estimate of β_{11}.

The regression models are as follows:

Models (1) and (2)[1]

$$\text{Narratives'Characteristics}_{it} = \boldsymbol{\beta_0} + \boldsymbol{\beta_1 \log}\,(\mathbf{1} + \boldsymbol{nr_{words}})_{it} + \boldsymbol{\beta_2 ROA_{it}} + \boldsymbol{\beta_3 Ch\_ROA_{it}} + \boldsymbol{\beta_4 Ret_{it}} + \boldsymbol{\beta_5 EY_{it}} + \boldsymbol{\beta_6 Ch\_EY_{it}} + \boldsymbol{\beta_7 Size_{it}} + \boldsymbol{\beta_8 MB_{it}} + \boldsymbol{\beta_9 BusSeg_{it}} + \boldsymbol{\beta_{10} Loss_{it}} + \boldsymbol{\beta_{11} finacialcrisis_{it}} + \textbf{firm-fixed effects} + \textit{year fixed effects} + \varepsilon_{it}$$

Models (3) and (4)[2]

$$Tone_{it} = \boldsymbol{\beta_0} + \boldsymbol{\beta_1 \log}\,(\mathbf{1} + \boldsymbol{nr_{words}})_{it} + \boldsymbol{\beta_2 ROA_{it}} + \boldsymbol{\beta_3 Ch\_ROA_{it}} + \boldsymbol{\beta_4 Ret_{it}} + \boldsymbol{\beta_5 EY_{it}} + \boldsymbol{\beta_6 Ch\_EY_{it}} + \boldsymbol{\beta_7 Size_{it}} + \boldsymbol{\beta_8 MB_{it}} + \boldsymbol{\beta_9 BusSeg_{it}} + \boldsymbol{\beta_{10} Loss_{it}} + \boldsymbol{\beta_{11} finacialcrisis_{it}} + \boldsymbol{\beta_{12} roacrisis_{it}} + \beta_{12} Ch\_ROAcrisis_{it} + \textbf{firm-fixed effects} + \boldsymbol{\varepsilon_{it}}$$

The dependent variable is the relevant narrative characteristic of the CEO Letter, more specifically, for models (1) and (2) we use tone, uncertainty, and causality. Models (3) and (4) are solely applied to tone.

We run panel data regression models on tone determinants as suggested by Li (2008, 2010). The length of each CEO letter (or section) is measured by the logarithm of the number of words. ROA and change in ROA are proxies for profitability. The interaction term of ROA and crisis measures the impact that turbulent economic cycles have on the relationship between firm's performance and the reported tone. The variables Return, Earnings yield and Earnings yield change aim to represent market and financial performance. Market-to-book ratio (MB) is a proxy for growth opportunities whereas the business segments control for the firm's operating complexity. Loss controls for the potential decrease in earnings. Financial crisis variable controls for the impact of the crisis. Each variable is defined in Table 14 in the appendix.

[1] The regression model (1) includes all right-sided variables, except the term of financial crisis, whereas the regression model (2) includes all variables in bold. Narratives' characteristics refer to tone, uncertainty, or causality.

[2] Models (3) and (4) are only applied to Tone. The regression model (3) contains all variables presented in bold while model (4) includes all variables excluding the interaction term *roacrisis* that was used in model (3).

The dummies $dS1_{ijt}$, $dS2_{ijt}$ and $dS3_{ijt}$ control for: sections 1, 2 and 3 of the CEO Letter of firm i in year t. The log $(1 + nr_{words})_{it}$ is the logarithm of the total number of words of CEO Letter. ROA is return on assets; *Ch_ROA* is the change in the return on assets; *roacrisis* is the product of the return on assets and the financial crisis variable; I *Ret* is the 12–month stock return. *EY* is the earnings yield; *Ch_EY* is the change in the earnings yield; *Size* is the logarithm of market capitalization at fiscal year-end; *MB* is the *market-to-book* ratio;

BusSeg is the log(1 + number of business segments); *Loss* is a dummy variable that takes the value 1 if the *EY* is lower than zero, or zero otherwise; *financialcrisis* is the indicator variable that takes the value of one if the CEO Letter is published in 2008, 2009 or 2010 and zero otherwise.

4. Results

Table 2 presents the descriptive statistics for our sample.

Table 2 – Descriptive statistics

PANEL A - WHOLE TEXT TONE MEASURES								
Variable Name	N	Mean	St. Dev.	Min.	P25	Median	P75	Max
No. of words	4,222	1713.243	1039.920	187.000	977.000	1455.000	2232.000	6094.000
Tone	4,222	0.682	0.244	-1.000	0.577	0.736	0.846	1.000
No. of positive words	4,222	57.901	39.991	0.000	30.000	49.000	77.000	262.000
No. of negative words	4,222	9.991	9.430	0.000	4.000	7.000	13.000	86.000
No. of uncertainty words	4,222	8.239	8.040	0.000	3.000	6.000	11.000	94.000
No. of causal words	4,222	13.897	10.549	0.000	6.000	11.000	19.000	86.000

PANEL B - TONE MEASURES BY SECTION								
Variable Name	N	Mean	St. Dev.	Min.	P25	Median	P75	Max
No. of words_S1	4,222	171.633	114.881	5.000	90.000	143.000	230.000	1507.000
Tone_S1	4,222	0.668	0.434	-1.000	0.500	0.833	1.000	1.000
No. of positive words_S1	4,222	7.179	6.212	0.000	3.000	6.000	10.000	48.000
No. of negative words_S1	4,222	1.104	1.642	0.000	0.000	1.000	2.000	13.000
No. of uncertainty words_S1	4,222	0.678	1.136	0.000	0.000	0.000	1.000	12.000
No. of causal words_S1	4,222	1.478	1.823	0.000	0.000	1.000	2.000	16.000
No. of words_S2	4,222	1390.444	835.593	4.000	797.000	1190.000	1799.000	4854.000
Tone_S2	4,222	0.675	0.252	-1.000	0.571	0.730	0.846	1.000
No. of positive words_S2	4,222	45.728	31.595	0.000	24.000	39.000	61.000	222.000
No. of negative words_S2	4,222	8.111	7.731	0.000	3.000	6.000	11.000	68.000
No. of uncertainty words_S2	4,222	6.679	6.692	0.000	2.000	5.000	9.000	87.000
No. of causal words_S2	4,222	11.369	8.559	0.000	5.000	9.000	15.000	69.000
No. of words_S3	4,222	151.167	119.130	5.000	71.000	123.000	198.000	1503.000
Tone_S3	4,222	0.639	0.467	-1.000	0.333	0.846	1.000	1.000
No. of positive words_S3	4,222	4.994	4.819	0.000	1.000	4.000	7.000	42.000
No. of negative words_S3	4,222	0.775	1.389	0.000	0.000	0.000	1.000	24.000
No. of uncertainty words_S3	4,222	0.883	1.575	0.000	0.000	0.000	1.000	25.000
No. of causal words_S3	4,222	1.050	1.515	0.000	0.000	1.000	2.000	17.000

PANEL C - TONE DETERMINANTS								
Variable Name	N	Mean	St. Dev.	Min.	P25	Median	P75	Max
ROA	16,888	-0.021	0.228	-1.582	-0.057	0.035	0.086	0.751
Change in ROA	16,888	0.011	0.183	-0.877	-0.033	0.008	0.046	1.927
Return	16,888	0.084	0.552	-0.886	-0.287	0.027	0.341	2.768
Earnings Yield	16,888	-0.020	0.240	-2.342	-0.054	0.045	0.085	0.486
Change in Earnings Yield	16,888	0.105	4.277	-44.400	-0.029	0.008	0.050	270.958
Market Value	16,564	959.453	3015.928	0.820	18.630	79.920	534.750	31500.000
Market to Book Ratio	16,888	2.806	3.948	-16.499	1.052	1.881	3.365	39.885
Business Segment	16,888	0.495	0.191	0.301	0.301	0.477	0.602	1.041
Loss	16,888	0.344	0.475	0.000	0.000	0.000	1.000	1.000
Financial Crisis	16,888	0.316	0.465	0.000	0.000	0.000	1.000	1.000

The sample comprises firm-years for firms listed in the London Stock Exchange with fiscal-year-ends between January 2005 and December 2014. Source: market and accounting data is extracted from DataStream.

The median firm presents a slightly more optimistic tone in the whole text approach (around 0.736) than in section 2 given by the weighted approach. However, both sections 1 and 3 are more optimistic than the whole text tone metric. This is in line with prior literature that says that the introduction and conclusion parts of the CEO Letters tend to be more optimistic than the others due to the recency effect and peak-end rule theory that suggests that readers recall better what they read lastly.

At the whole text-level we can infer that the median firm has a significantly higher number of positive words than negative words (49 and 7, respectively). Regarding the uncertainty and causal words, the median firm contains few of these words then compared to the total number of words. The median firm contains in its CEO Letters around 6 uncertainty words and 11 causality words in a total of 1,455 words.

In sections' approach, we observe that sections 1 and 3, present a more optimistic tone (0.833 and 0.846, respectively) when compared to section 2 (has a tone value of 0.730) that normally regards firm's performance and expectations (and thus, containing a more neutral language and lower readability). We can also state that section 3 – the conclusions part – is the more optimistic part of the CEO Letter of the median firm. This, as suggested by existing literature, shows that the final parts are normally more optimistic when compared to the reminiscent body of the CEO Letter.

The CEO Letter of the median firm, for the sections' approach, barely contains any uncertainty and causality words for sections 1 and 3. However, section 2 of the median firm's CEO Letter contains a small, yet a valid number of uncertainty and causality words (5 and 9, respectively). This is expectable, as this section usually contains information about firm's current performance and expectations regarding the future wealth of the company itself, thus it is predictable that

managers use more causal words (to blame others in case of poor performance) and uncertainty words (to explain firm's prospects).

Regarding the uncertainty words, only Section 2 has a value, which is 5 uncertainty words. Lastly, Sections 1, 2 and 3 have, for the median firm's CEO Letters, 1, 9 and 1 causal words, respectively.

The median firm has a market capitalization of 79.920, a market to book ratio of 1.881; a 12–month stock return of 2.7%; an earnings yield of 4.5%; a ROA of 3.5% and both have changes of, approximately, 0.8%.

Results are presented as follows. For tone means' differences – see tables 2 and 3, for tone regressions - models (1), (2), (3), (4) – see tables 4, 5, 6 and 7; for uncertainty – models (1) and (2) – see tables 8 and 9; and lastly, for causal reasoning – models (1) and (2) – see tables 10 and 11.

4.1. Tone across sections

Table 3 compares the tone means between sections and the whole CEO Letter using an ANOVA test. Results show that the average tone varies across sections (statistically significant at 1% confidence level), which in this case means that there are indeed differences in tone between sections of a CEO Letter and its different sections.

Table 3 – ANOVA test for measuring tone within CEO Letters

Differences between the tone means					
Source	Sum of Squares	Degrees of freedom	Mean Square	F	Prob > F
Between groups	5.441	3.000	1.814	13.040	0.000
Within groups	2,799.976	20124.000	0.139		

Additionally, we also applied a t-test for the dummies that represent each section (i.e., comparing: dS1 with dS2, dS1 with dS2 and dS2 with dS3) with similar results.

Finally, we tested these differences using the following regression model:

$Tone_{ijt} = \beta_0 + \beta_1 dS1_{ijt} + \beta_2 dS2_{ijt} + \beta_3 dS3_{ijt} + firm\text{-}fixed\ effects + year\ fixed\ effects + \varepsilon_{it}$, where dummies dS1, dS2 and dS3, represent, respectively, section 1, 2 or 3 of the CEO Letter of firm *i* in year *t*, section *j*.

Table 4 presents the results between the differences of tone means of the three sections of CEO Letters and the CEO Letter as whole.

Table 4 – Differences between tone means of CEO Letters' sections and the whole text tone

Differences between tone means		
Intercept	0.6923	***
	(0.0000)	
Dummy Section 1	-0.0136	**
	(0.0290)	
Dummy Section 2	-0.0075	*
	(0.0620)	
Dummy Section 3	-0.0433	***
	(0.0000)	
Firm fixed effects	Yes	
Year fixed effects	Yes	
F-statistic	30.3300	
Adjusted R-Squared	27.75	

Note: Sample contains 20,128 observations. P-values reported in parenthesis are computed using robust standard errors to obtain unbiased OLS coefficients. OLS estimates and its p-values in parenthesis. *p-value<0.1; **p-value<0.05; ***p-value<0.01.

The results suggest that there are statistically significant differences between the tone of the sections of a CEO Letters and the tone of the CEO Letter as a whole. These results also show that the sections are different between themselves, which then leads us to accept both our hypothesis H1 and H2.

The negative and statistically significant coefficients of the three dummies, suggest that the average tone of the sections of a given CEO Letter is less optimistic than the average tone of the CEO Letter as a whole, confirming our hypothesis that the tone that derives from each section is different from the tone of the CEO Letter as a whole. This is consistent with existing literature that states

that methods that do not attribute any weight and do not differentiate the importance of each CEO Letter's section, tend to lead to biased results.

The fact that the coefficients of the three dummies are all different between themselves, show us that each section has its own characteristics and not all have the same impact in building up the tone of the CEO Letter and thus, we can state that we confirm our hypothesis that tone's sections are indeed different between each other.

Table 5 presents the regression models on tone on the CEO Letters as a whole and on its 3 sections, individually. Model (1) is our base model and includes solely our control variables and firm and year fixed effects.

Table 5 – Model (1): OLS coefficients estimates of tone

Model (1) - Tone				
	Whole Text	Section 1	Section 2	Section 3
Intercept	0.6568 ***	0.3262 ***	0.6126 ***	-0.0785
	(0.0000)	(0.0010)	(0.0000)	(0.3930)
log(nr_words)	-0.0127	0.0401 ***	-0.0046	0.1295 ***
	(0.1700)	(0.0050)	(0.6610)	(0.0000)
ROA	0.0008	0.0413	0.0089	0.0689
	(0.9830)	(0.6010)	(0.8220)	(0.4000)
Change in ROA	0.0559 **	0.1091 **	0.0491 *	0.0352
	(0.0200)	(0.0310)	(0.0560)	(0.4910)
Return	0.0050	-0.0042	0.0002	0.0343 **
	(0.5010)	(0.7970)	(0.9800)	(0.0460)
Earnings Yield	0.0352	0.1247 **	0.0299	-0.0607
	(0.1230)	(0.0120)	(0.2020)	(0.2790)
Change in EY	0.0006	-0.0009	0.0002	0.0011
	(0.2730)	(0.3830)	(0.7490)	(0.2320)
Size	0.0001 ***	0.0001 ***	0.0001 ***	0.0001 **
	(0.0000)	(0.0000)	(0.0000)	(0.0130)
Market to Book Ratio	0.0024 **	0.0083 ***	0.0026 **	0.0005
	(0.0390)	(0.0080)	(0.0450)	(0.8490)
Business Segments	-0.0581 *	-0.0240	-0.0749 **	-0.0386
	(0.0600)	(0.7110)	(0.0290)	(0.6230)
Dummy Loss	-0.0299 **	-0.0637 **	-0.0320 **	-0.0131
	(0.0320)	(0.0310)	(0.0300)	(0.6680)
Firm-fixed effects	Yes	Yes	Yes	Yes
Year-fixed effects	Yes	Yes	Yes	Yes
F Statistic	18.300	9.580	14.780	7.760
Adjusted R-Squared	45.450%	24.820%	42.030%	19.070%

Note: Sample contains 4,222 observations. P-values reported in parenthesis are computed using robust standard errors to obtain unbiased OLS coefficients. OLS estimates and its p-values in parenthesis. *p-value<0.1; **p-value<0.05; ***p-value<0.01.

The different regressions show us that there are significant differences between sections, for example, while the length of the CEO letter which is

represented by the logarithm of the number of words plays a statistically significant role in explaining the reported tone of the CEO Letter for Sections 1 and 3, it does not for Section 2 and for the CEO Letter *per si*.

Comparing which tone determinants are statistically significant from one section to another we can state that section 3 is the one with less statistically significance. This is expectable as section 3 is the part of the CEO Letter that incorporates closing messages and therefore, assuming that this section is statistically significant and important in determining firm's discourse characteristics would make our results somehow biased. This suggests that there are indeed differences within CEO Letters' sections and they do not provide an equal informative ability.

According to the results, firms with improving performance will disclose a more positive tone, which is consistent with literature that defends that firms that perform well, tend to focus a bit more on positive news and do not dwell on negative news. Furthermore, we observe a positive and significant effect of a change in firm's current performance on the reported tone, except for section 3, which is understandable as this section does not contain any technical information nor performance related commentary.

The significant and positive coefficient of the market to book ratio – at the whole text level - suggests that firms with more growth opportunities will have a more positive tone in its CEO Letters. This key indicator is only influential in the reported tone for sections 1 and 2, confirming again our previous explanations that section 3 does not play a crucial role in shaping firm's discourse.

Loss variable indicates that when firm's face losses, the tone will necessarily be less optimistic.

Table 6 presents the regression models on tone on the CEO Letters as a whole and on its 3 sections, individually, including the dummy variable Financial Crisis. Our interest in this variable is to assess if financial crisis led to a decrease in the tone reported in CEO Letters.

Table 6 – Model (2): OLS coefficients estimates of tone

Model (2) - Tone				
	Whole Text	Section 1	Section 2	Section 3
Intercept	0.6835 ***	0.3629 ***	0.6357 ***	-0.0800
	(0.0000)	(0.0000)	(0.0000)	(0.3600)
log(nr_words)	-0.0143	0.0373 **	-0.0059	0.1294 ***
	(0.1260)	(0.0100)	(0.5740)	(0.0000)
ROA	-0.0049	0.0306	0.0052	0.0587
	(0.8960)	(0.6950)	(0.8950)	(0.4710)
Change in ROA	0.0555 **	0.1086 **	0.0479 *	0.0423
	(0.0180)	(0.0300)	(0.0570)	(0.4110)
Return	-0.0044	-0.0178	-0.0078	0.0251
	(0.5270)	(0.2480)	(0.2930)	(0.1180)
Earnings Yield	0.0470 **	0.1460 ***	0.0404 *	-0.0581
	(0.0390)	(0.0030)	(0.0820)	(0.2960)
Change in EY	0.0005	-0.0010	0.0001	0.0011
	(0.3470)	(0.2970)	(0.8550)	(0.2460)
Size	0.0001 ***	0.0001 ***	0.0001 ***	0.0001 ***
	(0.0000)	(0.0000)	(0.0000)	(0.0080)
Market to Book Ratio	0.0028 **	0.0090 ***	0.0029 **	0.0009
	(0.0170)	(0.0030)	(0.0250)	(0.7300)
Business Segments	-0.0912 ***	-0.0851	-0.1042 ***	-0.0527
	(0.0020)	(0.1820)	(0.0020)	(0.4830)
Dummy Loss	-0.0308 **	-0.0656 **	-0.0326 **	-0.0135
	(0.0280)	(0.0270)	(0.0280)	(0.6590)
Financial Crisis	-0.0672 ***	-0.0795 ***	-0.0656 ***	-0.0344 ***
	(0.0000)	(0.0000)	(0.000)	(0.0500)
Firm-fixed effects	Yes	Yes	Yes	Yes
Year-fixed effects	No	No	No	No
F Statistic	25.140	13.910	21.090	12.890
Adjusted R-Squared	44.55%	24.04%	41.43%	19.02%

Note: Sample contains 4,222 observations. P-values reported in parenthesis are computed using robust standard errors to obtain unbiased OLS coefficients. OLS estimates and its p-values in parenthesis. *p-value<0.1; **p-value<0.05; ***p-value<0.01.

Results show that firms adopt a more negative tone due to financial crisis. However, looking at the positive and significant relationship of the change in firm's performance (Change in ROA) and tone, we can infer that firm's with improving performance will disclose information in line with their performance. This is consistent with the study of Patelli and Pedrini (2014) who defend that during crisis incentives to distort information are lower in the sense that poor performers will not be as concerned in disclosing bad news than if it was in a

year of economic expansion because during downturns bad news are somehow granted.

Again, we can observe that among all sections, section 3 is the one that reports lower statistically significance independently of the tone determinants that we are referring to. This is due to the lack of informativeness that section 3 provides when compared to sections 1 and 2.

The fact that some tone determinants are statistically significant in some parts of the CEO Letter and in others is not, is an indicator that sections are different between themselves and not all carry the same informativeness and power in predicting the reported tone.

Table 7 adds an interaction term between an accounting measure of performance (*ROA*) and *Financial Crisis* dummy variable. This interaction term is included to differentiate the effects that crisis versus non-crisis years have on the relationship between ROA and tone, moreover, if there is a significant difference in their relationship depending on whether we are facing an economic crisis or not.

Table 7 – Model (3): OLS coefficients estimates of tone

Model (3) - Tone				
	Whole Text	Section 1	Section 2	Section 3
Intercept	0.6835 ***	0.3640 ***	0.6357 ***	-0.0802
	(0.0000)	(0.0000)	(0.0000)	(0.3590)
log(nr_words)	-0.0142	0.0373 **	-0.0059	0.1294 ***
	(0.1280)	(0.0100)	(0.5740)	(0.0000)
ROA	0.0045	0.0573	0.0058	0.0530
	(0.9090)	(0.4910)	(0.8900)	(0.5500)
Change in ROA	0.0555 **	0.1086 **	0.0479 *	0.0423
	(0.0180)	(0.0290)	(0.0570)	(0.4120)
Return	-0.0044	-0.0179	-0.0078	0.0251
	(0.5220)	(0.2440)	(0.2930)	(0.1180)
Earnings Yield	0.0482 **	0.1494 ***	0.0405 *	-0.0589
	(0.034)	(0.0030)	(0.0820)	(0.2920)
Change in EY	0.0005	-0.0011	0.0001	0.0011
	(0.3600)	(0.2830)	(0.8560)	(0.2430)
Size	0.0001 ***	0.0001 ***	0.0001 ***	0.0001 ***
	(0.0000)	(0.0000)	(0.0000)	(0.0080)
Market to Book Ratio	0.0028 **	0.0091 ***	0.0029 **	0.0009
	(0.0170)	(0.0030)	(0.0260)	(0.7320)
Business Segments	-0.0918 ***	-0.0868	-0.1042 ***	-0.0524
	(0.0020)	(0.1740)	(0.0010)	(0.4860)
Dummy Loss	-0.0302 **	-0.0641 **	-0.0326 **	-0.0138
	(0.0320)	(0.0300)	(0.0280)	(0.6510)
Financial Crisis	-0.0674 ***	-0.0803 ***	-0.0656 ***	-0.0342 *
	(0.0000)	(0.0000)	(0.000)	(0.0500)
ROA x Crisis	-0.0222	-0.0628	-0.0015	0.0134
	(0.5710)	(0.4570)	(0.9720)	(0.8800)
Firm-fixed effects	Yes	Yes	Yes	Yes
Year-fixed effects	Yes	No	No	No
F Statistic	23.880	13.430	19.810	11.820
Adjusted R-Squared	44.540%	24.040%	41.410%	19.000%

Note: Sample contains 4,222 observations. P-values reported in parenthesis are computed using robust standard errors to obtain unbiased OLS coefficients. OLS estimates and its p-values in parenthesis. *p-value<0.1; **p-value<0.05; ***p-value<0.01.

As presented, the interaction term between ROA and Crisis does not display a statistically significant effect and thus, we cannot accept our hypothesis that in Financial Crisis years the relationship between ROA and tone is different than in non-crisis years, moreover, we cannot state, for our sample, that firms report more accurate information during the economic crisis as reported by Patelli and Pedrini (2014). However, our results suggest that sections are not equal between themselves, giving robustness to our expectations. Comparing all sections, we can see that section 3 has almost no statistical significance when compared to the others. Financial Crisis led to the adoption of a more negative tone by firm's managers and the higher the losses faced by companies, the more pessimistic would be the firm's discourse.

Table 8 includes now an interaction term between *Changes in ROA* and *Financial Crisis* dummy variable. This interaction term is included to study the effects that crisis versus non-crisis years have on the relationship between Changes in ROA and tone, moreover, if there is a significant difference in their relationship depending on whether we are facing an economic crisis or not.

Table 8 – Model (4): OLS coefficients estimates of tone

Model (4) - Tone				
	Whole Text	Section 1	Section 2	Section 3
Intercept	0.6830 ***	0.3626 ***	0.6345 ***	-0.0803
	(0.0000)	(0.0000)	(0.0000)	(0.3590)
log(nr_words)	-0.0143	0.0373 **	-0.0058	0.1294 ***
	(0.1260)	(0.0100)	(0.5760)	(0.0000)
ROA	-0.0069	0.0293	0.0018	0.0571
	(0.8540)	(0.7130)	(0.9630)	(0.4910)
Change in ROA	0.0633 **	0.1139 **	0.0611 **	0.0484
	(0.0140)	(0.0410)	(0.0240)	(0.4050)
Return	-0.0044	-0.0178	-0.0078	0.0251
	(0.5270)	(0.2480)	(0.2930)	(0.1180)
Earnings Yield	0.0475 **	0.1463 ***	0.0412 *	-0.0578
	(0.0360)	(0.0030)	(0.0760)	(0.2980)
Change in EY	0.0005	-0.0010	0.0001	0.0011
	(0.3520)	(0.2980)	(0.8580)	(0.2470)
Size	0.0001 ***	0.0001 ***	0.0001 ***	0.0001 ***
	(0.0000)	(0.0000)	(0.0000)	(0.0080)
Market to Book Ratio	0.0028 **	0.0090 ***	0.0029 **	0.0009
	(0.0170)	(0.0030)	(0.0250)	(0.7300)
Business Segments	-0.0909 ***	-0.0849	-0.1037 ***	-0.0525
	(0.0020)	(0.1830)	(0.0020)	(0.4850)
Dummy Loss	-0.0300 **	-0.0650 **	-0.0313 **	-0.0129
	(0.0330)	(0.0270)	(0.0350)	(0.6730)
Financial Crisis	-0.0678 ***	-0.0800 ***	-0.0667 ***	-0.0349 **
	(0.0000)	(0.0000)	(0.0000)	(0.0460)
Change in ROA x Crisis	0.0530	0.0359	0.0895	0.0413
	(0.3440)	(0.7630)	(0.1060)	(0.7700)
Firm-fixed effects	Yes	Yes	Yes	Yes
Year-fixed effects	Yes	No	No	No
F Statistic	23.160	12.890	19.460	11.840
Adjusted R-Squared	44.540%	24.020%	41.450%	19.000%

Note: Each sample contains 4,222 observations. P-values reported in parenthesis are computed using robust standard errors to obtain unbiased OLS coefficients. OLS estimates and its p-values in parenthesis. *p-value<0.1; **p-value<0.05; ***p-value<0.01.

Our results suggest that there are no statistically significant differences in the relationship between the variation in firm's ROA and the reported tone in CEO Letters from crisis to non-crisis years.

In our model, the change in ROA variable reports a positive and statistically significant association with the tone, indicating that when firms' current

performance varies positively, so will the tone. Again, this relationship holds for the whole text level and all sections, excluding section 3. This together with the fact that most tone determinants lose their statistical significance only in section 3 is clear evidence this section is not important in accurately shaping firm's tone and thus, taking it into consideration when assessing the reported tone, might bias our results.

Financial Crisis negatively impacts the reported tone and so does the loss variable. Each indicating that financial crisis years lowers the optimism present in CEO Letters and that losses also make firms adopt a more negative tone.

Table 9 presents the regression models on Uncertainty from the CEO Letters. Moreover, Model (1) is our base model.

Table 9 – Model (1): OLS coefficients estimates of Uncertainty

Model (1) - Uncertainty				
	Whole Text	Section 1	Section 2	Section 3
Intercept	-4.6295 ***	-1.1596 ***	-4.0441 ***	-0.9223 ***
	(0.0000)	(0.0000)	(0.0000)	(0.0000)
log(nr_words)	0.8917 ***	0.3224 ***	0.8079 ***	0.2737 ***
	(0.0000)	(0.0000)	(0.0000)	(0.0000)
ROA	-0.1498 *	-0.0527	-0.0724	-0.2302 ***
	(0.0810)	(0.4660)	(0.4140)	(0.0050)
Change in ROA	0.0159	-0.0517	-0.0193	0.1000 *
	(0.7760)	(0.3010)	(0.7370)	(0.0740)
Return	-0.0097	0.0249	-0.0088	0.0044
	(0.6070)	(0.1460)	(0.6630)	(0.8120)
Earnings Yield	-0.0412	-0.0455	-0.0061	-0.0217
	(0.4130)	(0.3760)	(0.9090)	(0.7030)
Change in EY	-0.0004	-0.0003	0.0000	0.0027 ***
	(0.7120)	(0.7090)	(0.9800)	(0.0000)
Size	0.0000 ***	0.0000	0.0000	0.0000
	(0.0000)	(0.3350)	(0.6750)	(0.2700)
Market to Book Ratio	0.0042	0.0032	0.0030	-0.0013
	(0.2300)	(0.2280)	(0.3650)	(0.6820)
Business Segments	-0.0710	-0.1101	-0.0360	-0.0636
	(0.4270)	(0.1650)	(0.7030)	(0.4750)
Dummy Loss	-0.0528	-0.0252	0.0710 **	0.0135
	(0.1120)	(0.3920)	(0.0410)	(0.6840)
Firm-fixed effects	Yes	Yes	Yes	Yes
Year-fixed effects	Yes	Yes	Yes	Yes
F Statistic	98.600	25.250	55.120	29.080
Adjusted R-Squared	63.770%	25.300%	60.730%	27.120%

Note: Sample contains 4,222 observations. P-values reported in parenthesis are computed using robust standard errors to obtain unbiased OLS coefficients. OLS estimates and its p-values in parenthesis. *p-value<0.1; **p-value<0.05; ***p-value<0.01.

Overall, our results show that between sections, only section 3 seems to have a greater impact in estimating the use of uncertainty words in CEO Letters. We can see that the whole text results also report some statistically significant tone determinants, and this might be due to the influence of section 3 because looking separately at sections 1 and 2 we can observe that both do not seem to have an important role in estimating the reported tone as most tone determinants do not report any statistically significance. This suggests that there are indeed differences between sections.

Table 10 presents the regression models on Uncertainty from the CEO Letters. We add to this model the dummy variable Financial Crisis to measure the effect of the financial crisis in the use of uncertainty words.

Table 10 – Model (2): OLS coefficients estimates of Uncertainty

Model (2) - Uncertainty				
	Whole Text	Section 1	Section 2	Section 3
Intercept	-4.5493 ***	-1.1412 ***	-3.9725 ***	-0.8926 ***
	(0.0000)	(0.0000)	(0.0000)	(0.0000)
log(nr_words)	0.8933 ***	0.3227 ***	0.8099 ***	0.2734 ***
	(0.0000)	(0.0000)	(0.0000)	(0.0000)
ROA	-0.1284	-0.0502	-0.0546	-0.2235 ***
	(0.1310)	(0.4840)	(0.5350)	(0.0060)
Change in ROA	0.0079	-0.0555	-0.0247	0.0980 *
	(0.8870)	(0.2650)	(0.6650)	(0.0790)
Return	-0.0123	0.0195	-0.0117	0.0079
	(0.5010)	(0.2430)	(0.5420)	(0.6600)
Earnings Yield	-0.0428	-0.0466	-0.0096	-0.0210
	(0.3940)	(0.3630)	(0.8580)	(0.7090)
Change in EY	-0.0003	-0.0004	0.0001	0.0027 ***
	(0.7610)	(0.6070)	(0.9630)	(0.0000)
Size	0.0000	0.0000	0.0000	0.0000
	(0.5050)	(0.3270)	(0.5110)	(0.2890)
Market to Book Ratio	0.0035	0.0026	0.0024	-0.0015
	(0.3170)	(0.3250)	(0.4710)	(0.6540)
Business Segments	-0.0713	-0.0726	-0.0345	-0.0689
	(0.4110)	(0.3410)	(0.7050)	(0.4180)
Dummy Loss	0.0539	-0.0257	0.0722 **	0.0133
	(0.1050)	(0.3830)	(0.0380)	(0.6880)
Financial Crisis	0.0691 ***	0.0284	0.0671 ***	0.0284
	(0.0010)	(0.1240)	(0.0020)	(0.1580)
Firm-fixed effects	Yes	Yes	Yes	Yes
Year-fixed effects	No	No	No	No
F Statistic	165.900	42.430	92.950	49.120
Adjusted R-Squared	63.640%	25.210%	60.630%	27.220%

Note: Sample contains 4,222 observations. P-values reported in parenthesis are computed using robust standard errors to obtain unbiased OLS coefficients. OLS estimates and its p-values in parenthesis. *p-value<0.1; **p-value<0.05; ***p-value<0.01.

Results seem to indicate that Financial Crisis led to an increase on the use of uncertainty words for the CEO. However, this positive and statistically significant relationship only holds at the whole text level and for section 2, which is not surprising as this section normally mentions firm's performance, expectations, and growth opportunities. Being in the mid of a financial crisis, it is expectable that this section reports an increase in the number of uncertainty words. This result is a clear evidence that the content of sections is very important and determinant when assessing firm's discourse characteristics, which also gives robustness to our statement that sections of the same CEO Letter are different between each other.

Table 11 presents the regression models on Causality from the CEO Letters.

Table 11 – Model (1): OLS coefficients estimates of Causality

Model (1) - Causality				
	Whole Text	Section 1	Section 2	Section 3
Intercept	-4.2274 ***	-1.5800 ***	-3.8580 ***	-0.9876 ***
	(0.0000)	(0.0000)	(0.0000)	(0.0000)
log(nr_words)	0.9315 ***	0.4800 ***	0.8763 ***	0.3285 ***
	(0.0000)	(0.0000)	(0.0000)	(0.0000)
ROA	0.0036	-0.1508 *	0.0329	0.0689
	(0.9600)	(0.0580)	(0.6520)	(0.3650)
Change in ROA	-0.0482	-0.0363	-0.0463	-0.0413
	(0.3210)	(0.5160)	(0.3600)	(0.4210)
Return	0.0070	0.0054	0.0145	-0.0025
	(0.6610)	(0.7730)	(0.3780)	(0.8960)
Earnings Yield	-0.0085	0.0494	-0.0215	0.0357
	(0.8490)	(0.3850)	(0.6430)	(0.4910)
Change in EY	-0.0007	0.0008	-0.0009	0.0020 **
	(0.3990)	(0.4810)	(0.2740)	(0.0190)
Size	-0.0001 ***	0.0000	-0.0001 **	-0.0001 **
	(0.0040)	(0.1280)	(0.0210)	(0.0270)
Market to Book Ratio	0.0056 *	0.0073 **	0.0039	0.0001
	(0.0510)	(0.0350)	(0.1380)	(0.9760)
Business Segments	0.0416	0.0064	0.0586	0.1092
	(0.5490)	(0.9430)	(0.4240)	(0.2140)
Dummy Loss	0.0391	-0.0284	0.0476 *	0.0582 *
	(0.1420)	(0.3910)	(0.0860)	(0.0790)
Firm-fixed effects	Yes	Yes	Yes	Yes
Year-fixed effects	Yes	Yes	Yes	Yes
F Statistic	128.930	48.410	91.940	37.960
Adjusted R-Squared	72.150%	40.210%	69.410%	31.300%

Note: Sample contains 4,222 observations. P-values reported in parenthesis are computed using robust standard errors to obtain unbiased OLS coefficients. OLS estimates and its p-values in parenthesis. *p-value<0.1; **p-value<0.05; ***p-value<0.01.

The results from table 11 suggest that most determinants used to measure the use of uncertainty words in firm's CEO Letters do not load a statistically significant effect. For example, the coefficients for the Loss variable tell us that only sections 2 and 3 report a statistically significant association with the use of uncertainty words, which, in other words, means that firms that face losses tend to use more causal words. This is in line with prior literature that states that bad performers tend to blame external causes for their poor results and take the credit for the good ones. However, we can see that this positive and statistically significant relationship does not hold for the whole text level nor for section 1. This, as previously mentioned, shows consistency in our results in the sense that we can infer that sections are different between themselves as the effects of each determinant are not equal for all sections/approaches.

Table 12 presents the regression models on Causality from the CEO Letters. This model incorporates the dummy variable *Financial Crisis* with the aim of assessing if firms use causal words more frequently in financial crisis periods.

Table 12 – Model (2): OLS coefficients estimates of Causality

Model (2) - Causality				
	Whole Text	Section 1	Section 2	Section 3
Intercept	-4.2506 ***	-1.6306 ***	-3.8697 ***	-0.9746 ***
	(0.0000)	(0.0000)	(0.0000)	(0.0000)
log(nr_words)	0.9331 ***	0.4816 ***	0.8772 ***	0.3292 ***
	(0.0000)	(0.0000)	(0.0000)	(0.0000)
ROA	-0.0006	-0.1542 *	0.0288	0.0610
	(0.9940)	(0.0540)	(0.6920)	(0.4180)
Change in ROA	-0.0381	-0.0284	-0.0357	-0.0416
	(0.4330)	(0.6110)	(0.4810)	(0.4150)
Return	0.0066	0.0101	0.0116	-0.0030
	(0.6650)	(0.5770)	(0.4590)	(0.8700)
Earnings Yield	-0.0215	0.0389	-0.0338	0.0363
	(0.6260)	(0.4910)	(0.4610)	(0.4790)
Change in EY	-0.0007	0.0007	-0.0008	0.0018 **
	(0.4130)	(0.4950)	(0.3050)	(0.0380)
Size	-0.0001 **	0.0000	-0.0001 *	-0.0001 **
	(0.0120)	(0.2150)	(0.0500)	(0.0470)
Market to Book Ratio	0.0051 *	0.0069 **	0.0035	-0.0003
	(0.0720)	(0.0470)	(0.1710)	(0.9250)
Business Segments	0.0742	0.0332	0.0881	0.1264
	(0.2720)	(0.6980)	(0.2160)	(0.1340)
Dummy Loss	0.0399	-0.0274	0.0480 *	0.0583 *
	(0.1340)	(0.4050)	(0.0840)	(0.0790)
Financial Crisis	0.0007	0.0040	0.0023	-0.0307
	(0.9680)	(0.8420)	(0.8930)	(0.1310)
Firm-fixed effects	Yes	Yes	Yes	Yes
Year-fixed effects	No	No	No	No
F Statistic	217.440	82.220	152.450	64.000
Adjusted R-Squared	72.060%	40.160%	69.330%	31.220%

Note: Sample contains 4,222 observations. P-values reported in parenthesis are computed using robust standard errors to obtain unbiased OLS coefficients. OLS estimates and its p-values in parenthesis. *p-value<0.1; **p-value<0.05; ***p-value<0.01.

Despite our predictions that Financial Crisis would lead to an increase in the use of causal words, our results do not confirm so.

We can clearly affirm that sections are not homogenous between themselves as for example, the coefficient *ROA*, which is a proxy for current firm's performance, only loads a statistically significant value for section 1 of the CEO Letter. This means that the content present in the remaining sections is somehow irrelevant when measuring the relationship between *ROA* and the use of causal words.

In other words, and accordingly to existing literature, not all sections of financial narratives deliver the same informativeness and the fact that some of the determinants used in our regression load statistically significant values only for some sections, lead us to confirm that sections are not equal between each other.

4.2. Further analysis

Our tests aim at capturing the relationship between tone and performance. According to Gentry and Shen's (2010) firm's financial performance is not a single unidimensional construct that can be understood and measured in many ways, which raises the question of the sensitivity of this relationship to the performance proxy-measure used. In this section, we focus on an alternative approach focused on market performance measure by stock returns, in opposition of the previous accounting-based measures.

In Model (3) of tone regressions, we used *ROA* as our measure of firm performance to assess if the relationship between firm's performance and the reported tone present in CEO Letters change depending on whether we are living in a Financial Crisis year. Despite both accounting and market measures being both valid to estimate firm's performance, there has always been an ongoing debate between which one to choose.

Thus, as we did not capture any significant effect on the power of ROA in Financial Crisis periods, we will test again Model (3) twice, but this time, with an interaction term between market measures and the Financial Crisis dummy variable. In the first regression, we will use an interaction term between Return and Financial Crisis. In the following, we use Loss and Financial Crisis as our interaction term. Although both are almost equal to Model (3) of tone regressions, we call these additional regressions, Models (5) and (6).

Table 13 presents the regression models on tone on the CEO Letters as a whole and on its 3 sections, individually. These regressions include not only the dummy variable *Financial Crisis* - that takes the value of 1 if the year refers to 2008, 2009 or 2010, and zero otherwise but also an interaction term between Return and *Financial Crisis* dummy variable and lastly, our control variables. This interaction term is included to study the effects that crisis versus non-crisis years have on the relationship between Return and tone, moreover, if there is a significant difference in their relationship depending on whether we are facing an economic crisis or not. This regression is also included in order to assess whether market measures provide additional information when comparing to accounting measures.

Table 13 – Model (5): OLS coefficients estimates of Tone using an interaction term *Return x Crisis*

Model (5) - Tone				
	Whole Text	Section 1	Section 2	Section 3
Intercept	0.6822 ***	0.3624 ***	0.6345 ***	-0.0801
	(0.0000)	(0.0000)	(0.0000)	(0.3600)
log(nr_words)	-0.0139	0.0379 **	-0.0055	0.1293 ***
	(0.1350)	(0.0000)	(0.5980)	(0.0000)
ROA	-0.0037	0.0328	0.0064	0.0585
	(0.9210)	(0.6750)	(0.8700)	(0.4720)
Change in ROA	0.0535 **	0.1050 **	0.0459 *	0.0426
	(0.0230)	(0.0360)	(0.0680)	(0.4090)
Return	0.0119	0.0116	0.0086	0.0227
	(0.1850)	(0.5530)	(0.3670)	(0.2690)
Earnings Yield	0.0482 **	0.1481 ***	0.0417 *	-0.0583
	(0.0350)	(0.0030)	(0.0740)	(0.2940)
Change in EY	0.0006	-0.0009	0.0002	0.0011
	(0.2940)	(0.3360)	(0.7860)	(0.2500)
Size	0.0001 ***	0.0001 ***	0.0001 ***	0.0001 ***
	(0.0000)	(0.0000)	(0.0000)	(0.0080)
Market to Book Ratio	0.0028 **	0.0091 ***	0.0230 **	0.0009
	(0.0150)	(0.0030)	(0.0260)	(0.7320)
Business Segments	-0.0921 ***	-0.0867	-0.1051 ***	-0.0526
	(0.0020)	(0.1740)	(0.0010)	(0.4840)
Dummy Loss	-0.0283 **	-0.0611 **	-0.0301 **	-0.0138
	(0.0450)	(0.0390)	(0.0420)	(0.6500)
Financial Crisis	-0.0628 ***	-0.0717 ***	-0.0612 ***	-0.0350 *
	(0.0000)	(0.0000)	(0.000)	(0.0520)
Return x Crisis	-0.0358 **	-0.0645 **	-0.0359 **	0.0052
	(0.0100)	(0.0280)	(0.0150)	(0.8590)
Firm-fixed effects	Yes	Yes	Yes	Yes
Year-fixed effects	No	No	No	No
F Statistic	23.650	13.060	18.800	11.810
Adjusted R-Squared	44.690%	24.180%	41.560%	19.000%

Note: Sample contains 4,222 observations. P-values reported in parenthesis are computed using robust standard errors to obtain unbiased OLS coefficients. OLS estimates and its p-values in parenthesis. *p-value<0.1; **p-value<0.05; ***p-value<0.01.

According to the results presented in table 13 we can state that Financial Crisis has an important and significant effect on the reported tone of firm's CEO Letters, having a negative impact on it.

Changes in ROA seem to impact positively tone, contrary to ROA which does not seem to impact significantly tone. As expected, in the existence of losses, firms tend to report more negatively.

Despite the positive relationship between returns and tone, this is not a significant association. However, with this regression, we aim to understand how the relationship between Return and tone in crisis versus non crisis settings is. The interaction term *Return x Crisis* tells us that there is indeed a statistically significant effect on this relationship depending on the year that we are at. If we refer to Financial Crisis years (in this study, we considered 2008, 2009 and 2010), if returns increase, so does tone. This result is in line with prior literature that suggests that firms have more incentives to disclose accurate information that is aligned to their performance as bad news would not be surprising in such an economic context.

Table 14 presents the regression models on tone on the CEO Letters as a whole and on its 3 sections, individually. These regressions are different from the previous in the sense that they include an interaction term between Loss and *Financial Crisis* dummy variable. Our sole purpose is to study the effects that crisis versus non-crisis years have on the relationship between Loss and tone, moreover, if there is a significant difference in their relationship depending on whether we are facing an economic crisis or not. This regression is also included in order to assess whether market measures provide additional information when comparing to accounting measures.

Table 14 – Model (6): OLS coefficients estimates of Tone using an interaction term *Loss x Crisis*

Model (6) - Tone				
	Whole Text	Section 1	Section 2	Section 3
Intercept	0.6836 ***	0.3671 ***	0.6356 ***	-0.0828
	(0.0000)	(0.0000)	(0.0000)	(0.3450)
log(nr_words)	-0.0143	0.0374 ***	-0.0059	0.1292 ***
	(0.1260)	(0.0000)	(0.5720)	(0.0000)
ROA	-0.0049	0.0307	0.0052	0.0586
	(0.8960)	(0.6940)	(0.8950)	(0.4710)
Change in ROA	0.0555 **	0.1083 **	0.0479 *	0.0425
	(0.0180)	(0.0300)	(0.0570)	(0.4090)
Return	-0.0044	-0.0185	-0.0077	0.0256
	(0.5230)	(0.2300)	(0.2940)	(0.1110)
Earnings Yield	0.0472 **	0.1493 ***	0.0403 *	-0.0604
	(0.0380)	(0.0030)	(0.0830)	(0.2780)
Change in EY	0.0005	-0.0011	0.0001	0.0011
	(0.3510)	(0.2720)	(0.8530)	(0.2300)
Size	0.0001 ***	0.0001 ***	0.0001 ***	0.0001 ***
	(0.0000)	(0.0000)	(0.0000)	(0.0070)
Market to Book Ratio	0.0028 **	0.0091 ***	0.0029 **	0.0009
	(0.0170)	(0.0030)	(0.0260)	(0.7340)
Business Segments	-0.0912 ***	-0.0866	-0.1041 ***	-0.0516
	(0.0020)	(0.1740)	(0.0020)	(0.4920)
Dummy Loss	-0.0312 **	-0.0763 **	-0.0322 **	-0.0060
	(0.0420)	(0.0160)	(0.0470)	(0.8580)
Financial Crisis	-0.0676 ***	-0.0899 ***	-0.0652 ***	-0.0272
	(0.0000)	(0.0000)	(0.0000)	(0.1840)
Loss x Crisis	0.0012	0.0308	-0.0013	-0.0215
	(0.9460)	(0.4040)	(0.9450)	(0.5650)
Firm-fixed effects	Yes	Yes	Yes	Yes
Year-fixed effects	No	No	No	No
F Statistic	24.080	13.480	20.100	11.830
Adjusted R-Squared	44.530%	24.040%	41.410%	19.010%

Note: Sample contains 4,222 observations. P-values reported in parenthesis are computed using robust standard errors to obtain unbiased OLS coefficients. OLS estimates and its p-values in parenthesis. *p-value<0.1; **p-value<0.05; ***p-value<0.01.

Results shown in Table 14 allow us to affirm that Financial Crisis has a significant and negative impact on the reported tone in CEO Letters as a whole and at the section-level, excluding for Section 3. Business segments also seem to negatively affect tone, only showing statistical significance in the middle section and at the whole-text level.

The Market to Book Ratio, used to evaluate a firm's current market value relative to its book value, is also a significant determinant of tone in both whole-text and section-levels, except for the last section of the Letter. The length of CEO Letters is significant in helping explaining tone, having a positive association with it. Despite in this model ROA not showing a significant effect on tone, its

changes seem to impact positively the reported tone for almost all sections. Size, which is the logarithm of the market capitalization, has a positive and statistically significant effect on the Narrative Characteristic studied, i.e., tone.

In this model, our objective was to study the relationship between Losses and tone in crisis versus non-crisis periods. Moreover, we aimed to assess whether this relationship would change depending on the year of analysis. However, according to the lack of significance in the interaction term Loss x Crisis, this suggests that there is no difference between the two scenarios in the relationship between tone and losses, contrary to our expectations that in crisis periods firms would much likely have a negative association between these two last mentioned variables whereas in non-crisis periods it would be expectable that this association would be positive. This lack of significance may be because both variables are somehow overlapped, and their individual effect is hard to distinguish from one another, thus, it means that Loss does not add any relevant information to the Financial Crisis variable when predicting tone.

5. Conclusions

We examine several aspects about CEO Letters. We assess differences between tone measurement procedures and tone differences between sections of the same CEO Letter. Not only do we study tone but also the use of uncertainty and causal words. Moreover, we assess what are the effects of the Financial Crisis on companies' financial discourse. Our study also focuses on the relationship between accounting measures of firm's performance and tone to predict whether this relationship is impacted by the Financial Crisis, that is, we study if there is a change in the signal of this relationship between crisis and non-crisis periods.

According to our ANOVA results and the regression that we conducted, having the three dummy variables representing each section of the CEO Letters, we conclude that in fact the whole text approach and sections approach are

different from one another. Results also evidence that sections of CEO Letters display different tones between each other, which leads us to affirm that there are indeed textual differences between them. This heterogeneity among sections can be easily observed across almost all regressions conducted in our study, especially in tone regressions that show that section 3, which is a closing section and not particularly relevant in terms of estimating accurately the reported tone, does not have many statistically significant tone determinants, whereas some of these tone determinants are significant for other sections of the CEO Letter. This is a key indicator that the content of sections itself is a very important factor that should be considered when assessing firm's reported tone, mainly because considering sections that are not technical nor related to management' overview or predictions, can bias our conclusions. In other words, the conclusion parts of CEO Letters, for example, adds no additional information as all related

performance commentary and expectations are included in previous sections, and as such, should not have statistical significance when estimating tone.

Regarding Financial Crisis, our results show us that there is indeed a major effect of crisis on firm's discourse. We study CEO Letters published from 2005 to 2014, being 2008, 2009 and 2010 the Financial Crisis years that we considered in our study. We conduct an analysis on a sample comprising 4,222 CEO Letters, including only non-financial institutions. Tone is explained by the Financial Crisis. However, while there is indeed an increase in the use of uncertainty words during Financial Crisis years, the same cannot be stated for causal words for which we did not find any significant effect resulting from Financial Crisis.

With our study we wanted to assess if firms engage in impression management and more specifically, which years (financial crisis versus non-financial crisis years) are more likely for this to happen. The positive association between Change in ROA and tone shows us that firm's with increasing performance, report accurate information. However, according to our results, the interaction terms used, both in Model (3) and (4) of the tone regressions, show that financial crisis itself is our strongest variable in explaining tone and that accounting measures of performance do not deliver any incremental information to the financial crisis variable. This leads us to conclude that there are no differences between financial crisis and non-financial crisis years in the relationship between both accounting measures (ROA and Change in ROA), and tone. Suggesting that what dictates the tone is the financial crisis and not if there was a positive change in ROA, for example.

Even though our objective is to understand the effects of financial crisis on our companies' CEO Letters, we also study which firm dimension influences each sentiment. The tone determinants used control for profitability, market and financial performance, size, growth opportunities, firm's operating complexity

and we also add firm fixed effects (and year fixed effects in the absence of the Financial Crisis dummy variable – i.e., for models (1)).

Tone is positively and significantly associated with profitability – change in ROA reports a positive and significant coefficient, and the same for size and market to mook ratio. In other words, this means that firms adopt a more positive tone when profitability and size increase and, when the firm is more valuable.

As for the use of uncertainty words, there seems to exist a slightly significant and negative association between firm's current performance – ROA - and the use of these words. Also, there is a negative and significant association between loss and uncertainty words. This suggests that firms tend to use more frequently uncertainty words when they are performing poorly.

Regarding causality words, results depend on the section of the CEO Letter. For section 1, there seems to exist a negative and significant association between firm's performance – ROA – and the use of causality words, suggesting that firms adopt more frequently this type of speech when their performance is questionable.

Lastly, we also study the information power of market measures of firm's performance to see whether we get different results than we had for Models (3) and (4) of the tone regressions, using solely accounting measures of performance (ROA and Change in ROA).

Looking at the results in Model (5) - includes in the regressions the interaction term Return x Crisis – it suggests, above all, that market measures, in our sample, add incremental information to Financial Crisis variable when compared to accounting measures for which we did not capture any significant effect. In this model, firms report a positive and significant association between tone and firm's (accounting) performance – given by the ROA coefficient – which suggests that firms adopt a more positive tone if they perform better. However, the negative

and significant coefficient of Loss, suggests that firms do the exact opposite if they face losses.

Regarding Model (6), which includes the interaction term Loss x Crisis, we observe that Loss variable does not add any incremental information to the one given by the Financial Crisis indicator variable, suggesting that both variables are possibly overlapped and thus we cannot differentiate the effects from one another. As of the effects of each tone determinant in this model, these are quite like the ones reported in Model (5).

Our study faces some limitations, though. In our assessment of tone, we used a Bag of words technique which does not consider the position in which words appear throughout a text. However, using approaches that take the position of words into account would be hard to replicate and involve machine learning know-how.

One possible recommendation for future research would be to include a measure that captures CEO's characteristics as we are certainly aware that personality traits have also a reasonable importance in shaping firm's reported tone. Above all, most results reflect our expectations.

References

Abdelfattah, T., Bassyouny, H. & Tao, L. 2020. Beyond Narrative Disclosure Tone: The Upper Echelons Theory Perspective. ***International Review of Financial Analysis,*** 70

Abrahamson, E. & Amir, E. 1996. The Information Content of the President's Letter to Shareholders. ***Journal of Business Finance & Accounting,*** 23:8, 1157–1182

Aerts, W., Thewissen, J. & Yan, B. 2019. The informativeness of impression management – financial analysts and rhetorical style of CEO letters. ***Pacific Accounting Review,*** 31:3, 462-496

Aerts, W. & Yan, B. 2017. Rhetorical Impression Management in the Letter to Shareholders and Institutional Setting: A Metadiscourse Perspective. ***Accounting, Auditing & Accountability Journal,*** 30:2, 404-432

Alves, P., El–Haj, M., Rayson, P., Walker, M., & Young, S. 2016. ***Heterogeneous Narrative Content in Annual Reports Published as PDF Files: Extraction, Classification, and Incremental Predictive Ability.*** Working Paper, Lancaster University Management School, Lancaster University, UK

Amernic, J., Craig, R. & Tourish, D. 2010. ***Measuring and assessing tone at the top using annual report CEO letters***. The Institute of Chartered Accountants of Scotland

Argamon, S., Koppel, M., Pennybaker, J. W. & Schler, J. 2009. Automatically Profiling the Author of an Anonymous Text. **Communications of the ACM**, 52:2, 119–123

Babu, M., Imam, S. & Tan, K. L. 2019. The Use of Financial Information by Financial Analysts during a Financial Crisis. ***Strategic Change,*** 28:5, 369-379

Baddeley, A. & Hitch, G. 1993. The recency effect: Implicit Learning with Explicit Retrieval? ***Memory and Cognition,*** 21:2, 146-155

Bertrand, M. & Schoar, A. 2003. Managing with Style: The Effect of Managers on Firm Policies. ***The Quarterly Journal of Economics,*** 118:4, 1169-1208

Bliss, B. A., Doran, J. S., Peterson, D. R. & Price, S. M. 2012. Earnings Conference Calls and Stock Returns: The Incremental Informativeness of Textual Tone. ***Journal of Banking & Finance,*** 36:4, 992-1011

Bollen, L. H. H., Hassink, H. F. D. & Keusch, T. 2012. Self-serving Bias in Annual Report Narratives: An Empirical Analysis of the Impact of Economic Crises. ***European Accounting Review,*** 21:3, 623-648

Boudt, K. & Thewissen, J. 2015. Not All Words Are Equal: Sentiment Dynamics and Information Content Within CEO Letters. Working Paper No. 74

Boudt, K. & Thewissen, J. 2019. Jockeying for position in CEO Letters: impression management and sentiment analytics. ***Financial Management Association,*** 48:1, 77-115

Bournois, F. & Point, S. 2006. A letter from the president: seduction, charm and obfuscation in French CEO letters. ***Journal of Business Strategy,*** 27:6, 46-55

Bozec, R., Dia, M. & Bozec, Y. 2010. Governance – performance relationship: A Re-examination Using Technical Efficiency Measures. ***British Journal of Management***, 21, 684–700

Brunnermeier, M. K. 2009. Deciphering the Liquidity and Credit Crunch 2007–2008. ***Journal of Economic Perspectives,*** 23:1, 77-100

Buchholz, F., Jaeschke, R. & Lopatta, K. 2018. The use of optimistic tone by narcissistic CEOs. ***Accounting, Auditing & Accountability Journal,*** 31:2, 531-562

Bushee, B. 2001. Do Institutional Investors Prefer Near-Term Earnings Over Long-Run Value? ***Contemporary Accounting Research***, 18:2, 207–246

Costa, G. A., Oliveira, L. C., Rodrigues, L. L. & Craig, R. 2013. Factors associated with the publication of a CEO Letter. ***Corporate communications: An International Journal***, 18:4, 432-450

Clatworthy, M. A. & Jones, M. J. 2003. Financial Reporting of good news and bad news: evidence from accounting narratives. ***Accounting and Business Research,*** 33:3, 171-185

Clatworthy, M. A. & Jones, M. J. 2006. Differential patterns of textual characteristics and company performance in the chairman's statement. ***Accounting, Auditing & Accountability Journal,*** 19:4, 493-511

Cunitz, A. R. & Glanzer, M. 1966. Two Storage Mechanisms in Free Recall. ***Journal of Verbal Learning and Verbal Behaviour,*** 5:4, 351–360

Cunningham, C. 2005. Section 404 Compliance and 'Tone at the Top'. ***Financial Executive***, 21:5, 6

Davis, A., Piger, J. & Sedor, L. 2012. Beyond the Numbers: Measuring the Information Content of Earnings Press Release Language. ***Contemporary Accounting Research,*** 29:3, 845–868

Davis, A. & Tama-Sweet, I. 2012. Managers' Use of Language across Alternative Disclosure Outlets: Earnings Press Releases versus MD&A. ***Contemporary Accounting Research***, 29:3, 804– 837

DeFond, M. L. & Park, C. W. 1997. Smoothing Income in Anticipation of Future Earnings. ***Journal of Accounting and Economics,*** 23:2, 115−139

Demers, E. & Vega, C. 2010. Soft Information in Earnings Announcements: News or Noise? Working Paper No. 951, Fontainebleau, France

El-Haj, M., Rayson, P. & Young, S. 2015. UK Annual Report Narratives Dataset: CFIE-FRSE May 2019. Lancaster University

El–Haj, M., Rayson, P., Alves, P. & Young, S. 2018. Towards a Multilingual Financial Narrative Processing System. ***The First Financial Narrative Processing Workshop: In Proceedings of the 11th Edition of the Language Resources and Evaluation Conference – Miyazaki, Japan. Ed.*** 52–58

El–Haj, M., Rayson, P. Walker, M., Young, S. & Simaki, V. 2019. In Search of Meaning: Lessons, Resources and Next Steps for Computational Analysis of Financial Discourse. ***Journal of Business Finance and Accounting,*** 46:3–4, 2065– 306

Fisher, I. E., Garnsey, M. R. & Hughes, M. E. 2016. Natural Language Processing in Accounting and Finance: A Synthesis of the Literature with a Roadmap for Future Research. ***Intelligent Systems in Accounting, Finance and Management,*** 23:3, 157-214

Francis, J. 2001. Discussion of Empirical Research on Accounting Choice. ***Journal of Accounting and Economics***, 31:1–3, 309−319

Gennaioli, N., Shleifer, A. & Vishny, R. 2012. Neglected Risks, Financial Innovation, and Financial Fragility. ***Journal of Financial Economics,*** 104:3, 452–468

Gentry, R. J. & Shen, W. 2010. The Relationship between Accounting and Market Measures of Firm Financial Performance: How Strong Is It? ***Journal of Managerial Issues,*** 22:4, 514-530

Gibbins, M., Richardson, A. & Waterhouse, J. 1990. The management of corporate Financial Disclosure: opportunism, ritualism, policies, and processes. ***Journal of Accounting Research,*** 28:1, 121-143

Guidry, F., Leone, A. J. & Rock, S. 1999. Earnings-Based Bonus Plans and Earnings Management by Business-Unit Managers. ***Journal of Accounting and Economics,*** 26:1-3, 113−142

Healy, P. M., & Wahlen, J. M. 1999. A Review of the Earnings Management

Literature and its Implications for Standard Setting. ***Accounting Horizons***, 13:4, 365−383

Hambrick, D. C. & Mason, P.A. 1984. Upper Echelons: The Organization as a Reflection of its Top Managers. ***Academy of Management Review,*** 9:2, 93-206

Heaton, J. B. 2002. Managerial optimism and corporate finance. ***Financial Management Association,*** 31:2, 33-45

Huang, A. H., Zhang, A. Y. & Zheng, R. 2014. Evidence on the Information Content of Text in Analysts' Reports. ***The Accounting Review,*** 89:6, 2151-2180

Huang, X., Teoh, S. H. & Zhang, Y. 2011. Tone Management. ***The Accounting Review,*** 89:3, 1083-1113

Lambert, R. A. 2001. Contracting Theory and Accounting. ***Journal of Accounting and Economics***, 32:1–3, 3–87

Laux, C. & Leuz, C. 2010. Did Fair-Value Accounting Contribute to the Financial Crisis? ***Journal of Economic Perspectives,*** 24:1, 93-118

Li, F. 2010. The information content of Forward-Looking Statement in Corporate Fillings – A Naïve Bayesian Machine Learning Approach. ***Journal of Accounting Research,*** 48:5, 1049-1102

Liu, P. & Nguyen, H. T. 2019. CEO characteristics and tone at the top of inconsistency. ***Journal of Economics and Business,*** 108

Loughran, T. & McDonald, B. 2011. When Is a Liability Not a Liability? Textual Analysis, Dictionaries, and 10-Ks. ***The Journal of Finance,*** 66:1, 35-65

Loughran, T. & McDonald, B. 2015. The Use of Word Lists in Textual Analysis. ***Journal of Behavioural Finance,*** 16:1, 1-11

Loughran, T. & McDonald, B. 2016. Textual Analysis in Accounting and Finance: A Survey. ***Journal of Accounting Research,*** 54:4, 1187-230

Magnan, M. L. 2009. Fair Value Accounting and the Financial Crisis: Messenger or Contributor? ***Accounting Perspectives,*** 8:3, 189-213

Malmendier, U. & Tate, G. 2005. CEO Overconfidence and Corporate Investment. ***The Journal of Finance,*** 60:6, 2661-2700

Meek, G., Robert, C. and Gray, S.1995. Factors influencing voluntary annual report disclosures by US, UK and Continental European multinational corporations. ***Journal of International Business Studies***, 26:3, 555-572

Merkl–Davies, D. M. & Brennan, N. M. 2007. Discretionary Disclosure Strategies in Corporate Narratives: Incremental Information or Impression Management. ***Journal of Accounting Literature,*** 27, 116–196

Patelli, L. & Pedrini, M. 2014. Is the optimism in CEO's Letters to shareholders sincere? Impression Management versus Communicative action during economic crisis. ***Journal of Business Ethics,*** 124, 19-34

Perols, J. L. & Lougee, B. A. 2011. The Relation Between Earnings Management and Financial Statement Fraud. ***Advances in Accounting,*** 27:1, 39-53

Rezaee, Z. 2005. Causes, Consequences, and Deterrence of Financial Statement Fraud. ***Critical Perspectives on Accounting,*** 16:3, 277-298

Scholes, M. S., Wilson, G. P., & Wolfson, M. A. 1992. Firms' Responses to Anticipated Reduction in Tax Rates: The Tax Reform Act of 1986. ***Journal of Accounting Research,*** Working Paper No. 4171, National Bureau of Economic Research

Shaffer, S. 2010. ***Fair Value Accounting: Villain or Innocent Victim – Exploring the Links Between Fair Value Accounting, Bank Regulatory Capital and Recent Financial Crisis***. Working Paper No. 10-01. Federal Reserve Bank of Boston

Thakor, A. V. 2012. The Financial Crisis of 2007–2009: Why Did It Happen and What Did We Learn? ***The Review of Corporate Finance Studies,*** 4:2, 155-205

Wang, G., Holmes, R. M., Oh, I. S. & Zhu, W. 2017. Do CEOs Matter to Firm Strategic Actions and Firm Performance? A Meta-Analytic Investigation Based on Upper Echelons Theory. ***Personnel Psychology,*** 69:4, 775-862

Yuthas, K., Rogers, R. & Dillard, J. F. 2002. Communicative Action and Corporate Annual Reports. ***Journal of Business Ethics***, 41, 141-157

Young, S. 1998. The Determinants of Managerial Accounting Policy Choice: Further Evidence for the UK. ***Accounting and Business Research***, 28:2, 133−143

Zhang, S. & Aerts, W. 2015. Management's Performance Justification and Failure to Meet Earnings Thresholds. ***Accounting and Business Research,*** 45:6–7, 841–868

Appendix

Table 15 – Variable's description

Variable name	Definition
No. of words	Number words present in the whole CEO Letter
Tone	Net tone of the whole CEO Letter: the difference between the number of positive and negative words scaled by their sum. Words are defined as positive and negative according to Loughran and McDonald's list
No. of positive words	Number of positive words present in the whole CEO Letter proposed by Henry's finance-specific words list
No. of negative words	Number of negative words present in the whole CEO Letter proposed by Henry's finance-specific words list
No. of uncertainty words	Number of uncertainty words present in the whole CEO Letter proposed by Loughran and McDonald words list
No. of causal words	Number of causal words present in the whole CEO Letter proposed by Loughran and McDonald words list
No. of words_Sx	Number words present in Section x, where Sx refers to Sections 1, 2 or 3 of the CEO Letter
Tone_Sx	Net tone of CEO Letter's Section x, where Sx refers to Sections 1, 2 or 3: the difference between the number of positive and negative words scaled by their sum. Words are defined as positive and negative according to Loughran and McDonald's list
No. of positive words_Sx	Number of positive words present in CEO Letter's Section x, where Sx refers to Sections 1, 2 or 3 proposed by Henry's finance-specific words list
No. of negative words_Sx	Number of negative words present in CEO Letter's Section x where Sx refers to Sections 1, 2 or 3 proposed by Henry's finance-specific words list
No. of uncertainty words_Sx	Number of uncertainty words present in Section x where Sx refers to Sections 1, 2 or 3 of the CEO Letter proposed by Loughran and McDonald words list
No. of causal words_Sx	Number of causal words present in Section x where Sx refers to Sections 1, 2 or 3 of the CEO Letter proposed by Loughran and McDonald words list
ROA	Return on assets: net income from accounting scaled by lagged total assets
Change in ROA	Change in Return on assets: net income change (the difference between the current year and the previous one) scaled by lagged total assets
Return	12-month raw stock return
Earnings Yield	Earnings per share scaled by lagged stock price
Change in Earnings Yield	Change in Earnings Yield: earnings per share variation (the difference between the current and the previous one) scaled by lagged stock price
Market Value	Market value of Equity at fiscal year-end
Market to Book Ratio	Market to Book Ratio: Firm's market value scaled by its book value
Business Segment	Log (1+number of business segments)
Loss	Indicator variable that takes the value one if the earnings yield is lower than zero in the current year and zero otherwise
Financial Crisis	Indicator variable that takes the value one if year is 2008, 2009 or 2010 and zero otherwise
ROA x Crisis	Interaction term between the Return on assets of the current year and the indicator variable Financial Crisis
Change in ROA x Crisis	Interaction term between the Change in Return on assets of the current year and the indicator variable Financial Crisis
Return x Crisis	Interaction term between Return of the current year and the indicator variable Financial Crisis
Loss x Crisis	Interaction term between the indicator variable Loss of the current year and the indicator variable Financial Crisis