UNIVERSIDADE CATÓLICA PORTUGUESA

Can Uncertainty Explain Residual Returns?

Evidence from UK non-financial firms

Rodrigo Ferreira Borges da Costa

Católica Porto Business School
April 2024

UNIVERSIDADE CATÓLICA PORTUGUESA

Can Uncertainty Explain Residual Returns?

Evidence from UK non-financial firms

Final Assignment in the modality of Dissertation presented to Universidade Católica Portuguesa to obtain the degree of Master in Finance

by

Rodrigo Ferreira Borges da Costa

under the supervision of

(PhD) Paulo Alexandre Pimenta Alves

Católica Porto Business School
April 2024

Acknowledgments

I would like to express my gratitude to the individuals and institutions whose support was invaluable during the writing of this work. First, I want to thank my parents for their encouragement, love and belief in my abilities. Without them, this whole journey wouldn't have been possible.

I'm also thankful to Católica Porto Business School for providing the best academic environment with resources, opportunities and guidance during my time there. At last, but as important as the others, I want to thank my supervisor, Prof. Paulo Alves (PhD) for his mentorship, patience and expertise.

Resumo

Este estudo analisa o impacto da incerteza nas secções do CEO e do Presidente dos relatórios anuais de empresas não financeiras do Reino Unido sobre os retornos residual das ações, não explicados pelo modelo de 3 Fatores de Fama-French. O estudo incorpora dados de 1344 empresas cotadas na Bolsa de Valores de Londres durante o período de 2011-2016.

Ao aplicar uma metodologia que usa regressão linear e análise textual, este estudo preencher uma lacuna na literatura na falta de estudo sobre incerteza e os seus impactos. Contrariamente ao que era inicialmente esperado, estes resultados sugerem que a incerteza como medida de narrativas financeiras não impacta os resíduos anuais em nenhuma das secções analisadas, não sendo, portanto, um explicador fiável do desempenho das ações.

Os resultados também desafiam anteriores noções sobre a influência dos estilos de narrativa dos executivos no comportamento do mercado e fornecem conhecimentos para os investidores sobre o peso da incerteza nos relatórios das empresas no mercado. O estudo sublinha a necessidade de mais pesquisa neste campo, especificamente para diferentes secções de relatórios e períodos temporais. Este trabalho contribui para a literatura testando o poder de explicar da incerteza e fornece implicações práticas para a forma como investidores e analistas interpretarão o aspeto da incerteza na divulgação corporativa.

Palavras-chave: narrativas financeiras; incerteza; relatório anual; retornos residuais; tomada de decisão de investidores; Bolsa de Valores de Londres

Número de palavras: 7837

Abstract

This study analyses the impact of textual uncertainty in the CEO and Chairman sections of UK non-financial firms' annual reports on residual stock returns not explained by the Fama-French 3-Factor model. The study incorporates data from 1344 companies listed on the London Stock Exchange over the 2011-2016 period.

By applying a methodology that uses Ordinary Least Squares regressions and textual analysis, this study aims to bridge a gap in the literature about uncertainty and its impacts. Opposite to what was expected initially, these findings suggest that uncertainty as a measure of financial narratives does not impact annual residuals in any of the sections analysed, thus, it is not a reliable explanator of stock performance. The findings also challenge previous notions about executive narrative styles' influence on market behaviour and provide insights for investors about the weight of uncertainty in corporate reports on the market.

The study underscores the necessity for further research in this field, specifically for different report sections and time frames. This work contributes to the academic literature by testing the predictive power of uncertainty and provides practical implications for the way investors and analysts will interpret the uncertainty aspect of corporate disclosure.

Keywords: financial narratives; uncertainty; annual report; residual returns; investor decision making; London Stock Exchange

Number of words: 7837

Table of Contents

Acknowledgments iii

Resumo v

Abstract vii

List of tables xii

1. Introduction 1
2. Literature Review 5
    2.1. Textual Analysis in Finance 5
    2.2. Uncertainty Analysis in Corporate Communications 7
    2.3. Investor Reaction to Financial Reports 9
    2.4. Narrative in Executive Communications 12
3. Research Question and Hypothesis 14
4. Data and Methodology 16
    4.1. Sample selection and data collection 16
    4.2. Methodology 20
5. Results 22
6. Conclusion 27

References 29

Appendix 34

List of Tables

Table 1: Sample by industry 18

Table 2: Descriptive statistics – uncertainty measures 19

Table 3: OLS coefficients estimates – individual regressions 23

Table 4: OLS coefficients estimates – section measure 24

Table 5: OLS coefficients estimates – narrative measure 24

Table 6: Variables Definition 34

1. Introduction

Financial narratives are, in a simple way, the qualitative information that companies provide to the public, for example, in their annual reports. This data has been overlooked in the past mainly due to the complexity of measuring, understanding and assessing the immense textual data available to the public accurately. Past literature suggests that investors no longer respond to financial statement releases with large changes, as they once did, (Brown & Tucker, 2011). This lack of announcement effect is not due to financial statements becoming less informative, but rather from investors missing subtle and important signals from company reports at the time of their release (Cohen et al., 2020), mainly due to the increased complexity and length of these reports.

More research regarding financial narratives and how a company's underlying financial performance affects it is focused on the chairman's statement of the most profitable and unprofitable UK firms. It found the presence of "impression management techniques", where management is more or less propense to associate themselves with the firm's results, depending on its financial performance (Clatworthy & Jones, 2006). Although these findings have limitations, they provide good insights into the importance of analysing qualitative financial data.

This research aims to evaluate how the uncertainty, as a measure of text narrative, for the CEO and Chairman sections can explain the annual returns not explained by the Fama-French 3-factor model. This paper also proposed four other variables that are variations from the uncertainty variable that aim to capture a fuller picture of the hidden impacts of this narrative on the market. While textual changes' impact on stock performance has been studied in

previous research, understanding the relationship between the CEO and Chairman's uncertainty in their statements and its explanatory power on stock returns remains a significant gap in the literature. More specifically, unifying the same variable for different sections of an annual report and the change over time. The analysis method chosen has been used before when it comes to studying the communication style in executive members' communications and their impact on market behaviour.

The period of analysis in this study goes from 2011 to 2016 and it analyses a total of 1,344 non-financial companies from the London Stock Exchange. The data sources used are the following: i) Refinitiv Datastream, used to retrieve market, accounting and other firm characteristics; ii) the database available on the UK data service that provides the Fama-French 3 factors for the UK market made available by the University of Exeter and iii) financial narratives data obtained from El-Haj et al. (2020).

It is important to note that the number of observations declined at each step of the matching and formatting process of the data. This is mainly due to: i) the annual report of a company could not be obtained or ii) text extraction issues that result in unusable output.

The regression model employed uses the calculated annual residual returns as the dependent variable while uncertainty and its derived measures for the CEO, Chairman and merged sections are used as the independent variable across each section.

Contrary to what was initially thought, the uncertainty measures do not have a strong impact on the annual residual returns on any section that was analysed. Thus, this measure does not help investors assess a company’s value. Additionally, the other four proposed measures did not show any explanatory relevance on returns for each of the sections.

Therefore, we conclude that an investor looking to obtain a better understanding of a company's performance should not consider uncertainty when looking for a more informed decision as it does not provide any significant predictive value.

This research builds upon the discoveries made by Cohen et al. (2020) regarding the significance of textual alterations in financial reports and their influence on the underlying stock returns, also, valuable insights into the predictive power of different narrative sections in annual reports are provided by Alves et al. (2016).

The remainder of this dissertation is structured as follows: Chapter 2 provides an insightful review of the literature, covering the theoretical framework on financial narratives, ways to measure them and their impact on capital markets. Chapter 3 presents the research question and the respective hypothesis. Chapter 4 covers the methodology used, including data collection and sample selection. Chapter 5 analysis the results and Chapter 6 offers concluding observations.

2. Literature Review

2.1. Textual Analysis in Finance

The way executives and managers speak about their company provides several linguistic features that can be identified and captured to better assess the firm's underlying risk. Features such as sentiment, whether it is positive, negative or neutral, uncertainty, forward-looking, casual reasoning, similarity of words, complexity or readability are aspects that can impact the perception of a firm's risk profile. For example, when a company's disclosure is favourable, the firm's risk, which can be a proxy for stock return volatility, declines significantly, while unfavourable disclosures are accompanied by significant increases in risk measures (Kothari et al., 2009).

The relationship between annual report readability, firm's performance and earnings persistence is a prominent topic of study in this field, mainly from the SEC plain English disclosure regulation of corporate disclosures that attempt to make firms' reports standardized and easier to read. Considering that annual reports contain detailed quantitative data on historical performance, executives choose to be opportunistic and obfuscate information regarding future performance. Hence, disclosure readability is strategically used by executives to hide negative information from investors (Li, 2008a).

Early research makes use of the Fog index, going as far as defining it as a "measure of financial statement readability"(Biddle et al., 2009; Dougal et al., 2012; Lawrence, 2013; Li, 2008b), which is a readability test for English writing. It is a commonly applied readability measure but has shown to be poorly specific regarding financial applications. This index uses two components, which are either non-specific or difficult to measure, this makes using document file size a

simpler proxy for readability that commonly outperforms the Fog Index (Loughran & McDonald, 2014). Words such as *Company*, *Operation* and *Revenue* are considered complex words according to the index and yet, these are words commonly found in a financial report.

Another readability measure was developed as an upgrade to the Fog Index, the Bog Index. This index is a better way to measure readability and style of a financial document since, opposing to the Fog index, which uses sentence length and syllable count to measure readability, here, the key feature of this index is a graded 200-thousand-word dictionary where each word has a grade from easy to difficult to read depending on the word's frequency and ease of understanding. This index is proposed as a new measure for readability in financial reporting, and it is positively correlated with post-report filing stock return volatility where less readable fillings are associated with greater future volatility (Bonsall et al., 2017). Although the Bog index has the largest association with future stock volatility, it fails with industry-specific words. Words like *Immunochemical*, *Iodide*, *Oxytocin* and many others are words that get high bog scores when, in an industry context, these words are not complex.

From the limitations that arise from using either one of the readability measures, the Fog and Bog index, one can conclude that these can lead to inaccuracies when used for the purpose of this research since we aim to find a relationship between a broader non-industry-specific firms' annual reports narratives and stock returns with a measure more complex than readability.

2.2. Uncertainty Analysis in Corporate Communications

So far, a brief overview of the linguistic features that can be detected and measured in financial documents, as well as a comment on the most used readability measures and their limitations was done.

A good approach is a dictionary-based approach that focuses on word count. Previous research suggested negative word counts to measure certain narratives in a text can be effective, as highlighted by significant correlations with financial variables. Examples include Antweiler & Frank (2004), Engelberg et al. (2008), Tetlock (2007) and Tetlock et al. (2008). New research suggests an alternative negative word list for specific use in financial texts that reflects the uncertainty.

There are several different sources used for word classifications. A commonly used source is the Harvard Psychosociological Dictionary, highlighting the Harvard-IV-4 TagNeg (H4N) file. A feature of this list for research is that the researcher cannot pick or choose which words have negative implications. The English language has many meanings so using a word categorization scheme that is limited to one discipline's dialect might not translate effectively into other disciplines. The list of negative words provided by the H4N may, in some contexts, accurately capture the sentiment of a text, but might not capture it as well in others. The question raised by Loughran and McDonald (2011) is whether a wordlist developed for psychology translates into the realm of business.

While the authors recognize that measuring narrative characteristics using word classification is hard, the H4N list misclassifies a significant number of words in a financial context. Misclassifying words that are not directly correlated with the variables considered adds noise to the measurement of narratives, for example, *taxes* or *liabilities* are words classified as negative but are common in a financial context. A similar limitation from the previous readability measures is

found when using non-business-specific word lists, where misclassifying words can have a big impact on this study's results.

Loughran and McDonald (2011)'s initial findings suggest that almost three-fourths of the H4N word list classifies as negative words that, in a financial context, are not. Words like *tax*, *cost*, *capital*, *board*, *foreign* and *vice* appear on the H4N list but are frequently used in companies' reports to name a *board* of directors or a company's *vice*-president. Other words present on the Harvard's list, such as *mine*, *cancer*, *crude*, *tire* and others, are likely to identify a specific industry of a company rather than reveal negative events.

Loughran and McDonald (2011) developed Fin-Neg, a list of more than 2000 words that have negative implications in a financial sense and based on usage frequency. This list includes not only words already present in the H4N but also includes words more frequently used in financial documents that are less industry-specific and have more negative implications for sentiment, like *felony*, *litigation*, *restated* and *misstatement*. In addition, Loughran and McDonald (2011) created five more wordlists, one of them being uncertainty, or *Fin-Unc* by the authors.

It is still a topic of discussion whether narrative characteristics should be captured in the entire document or a specific part. While the MD&A section does not give a reliable narrative score that has a discernible impact on a company's report file date excess returns, sections like the CEO and Chairman statements can predict future earnings above contemporaneous earnings (Alves et al., 2016). Also, when looking for a relation between both word lists and filing returns, for the Harvard list, increments on H4N words on documents are not indicative of fewer report filing returns, while for the proposed word lists this is a strong pattern.

Aiming to further expand the knowledge in the field of financial narratives and how different narrative classification methods can predict companies'

unexpected returns. This paper proposed using the additional word list, *Fin-Unc*, made available by Loughran and McDonald (2011). Although Alves et al. (2016) did not test the predictive power of uncertainty in business reviews signed by the CEO and in the letter to the shareholders signed by the Chairman, the creators of *Fin-Unc* found evidence that this word list is related to market reactions around report filing date, trading volume, stock return volatility and especially unexpected earnings.

It is important to test the capabilities of uncertainty as an explanatory variable of unexpected returns. Capturing the predictive power of uncertainty for the CEO and Chairman sections on unexplained returns will give us insights into the usability of this measure and how it relates to the financial market. Additionally, identifying the relationship between uncertainty from the CEO and the Chairman sections can provide a better understanding of the sentiment in executive communications and the potential impact on stock prices. It can provide valuable insight into the correlation between market behaviour and executive and manager narratives.

2.3. Investor Reaction to Financial Reports

Literature has indicated that financial reports and disclosure are relevant and important means for management to communicate how a firm is performing to investors (Healy & Palepu, 2001).

Investors have reacted to information in firms' reports differently throughout time. As the amount of information produced increases, mainly because the cost of production and dissemination of information is decreasing, it leads to more complexity in the search and processing of information. If investors don't keep

up with the increased complexity, disclosed information may not be fully incorporated into prices.

When taking a financial narrative approach, by using methodologies to detect changes in documents, research finds that changes in documents impact stock prices in a significant way, but the full impact of the news gets gradually incorporated into prices, eventually, it gets fully impounded into stock prices and firm operations (Cohen et al., 2020). In contrast to other studies that defended corporate documents are less informative and hence are not useful to investors in the current capital markets situation, Cohen et al. (2020) defend that investors fail to account for or are inattentive to the rich information present in document changes.

A part of a financial report that investors find valuable and useful is the Management Discussion and Analysis or MD&A. As the MD&A section of a report must include certain topics, it is an easy way for investors to get the information they need to make their investment decisions. Investors respond stronger to a firm's annual report filings when there are noticeable changes to the MD&A section, as it provides information that helps predict future cash flows (Brown & Tucker, 2011). Although it can be seen as a "gold mine" for investors, managers have the flexibility to choose how in-depth they discuss a topic. This is intended as a way for managers to give their personal views of the firm to investors. Topics covered in the MD&A section help investors assess the firm's past and current performance as well as their financial condition, therefore, if this section is unchanged from previous years, it is unlikely to serve its purpose.

Studies find a trend where MD&A sections of firms' annual reports become lengthier but contain less useful information mirroring a growing sckepticism amongst investors about the credibility of these sections (Kothari, Li, & Short, 2009). Managers take advantage of many words to disclose standard information, with little firm or fiscal period-specific content. It is visible the surge

of impression management technics (Clatworthy & Jones, 2006) as discussed before. Also, over time, price responses to MD&A modifications have become weaker, suggesting a decline in the overall usefulness of this section, despite regulations, like the SEC 2003 guidance, on improving the MD&A.

Kothari, Li and Short (2009) offer a wider view of how the content of disclosure impacts investor reactions, particularly about the cost of capital and return volatility. They focus on the different impacts that different sources like management, analysts and the business press have on investors. They find markets often devalue management statements, which suggests scepticism regarding their credibility. This comes on par with the findings of Brown and Tucker about MD&A sections becoming more redundant. The lack of trust in management disclosures results from management biases when presenting a company's performance, tending to be overly favourable. In contrast, disclosures from the business press have a greater impact on investor reactions.

The content of a disclosure is also relevant for investors. While a favourable disclosure, detected using content analysis (Kothari et al., 2009), is reflected by a decline in both cost of capital and stock return volatility, unfavourable disclosures correlate with an increase in risk measures. This implies that not only the presence of changes in a section of a financial report shapes investor and market behaviour but also the nature of the content, whether it's favourable or not impacts investors.

The trend of lengthier but less informed MD&A sections directly correlates to Kothari et al. (2009)'s findings. As MD&A changes see a decrease in credibility and become less informative, it explains the systematic decline in this section's influence on investor decisions.

The findings presented so far point towards an emerging trend, where the once "information spaces" like the MD&A are less impactful to investors, caused by the systematic lengthier sections with less information, leading to a shift in the

perimeter of investor influence and information processing. This trend, combined with a lack of trust towards management narratives highlights the importance of steering the focus on alternative sections of a financial report, expecting richer and more credible information, like the Chief Executive Officer (CEO) and the Chairman's statements.

2.4. Narrative in Executive Communications

Facing the complexity of corporate financial reporting, key executives' narratives have an important role in the way they shape the investor and the market.

Content analysis of disclosure is a vast field of study for research. These usually focus on reading ease, themes or visuals, performance comparisons and choice of earnings number, amongst others (Merkl-Davies & Brennan, 2007). Isolating linguistic features for content analysis falls in the theme category, more specifically by trying to assert meaning without using all aspects that are part of the linguistic feature like voice or body language, which can be problematic to solve.

The use of uncertainty analysis in financial reports is not yet well explored but it is seeing a significant increase in attention provided by new research. Although it is a complex linguistic feature to measure using word classification methods, as highlighted by Loughran and McDonald (2011), extracting linguistic features and linking it to stock returns, volatility and firm performance is a common route for research (Demers & Vega, 2008; Henry, 2008; Mckay Price et al., 2011; Sadique et al., 2008), suggesting that there is relevant information on managers' words of choice held by statistically significant results. Therefore, despite not capturing

the full picture of "uncertainty", its analysis is relevant in corporate communications as it impacts market behaviour.

Regarding investor preference, studies indicate that firms that keep information "in the dark" are riskier for uninformed investors and so demand for these firms' shares decreases (Diamond & Verrecchia, 1991; Easley & O'Hara, 2004). Consistent with the findings of Li (2008) and the Securities Exchange Commission (SEC) (2007), Lawrence's (2013) findings indicate that firms with less readable and longer annual reports are less transparent and have a lower quality financial disclosure.

To individual investors, effective communication is key as lengthier and less readable reports take more time and effort to obtain relevant information. Coupled with investor's limited attention (Hirshleifer & Teoh, 2003) and the fact that noisy communication decreases an investor's ability to acquire information (Indjejikian, 1991), it is clear that disclosure clarity plays a significant role in a firm's disclosure quality as well as the amount of information investors can gather.

If the CEO and Chairman sections exhibit a lack of clarity and conciseness, this can inevitably influence the investor's perspective on the firm and its investment decisions, therefore impacting the firm's stock performance.

3. Research Question and Hypothesis

Empirical evidence indicates that a way to predict a company's future performance can be by analysing the published annual report published. Since each section of an annual report has its distinctive predictive capacity regarding the firm's performance (Alves et al., 2016), it is vital to analyse different annual report sections, how they link and test their predictive capacity. It is important to mention that multiple authors are involved in making these reports, thus the linguistic style of reporting is subject to variations (Argamon et al., 2009). Also, companies with more robust governance structures are typically associated with higher quality disclosures, as evidenced by García Osma & Guillamón-Saorín (2011), hence the type of governance in place may affect the quality of disclosure.

The main objective of this study is to evaluate how uncertainty as a narrative measure is indicative of residual returns obtained from the Fama-French 3-factor model in the CEO and Chairman sections. We compare the narrative characteristics with the annual residual from a security. Additionally, because changes in narrative can have predictive power on residual returns (Cohen et al., 2020), it is proposed four new narratives that are explained below. The paper focuses on a specific narrative, uncertainty, and how variations of this narrative serve as a meaningful attribute that can explain the unexplained residual returns by the model.

The narrative measures analysed in this study are: uncertainty, for both the CEO and Chairman sections; *uncertainty_sum*, the sum of the uncertainty measures for the same firm; and, finally, the change over the years analysed on all these variables.

The sections used were selected based on their relevance and on the freedom of writing each section's authors have when compared to more regulated or standardized ones like the MD&A section.

The MD&A section was excluded since studies show it no longer can predict future returns, caused by lengthier and less detailed sections, making narrative measurements less reliable since the text tends to become neutral. As mentioned before, this emerging trend, incorporated with the prevailing mistrust towards management narratives, leads to a less impactful section for investors. Hence, a shift of focus towards alternative sections of a financial report that are expected to have richer information is likely.

One section that is proposed to contain rich information is the CEO section since it is tied to the thoughts of top management regarding the company's performance, so there is a direct link of information between this section and the MD&A which may be a good indicator of good predictability on expected returns. The Chairman's statement section was included to complement the CEO section and to provide a wider view of uncertainty's predictive power.

A relevant insight about the Chairman section is that the Chairman usually does not have an active role in daily operations and previous research indicates that it is a good predictor for financial distress in firms (Smith & Taffler, 2000). Therefore, studying these sections together and their relation regarding the specific narrative measure in question is relevant.

To investigate the impact of uncertainty and its derived measures on predicting future returns and on its importance for investors, the below hypotheses were formulated:

Hypothesis 1: *The uncertainty in the discourse by top executives in the annual report can explain annual residual returns.*

H_1a: the uncertainty words present in the CEO section can help explain annual residual returns.

H_1b: the uncertainty words present in the Chairman section can help explain residual returns.

Another hypothesis is that uncertainty should not be considered by itself but using uncertainty of the top executives of the company as one whole variable. Here, we test how impactful these sections together have when compared to separate.

Hypothesis 2: *Linking uncertainty measures in companies' annual reports can explain residual returns.*

H_2a: the sum of uncertainty between the CEO and Chairman sections can explain residual returns.

Finally, the change in uncertainty over the time frame of the study could have some predictability power regarding residual returns.

Hypothesis 3: *Changes over the year in uncertainty in companies' annual reports can explain residual returns.*

H3a: the change in uncertainty in the CEO section can explain residual returns.

H3b: the change in uncertainty in the Chairman section can explain residual returns.

H3c: the change in the sum of uncertainty between the CEO and Chairman sections can explain residual returns.

4. Data and Methodology

4.1. Sample selection and data collection

The sample used in this study comprises non-financial companies listed on the London Stock Exchange from 2011 to 2016. This period was selected because it covers the 5 most recent years, considering that the data available to calculate the Fama French 3-factor model for the UK market was only available until 2016.

Financial firms had to be excluded since they have unique characteristics as identified by Athanasakou & Hussainey (2014) and later by Hassanein et al. (2019). Companies were excluded if monthly returns could not be calculated from lack of data or if their residuals were not sufficiently significant statistically.

The original sample consists of 1,371 companies, accounting for 6,105 observations of annual residuals. After merging the residuals with the corresponding narrative measures and applying a trimming method to remove outliers by eliminating observations that are away from the mean more than three times the standard deviation, the final sample size is now 1,344 companies or 4,948 observations, representing the total number of unique annual reports analyzed.

A python code was developed to calculate the sum of uncertainty when the report had both sections and to calculate the change on uncertainty over the years, controlling for year and company. Leaving the final data with 6 independent variables.

The industries of each of the 1,344 companies from the final data vary. The distribution of industries across the sample is presented below, in table 1.

Table 1: Sample by industry

	Freq.	Percent
Basic materials	167	12,4%
Consumer Discretion	293	21,8%
Consumer Staples	66	4,9%
Energy	137	10,2%
Health Care	117	8,7%
Industrials	311	23,1%
Technology	186	13,8%
Telecommunications	39	2,9%
Utilities	28	2,1%
Total	1344	100%

From 2011 to 2016, it is visible a strong presence of companies in the consumer discretion and industrials sectors. In contrast, the sectors of utilities and telecommunications see a less populated market.

The monthly factors for each year of the sample used to calculate the Fama-French 3-Factor model were retrieved from the database from a UK data service made available by the Finance and Accounting Department of the University of Exeter. To compute the returns of the companies, the necessary data was retrieved from Refinitiv Datastream.

Using annual reports from UK companies to compute narratives has proved to be a challenge. Although UK firms must follow what the IASB Regulation imposes, each firm exercises significant variance regarding the structure and the content of their reports, as documented by El-Haj et al. (2020). In extent, the annual reports of these firms are available in PDF format, posing one additional challenge when retrieving and processing the information. Additionally, the variations in content and structure can happen even between the same company's reports from different years. To overcome these challenges, the authors of the paper cited develop an automated procedure to extract and

categorize the narrative component of glossy annual reports in PDF file. The output data of his procedure is utilized in this study.

With this, the annual data regarding uncertainty for the CEO and Chairman sections analyzed were obtained from El-Haj et al. (2020) study and transformed logarithmically to address distributional properties. In table 2, there are the descriptive statistics of this data.

Table 2: Descriptive statistics - uncertainty measures

Variable name	N	Mean	Std.Dev.	Min.	Median	Max.
residual_return	4,948	0.043659	0.040085	0.000048	0.035533	0.376834
uncertainty_CEO	2,539	0.004520	0.003130	0	0.003912	0.021698
uncertainty_Chair	4,098	0.004774	0.003069	0	0.004370	0.015855
uncertainty_sum	1,842	0.009614	0.004422	0.001180	0.008983	0.031553
uncertainty_CEO_change	1,651	0.000021	0.003221	-0.018012	0	0.019275
uncertainty_Chair_change	2,746	-0.000050	0.003510	-0.014039	0	0.015038
uncertainty_sum_change	1,005	-0.000028	0.004433	-0.021173	-0.000099	0.017383

Considering that a total number of 4,948 annual reports was examined, no variables were observed for the entire sample, as shown by the inferior number of observations from the uncertainty variables when compared to the *residual_return* variable. This is due to not all annual reports analyzed having the CEO section or the Chairman. The number of annual reports that have both sections is 1,842. The average firm's annual report has a similar uncertainty score in the Chairman section (0.0047) when compared to the CEO section (0.0045). Also, throughout the years, we see that the CEO section tends to increase the number of uncertainty words when compared to the negative trend in the Chairman section, as seen by their respective signal. The CEO section registered the largest positive variation in uncertainty (0.01928), while the top executive comment had the largest lowest variation (-0.02117). Finally, the change in uncertainty from the overall top executive communications tends to also change negatively.

4.2. Methodology

This paper studies whether uncertainty as a narrative measure can be a good explanation for residual returns obtained from an Asset Pricing Model. More precisely, it aims to determine if uncertainty and its derived variables can benefit investors when trying to predict the expected returns of a company.

By selecting the Fama-French 3-factor model (Fama & French, 1992), adds significant factors that influence stock returns, such as market risk, size, and value. This model is proven to be useful in explaining the majority of returns, enabling investors to make better-informed investment decisions by allowing them to get a precise understanding of the risks and opportunities associated with a certain company.

Over the years, this model has proven to be efficient by gathering extended empirical support. The model being easily customized also helped this study since adopting other supplementary factors is relatively simple.

By adopting this model it is possible to compute the residual returns that can not be accounted for by the three factors of the Fama-French 3-Factor model (market risk, size and value). The definition of residual returns is the difference between the real return of a security and the expected return computed by the model. Therefore, this expected return can be expressed as the sum of the expected returns derived from the Fama-French 3-Factor model and the expected return, represented by an error term (Fama & French, 1992). The use of the model is not the focus of the study, it is simply a way to a means.

The following regression was performed monthly for the companies in the previous sample:

$$R_{it} - R_{ft} = (R_{mt} - R_{ft}) + SMB_{it} + HML_{it} + \varepsilon_{it}$$

By running an Ordinary Least Squares (OLS), we computed the residual return for each company i in month t. Considering that the narratives data is derived from annual reports, the monthly residuals were then converted to annual residuals by computing the annual mean of the residuals considering the fiscal year end month. From the initial 6,105 annual observations, after matching with the narratives, the final sample was of 4,948 annual residual returns.

To mitigate the effect of the direction of the return and to better analyse the impact of the magnitude of these returns, it is considered the absolute value of the annual residual return.

Eleven OLS regressions were performed to test the hypothesis presented previously. Its objective is to assess the impact of uncertainty and its derived narratives on residual returns. The absolute value of the annual residuals was regressed on the uncertainty of the CEO and Chairman, their sum and the changes over time while controlling for both year and industry fixed effects.

Below is the model employed for all regressions:

$$| Residual\ Return_{it} | = \alpha + Narrative_{it} + fixed\ effects + \varepsilon_{it}$$

It is important to note that the variable | *Residual Return* | remains constant for all regression models for firm i in year t. This regression differs only in terms of the independent variables, which vary depending on the uncertainty measure being tested.

The first step involved running individual regressions for each of the narrative measures: uncertainty, change in uncertainty and sum of executive uncertainty, for each of the sections: CEO section and the Chairman section. After all independent variables were tested, the end result was 5 separate regressions where: model [1] tested individually the uncertainty of each section for firm i in

year *t*; model [2] tested the executives' uncertainty for firm *i* in year *t*; model [3] tested individually the change in uncertainty of each variable for firm *i* in year *t*.

The next three additional regressions are as follows: model [4] tested the uncertainty and change in uncertainty of each section for firm *i* in year *t*; model [5] tested the uncertainty and change in uncertainty in the complete top executive comments for firm *i* in year *t*. These regressions aim to capture the uncertainty at a given time and its dynamic change, providing a more comprehensive analysis of the impact of the narrative on residual returns.

Finally, model [6] tested the uncertainty for both sections for firm *i* in year *t*; and model [7] tested the change in uncertainty for both sections for firm *i* in year *t*.

To facilitate, definitions of each variable can be found in table 6 in the Appendix.

5. Results

It is important to remember that the dependent variable in all these regressions, the absolute value of annual residuals, remained constant. As mentioned before, could be obtained the annual residual returns using the Fama-French 3-factor model. The data contained 4,948 observations with a mean value of 0,043 and a standard deviation of 0.04. Additionally, the median is 0,035 with a minimum value and maximum value of 0 and 0.37 respectively.

This study's approach to testing the hypothesis is to implement different regression models, changing the independent variables in each of them in order to gain a comprehensive understanding of the impact of these variables.

Below is a series of tables presenting each regression performed. Each table is grouped in a logical way to facilitate the testing the hypothesis.

Table 3: OLS coefficients estimates – individual regressions

	Residual_return					
	[1a]	[1b]	[2a]	[3a]	[3b]	[3c]
Const	0.0407***	0.044***	0.0613***	0.0429***	0.043***	0.042***
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
uncertainty_CEO	-0.0002					
	(0.7431)					
uncertainty_Chair		-0.0002				
		(0.5956)				
uncertainty_sum			-0.0047***			
			(0.0023)			
uncertainty_CEO_change				-0.0006		
				(0.3439)		
uncertainty_Chair_change					0	
					(0.8950)	
uncertainty_sum_change						-0.0006
						(0.4079)
industry fixed effects	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*
Year fixed effects	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*
Number of observations	2,539	4,098	1,842	1,651	2,746	1,005
R squared	0.0002	0.0003	0.0057	0.001	0	0.0007

P–values reported in parentheses are computed using the robust standard errors to obtain unbiased OLS coefficients. * < 0.1; ** < 0.05; *** < 0.01.

Table 4: OLS coefficients estimates – section measure

	Residual_return		
	[4a]	[4b]	[5a]
Const	0.0406***	0.0423***	0.0535***
	(0.0000)	(0.0000)	(0.0000)
uncertainty_CEO	-0.0002		
	(0.4761)		
uncertainty_Chair		0.0002	
		(0.5867)	
uncertainty_sum			-0.0026
			(0.2170)
uncertainty_CEO_change	-0.0006		
	(0.3060)		
uncertainty_Chair_change		-0.0001	
		(0.7838)	
uncertainty_sum_change			-0.0005
			(0.5010)
industry fixed effects	*Yes*	*Yes*	*Yes*
Year fixed effects	*Yes*	*Yes*	*Yes*
Number of observations	1,651	2,746	1,005
R squared	0.0012	0.0001	0.0023

P–values reported in parentheses are computed using the robust standard errors to obtain unbiased OLScoefficients. * < 0.1; ** < 0.05; *** < 0.01.

Table 5: OLS coefficients estimates – narrative measure

	Residual_return	
	[6a]	[7a]
Const	0.0430***	0.0431***
	(0.0000)	(0.0000)
uncertainty_CEO	-0.0005	
	(0.1405)	
uncertainty_Chair	-0.0002	
	(0.5092)	
uncertainty_CEO_change		0
		(0.9288)
uncertainty_Chair_change		-0.0007
		(0.1151)
industry fixed effects	*Yes*	*Yes*
Year fixed effects	*Yes*	*Yes*
Number of observations	1,651	2,746
R squared	0.0012	0.0001

P–values reported in parentheses are computed using the robust standard errors to obtain unbiased OLS coefficients. * < 0.1; **< 0.05; *** < 0.01.

To assess the impact the uncertainty of annual reports' top management sections has on annual residual returns, first we break down each of the regressions and models used.

Lets first consider table 3. Here, the outcome of regression model [1a] and model [1b], indicates that uncertainty for the CEO and Chairman sections cannot independently account for the annual residuals. This is not a strange result especially when considering other studies that analyze the impact of uncertainty also found no significant explanatory power (Loughran & Mcdonald, 2011). Looking at table 5 model [6a], we find similar results. Here when mixing the CEO and Chairman sections we find no statistical significance for these measures to account for annual residual returns. Therefore, it is visible that no section's uncertainty has the ability to explain these excess returns. Concluding that hypotheses H1a and H1b are rejected based on these results.

We also investigate the relation between the uncertainty of the CEO and Chairman sections as one and residual returns. The regression model [2a], visible in table 3 indicates a moderate impact of this narrative on residual returns. These findings suggest ambiguity since the significant impact does not repeat or show up again in other regressions and the coefficient derived from using *uncertainty_sum* is low. Therefore, this result does not provide enough robustness to explain in a significant way residual returns. Suggesting rejecting hypothesis H2a.

Another branch of this study is the impact of changes in uncertainty over time on residual returns. The individual regression, seen on model [3a], model [3b] and model [3c] in table 3 indicates that there is no correlation between the change in uncertainty across the CEO and Chairman sections of the report and residual returns, neither for the impact of the *uncertainty_sum* on residuals. This suggests that uncertainty as a measure of narrative does not provide significant value

when searching for explanations for unexpected returns. We can conclude that, in contrast to the findings of Cohen et al. (2020), changes over time are not a significant analysis for the uncertainty measure.

In table 5 model [7a], regressing the change in CEO and Chairman's uncertainty with residual returns, indicates no relation and no explanatory power from these variables. Thus rejecting hypotheses H3a, H3b and H3c.

Finally, when matching the *uncertainty* of a specific section with the change over time of that variable, we find no significant explanatory power. These results, visible in table 4 models [4a], model [4b] and model [5a], suggest that the uncertainty for a given section and the change in uncertainty at a given time neither individually nor collectively provide substantial robustness as a predictor for the financial performance of the firm.

All models observed have R-squared values consistently positive which indicates that the model applied was appropriate and the results are reliable. Additionally, the model with the strong relationship between the variables has a negative coefficient, which indicates a negative correlation between the variable and residual returns.

6. Conclusions

From the increasing interest in qualitative data's impact on the market in literature, more specifically about the impact of financial narratives on a firm's performance, this dissertation's objective was to investigate the extent to which uncertainty, as a narrative measure, has any explanatory power for residual returns in firm valuation models, facilitating the decision making by investors and by reducing information asymmetry.

This study hypothesized that the uncertainty on the CEO and Chairman review, their sum and the change over the years of these variables have an impact on residual returns and can impact investor decisions. The sample used consisted of 1344 non-financial companies from the London Stock Exchange.

Our study found no statistically significant variables across all models. The results do not support the notion that analysing the uncertainty of annual reports can help investors make better decisions. Although the uncertainty variable, which represents the sum of the uncertainty between the CEO and the Chairman, had a moderate impact on residual returns, by analysing the other regressions, the lack of statistically significant variables across this variable allows us to conclude that this finding is not significantly useful for investors to take advantage.

In summary, while this dissertation aimed to analyse the potential impact of uncertainty as a narrative measure on residual returns, the results suggest that, in the specific context of the analysis, uncertainty does not provide significant value to mitigate information asymmetry between companies and investors, contrary to what was hypothesised.

Further research should aim to investigate the impact of change and the impact of aggregating similar narratives from different sections of annual reports and check their explanatory power to shed light on the impact of variations in

variables. A limitation of this dissertation is the unavailability of acquiring the factors needed for the Fama-French 5-factor model for the UK market, which could have provided better insights. Researchers would also benefit from analysing a different geographical scope to capture different dynamics as well as a different timeframe.

References

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. http://ucrel.lancs.ac.uk/cfie/wmatrix-import-software.php.Thedatasetofannualreportscoresisavailablea

Antweiler, W., & Frank, M. Z. (2004). Is all that talk just noise? The information content of Internet stock message boards. Journal of Finance, 59(3), 1259–1294. https://doi.org/10.1111/j.1540-6261.2004.00662.x

Argamon, S., Koppel, M., Pennebaker, J. W., & Schler, J. (2009). Automatically profiling the author of an anonymous text. Communications of the ACM, 52(2), 119–123. https://doi.org/10.1145/1461928.1461959

Biddle, G. C., Hilary, G., & Verdi, R. S. (2009). How does financial reporting quality relate to investment efficiency? Journal of Accounting and Economics, 48(2–3), 112–131. https://doi.org/10.1016/J.JACCECO.2009.09.001

Bonsall, S. B., Leone, A. J., Miller, B. P., & Rennekamp, K. (2017). A plain English measure of financial reporting readability. Journal of Accounting and Economics, 63(2–3), 329–357. https://doi.org/10.1016/j.jacceco.2017.03.002

Brown, S. V., & Tucker, J. W. (2011). Large-Sample Evidence on Firms' Year-over-Year MD&A Modifications. Journal of Accounting Research, 49(2), 309–346. https://doi.org/10.1111/j.1475-679X.2010.00396.x

Clatworthy, M. A., & Jones, M. J. (2006). Differential patterns of textual characteristics and company performance in the chairman's statement. Accounting, Auditing and Accountability Journal, 19(4), 493–511. https://doi.org/10.1108/09513570610679100

Cohen, L., Malloy, C., & Nguyen, Q. (2020). Lazy Prices. Journal of Finance, 75(3), 1371–1415. https://doi.org/10.1111/jofi.12885

Demers, E., & Vega, C. (2008). Soft Information in Earnings Announcements: News or Noise? Soft Information in Earnings Announcements: News or Noise? *. http://www.sssrn.com.

Diamond, D. W., & Verrecchia, R. E. (1991). Disclosure, Liquidity, and the Cost of Capital. In Source: The Journal of Finance (Vol. 46, Issue 4).

Dougal, C., Engelberg, J., García, D., & Parsons, C. A. (2012). Journalists and the stock market. In Review of Financial Studies (Vol. 25, Issue 3, pp. 640–679). https://doi.org/10.1093/rfs/hhr133

Easley, D., & O'Hara, M. (2004). Information and the cost of capital. In Journal of Finance (Vol. 59, Issue 4, pp. 1553–1583). https://doi.org/10.1111/j.1540-6261.2004.00672.x

El-Haj, M., Alves, P., Rayson, P., Walker, M., & Young, S. (2020). Retrieving, classifying and analysing narrative commentary in unstructured (glossy) annual reports published as PDF files. Accounting and Business Research, 50(1), 6–34. https://doi.org/10.1080/00014788.2019.1609346

Engelberg, J., Andersen, T., Barberis, N., Gao, P., He, Z., Hertzberg, A., Jagannathan, R., Kondo, J., Korajczyk, R., Matsa, D., Parker, J., Williams,

J., & Walther, B. (2008). Costly Information Processing: Evidence from Earnings Announcements *. https://ssrn.com/abstract=1107998

Fama, E. F., & French, K. R. (1992). The Cross-Section of Expected Stock Returns. The Journal of Finance, 47(2), 427–465. https://doi.org/10.1111/j.1540-6261.1992.tb04398.x

García Osma, B., & Guillamón-Saorín, E. (2011). Corporate governance and impression management in annual results press releases. Accounting, Organizations and Society, 36(4–5), 187–208. https://doi.org/10.1016/J.AOS.2011.03.005

Healy, P. M., & Palepu, K. G. (2001). Information asymmetry, corporate disclosure, and the capital markets: A review of the empirical disclosure literature. Journal of Accounting and Economics, 31(1–3), 405–440. https://doi.org/10.1016/S0165-4101(01)00018-0

Henry, E. (2008). Are investors influenced by how earnings press releases are written? Journal of Business Communication, 45(4), 363–407. https://doi.org/10.1177/0021943608319388

Hirshleifer, D., & Teoh, S. H. (2003). Limited attention, information disclosure, and financial reporting. Journal of Accounting and Economics, 36(1-3 SPEC. ISS.), 337–386. https://doi.org/10.1016/j.jacceco.2003.10.002

Indjejikian, R. J. (1991). The Impact of Costly Information Interpretation on Firm Disclosure Decisions. In Source: Journal of Accounting Research (Vol. 29, Issue 2).

Kothari, S. P., Li, X., & Short, J. E. (2009). The effect of disclosures by management, analysts, and business press on cost of capital, return volatility, and analyst forecasts: A study using content analysis.

Accounting Review, 84(5), 1639–1670. https://doi.org/10.2308/accr.2009.84.5.1639

Lawrence, A. (2013). Individual investors and financial disclosure. Journal of Accounting and Economics, 56(1), 130–147. https://doi.org/10.1016/J.JACCECO.2013.05.001

Li, F. (2008a). Annual report readability, current earnings, and earnings persistence. Journal of Accounting and Economics, 45(2–3), 221–247. https://doi.org/10.1016/J.JACCECO.2008.02.003

Li, F. (2008b). Annual report readability, current earnings, and earnings persistence. Journal of Accounting and Economics, 45(2–3), 221–247. https://doi.org/10.1016/j.jacceco.2008.02.003

Loughran, T., & Mcdonald, B. (2011). When Is a Liability Not a Liability? Textual Analysis, Dictionaries, and 10-Ks. Journal of Finance, 66(1). 35–65. https://doi.org/10.1111/j.1540-6261.2010.01625.x

Loughran, T., & Mcdonald, B. (2014). Measuring readability in financial disclosures. Journal of Finance, 69(4), 1643–1671. https://doi.org/10.1111/jofi.12162

Loughran, T., & Mcdonald, B. (2016). Textual Analysis in Accounting and Finance: A Survey. Journal of Accounting Research, 54(4), 1187–1230. https://doi.org/10.1111/1475-679X.12123

Mckay Price, S., Doran, J. S., Peterson, D. R., & Bliss, B. A. (n.d.). Earnings conference calls and stock returns: The incremental informativeness of textual tone.

Merkl-Davies, D. M., & Brennan, N. M. (2007). Discretionary Disclosure Strategies in Corporate Narratives: Incremental Information or Impression Management? https://ssrn.com/abstract=1089447

Sadique, S., In, F., & Veeraraghavan, M. (n.d.). The Impact of Spin and Tone on Stock Returns and Volatility: Evidence from Firm-issued Earnings Announcements and the Related Press Coverage https://ssrn.com/abstract=1121231

Smith, M., & Taffler, R. J. (2000). The chairman's statement A content analysis of discretionary narrative disclosures. In Accounting Auditing & Aaccountability Journal (Vol. 13, Issue 5). MCB University Press.

Tetlock, P. C. (2007). Giving content to investor sentiment: The role of media in the stock market. Journal of Finance, 62(3), 1139–1168. https://doi.org/10.1111/j.1540-6261.2007.01232.x

Tetlock, P. C., Saar-Tsechansky, M., & Macskassy, S. (2008). More Than Words: Quantifying Language to Measure Firms' Fundamentals. Journal of Finance, 63(3), 1437–1467. https://doi.org/10.1111/j.1540-6261.2008.01362.x

Appendix

Table 6: variables definition

Variable name	Definition	Source
Dependent variable		
Residual_return	Absolute value of the residual annual return for each company, derived from the Fama-French 3-Factor model	Thomson Reuters Datastream | Database Finance and Accounting Department - University of Exeter
Independent variables		
uncertainty_CEO	CEO review - Number of uncertainty words scaled by the total number of words of the section	El-Haj et al. (2020)
uncertainty_Chair	Chair's letter - Number of uncertainty words scaled by the total number of words of the section	El-Haj et al. (2020)
uncertainty_sum	CEO and Chairman section - Number of uncertainty words scaled by the total number of words of the sections	Derived from El-Haj et al. (2020)
uncertainty_CEO_change	CEO review - The yearly change in the number of uncertainty words scaled by the total number of words of the section	Derived from El-Haj et al. (2020)
uncertainty_Chair_change	Chair's letter - The yearly change in the number of uncertainty words scaled by the total number of words of the section	Derived from El-Haj et al. (2020)
uncertainty_sum_change	CEO and Chairman section - The yearly change in the number of uncertainty words scaled by the total number of words of the section	Derived from El-Haj et al. (2020)