UNIVERSIDADE CATÓLICA PORTUGUESA

The impact of textual features on cost of equity capital

The case of UK Annual Reports

Renato Filipe Peixoto Carneiro

Católica Porto Business School

2021

UNIVERSIDADE CATÓLICA PORTUGUESA

The impact of textual features on cost of equity capital

The case of UK Annual Reports

Masters' Dissertation
presented to Universidade Católica Portuguesa
in order to obtain a master's degree in Finance

by

Renato Filipe Peixoto Carneiro

Under supervision of Prof. Paulo Alves

Católica Porto Business School

2021

Acknowledgements

To Prof. Paulo Alves for his guidance in this research. I express my gratitude for his patience, advice, and readiness throughout this process.

To my family, for always believing in me and giving me the opportunity to enroll in such an amazing journey.

To my friends, for the motivation and support.

To Francisca Gomes, for the encouragement and enthusiasm.

Abstract

This thesis aims to analyze the impact of textual features of UK Annual Reports sections on the cost of equity. We analyze the impact of the Chairman Statement, Financial Review and Risk Section linguistic features separately, for the UK firms from 2012 to 2017. We estimate the cost of equity through the Price to Earnings Growth Model and the linguistic features are divided between readability and sentiment (tone). We find that disclosure content has an impact on cost of equity, while readability has is associated with no impact on cost of equity. We also find that positive tone is associated with lower cost of equity in all sections of the Annual Report.

Keywords: Cost of equity, Readability, Textual Tone, Disclosure, Annual Report.
JEL-Codes: G14, G32, M41

Resumo

O objetivo desta tese é o de analisar o impacto das características textuais de certas secções dos Relatórios Anuais do Reino Unido sobre o custo de capital próprio. Analisamos os impactos das características linguísticas da Declaração do Presidente, da Análise Financeira e da Seção de Risco separadamente, para empresas do Reino Unido de 2012 a 2017. Estimamos o custo de capital próprio por meio do "PEG model" e as características linguísticas são divididas entre legibilidade e conteúdo (tom). Descobrimos que o conteúdo de divulgação tem impacto no custo de capital, enquanto a legibilidade não está associada a nenhum impacto no custo de capital. Também descobrimos que o tom positivo está associado a um custo de capital próprio inferior em todas as seções do Relatório Anual.

Palavras-Chave: Custo de Capital Próprio, Legibilidade, Tom, Divulgação, Relatório Anual.

JEL-Codes: G14, G32, M41

Index

Index of tables

1. Introduction

Information is a pivotal factor in financial markets. In fact, an efficient allocation of resources in a capital market economy depends on the effectiveness of information and incentives. Credible disclosures are keen on mitigating potential problems of information (Healy & Palepu 2001).

The question on whether cost of capital decreases with enhanced disclosure remains a question (Botosan 2006). Firstly, studies relating cost of equity to disclosure used measures of disclosure based on quantity (Botosan 1997, as an example), excluding lexical measures. Despite the theoretical standpoint, the empirical results were not consensual. Moreover, studies such as Gietzmann & Ireland (2005) embraced the usage of measures of disclosure quality, instead of quantity, and further these measures turned to the lexical dimension (Li 2008). Moreover, Kothari et al (2009) state that mixed evidence in the prior literature is due to the assumption that disclosure quality is unidirectional and that the content of disclosures influences disclosure quality. Also, Johnstone (2016) state that the quality of the disclosure creates more certainty about the estimates of investors, however these estimates could lead to rationally higher costs of equity.

The usage of quantitative accounting information has been traditionally used and studied in accounting and finance. However, the vast majority of research focuses on quantitative market and accounting variables widely available in commercial databases. Only recently, research started to focus on the disclosures, which are a fundamental source of information for investors, as financial statements provide limited insights about a company. Additionally, the number of unstructured data present on corporate reports has increased, magnifying the need to study these types of data. In fact, Lewis & Young (2019) state that the median number of words on UK annual report narratives increased 122% (14.954

to 33.193 words) from 2003 to 2016. This increase justifies the growing body of research focusing on textual analysis.

Textual analysis can be divided in different categories, we focus on the readability and sentiment analysis. Readability differentiates from other textual analysis methods, as it attempts to measure the ability of the reader to decipher the message of the text. Sentiment analysis tries to quantify and extract the meaning from text (Loughran & McDonald 2016).

Given the mixed evidence on the disclosure and cost of equity relationship, we decided study this relationship using measures of readability as a disclosure measure. We also use measures of textual tone, given the literature mixed evidence often attributed to disclosure measure on quality and quantity. Therefore, we study both types of measures since both effects can be coexistent.

This dissertation assesses if different textual analysis properties influence cost of equity. Particularly, we examine if readability and sentiment impact cost of equity capital. We also test this hypothesis for non-profitable firms.

For this purpose, we analyze the Chairman Statement, Financial Review (often associated to the CFO) and Risk section of UK annual reports from 2012 to 2017. We calculate readability using Fog index and the natural logarithm of total words, as used by Li (2008). We measure two different types of sentiment, net positive and uncertainty tone, based on Loughran & McDonald (2011) word lists.

The remainder of this work is organized as follows: Section 2 provides the theoretical framework surrounding the relation between disclosure and cost of equity, as well as sentiment analysis. Section 3 provides some insights about the linguistic features and objectives of the different UK annual reports sections. Section 4 refers to the research hypothesis. Section 5 describes the model development, the estimation of cost of equity, the estimation of readability and

tone measures and the descriptive statistics. Section 6 reports and discuss the results. Section 7 presents the conclusion.

2. Literature Review

2.1 The relation between cost of equity and disclosure

The relationship between cost of equity capital and disclosures has been subject of discussion among theorists and empiricists. However, no consensus has been found about this relation, or its causes. Literature relating disclosure and cost of equity often state a negative relationship between both variables. For instance, Botosan (1997) finds a negative relation between disclosure and cost of equity, particularly for firms with low analyst follow. The causes for the negative relation are often explained by two different streams of research (Botosan 2006). One stream of the literature states that disclosure reduces estimation risk of investors, related to the uncertainty of the parameters on a security's returns (Lambert et al 2007) (Brown 1979) (Easley & O'hara 2004) (Clarkson et al 1996) (Klein & Bawa 1976). The other stream is related to disclosure reducing information asymmetry, consequently reducing transaction costs and cost of equity capital (Diamond & Verrecchia 1991) (Welker 1995) (Amihud & Mendelson 1986).

The literature related to estimation risk, theorizes that investors use the information available to estimate parameters. Therefore, their confidence on the estimations, depends on the attributes of the information set. The main questions present in this segment of literature is if this risk is diversifiable or not by investors and if this risk is priced. Botosan et al (2004) state that this risk is not captured by traditional market betas, such as Capital Asset Pricing Model (CAPM), but is non-diversifiable and consequently, should be priced. Also, Easley & O'hara (2004) create an "expectations equilibrium model" that relates informational structure (public and private information) and cost of capital, concluding that private information is a form of systematic risk and as a result, not diversifiable. The study also concludes that investors require compensation

for bearing this risk, hence this risk is priced. However, questions remain about the diversifiability of estimation risk (Clarkson et al 1996). For instance, studies such as Hughes et al (2007), argue that the finite number of assets assumed in the "expectations equilibrium model" model makes the price effect non-diversifiable. Lambert et al (2007) find that disclosure quality affects the assessment of the market cash flows by investors, and this effect is not diversifiable. This verifies the theory of Easley & O'hara (2004), stating that information risk is priced, and uninformed traders require a premium to trade against informed traders.

From the perspective of information asymmetry and transaction costs, the general theory is that the lower is the information about a firm, the higher the transaction costs tend to be. By applying Akerlof (1978) "lemons problem theory", Welker (1995) states that the fact that uninformed investors react less perfectly to new information and thus protecting the price of a security leads to decreases in liquidity. Diamond & Verrecchia (1991) state that, by revealing more information and thus reducing information asymmetry, firms experience a reduction in the cost of capital, due to heightened levels of demand for its securities, resulting from increases in liquidity. Other stream of the empirical literature use measures of private information (PIN) to study the relationship between information asymmetry and disclosure quality. Brown & Hillegeist (2007), for example, use PIN measures and find that disclosure quality is negatively associated with information asymmetry, mainly caused by a decrease in the frequency of private information events. However, these measures of PIN find no consensus in the literature (Duarte & Young (2009), as an example).

Botosan & Plumlee (2002) studies the association between cost of capital and annual reports disclosures, timely disclosures, and investor relations activities. The authors find a negative relation with annual report disclosures (as expected), but a positive relationship with timely disclosures. The study explains the

unexpected positive impact of timely disclosures alleging that these events increase stock price volatility leading to higher cost of equity. Gietzmann & Ireland (2005) argue that the positive, and contrary to theory, relation between timely disclosures and cost of equity capital found in Botosan & Plumlee (2002) is due to problems with the measurement of disclosure. Hence, Gietzmann & Ireland (2005) construct a measure of disclosure based on quality, instead of quantity, on strategic disclosures. The authors argue that the quality and quantity of disclosure are different dimensions of disclosure. Furthermore, the study provides an example that a firm could disclose large sums of information and most of it be uninformative, magnifying the need to study the quality of disclosures. Botosan (1997) acknowledge the importance of studying disclosure quality, however, the author state that it is difficult to assess and therefore assume that quantity and quality are positively related.

The development of computer power and linguistic processes allowed to further study disclosure impacts. Content analysis is a process that allows to objectively characterize the message transmitted by descriptive information. There are some examples of studies that use content analysis to study disclosure relation to the cost of capital, however, these techniques are more commonly used in the non-financial disclosure literature. This type of literature uses content analysis to measure different types of disclosure, such as corporate socially responsible, environmental, and intellectual capital disclosures. For example, Michaels & Grüning (2017) use content analysis to measure corporate social responsibility disclosures on 264 German companies and relate it to the cost of capital. Boujelbene & Affes (2013) in a study that relates intellectual capital disclosures and cost of equity on French firms, uses content analysis to measure intellectual capital disclosure. Also, Mangena et al (2010) use this analysis to measure voluntary and intellectual capital disclosures on a study of 126 UK firms.

2.2 Readability

Linked with disclosure quality, language characteristics may influence the cost of equity capital. Readability, as defined by Loughran & McDonald (2014), is "the ability of investors and analysts to integrate valuation relevant information from the business documents in stock prices". There are two possible ways for language to influence the cost of equity, as mentioned by Rjiba et al (2021). Firstly, the usage of more complex narratives enhances information asymmetry problems. Lawrence (2013) state that higher quality disclosures reduce less sophisticated investor's information disadvantage. Also, Miller (2010) find evidence that more complex fillings have a limited effect on large investors but generate lower consensus and trading activity among small traders. Usage of complex narratives can be used by managers to opportunistically hide bad performance or news from investors (Li 2008). In fact, Lo et al (2017) find that firms that most likely used earnings management techniques to meet or beat earnings, provide more complex disclosures. Also, Chakrabarty et al (2018) state that managers with higher risk incentives (measured by option's vega) on stock compensation, strategically issue less readable disclosures.

Secondly, more complex annual reports increase information risk, since it enhances uncertainty about the future of the company and the parameters of the stock return's distribution (Rjiba et al 2021). It also creates disagreement among investors on how to interpret public signals. Bloomfield & Fischer (2011) analyze the impact of disagreement among investors and its effect on the cost of capital. In fact, they find that opinion divergence decreases the precision of investors beliefs and as a result, increases the cost of capital. Furthermore, Lehavy et al (2011) find that firms with less readable 10-K fillings are associated with greater overall uncertainty in analyst earnings forecasts and increasing demand for analyst services.

Bloomfield (2002) refutes the efficient market hypothesis (EMH) by creating the theory of incomplete revelation hypothesis (IRH). This theory states that the cost incorporated in extracting useful statistics from public data keep the markets from fully revealing the true meaning of those statistics. Li (2008) hypothesis of opportunistic management behavior is inspired by the IRH. The author states that the less complete reaction to information, give managers the incentive to obfuscate information. In fact, the study finds that the length and readability of annual reports are associated with current and future earnings. However, Bloomfield (2008) provides potential explanations for the relationships observed by Li (2008). The author theorizes that, one possible explanation for the negative relation between readability and profitability, is that managers try to hide bad news from investors by issuing more complex narratives (obfuscation). However, the author also provides other explanations that contrast with the management obfuscation explanation. The "ontology" explanation assumes that losses and transitory income are more difficult to describe, therefore disclosures are naturally more complex. The "management by exception" state that managers provide more complete explanations for poor performance, leading to greater length and fogginess. Concluding, in theory, readability can be a consequence of the circumstances of the firm, or of managerial intervention.

2.3 Cost of equity and disclosure content

Kothari et al (2009) state that the mixed evidence in the disclosure and cost of capital, or other capital market variables, relationship is due to the assumption that it is unidirectional. One possible explanation, developed by the same study, is that measures of disclosure quality are likely influenced by disclosure content. Moreover, Li (2006) suggests that the information conveyed in the form of text can provide a more independent test of market efficiency. Also, Johnstone (2016)

states that the common idea that good financial reporting reduces cost of equity, by reducing uncertainty is misleading. The same study argues that greater certainty about the information on asset pricing models can create estimates that lead to a higher cost of equity. In other words, more information leads to more accurate forecasts, however, these forecasts will sometimes bring rationally higher cost of capital. Therefore, is not only necessary to study the precision of what is being said, but also the content of the narrative.

The process of content analysis enables to analyze market reactions to qualitative information by quantifying document tone (Jegadeesh & Wu 2013). The most important work in sentiment analysis is based on sources such as corporate disclosures, news articles or analyst reports and even internet-posting (Kearney & Liu 2014).

2.3.1 Positive tone

The first line of studies on textual tone argued that tone is informative to investors and allows for investors to infer managers private information. These studies suggest that managers can use tone to disclose important information that is not present in the quantitative form, reducing information asymmetry (Demers & Vega 2014) (Feldman 2008) (Li 2010) (Schleicher & Walker 2010). Luo & Zhou (2020) divide the determinants of financial performance and tone relation in two segments.

The first class of studies states that textual tone is directly related to the firm's current and future performance. In fact, some studies state that a more positive tone on disclosures is associated with higher financial performance. These studies often analyze the Manager Discussion and Analysis (MD&A) section on Annual reports. Li (2010) finds on the MD&A section of the 10-K and 10-Q

statements that a more positive tone is related to the current performance of the firm. Davis & Tama-Sweet (2012) find, in the same section of 10-K and 10-Q filings, that tone is positively associated with future performance. Wu et al (2020) also find that the tone of this section on Chinese annual reports is positively related to firm value.

The second segment of studies states that tone is associated with firm's fundamentals, such as size, market to book value, age, operational complexity, etc. In fact, these variables are often used as control variables for textual tone. Huang (2014) document that business complexity (number of business and geographic segments) has a negative effect on the tone of earnings announcements. Li (2010) also finds that older firms , firms with lower market-to-book values and smaller size present a more positive tone. Bakarich (2019) show that the corporate life cycle of a company affects the textual characteristics of 10-K's. More specifically, the narratives of a company become more optimistic, as firms progress from the introduction to the maturity stage, but less optimistic as firms progress to the declining phase.

The above results assume that managers interpret the tone of disclosures as an unbiased signal of managers private information. However, other contrary lines of theory state that tone can be used by managers to deceive investors (Huang 2014) (Arslan-Ayaydin et al 2016) (Price 2012). Although a positive tone is a useful tool to understand present and future performance, or to convey information beyond quantitative information, it also can be used to hide bad news and expectations by managers, to influence the investor (Huang 2014). Thus, tone has the power to both, increase and reduce information asymmetry, depending on the reliability of the statements. Davis & Tama-Sweet (2012) state that, even if market participants process language efficiently, it is possible that managers still strategically disclose language to take advantage of inefficiencies at the individual firm level. Zhou et al (2018) studies the relationship between

stock crash risk and annual report's tone on Chinese firms. The study finds that when the tone of the document is overly optimistic, the risk of a stock crash risk in the next 365 days significantly increases. Therefore, positive tone does not necessarily reflect good expectations of the firm's future, but rather the release of false good news. Furthermore, Huang (2014) finds, in a study of US annual earnings press releases, that tone is negatively related with future annual earnings, concluding that managers use tone to deceive investors on future performance. Also, Kashanipoor (2020) find that the usage of positive general words is correlated with a decrease in the future return on assets. In a study of the UK Interim Management Statements (IMS) of the first and third quarters, Rahman (2019) states that the IMS discretionary tone of the third quarter is positively aligned with actual earnings, while IMS discretionary tone of first quarter is not. This could mean that, since investors and managers have less information at the start of the year, managers are more likely to mislead investors on expected annual results. Other relevant streams of the tone management literature focus on the association between tone manipulation and manager's stock-based compensation (Arslan-Ayaydin 2016) (Huang 2014). For example, Arslan-Ayaydin (2016) finds that managers use more optimistic tone when their stock-based compensation portfolio is more sensitive to stock price volatility. Also, corporate governance characteristics, such as CEO power, are important drivers of tone management, associated with more optimistic tone on earnings announcements (DeBoskey 2019).

Tone can be viewed as a mean to reduce information asymmetry, as it allows for managers to disclose their insights on the firm's activities. However, this assumption is not a panacea, as tone management literature states that managers can use these means of communication to deceive investors about the firm's performance in an easier way than they do with accounting numbers, due to the less regulation and increased elasticity of words.

2.3.2 Negative and uncertainty tone

Although positive words have been identified with statistically significant effects in the literature above, Loughran & McDonald (2011) find little incremental information in positive word lists. Much of the literature uses the simple positive/negative distinction on information. Some papers also use the net measure of both types of words. However, both types of words have different characteristics. Negative words are unambiguous, while positive words are more commonly used on both, positive and negative statements (Loughran & McDonald 2016). Rjiba et al (2021) use negative tone as a proxy for bad news and finds that the adverse effect of lower readability on the cost of equity capital is magnified by the percentage of negative words. Also, Schleicher & Walker (2010) find that the main vehicle managers use to bias tone is the change of negative statements, not the number of positive statements.

Furthermore, similarly to negative tone, uncertainty tone might change the relation between disclosure and cost of capital. Li (2006) finds that increases in the usage of textual risk words lead to lower future earnings. Prior literature also states that uncertainty tone is negatively associated with stock returns (Loughran & McDonald 2011) and positively related with stock return volatility (Demers & Vega 2014). Barakat et al (2019) states that the findings of the prior literature on uncertainty tone indicate that uncertainty (ambiguous) words reveal distrust in a firm's fundamentals which leads to higher discount rates and volatility. Rjiba et al (2021), in a study of the US annual reports, finds that the effect of textual complexity on the cost of equity is more pronounced when the annual report is ambiguous (uncertainty tone). The literature suggests that the usage of uncertainty tone makes investors' confidence in their predictions of expected performance and payoffs diminish, therefore, increasing information risk. However, Barakat et al (2019) poses a different hypothesis for the consequences of uncertainty tone. The study states that, in an environment of high uncertainty,

the ambiguous tone could be conditionally optimistic and that risks might not be as bad as suggested. They applied this argument for the nature of operational risk announcements, where investors could interpret uncertainty in bad news as good news. We apply these claims to the case of risk statements, where tone must be more uncertain and negative.

3. UK annual reports context

In the last two decades, some section areas of annual reports experienced some developments. Management performance commentary, forward-looking commentary and business strategy were some of the areas where enhanced narrative disclosure was recommended. UK firms Annual Reports are considered to give significant discretion for managers to tell the firm's processes of value creation through their own words (Athanasakou et al 2020). The fact that UK only requires corporate reporting on core principals, allows preparers to have almost unlimited discretion when including non-mandatory content, as mentioned by Lewis & Young (2019). This is a major difference in corporate reporting from the US 10-K fillings, which are shaped by a more standardized nature. Therefore, and given the fact that the UK is a major stock market, we consider it to have a favorable study environment to analyze the linguistic features of annual reports.

We further developed our analysis in different sections of the annual report, focusing on three different narrative sections, the Chairman Statement, Financial Review and Risk Report. We chose these sections, because of their different purposes and language features, but also by their data availability.

The Chairman Statement is a voluntary disclosure section of the annual report. It is used to provide insights about the company's activities and performance, being one of the most widely read features of the annual report, as noted by Clatworthy & Jones (2006).

The financial review is part of the management's commentary section. The financial review objective is to review the financial aspects of the company and is signed by the CFO of the company (El-Haj et al 2020).

The risk report is a narrative, mandatory report that contains the description of the principal risk and uncertainties facing the company. These types of disclosures require that directors assess the prospects of the company and state the viability of these prospects (El-Haj et al 2020).

Since the objectives of each section are different, it is expected that the linguistic features vary between them. The Chairman Statement, as voluntary disclosure, is built to provide an accessible overview of the firm's performance. Therefore, it is expected that the average chair letter presents higher levels of readability, compared with other, more compliant, statements. El-Haj et al (2020) state that, by using the Fog index as a measure of readability, the Chair Letter require 1.9 years less education than the Governance Statement, which is shaped by compliance considerations. Also, since this section is performance-focused, it is associated with a positive tone. The same study finds that the average Chair letter tone is five times more positive than the corresponding governance statement.

The management's commentary section compiles a variety of different sections, including the financial review. It typically features sections like the CEO review, the review of operations, the business, and the financial reviews. This section is expected to be associated with a more positive tone than compliance sections, just like the Chairman Statement. El-Haj et al (2020) finds that, while the average chair letter is five times more positive than the corresponding governance statement, the average management's commentary section is over four times more positive than the governance statement.

The risk report section allows investors to assess the firm's risks and traduce this information to the stock prices Campbell et al (2014). We expect that the average tone of this section is less positive than the tone of the sections above described, or even negative since it is not a performance-based section and in fact, assesses the viability of the performance prospects. However, we expect higher levels of uncertainty tone in this section.

4. Research Hypothesis

The relation between disclosure and cost of equity is still subject of discussion among empiricists. However, many of the theorists' state that this relation is negative. Information risk is attributed as one of the causes for disclosure to reduce the cost of equity. Another attributed cause is information asymmetry. However, measures of disclosure and cost of equity are not a panacea, and empiricists find different preferences on how to calculate these variables.

In both theories that try to explain the negative relation between the cost of equity and disclosure, empiricists and theorists fail to find consensus. In the estimation risk literature, theories on the non-diversifiability of information risk fail to find consensus (Hughes et al 2007). On the information asymmetry literature, PIN measures also fail to find consensus among empiricists (Duarte & Young 2009). We address this problem by attempting to relate cost of equity to readability measures. Therefore, we analyze cost of equity reaction to readability measures in different sections of an UK annual report. We analyze different sections of the annual report, given the distinguished linguistic features found in each of these sections.

H1: What is the relation between the cost of equity and readability measures?

H1-A: What is the relation between the cost of equity and readability in the chairman statement section?

H1-B: What is the relation between the cost of equity and readability on the financial review section?

H1-C: What is the relation between the cost of equity and readability on the risk section?

Some studies state that the mixed evidence found in the disclosure relation with cost of equity capital literature is due to problems with the measure of disclosure quality (Kothari et al 2009). Other studies disregard theories such as information risk and state that disclosure quality allows to make more accurate forecasts, however, these forecasts could lead to higher or lower cost of equity (Johnstone 2016). Therefore, these studies attribute the content of the disclosure as the main driver of the cost of equity.

Studies like Demers & Vega (2014) and Li (2010) suggest that tone allows for investors to infer managers' private information. It also lets managers disclose information that is not present in the quantitative form. These features allow to reduce information asymmetry. However, these arguments assume that managers do not bias tone in their favor, manipulating the investor's perception of the company's activities (Huang 2014). Therefore, tone could reduce or increase information asymmetry, as well as returns, depending on the trustworthiness of the narrative.

Prior literature states that uncertainty tone creates distrust in the investors about a firm's earnings and cash flow distribution. This leads to higher discount rates for investors (Barakat et al 2019). In fact, one of the causes under study for the negative relation between disclosure and cost of equity is the pricing of information risk.

In our continuation of the study of the tone relation to the cost of equity, we hypothesize the relation between cost of equity and tone (net and uncertainty).

H2: What is the relation between the cost of equity and content?

H2-A: What is the relation between the cost of equity and content on the Chairman Statement Section?

H2-B: What is the relation between the cost of equity and content on the Financial Review Section?

H2-C: What is the relation between the cost of equity and content on the Risk Section?

5. Methodology

5.1. Dependent Variable - Cost of Equity Capital measure

Cost of capital is an ex-ante variable, which makes it not directly observable since it is based on market expectations. Botosan (2006) derive two classes of methods that allow for the estimation of the variable. One class of methods is based on predetermined priced risk factors to derive explicit measurements, for example, the Capital Asset Pricing Model (CAPM) and the Fama & French (1993) three-factor model. However, the author states that this set of measures limit the empirical relation between the cost of equity and disclosure to the set of predetermined factors used in the estimation. The second class of methods approaches the estimation by estimating the internal rate of return that equates the market's expectation of future cash flows to the current stock price. This class of methods are not a function of predetermined price risk factors and some of the models produce estimates that capture risk variations. While these models pose some advantages, there is no optimal way of estimating cost of equity capital in a disclosure environment. However, Cooper (2006) states that the absolute values of cost of capital should not be the main concern, but the variability of this estimate across firms. Therefore, we use the Price to Earnings Growth Model (PEG model) from Easton (2004) to estimate cost of equity. The author of the model states that the model is recommended to study the relation between disclosures and cost of equity capital. This model is associated with the second class of methods above mentioned. Following Mangena et al (2016), the PEG model requires less and simpler data to estimate, since it only requires share prices and analyst forecasts, which makes it a viable model to employ.

$$coeq = \sqrt{\frac{ep\ s2 - ep\ s1}{P_0}}$$

The estimation of cost of equity $coeq$, depends on variable $ep\ s2$ related to the 2 year ahead earnings forecast, $ep\ s1$ related to the 1 year ahead analyst forecast and P_0 which is the current share price. Both, $ep\ s2$ and ep s1 must be positive and $ep\ s2$ must be greater than $ep\ s1$ (Botosan & Plumlee (2005)). To compute these estimates, we extract from I/B/E/S Database the analysts forecast median for the 2 and 1 year ahead forecasts. We also use the stock price at the year-end of each firm, extracted from DATASTREAM.

5.2. Explanatory Variables

5.2.1.Textual tone measures

In our textual analysis, we used a "bag-of-words" technique to calculate net and uncertainty tone. We used Loughran & McDonald (2011) word list to determine the number of positive, negative and uncertainty words in each analyzed section. This word list's main advantage is the fact that it is based on the most likely interpretation of a word in a business context Loughran & McDonald (2016). The dataset used to obtain the number of words of each section in the UK annual reports is available at https://doi.org/10.17635/lancaster/researchdata/296, constructed using the procedures on El-Haj et al (2020).

The approach taken to quantify net and uncertainty tone is associated with the one used by Barakat et al (2019). We measure net tone as the spread between positive and negative words, relative to the total number of sentiment words. We also measure uncertainty tone as the number of uncertainty words, relative to the

total number of sentiment words. Sentiment words are the sum of positive, negative and uncertain words.

$$NetTone_{j,t} = 100 \cdot \frac{Positive\ words - Negative\ words}{Total\ sentiment\ words}$$

$$UncTone_{j,t} = 100 \cdot \frac{Uncertainty\ words}{Total\ sentiment\ words}$$

We compute these two measures of tone for the Chairman statement, Financial Review and Risk section of annual reports.

Table 1 *– Descriptive Statistics: Tone analysis*

	Variable	N	Mean	Std. Dev	Min	Perc.5	Median	perc 95	Max
Chairman	Net	4,887	43.614	30.183	-91.667	-11.87	48.333	85.294	97.5
Statement	Uncertainty	4,887	14.533	8.9026	1.3889	3.8462	12.821	31.67	66.667
Financial	Net	3,232	4.3834	32.451	-91.667	-49.326	4.4127	59.482	96.875
Review	Uncertainty	3,232	23.631	13.724	1.2987	6.0606	20.856	50	91.667
Risk	Net	3,430	-10.88	17.273	-75	-35.914	-11.94	17.857	81.481
Section	Uncertainty	3,430	51.232	13.874	3.7037	30.435	50	77.311	94.737

Table 1, shows the descriptive statistics for disclosure content proxies. We document that net tone mean and median are positive for the chairman statement (43.614 and 48.333, respectively) and for the financial review (4.3834 and 4.4127), however negative for the risk statement (-10.88 and -11.94, respectively). These results are consistent with the ones found in El-Haj et al (2020), where it is found that tone is more positive for sections that are performance related, and less positive for sections that are shaped by compliance considerations. The table also shows that the mean uncertainty tone is more pronounced for the risk statement section (51.232), than for the Financial Review and Chairman Statement (23.631 and 23.631, respectively). These results were expected since the risk section

objective is to disclose the main risks and uncertainties that the firm faces in its activities.

5.2.2.Readability measure

To measure readability, we estimate two different measures as used by Li (2008). To estimate this factor, we use Gunning's fog-index and the logarithm of the total number of words.

The fog-index is a well-known, simple formula that measures readability. This index measures readability as a function of words per sentence and syllables per word. Complex words are defined as words with more than two syllables.

$$FOG\ Index_{j,t} = 0.4(average\ number\ of\ words\ per\ sentence + percent\ of\ complex\ words)$$

Fog index values are categorized as follows: Fog ≥ 18 – unreadable; 14-18 – difficult; 12-14 – ideal; 10-12 – acceptable; 8-10 – childish.

There are not perfect readability measures for business and financial texts. The usage of Fog index as a readability measure for financial disclosure has been subject to critics. Loughran & McDonald (2014) state that the problem with the Fog index is due to the factor of complex words. The study states that this component is misleading and that many of the complex words considered in the Fog index are not complex for the average investor and are of common usage on corporate statements. They also state that the element of words per sentence is correlated with other measures of readability. Loughran & McDonald (2014) propose as a substitute for fog index, the usage of the size of the filling (in megabytes). However, Bonsall IV et al (2017) state that this is a noisy proxy for readability since it considers elements unrelated to text, such as images.

For this reason and given the absence of better available proxies of readability, we employ the Fog index.

We also use the logarithm of total words. The central idea of using this variable to measure readability is that longer documents are more deterring and require a higher cost of information processing.

$$Log(Total\ words)_{j,t}$$

Table 2 *- Descriptive statistics: Readability analysis*

	Variable	N	Mean	Std.Dev	Min	Perc.5	Median	perc.95	Max
Chairman	Fog	4,740	20.082	1.8058	16.45	17.233	19.963	23.253	24.562
Statement	Total Words	4,740	1102.7	860.76	53	384	922	2298.9	26805
Financial	Fog	3,079	20.55	1.7772	17.184	17.833	20.376	23.793	25.357
Review	Total Words	3,079	2306.5	1881.1	25	493	1846	5589	20961
Risk	Fog	3,102	20.947	2.2775	17.456	18.026	20.595	25.555	31.056
Statement	Total Words	3,102	1753.7	1530.7	86	402	1338.5	4600.2	21882

Table 2 shows the descriptive statistics for readability measures. For the Fog index category, we registered a mean of 20.082; 20.55; 20.945 and a median of 19.963; 20,376 and 20.595 for the Chairman Statement, Financial Review and Risk statement. respectively. The mean values observed for Fog index by Li, F. (2008) on the whole annual report and MD&A statement (19.39 and 18.23, respectively) are smaller than the values observed on our sample. Our mean and median are also higher than the values found in Lo, K. (2017) for the MD&A section (18.020 and 17.907, respectively). However, the value of Fog for the Chairman Statement is similar to the one identified by El-Haj, M. et al (2020) on the Chair letter (19.7). The mean values are interpreted as unreadable by the standard interpretation of the index, which is common for the annual report's Fog index (Loughran & McDonald (2014)). For the Total Words we observe mean values of 1102.7; 2306.5

and 1753.7 for the Chairman Statement, Financial Review and Risk Section, respectively. We use Li (2008) MD&A section as a benchmark, for section total words, and note that the mean of words in that section is significantly higher than our section values (4665 words). However, as mentioned in section 3, the MD&A section corresponds to a group of statements from different managerial groups, so it is expected that this section presents higher values of total words.

5.3. Control Variables

We use as control variables, measures that can potentially be determinants of the cost of equity. **ANLST** indicates the *number of analysts forecasts* for the firm in the previous year. This variable constitutes a proxy for the information environment of the firm. Botosan (1997) states that the relation between disclosure and cost of equity is negative, particularly for firms with low analyst following. Therefore, we expect that firms with more analysts following have a greater informational environment and present lower values of the cost of equity. **Log(TA)** displays the *natural logarithm of total assets* (£), as a proxy for firm size. We expect to find a negative relation between size and cost of equity capital since bigger firms are associated with greater levels of information and monitoring by analysts (Botosan (1997)). *Market to Book value* (**MTBV**), deducted as the market value of a firm divided by its book value. We expect to find a negative relation between MTBV and cost of equity as prior literature suggests that firms with lower growth opportunities face a higher cost of equity (Easton (2004)) (Cheng et al 2006). **Beta**, as prior literature, indicates a positive relationship between market beta and cost of equity (Easton2004) (Cheng et al 2006). *Return on Assets* (**ROA**), as a measure of the profitability of the firm. We expect a negative relation between ROA and cost of equity, since profitability enhances returns, leading to higher payoffs for investors. Dummy for negative ROA (**NegROA**) firms. *Debt to common equity* as

a measure of *leverage* (**LVRG**). We expect leverage to be an indicator of (credit) risk, therefore being associated with higher levels of cost of equity (Cheng et al 2006). *Earnings Per share* (**EPS**) since our measure of the cost of equity is based on the analyst's estimates of earnings per share. **Volatility** as Botosan & Plumlee (2002) state that stock price volatility may increase the cost of equity.

Table 3 shows that the mean (median) **cost of equity** is 10.466% (8.88%). These values are lower than the ones found by Easton (2004) (11.3% median value). Moreover, the mean (median) number of **analysts following** the firm is 7.4344 (5) while the maximum number of analysts following a firm, in our sample, is 28. We observe that the mean and median values of our sample are higher than the ones on the Gietzmann, M., & Ireland, J. (2005) sample (6.173 mean and a median of 4, respectively). Furthermore, the mean and median values of **beta** in our sample are 0.75919 and 0.73, respectively. Our values are similar to the ones found in Cheng, C. A. (2006) with 0.833 (mean) and 0.692 (median). Our sample shows that the **market to book value** mean (2.9371) is equivalent to the Cheng, C. A. et al (2006)'s sample (0.379 book to market value, which is equivalent to 2.639 market to book value). This means that the average market value is almost three times higher than the book value of the same firm. The variable **ROA** indicates that the mean values found in our sample are positive (6.7261). Furthermore, our sample shows that the maximum value of **log(total assets)** 18.411, while the mean value is 12.712. Our sample mean values of log(total assets) are lower than the values found in Ahmed, A. H. et al (2019) (14.260).

Table 3 - *Descriptive Statistics*

Variables	Mean	Std.Dev	Min	perc.5	Median	perc.95	Max
COEQ	0.10466	0.058604	0.012883	0.044582	0.088838	0.22646	0.44721
ANLST	7.4344	6.8764	1	1	5	22	29
BETA	0.75919	0.52113	-0.63	-0.06	0.73	1.67	2.31
EPS	0.28882	0.36561	0	0	0.16	1.09	2.49
LVRG	55.529	86.601	-190.86	0	30.49	213.09	679.06
ROA	6.7261	7.5206	-33.12	-4.95	6.37	19.07	32.53
VOLATILITY	0.35962	0.13797	0.1225	0.1844	0.3342	0.6205	0.9362
MTBV	2.9371	2.7181	-5.57	0.6	2.18	7.91	23.34
NegROA	0.10646	0.30849	0	0	0	1	1
TA	2137800	7293000	5768	19555	285100	8488000	99056000
N	2339						

5.4. Estimation model

As mentioned in the literature review, the relation between cost of equity capital and disclosure can be separated in different factors of disclosure. Hence, we applied two different models, one that estimates cost of equity relation to readability (2) and other two that estimate cost of equity relation to disclosure content (3)(4). Equation (1) is composed of only control variables, to use as a benchmark. Equation (2) uses length, measured as the logarithm of total words, as a readability measure, to estimate the impact of readability on cost of equity. We also keep the logarithm of total words for model robustness purposes, to detect if the effect of variables of disclosure content remains even with readability variables. Model (3) uses uncertainty tone as the first measure of disclosure content. In model (4) we use net tone instead of uncertainty as a measure of disclosure content. Model (5) uses length and both measures of disclosure content, to evaluate if both (positive/negative and uncertainty) effects are concurrent.

All equations above, minus the equation 1, are estimated for the three sections mentioned in section 3.

$$(1): Coeq_{j,t} = \alpha + \beta_{1,t} LogTA + \beta_{2,t} ANLST + \beta_{3,t} BETA + \beta_{4,t} MTBV + \beta_{5,t} LVRG + \beta_{6,t} ROA + \beta_{7,t} EPS + \beta_{8,t} NegROA + \beta_{9,t} VOLATILITY + \varepsilon_{j,t}$$

$$(2): Coeq_{j,t} = \alpha + \beta_{1,t} LogTA + \beta_{2,t} ANLST + \beta_{3,t} BETA + \beta_{4,t} MTBV + \beta_{5,t} LVRG + \beta_{6,t} ROA + \beta_{7,t} EPS + \beta_{8,t} NegROA + \beta_{9,t} VOLATILITY + \boldsymbol{\beta_{10,t} LogTotWords} + \varepsilon_{j,t}$$

$$(3): Coeq_{j,t} = \alpha + \beta_{1,t} LogTA + \beta_{2,t} ANLST + \beta_{3,t} BETA + \beta_{4,t} MTBV + \beta_{5,t} LVRG + \beta_{6,t} ROA + \beta_{7,t} EPS + \beta_{8,t} NegROA + \beta_{9,t} VOLATILITY + \boldsymbol{\beta_{10,t} LogTotWords} + \boldsymbol{\beta_{11,t} Uncertainty} + \varepsilon_{j,t}$$

$$(4): Coeq_{j,t} = \alpha + \beta_{1,t} LogTA + \beta_{2,t} ANLST + \beta_{3,t} BETA + \beta_{4,t} MTBV + \beta_{5,t} LVRG + \beta_{6,t} ROA + \beta_{7,t} EPS + \beta_{8,t} NegROA + \beta_{9,t} VOLATILITY + \boldsymbol{\beta_{10,t} LogTotWords} + \boldsymbol{\beta_{11,t} NET} + \varepsilon_{j,t}$$

$$(5): Coeq_{j,t} = \alpha + \beta_{1,t} LogTA + \beta_{2,t} ANLST + \beta_{3,t} BETA + \beta_{4,t} MTBV + \beta_{5,t} LVRG + \beta_{6,t} ROA + \beta_{7,t} EPS + \beta_{8,t} NegROA + \beta_{9,t} VOLATILITY + \boldsymbol{\beta_{10,t} LogTotWords} + \boldsymbol{\beta_{11,t} NET} + \boldsymbol{\beta_{12,t} Uncertainty} + \varepsilon_{j,t}$$

6. Results

Table 4 reports the OLS regression for the equation 1, which is the baseline model with the control variables. To estimate this regressions, we introduced year and industry fixed effects to control for heteroskedasticity. We also note that the adjusted explanatory power is 29.007%.

On the control variables, we find statistical significance on the 1% level, for the **number of analysts forecasts (ANLST)**, with a positive impact on cost of equity. This is inconsistent with the idea that greater informational environment is associated with lower cost of equity (Botosan 1997). **Leverage (LVRG)** registered significance at the 1% level and a positive coefficient. This is consistent with the idea that higher leverage traduces in higher credit risk for the company (Cheng et al 2006). **Return on assets (ROA)** also shows significance at the 1% level for all models. The coefficient is negative, consistent with the idea that more profitable firms have lower costs of equity. **Volatility** shows significance at the 1% level and a positive coefficient, consistent with the hypothesis that stock return volatility increases cost of equity (Botosan & Plumlee 2002). **MTBV** consistent with the idea that firms with more growth opportunities are considered to have lower costs of equity, reports negative coefficients at the 1% statistical significance level. **Log(TA)** also finds statistically significant results (1% level), that larger firms experience lower cost of equity, consistent with the theory that larger firms have greater informational contexts (Botosan 1997). Lastly, we find significance for the **NegROA** dummy at the 1% significance in all the estimated models. We do not report any statistical significance for the **BETA** and **EPS** variables.

Table 4 - *Results for equation 1*

*,**,***, indicate that a coefficient is statistically significant at the 10%, 5% and 1% level, or better. The values below each coefficient refer to the correspondent standard error.

Intercept	0.17115	***
	0.01728	
ANLST	0.00082	***
	0.00031	
BETA	-0.00153	
	0.00265	
EPS	-0.00291	
	0.00363	
LVRG	0.00009	***
	0.00002	
ROA	-0.00115	***
	0.00021	
VOLATILITY	0.11872	***
	0.01008	
MTBV	-0.00355	***
	0.00051	
NegROA	0.01870	***
	0.00457	
Log(TA)	-0.00790	***
	0.00124	
Industry fixed-effects	Yes	
Year fixed-effects	Yes	
Adjusted R2	0.29007	

Table 5 shows the OLS regressions made to estimate the impact of readability on the cost of equity (H1), through equation (2). We include industry and year fixed effects to control for heteroskedasticity. The adjusted explanatory power of the models ranges from 28.119% to 29.704%.

In these regressions, we find no statistical significance for the variable of readability (**logarithm of total words).** We interpret these results as readability, measured through the length of each section, not having an impact on the cost of equity capital. Therefore, we reject the first hypothesis **H1.**

Table 5 - *Results for equation 2*

*,**,***, indicate that a coefficient is statistically significant at the 10%, 5% and 1% level, or better. The values below each coefficient refer to the correspondent standard error.

	Chairman Statement		Financial Review		Risk Section	
Intercept	0.14395	***	0.16311	***	0.14708	***
	0.02494		0.02503		0.02306	
ANLST	0.00083	**	0.00094	***	0.00047	
	0.00033		0.00036		0.00034	
BETA	-0.00387		-0.00092		-0.00189	
	0.00285		0.00314		0.00300	
EPS	-0.00712	*	-0.00335		-0.00810	**
	0.00407		0.00421		0.00405	
LVRG	0.00008	***	0.00008	***	0.00007	***
	0.00002		0.00002		0.00002	
ROA	-0.00116	***	-0.00111	***	-0.00063	**
	0.00023		0.00025		0.00024	
VOLATILITY	0.11727	***	0.11954	***	0.12914	***
	0.01096		0.01247		0.01200	
MTBV	-0.00303	***	-0.00312	***	-0.00337	***
	0.00055		0.00063		0.00060	
NegROA	0.01833	***	0.02185	***	0.01838	***
	0.00489		0.00532		0.00524	
Log(TA)	-0.00707	***	-0.00727	***	-0.00415	***
	0.00135		0.00155		0.00151	
Log(Words)	0.00199		-0.00119		-0.00034	
	0.00229		0.00209		0.00210	
Industry fixed-effects	Yes		Yes		Yes	
Year fixed-effects	Yes		Yes		Yes	
Adjusted R2	0.28502		0.29704		0.28119	

To test if disclosure content impacts the cost of equity (H2), we estimate other sets of models that add two different types of tone, for the three sections mentioned in section 3. On the table 6 and 7, we report the OLS regressions to equation 3 and 4. These regressions investigate the impact of net positive and uncertainty tone separately. We include year and industry fixed effects to control for heteroskedasticity in both types of models. The explanatory power of the models used to estimate equation (3) ranges between 28.039% and 29.875%. For the model on table 7, used to estimate equation (4), the range of adjusted explanatory power ranges between 28.229% and 30.642%.

Table 6 - *Results for equation 3*

*,**,***, indicate that a coefficient is statistically significant at the 10%, 5% and 1% level, or better. The values below each coefficient refer to the correspondent standard error.

	Chairman Statement		Financial Review		Risk Section	
Intercept	0.16040	***	0.13622	***	0.15572	***
	0.02687		0.02563		0.02493	
ANLST	0.00092	***	0.00084	**	0.00044	
	0.00035		0.00036		0.00034	
BETA	-0.00384		-0.00145		-0.00194	
	0.00297		0.00315		0.00300	
EPS	-0.00674		-0.00353		-0.00799	**
	0.00426		0.00425		0.00405	
LVRG	0.00008	***	0.00008	***	0.00007	***
	0.00002		0.00002		0.00002	
ROA	-0.00110	***	-0.00108	***	-0.00061	**
	0.00024		0.00026		0.00024	
VOLATILITY	0.11746	***	0.12148	***	0.13036	***
	0.01152		0.01254		0.01205	
MTBV	-0.00322	***	-0.00301	***	-0.00340	***
	0.00058		0.00063		0.00060	
NegROA	0.01900	***	0.02010	***	0.01999	***
	0.00508		0.00536		0.00529	
Log(TA)	-0.00742	***	-0.00711	***	-0.00388	**
	0.00142		0.00158		0.00152	
Log(Words)	0.00024		0.00189		-0.00138	
	0.00249		0.00226		0.00224	
Uncertainty	0.00030	*	0.00006		-0.00010	
	0.00017		0.00010		0.00010	
Industry fixed-effects	Yes		Yes		Yes	
Year fixed-effects	Yes		Yes		Yes	
Adjusted R2	0.28039		0.29875		0.28118	

For the equation 3 we note that uncertainty shows statistical significance on the 10% level for the Chairman Statement. The coefficient of this variable is positive. We interpret this result as uncertainty tone being associated with higher levels of cost of equity. Table 6 do not report statistical significance for uncertainty tone in any of the remaining models corresponding to the other sections mentioned.

The results on table 7, referring to equation 4 show that the variable net tone shows statistical significance on the 1% level for the Chairman Statement and Financial Review models. Moreover, the same variable, shows significance at the 10% level for the Risk Section model. The coefficients of net tone are negative for

all sections. We interpret these results as net positive tone being associated with lower level of cost of equity. This could mean that positive tone is interpreted as good news by investors.

The results of equation 3 and 4 allow us to accept H2. We conclude that tone impacts cost of equity. Furthermore, we note that the impact of uncertainty tone in the Chairman Statement is associated with higher cost of equity, while the impact of net tone is negative for all sections.

Table 7 - *Results for equation 4*

*,**,***, indicate that a coefficient is statistically significant at the 10%, 5% and 1% level, or better. The values below each coefficient refer to the correspondent standard error.

	Chairman Statement		Financial Review		Risk Section	
Intercept	0.21000	***	0.14645	***	0.14743	***
	0.02717		0.02542		0.02311	
ANLST	0.00088	**	0.00078	**	0.00049	
	0.00035		0.00036		0.00034	
BETA	-0.00341		-0.00202		-0.00183	
	0.00293		0.00313		0.00300	
EPS	-0.00763	*	-0.00469		-0.00829	**
	0.00419		0.00423		0.00405	
LVRG	0.00007	***	0.00008	***	0.00007	***
	0.00002		0.00002		0.00002	
ROA	-0.00093	***	-0.00094	***	-0.00062	**
	0.00023		0.00026		0.00024	
VOLATILITY	0.11688	***	0.12133	***	0.13070	***
	0.01134		0.01243		0.01203	
MTBV	-0.00297	***	-0.00304	***	-0.00338	***
	0.00057		0.00062		0.00060	
NegROA	0.01308	***	0.01649	***	0.01932	***
	0.00506		0.00541		0.00528	
Log(TA)	-0.00742	***	-0.00683	***	-0.00395	***
	0.00140		0.00157		0.00151	
Log(Words)	-0.00381		0.00054		-0.00128	
	0.00251		0.00227		0.00214	
NetTone	-0.00041	***	-0.00020	***	-0.00014	*
	0.00005		0.00005		0.00008	
Year fixed-effects	Yes		Yes		Yes	
Industry fixed-effects	Yes		Yes		Yes	
Adjusted R2	0.30273		0.30642		0.28229	

Table 6 and 7 analyze the impact of uncertainty and net positive tone separately. We further develop our analysis on the disclosure content impact on cost of equity, by analyzing if the impact of net tone and uncertainty tone is concurrent. Therefore, we estimate the OLS regressions of equation 5.

The results of equation 5 show that positive net tone show statistical significance at the 1% level for the Chairman Statement and Financial Review. The same variable also shows statistical significance at the 10% level on the Risk Section. The coefficients for all sections are negative. Moreover, we note that uncertainty tone has no statistical significance for all the three analyzed sections. Concluding, the impact of net positive tone found in equation 5 is similar to the impact for equation 4. However, uncertainty tone loses the statistical impact of equation 3, when analyzing both types of tone concurrently.

Table 8 - *Results for equation 5*

*,**,***, indicate that a coefficient is statistically significant at the 10%, 5% and 1% level, or better. The values below each coefficient refer to the correspondent standard error.

	Chairman Statement		Financial Review		Risk Section	
Intercept	0.21259	***	0.14888	***	0.15632	***
	0.02736		0.02569		0.02491	
ANLST	0.00087	**	0.00078	**	0.00050	
	0.00035		0.00036		0.00034	
BETA	-0.00342		-0.00197		-0.00187	
	0.00293		0.00313		0.00300	
EPS	-0.00755	*	-0.00451		-0.00831	**
	0.00419		0.00424		0.00405	
LVRG	0.00007	***	0.00008	***	0.00007	***
	0.00002		0.00002		0.00002	
ROA	-0.00093	***	-0.00093	***	-0.00061	**
	0.00023		0.00026		0.00024	
VOLATILITY	0.11686	***	0.12061	***	0.13040	***
	0.01134		0.01248		0.01204	
MTBV	-0.00296	***	-0.00303	***	-0.00342	***
	0.00057		0.00062		0.00060	
NegROA	0.01292	**	0.01624	***	0.01964	***
	0.00507		0.00542		0.00529	
Log(TA)	-0.00744	***	-0.00692	***	-0.00386	**
	0.00140		0.00157		0.00151	
Log(Words)	-0.00380		0.00066		-0.00200	
	0.00251		0.00227		0.00227	
NetTone	-0.00043	***	-0.00021	***	-0.00014	*
	0.00006		0.00005		0.00008	
Uncertainty	-0.00014		-0.00007		-0.00009	
	0.00018		0.00010		0.00010	
Industry fixed-effects	Yes		Yes		Yes	
Year fixed-effects	Yes		Yes		Yes	
Adjusted R2	0.30259		0.30614		0.28225	

Overall, our results show that readability has no significant impact on the cost of equity. However, our results show that uncertainty and net tone impact the cost of equity. Furthermore, these results show that uncertainty tone is associated with higher levels of cost of equity, only for the Chairman Statement. Net tone reports a negative relation with the cost of equity, for all sections. Yet, when both uncertainty and net tone, are analyzed concurrently, the impact uncertainty tone on the Chairman Statement is lost, while the impact of net tone stays the same for all sections.

7. Robustness tests

As mentioned above in section 5.2.2, there is no perfect method to measure readability. Therefore, to test for the robustness of the results, we substitute the length of the document, for the Fog index. We estimate the OLS regressions for the same above models. Fog index method is explained in the section 5.2.2. We then estimate the following models:

$$(6): Coeq_{j,t} = \alpha + \beta_{1,t}LogTA + \beta_{2,t}ANLST + \beta_{3,t}BETA + \beta_{4,t}MTBV + \beta_{5,t}LVRG + \beta_{6,t}ROA + \beta_{7,t}EPS + \beta_{8,t}NegROA + \beta_{9,t}VOLATILITY + \boldsymbol{\beta_{10,t}Fog} + \varepsilon_{j,t}$$

$$(7): Coeq_{j,t} = \alpha + \beta_{1,t}LogTA + \beta_{2,t}ANLST + \beta_{3,t}BETA + \beta_{4,t}MTBV + \beta_{5,t}LVRG + \beta_{6,t}ROA + \beta_{7,t}EPS + \beta_{8,t}NegROA + \beta_{9,t}VOLATILITY + \boldsymbol{\beta_{10,t}Fog} + \boldsymbol{\beta_{11,t}Uncertainty} + \varepsilon_{j,t}$$

$$(8): Coeq_{j,t} = \alpha + \beta_{1,t}LogTA + \beta_{2,t}ANLST + \beta_{3,t}BETA + \beta_{4,t}MTBV + \beta_{5,t}LVRG + \beta_{6,t}ROA + \beta_{7,t}EPS + \beta_{8,t}NegROA + \beta_{9,t}VOLATILITY + \boldsymbol{\beta_{10,t}Fog} + \boldsymbol{\beta_{11,t}NET} + \varepsilon_{j,t}$$

$$(9): Coeq_{j,t} = \alpha + \beta_{1,t}LogTA + \beta_{2,t}ANLST + \beta_{3,t}BETA + \beta_{4,t}MTBV + \beta_{5,t}LVRG + \beta_{6,t}ROA + \beta_{7,t}EPS + \beta_{8,t}NegROA + \beta_{9,t}VOLATILITY + \boldsymbol{\beta_{10,t}Fog} + \boldsymbol{\beta_{11,t}NET} + \boldsymbol{\beta_{12,t}Uncertainty} + \varepsilon_{j,t}$$

Analyzing the equation 6 OLS estimates, present on table 9, we note that the impact of fog index shows statistical significance at the 1% level for the Risk Section. This variable shows a positive coefficient. We interpret this result as an increase in the complexity of the Risk section narrative to increase cost of equity. Fog index do not show any statistical significance for the Chairman Statement and Financial Review models.

Table 9 - *Results for equation 6*

*,**,***, indicate that a coefficient is statistically significant at the 10%, 5% and 1% level, or better. The values below each coefficient refer to the correspondent standard error.

	Chairman Statement		Financial Review		Risk Section	
Intercept	0.17189	***	0.15346	***	0.12031	***
	0.02376		0.02599		0.02287	
ANLST	0.00082	**	0.00092	**	0.00049	
	0.00033		0.00036		0.00034	
BETA	-0.00389		-0.00083		-0.00175	
	0.00285		0.00314		0.00299	
EPS	-0.00691	*	-0.00326		-0.00798	**
	0.00406		0.00421		0.00404	
LVRG	0.00008	***	0.00008	***	0.00007	***
	0.00002		0.00002		0.00002	
ROA	-0.00116	***	-0.00111	***	-0.00060	**
	0.00023		0.00025		0.00024	
VOLATILITY	0.11733	***	0.11894	***	0.12920	***
	0.01096		0.01247		0.01194	
MTBV	-0.00308	***	-0.00312	***	-0.00343	***
	0.00055		0.00063		0.00060	
NegROA	0.01848	***	0.02177	***	0.01800	***
	0.00489		0.00532		0.00522	
Log(TA)	-0.00710	***	-0.00745	***	-0.00467	***
	0.00135		0.00154		0.00146	
FOG	-0.00071		0.00017		0.00149	***
	0.00067		0.00078		0.00053	
Industry fixed-effects	Yes		Yes		Yes	
Year fixed-effects	Yes		Yes		Yes	
Adjusted R2	0.28516		0.29690		0.28516	

Table 10 shows that uncertainty tone has statistical significance at 10% level for the Chairman Statement model, and this impact is positive. We note that this result is similar to the one found in table 6. We also note that Fog index has a positive impact in the Risk Section at the 1% level significance.

Table 10 *- Results for equation 7*

*,**,***, indicate that a coefficient is statistically significant at the 10%, 5% and 1% level, or better. The values below each coefficient refer to the correspondent standard error.

	Chairman Statement		Financial Review		Risk Section	
Intercept	0.18142	***	0.14372	***	0.12266	***
	0.02510		0.02636		0.02335	
ANLST	0.00091	***	0.00087	**	0.00045	
	0.00035		0.00036		0.00034	
BETA	-0.00376		-0.00146		-0.00172	
	0.00297		0.00315		0.00299	
EPS	-0.00675		-0.00371		-0.00791	*
	0.00425		0.00425		0.00404	
LVRG	0.00008	***	0.00008	***	0.00007	***
	0.00002		0.00002		0.00002	
ROA	-0.00110	***	-0.00107	***	-0.00059	**
	0.00024		0.00026		0.00024	
VOLATILITY	0.11762	***	0.12209	***	0.13001	***
	0.01151		0.01254		0.01199	
MTBV	-0.00324	***	-0.00300	***	-0.00348	***
	0.00058		0.00063		0.00060	
NegROA	0.01902	***	0.02022	***	0.01952	***
	0.00507		0.00536		0.00526	
Log(TA)	-0.00745	***	-0.00688	***	-0.00463	***
	0.00142		0.00156		0.00147	
FOG	-0.00098		0.00016		0.00163	***
	0.00071		0.00079		0.00054	
Uncertainty	0.00033	*	0.00007		-0.00011	
	0.00017		0.00010		0.00009	
Industry fixed-effects	Yes		Yes		Yes	
Year fixed-effects	Yes		Yes		Yes	
Adjusted R2	0.28118		0.29843		0.28568	

Table 11 analyzes the impact of net tone with fog index as a robustness measure. Net tone has statistical significance at the 1% level for the Chairman Statement and Financial Review models. This variable also registers statistical significance at the 10% level for the Risk Section. These results are similar to the results found in table 7. Therefore, we confirm the roubustness of the above results on table 7.

Table 11 *- Results for equation 8*

*,**,***, indicate that a coefficient is statistically significant at the 10%, 5% and 1% level, or better. The values below each coefficient refer to the correspondent standard error.

	Chairman Statement		Financial Review		Risk Section	
Intercept	0.20116	***	0.14690	***	0.11532	***
	0.02483		0.02586		0.02294	
ANLST	0.00085	**	0.00079	**	0.00050	
	0.00035		0.00036		0.00034	
BETA	-0.00325		-0.00200		-0.00164	
	0.00293		0.00314		0.00299	
EPS	-0.00790	*	-0.00475		-0.00818	**
	0.00419		0.00422		0.00404	
LVRG	0.00007	***	0.00008	***	0.00007	***
	0.00002		0.00002		0.00002	
ROA	-0.00095	***	-0.00094	***	-0.00060	**
	0.00023		0.00026		0.00024	
VOLATILITY	0.11723	***	0.12136	***	0.13048	***
	0.01134		0.01243		0.01197	
MTBV	-0.00292	***	-0.00304	***	-0.00345	***
	0.00057		0.00062		0.00060	
NegROA	0.01307	***	0.01649	***	0.01887	***
	0.00506		0.00541		0.00526	
Log(TA)	-0.00743	***	-0.00679	***	-0.00465	***
	0.00140		0.00154		0.00147	
FOG	-0.00091		0.00015		0.00156	***
	0.00070		0.00078		0.00053	
NetTone	-0.00040	***	-0.00020	***	-0.00013	*
	0.00005		0.00005		0.00008	
Industry fixed-effects	Yes		Yes		Yes	
Year fixed-effects	Yes		Yes		Yes	
Adjusted R2	0.30248		0.30641		0.28647	

On table 12 we note that the coefficients of net tone are statistically significant for the models of Chairman Statement and Financial Review and are negative. The model of Risk Section has no significance for net tone. Furthermore, uncertainty tone reports no statistical significance for any of the models.

Concluding, the results on table 8 for Chairman Statement and Financial Review are robust. However, the statistical significance of the Risk section results is lost.

***Table 12** - Results for equation 9*

*,**,***, indicate that a coefficient is statistically significant at the 10%, 5% and 1% level, or better. The values below each coefficient refer to the correspondent standard error.

	Chairman Statement		Financial Review		Risk Section	
Intercept	0.20237	***	0.14981	***	0.12017	***
	0.02490		0.02626		0.02338	
ANLST	0.00084	**	0.00079	**	0.00050	
	0.00035		0.00036		0.00034	
BETA	-0.00326		-0.00195		-0.00163	
	0.00293		0.00314		0.00299	
EPS	-0.00783	*	-0.00459		-0.00822	**
	0.00419		0.00423		0.00404	
LVRG	0.00007	***	0.00008	***	0.00007	***
	0.00002		0.00002		0.00002	
ROA	-0.00095	***	-0.00093	***	-0.00059	**
	0.00023		0.00026		0.00024	
VOLATILITY	0.11720	***	0.12070	***	0.12985	***
	0.01134		0.01247		0.01198	
MTBV	-0.00291	***	-0.00303	***	-0.00349	***
	0.00057		0.00062		0.00060	
NegROA	0.01293	**	0.01624	***	0.01910	***
	0.00507		0.00542		0.00526	
Log(TA)	-0.00744	***	-0.00686	***	-0.00470	***
	0.00140		0.00155		0.00147	
FOG	-0.00086		0.00015		0.00163	***
	0.00070		0.00079		0.00054	
NetTone	-0.00041	***	-0.00021	***	-0.00013	
	0.00006		0.00005		0.00008	
Uncertainty	-0.00012		-0.00007		-0.00010	
	0.00018		0.00010		0.00009	
Industry fixed-effects	Yes		Yes		Yes	
Year fixed-effects	Yes		Yes		Yes	
Adjusted R2	0.30226		0.30612		0.28654	

To further test the results of our estimations, we add the interaction between negative ROA variable and measures of readability and tone. We add this interaction to test if these results are robust for firms with negative profitability.

$$(10): Coeq_{j,t} = \alpha + \beta_{1,t} LogTA + \beta_{2,t} ANLST + \beta_{3,t} BETA + \beta_{4,t} MTBV + \beta_{5,t} LVRG + \beta_{6,t} ROA + \beta_{7,t} EPS + \beta_{8,t} NegROA + \beta_{9,t} VOLATILITY + \boldsymbol{\beta_{10,t}} Log(words) + \boldsymbol{\beta_{11,t}} \boldsymbol{Uncertainty} + \boldsymbol{\beta_{12,t}} \boldsymbol{NegROA} * \boldsymbol{Log(words)} + \boldsymbol{\beta_{13,t}} \boldsymbol{NegROA} * \boldsymbol{Uncertainty} + \varepsilon_{j,t}$$

$$(11): Coeq_{j,t} = \alpha + \beta_{1,t}LogTA + \beta_{2,t}ANLST + \beta_{3,t}BETA + \beta_{4,t}MTBV + \beta_{5,t}LVRG + \beta_{6,t}ROA + \beta_{7,t}EPS + \beta_{8,t}NegROA + \beta_{9,t}VOLATILITY + \boldsymbol{\beta_{10,t}}Log(words) + \boldsymbol{\beta_{11,t}NET} + \boldsymbol{\beta_{12,t}NegROA} * \boldsymbol{Log(words)} + \boldsymbol{\beta_{13,t}NegROA} * \boldsymbol{NET} + \varepsilon_{j,t}$$

The analysis of table 13 shows that uncertainty has no statistical significance in the model.

Table 13 - *Results for equation 10*

*,**,***, indicate that a coefficient is statistically significant at the 10%, 5% and 1% level, or better. The values below each coefficient refer to the correspondent standard error.

	Chairman Statement		Financial Review		Risk Section	
Intercept	0.16739	***	0.13392	***	0.15798	***
	0.02766		0.02615		0.02571	
ANLST	0.00093	***	0.00084	**	0.00045	
	0.00035		0.00036		0.00034	
BETA	-0.00381		-0.00129		-0.00192	
	0.00298		0.00314		0.00301	
EPS	-0.00652		-0.00324		-0.00787	*
	0.00426		0.00425		0.00406	
LVRG	0.00008	***	0.00008	***	0.00007	***
	0.00002		0.00002		0.00002	
ROA	-0.00112	***	-0.00106	***	-0.00062	**
	0.00024		0.00026		0.00024	
VOLATILITY	0.11731	***	0.12187	***	0.13078	***
	0.01152		0.01253		0.01206	
MTBV	-0.00321	***	-0.00303	***	-0.00338	***
	0.00058		0.00063		0.00060	
NegROA	-0.04536		0.05302		-0.00834	
	0.06016		0.04958		0.05571	
Log(TA)	-0.00752	***	-0.00722	***	-0.00388	**
	0.00142		0.00158		0.00152	
Log(Words)	-0.00054		0.00012		-0.00155	
	0.00260		0.00010		0.00234	
Neg*Words	0.00901		-0.00077	**	0.00197	
	0.00858		0.00034		0.00635	
Uncertainty	0.00029		0.00223		-0.00013	
	0.00018		0.00236		0.00010	
Neg*Uncertainty	0.01191		-0.00210		0.00026	
	0.05396		0.00637		0.00030	
Industry fixed-effects	Yes		Yes		Yes	
Year fixed-effects	Yes		Yes		Yes	
Adjusted R2	0.28004		0.30038		0.28052	

On table 14, we report that net tone shows similar results for all sections, to the ones found in table 7.

Table 14 - *Results for equation 11*

*,**,***, indicate that a coefficient is statistically significant at the 10%, 5% and 1% level, or better. The values below each coefficient refer to the correspondent standard error.

	Chairman Statement		Financial Review		Risk Section	
Intercept	0.21051	***	0.14093	***	0.14784	***
	0.02816		0.02596		0.02397	
ANLST	0.00089	**	0.00073	**	0.00049	
	0.00035		0.00036		0.00034	
BETA	-0.00329		-0.00180		-0.00182	
	0.00293		0.00313		0.00301	
EPS	-0.00748	*	-0.00451		-0.00828	**
	0.00419		0.00423		0.00406	
LVRG	0.00008	***	0.00008	***	0.00007	***
	0.00002		0.00002		0.00002	
ROA	-0.00095	***	-0.00093	***	-0.00062	**
	0.00023		0.00026		0.00024	
VOLATILITY	0.11727	***	0.12151	***	0.13062	***
	0.01133		0.01241		0.01213	
MTBV	-0.00297	***	-0.00301	***	-0.00338	***
	0.00057		0.00062		0.00060	
NegROA	-0.02594		0.04669		0.01705	
	0.06077		0.04928		0.04346	
Log(TA)	-0.00747	***	-0.00676	***	-0.00396	***
	0.00140		0.00156		0.00152	
Log(Words)	-0.00417		0.00116		-0.00131	
	0.00263		0.00237		0.00225	
NegWords	0.00691		-0.00446		0.00032	
	0.00854		0.00635		0.00583	
NetTone	-0.00037	***	-0.00016	***	-0.00014	*
	0.00006		0.00005		0.00008	
Neg*NetTone	-0.00024	*	-0.00038	**	0.00001	
	0.00014		0.00017		0.00024	
Industry fixed-effects	Yes		Yes		Yes	
Year fixed-effects	Yes		Yes		Yes	
Adjusted R2	0.30364		0.30813		0.28126	

The results for readability are not robust for the Risk Section, being that when readability is measured through the length of the document it does not show any sign of statistical significance, for any section. However, fog index shows strong

statistical significance for the Risk section. Also, the results on uncertainty tone show that the uncertainty tone results are robust, and that uncertainty does not have an impact on cost of equity capital. The net tone results are robust for all analyzed models.

Overall, our robustness tests confirm, the results found on net tone. They also confirm that, when both types of tone measures (net tone and uncertainty tone) are measured concurrently, only the impact of the Chairman Statement and Financial Review remain significant. However, the robustness tests do not confirm the impact of uncertainty (both isolated and concurrently), nor readability on the cost of equity.

8. Conclusion

The objective of this work was to study the impact of UK annual reports linguistic features on the cost of equity capital. We test the impact of readability measures and textual tone on the cost of equity. This analysis was executed for three different sections of the annual reports, the Chairman Statement, Financial Review and Risk Section. Our results report no significance for the impact of readability on the cost of equity. Also, our results show that the impact of net tone on all sections is associated with lower levels of cost of equity. Therefore, we interpret that, positive tone of annual reports is associated with higher levels of cost of equity, on average. These results are robust, even after taking unprofitable firms into consideration, for all sections. We also report that uncertainty tone has no robust impact on the cost of equity capital.

Ultimately, we validate the hypothesis which states that disclosure content impacts cost of equity capital. More specifically, we conclude that the effect of positive tone is associated with lower cost of equity. Our results are in line with Johnstone (2016) assumption that dismisses the theories of information risk and state that the fundamental measure of disclosure, that drives the cost of equity, is content.

Further studies are needed to analyze the impact of other different types of tone and the usage of more appropriate measures of readability on financial narratives. The study on the impact of textual tone on cost of equity also needs to address other sources of disclosure.

9. References

Ahmed, A. H., Eliwa, Y., & Power, D. M. (2019). The impact of corporate social and environmental practices on the cost of equity capital: UK evidence. *International Journal of Accounting & Information Management*.

Akerlof, G. A. (1978). The market for "lemons": Quality uncertainty and the market mechanism. In *Uncertainty in economics* (pp. 235-251). Academic Press.

Amihud, Y., & Mendelson, H. (1986). Asset pricing and the bid-ask spread. *Journal of Financial Economics*, *17*(2), 223-249.

Arslan-Ayaydin, Ö., Boudt, K., & Thewissen, J. (2016). Managers set the tone: Equity incentives and the tone of earnings press releases. *Journal of Banking & Finance*, *72*, S132-S147.

Athanasakou, V., Eugster, F., Schleicher, T., & Walker, M. (2020). Annual report narratives and the cost of equity capital: UK evidence of a U-shaped relation. *European Accounting Review*, *29*(1), 27-54.

Bakarich, K. M., Hossain, M., & Weintrop, J. (2019). Different time, different tone: Company life cycle. *Journal of Contemporary Accounting & Economics*, *15*(1), 69-86.

Barakat, A., Ashby, S., Fenn, P., & Bryce, C. (2019). Operational risk and reputation in financial institutions: Does media tone make a difference?. *Journal of Banking & Finance*, *98*, 1-24.

Bloomfield, R. (2008). Discussion of "annual report readability, current earnings, and earnings persistence". *Journal of Accounting and Economics*, *45*(2-3), 248-252.

Bloomfield, R. J. (2002). The 'incomplete revelation hypothesis' and financial reporting.

Bloomfield, R., & Fischer, P. E. (2011). Disagreement and the Cost of Capital. *Journal of Accounting Research*, *49*(1), 41-68.

Bonsall IV, 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.

Botosan, C. A. (1997). Disclosure level and the cost of equity capital. *Accounting review*, 323-349.

Botosan, C. A. (2006). Disclosure and the cost of capital: what do we know?. *Accounting and business research*, *36*(sup1), 31-40.

Botosan, C. A., & Plumlee, M. A. (2002). A re-examination of disclosure level and the expected cost of equity capital. *Journal of accounting research*, *40*(1), 21-40.

Botosan, C. A., & Plumlee, M. A. (2005). Assessing alternative proxies for the expected risk premium. *The accounting review*, *80*(1), 21-53.

Botosan, C. A., Plumlee, M. A., & Xie, Y. (2004). The role of information precision in determining the cost of equity capital. *Review of Accounting Studies*, *9*(2-3), 233-259.

Boujelbene, M. A., & Affes, H. (2013). The impact of intellectual capital disclosure on cost of equity capital: A case of French firms. *Journal of Economics Finance and Administrative Science*, *18*(34), 45-53.

Brown, S. (1979). The effect of estimation risk on capital market equilibrium. *Journal of Financial and Quantitative Analysis*, 215-220.

Brown, S., & Hillegeist, S. A. (2007). How disclosure quality affects the level of information asymmetry. *Review of accounting studies*, *12*(2), 443-477.

Campbell, J. L., Chen, H., Dhaliwal, D. S., Lu, H. M., & Steele, L. B. (2014). The information content of mandatory risk factor disclosures in corporate filings. *Review of Accounting Studies*, *19*(1), 396-455.

Chakrabarty, B., Seetharaman, A., Swanson, Z., & Wang, X. (2018). Management risk incentives and the readability of corporate disclosures. *Financial Management*, *47*(3), 583-616.

Cheng, C. A., Collins, D., & Huang, H. H. (2006). Shareholder rights, financial disclosure and the cost of equity capital. *Review of Quantitative Finance and Accounting*, *27*(2), 175-204.

Cho, C. H., Michelon, G., Patten, D. M., & Roberts, R. W. (2014). CSR report assurance in the USA: an empirical investigation of determinants and effects. *Sustainability Accounting, Management and Policy Journal*.

Clarkson, P., Guedes, J., & Thompson, R. (1996). On the diversification, observability, and measurement of estimation risk. *Journal of Financial and Quantitative analysis*, 69-84.

Clatworthy, M. A., & Jones, M. J. (2006). Differential patterns of textual characteristics and company performance in the chairman's statement. *Accounting, Auditing & Accountability Journal*.

Cooper, S. (2006). Discussion of'Disclosure and the cost of capital: What do we know?'. *Accounting and Business Research*, *36*, 41-42.

Davis, A. K., & 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.

DeBoskey, D. G., Luo, Y., & Zhou, L. (2019). CEO power, board oversight, and earnings announcement tone. *Review of Quantitative Finance and Accounting*, *52*(2), 657-680.

Demers, E. A., & Vega, C. (2014). The impact of credibility on the pricing of managerial textual content. *Available at SSRN 1153450*.

Diamond, D. W., & Verrecchia, R. E. (1991). Disclosure, liquidity, and the cost of capital. *The journal of Finance*, *46*(4), 1325-1359.

Easley, D., & O'hara, M. (2004). Information and the cost of capital. *The journal of finance*, *59*(4), 1553-1583.

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.

Fama, E. F., & French, K. R. (1993). Common risk factors in the returns on stocks and bonds. *Journal of*.

Feldman, R., Govindaraj, S., Livnat, J., & Segal, B. (2008). The incremental information content of tone change in management discussion and analysis.

Feldman, R., Govindaraj, S., Livnat, J., & Segal, B. (2010). Management's tone change, post earnings announcement drift and accruals. *Review of Accounting Studies*, *15*(4), 915-953.

Gietzmann, M., & Ireland, J. (2005). Cost of capital, strategic disclosures and accounting choice. *Journal of Business Finance & Accounting*, *32*(3-4), 599-634.

Handa, P., & Linn, S. C. (1993). Arbitrage pricing with estimation risk. Journal of Financial and Quantitative Analysis, 81-100.

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.

Henry, E. (2008). Are Investors Influenced By How Earnings Press Releases Are Written? *The Journal of Business Communication (1973)*, *45*(4), 363–407

Henry, E., & Leone, A. J. (2016). Measuring qualitative information in capital markets research: Comparison of alternative methodologies to measure disclosure tone. *The Accounting Review*, *91*(1), 153-178.

Hughes, J. S., Liu, J., & Liu, J. (2007). Information asymmetry, diversification, and cost of capital. *The accounting review*, *82*(3), 705-729.

Jegadeesh, N., & Wu, D. (2013). Word power: A new approach for content analysis. *Journal of financial economics*, *110*(3), 712-729.

Johnstone, D. (2016). The effect of information on uncertainty and the cost of capital. *Contemporary Accounting Research*, *33*(2), 752-774.

Kashanipoor, M., Aghaee, M. A., & Mohseni Namaghi, D. (2020). Information disclosure tone and future performance. *Accounting and Auditing Review*, *26*(4), 570-594.

Kearney, C., & Liu, S. (2014). Textual sentiment in finance: A survey of methods and models. *International Review of Financial Analysis*, *33*, 171-185.

Klein, R. W., & Bawa, V. S. (1976). The effect of estimation risk on optimal portfolio choice. *Journal of financial economics*, *3*(3), 215-231.

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. *The Accounting Review*, *84*(5), 1639-1670.

Lambert, R., Leuz, C., & Verrecchia, R. E. (2007). Accounting information, disclosure, and the cost of capital. *Journal of accounting research*, *45*(2), 385-420.

Lawrence, A. (2013). Individual investors and financial disclosure. *Journal of Accounting and Economics*, *56*(1), 130-147.

Lehavy, R., Li, F., & Merkley, K. (2011). The effect of annual report readability on analyst following and the properties of their earnings forecasts. *The Accounting Review*, *86*(3), 1087-1115.

Lewis, C., & Young, S. (2019). Fad or future? Automated analysis of financial text and its implications for corporate reporting. *Accounting and Business Research*, *49*(5), 587-615.

Li, F. (2008). Annual report readability, current earnings, and earnings persistence. *Journal of Accounting and economics*, *45*(2-3), 221-247.

Li, F. (2010). The information content of forward-looking statements in corporate filings—A naïve Bayesian machine learning approach. *Journal of Accounting Research*, *48*(5), 1049-1102.

Lo, K., Ramos, F., & Rogo, R. (2017). Earnings management and annual report readability. *Journal of Accounting and Economics*, *63*(1), 1-25.

Loughran, T., & McDonald, B. (2014). Measuring readability in financial disclosures. *the Journal of Finance*, *69*(4), 1643-1671.

Luo, Y., & Zhou, L. (2020). Textual tone in corporate financial disclosures: a survey of the literature. *International Journal of Disclosure and Governance*, *17*(2), 101-110.

Mangena, M., Li, J., & Tauringana, V. (2016). Disentangling the effects of corporate disclosure on the cost of equity capital: A study of the role of intellectual capital disclosure. *Journal of Accounting, Auditing & Finance*, *31*(1), 3-27.

Mangena, M., Pike, R. H., & Li, J. (2010). *Intellectual capital disclosure practices and effects on the cost of equity capital: UK evidence*.

Michaels, A., & Grüning, M. (2017). Relationship of corporate social responsibility disclosure on information asymmetry and the cost of capital. *Journal of Management Control*, *28*(3), 251-274.

Miller, B. P. (2010). The effects of reporting complexity on small and large investor trading. *The Accounting Review*, *85*(6), 2107-2143.

Price, S. M., Doran, J. S., Peterson, D. R., & Bliss, B. A. (2012). Earnings conference calls and stock returns: The incremental informativeness of textual tone. *Journal of Banking & Finance*, *36*(4), 992-1011.

Rahman, S. (2019). Discretionary tone, annual earnings and market returns: Evidence from UK Interim Management Statements. *International Review of Financial Analysis*, *65*, 101384.

Reinganum, M. R., & Smith, J. K. (1983). Investor preference for large firms: new evidence on economies of size. *The Journal of Industrial Economics*, 213-227.

Rjiba, H., Saadi, S., Boubaker, S., & Ding, X. S. (2021). Annual report readability and the cost of equity capital. *Journal of Corporate Finance*, 101902.

Schleicher, T., & Walker, M. (2010). Bias in the tone of forward-looking narratives. *Accounting and business research*, *40*(4), 371-390.

Welker, M. (1995). Disclosure policy, information asymmetry, and liquidity in equity markets. *Contemporary accounting research*, *11*(2), 801-827.

Wu, D. X., Yao, X., & Guo, J. L. (2020). Is Textual Tone Informative or Inflated for Firm's Future Value? Evidence from Chinese Listed Firms. *Economic Modelling*, *94*, 513-525.

Zhou, B., Zhang, C., & Zeng, Q. (2018). Does the rhetoric always hide bad intention: annual report's tone and stock crash risk. *China Journal of Accounting Studies*, *6*(2), 178-205.