STOCKHOLM SCHOOL OF ECONOMICS
Department of Accounting
BE351 Degree Project in Accounting and Financial Management
Fall 2025

Does Readability Matter?

Annual Report Readability and Its Effect on Swedish Listed Firms' Implied Cost of Equity

Matilda Hellström (26048) and Saga Bolmgren (26002)

Abstract

This thesis examines the effect of annual report readability on firms' implied cost of equity, and whether linguistic tone moderates this relationship. We employ OLS regressions with industry and year fixed effects on 547 firm-year observations from Swedish listed firms between 2019 and 2024, using robust standard errors. Readability is measured through textual complexity, captured by the Fog Index, and report length in the narrative sections of annual reports. Our results show that firms with more textually complex annual reports face higher implied costs of equity, suggesting that narrative complexity is priced by investors. In contrast, longer reports are associated with lower implied costs of equity. Linguistic tone is measured as the share of negative and ambiguous words in the narrative sections. We find that negative tone moderates the relationship between complexity and the implied cost of equity, while ambiguous tone shows no such effect. These findings suggest that annual report readability carries economic consequences and exhibits value relevance in the Swedish market.

Keywords: *Annual report readability, implied cost of equity, linguistic tone, value relevance*
Supervisor: Katerina Hellström
Acknowledgements: We sincerely thank our supervisor, Katerina Hellström, for her valuable guidance and support throughout our process.

Table of Contents

1. Introduction

As famously quoted by the investor Warren Buffett, *"Never invest in a business you cannot understand"*. The statement captures an essential challenge in financial communication: although annual reports are meant to inform investors, their length and linguistic complexity can obscure the very information they are meant to convey (You and Zhang, 2009). In a period when investors rely more on firms' textual disclosures to interpret performance and risk, the way information is communicated may take on greater importance. In this context, the readability of corporate reports may influence how efficiently markets process information, raising the question: does readability matter?

In capital markets, corporate disclosures play a central role in reducing information asymmetry (Diamond and Verrecchia, 1991). Such disclosures consist of both quantitative figures and qualitative explanations, of which both have documented importance for when investors form their valuations (Loughran and McDonald, 2016). Among these disclosures, annual reports represent one of the most comprehensive communication tools available to firms, and their accounting information has long been shown to be value relevant (Rjiba et al., 2021). However, much less is known about whether the readability of narrative disclosures forms part of that value relevance. Given the increasing emphasis on qualitative elements in financial reporting, this raises the question of whether readability constitutes an important component of disclosure quality, and to what extent it facilitates more effective absorption of information by the market.

Although previous research on readability is relatively limited, the studies that do exist have yielded interesting insights into its potential economic implications. Rjiba et al. (2021) find that less readable annual reports are associated with a higher implied cost of equity, suggesting that readability is linked to firms' financing conditions. Additional evidence indicates that readability has implications beyond equity valuation, as lower readability in regulatory filings is associated with higher borrowing costs (Bonsall and Miller, 2017). Moreover, prior research shows that linguistic tone also carries informational value in narrative reporting, suggesting that tone may interact with readability in shaping market outcomes (Rjiba et al., 2021). While these studies highlight readability as an economically meaningful characteristic of corporate disclosures, existing evidence is largely confined to the US setting, leaving limited understanding of the role of readability in other reporting environments.

In the context of the Swedish market, previous studies on the value relevance of accounting information have paid limited attention to whether readability influences how effectively such information is absorbed by the market. The potential moderating role of linguistic tone has received even less focus and remains largely unexplored in the Swedish setting.

Against this background, the purpose of this thesis is to analyze whether the readability of annual reports influence the implied cost of equity for Swedish listed firms, and whether linguistic tone moderates the effect of readability. Specifically, we aim to answer the following research question:

Does the readability of annual reports affect firms' implied cost of equity, and is this relationship moderated by narrative tone?

To address our research question, we employ OLS regressions with industry and year fixed effects on Swedish listed firms from 2019 to 2024. Building on the empirical approach of Rjiba et al. (2021), we examine two dimensions of readability in our setting: textual complexity and report length. Textual complexity is measured using the Fog Index, a widely applied readability metric in disclosure research, while length is captured through total word count in the narrative sections of the annual report. The implied cost of equity serves as the dependent variable, estimated using the Easton (2004) PEG model and the Ohlson and Juettner-Nauroth (2005) model. Linguistic tone from the Loughran and McDonald (2011) dictionaries is included as a moderating variable. Standard firm-level controls are incorporated throughout.

Our study contributes to the literature by examining whether the relationship between annual report readability and the implied cost of equity also holds in the Swedish capital market. Additionally, by analyzing how readability, tone, and firm-specific characteristics relate to firms' implied cost of equity, we also contribute to the broader literature on the determinants of cost of equity. Improved understanding of how readability and tone influence the cost of equity is relevant for firms seeking to improve the effectiveness of their disclosures, for investors and analysts evaluating risk and performance, and for regulators interested in how disclosure quality affects market efficiency. In this way, our study offers new evidence on the economic consequences of narrative reporting in a market that has previously been largely unexplored.

We limit our study to Swedish listed firms that publish English versions of their annual reports. This ensures that all firms follow comparable regulatory requirements and allows us to conduct a consistent readability analysis. We focus exclusively on the narrative sections of the annual reports, excluding tables and numerical appendices.

Our results indicate that firms with more textually complex annual reports face higher implied costs of equity, suggesting that investors demand a higher risk premium when disclosures are harder to interpret. This effect exists in the overall sample but disappears once financial firms are excluded, indicating that the pricing of complexity may not be uniform across industries. We also find that the effect of complexity on the implied cost of equity is more pronounced in annual reports with stronger negative disclosure tone, whereas we find no corresponding effect for ambiguous tone. Furthermore, we find indications that longer reports are associated with lower implied costs of equity, which may reflect reduced information asymmetry when firms offer more contextual information. Together, our findings suggest that the readability of annual reports has value relevance, supporting the notion that readability matters.

The remainder of this thesis is structured as follows. Section 2 outlines the theoretical foundations for our hypotheses and reviews the relevant literature on readability, narrative tone, and the implied cost of equity. Section 3 develops the hypotheses and describes the construction of the variables and empirical models used to test them. Section 4 explains the data collection process and the resulting sample. Section 5 presents the main regression results and robustness tests, while section 6 discusses the findings in relation to the research question and prior literature. Finally, section 7 concludes the thesis by outlining the implications of our results and providing suggestions for future research.

2. Theory and Literature Review

This section introduces the theoretical and empirical foundations relevant to our study. We begin by outlining the Efficient Market Hypothesis and the concept of information asymmetry. We then review research on corporate disclosure, value relevance, and readability measures, along with linguistic attributes commonly used in financial text analysis. Finally, we summarize prior evidence on the cost of equity and external financing, with a particular focus on how disclosure complexity and tone influence investors' required returns.

2.1 The Efficient Market Hypothesis and Information Asymmetry

According to the Efficient Market Hypothesis, asset prices should fully reflect all available information, implying that disclosure practices or trading frictions should not systematically affect expected returns. In its weak form, prices incorporate all past trading data. The semi-strong form argues that prices adjust instantly to all publicly available information, such as financial statements, earnings announcements and annual reports. Finally, the strong form claims that even private or insider information is already reflected in market prices, implying that no investor can consistently outperform the market (Fama, 1970).

In this study, the semi-strong form is the most relevant, as it focuses on how public information, like annual reports, is absorbed by the market. Under this view, new disclosures should be reflected in prices as soon as they are released. However, the extent to which this happens may depend on how clearly the information is communicated. If the information in an annual report is presented in an overly complex way, it may take longer for investors to interpret and react to it (You and Zhang, 2009).

The concept of information asymmetry originates from Akerlof (1970), who demonstrated that when one party in a transaction possesses more information than the other, market inefficiencies arise. Applied to capital markets, this means that managers typically hold superior knowledge about a firm's true performance and risk compared to external investors. As a result, investors face uncertainty about the reliability of the information disclosed, which leads them to demand a higher return as compensation for this perceived risk (Healy and Palepu, 2001).

To mitigate such asymmetry, firms engage in corporate disclosure, providing both quantitative data and qualitative information in financial reports. Studies such as Diamond and Verrecchia (1991) as well as Healy and Palepu (2001) argue that increased disclosure transparency can reduce information asymmetry. Evidence also shows that markets price information risk. When a stock is more likely to be traded by informed investors, uninformed investors face a higher risk of transacting at a disadvantage. To compensate for this adverse-selection risk, they require a higher expected return (Easley et al., 2002).

However, prior research also shows that information asymmetry depends not only on how much firms disclose, but also on how easily investors can interpret the information. Neel (2017) provides evidence from the mandatory International Financial Reporting Standards (IFRS) adoption, showing that when firms apply a more standardized reporting framework, their financial statements become easier to compare across companies. The higher comparability reduces uncertainty for investors and illustrates that transparent and more consistent reporting through IFRS plays an important role in reducing information asymmetry.

Consistent with this view, evidence on readability in analyst reports show that disclosures lead to more positive market reactions because they reduce uncertainty about future performance.

This effect is strongest when information asymmetry is high, such as in firms with a larger share of uninformed investors, indicating that readability matters most in settings where investors have limited access to information (Hui et al., 2016).

2.2 Corporate Disclosure

Corporate disclosure can broadly be categorized along two dimensions: quantitative versus qualitative and mandatory versus voluntary. Quantitative disclosures refer to numerical data such as earnings, cash flows or leverage, while qualitative disclosures include narrative information like management commentary and risk discussions. Investors rely on both qualitative elements and quantitative data when forming valuations (Loughran and McDonald, 2016).

Mandatory disclosures, such as audited financial statements, are required by regulation and ensure comparability across firms. Voluntary disclosures, on the other hand, give managers discretion over how much and what level of complexity they communicate beyond the numbers. Verrecchia (2001) argue that voluntary disclosure reduces information asymmetry and can lower a firms' cost of capital, as more transparent communication improves investor understanding and reduces perceived risk.

The International Financial Reporting Standards (IFRS) provide a globally recognized framework for how firms present the financial components of their annual reports. IFRS standardizes the structure and comparability of the primary financial statements, but it places less emphasis on narrative reporting. As a result, differences in how firms communicate through the narrative sections become more important. Efretuei et al. (2022) note that IFRS adoption has increased the length and complexity of these sections, as firms must provide more extensive explanations alongside the numerical disclosures. This added complexity largely reflects informative content, such as detailed discussions of accounting judgments and fair value measurements, rather than attempts to obscure information.

2.3 Value Relevance

Narrative disclosures constitute a central component of corporate reporting and provide a substantial source of value-relevant information. These textual sections offer insights that complement financial statement numbers and can support more informed decision-making by investors (Rjiba et al., 2021). Research shows that both investors and analysts pay attention to narrative disclosures. Davis et al. (2012) show that net optimistic language in earnings press releases is associated with positive market reactions, suggesting that investors use qualitative language when forming valuations. Analysts similarly rely on narrative information. Muslu et al. (2008) show that more extensive forward-looking discussion in the Management's Discussion and Analysis (MD&A) section helps analysts produce more accurate earnings forecasts.

Another important aspect concerns the timing of market prices. Value relevance studies examine how accounting information is reflected in market valuations and therefore require share prices to be measured at a point when annual report information has been fully released and incorporated into the market. Kothari (2001) argues that prices gradually absorb accounting information, implying that prices measured too close to the fiscal year-end may not yet reflect firms' disclosures. Consistent with this reasoning, empirical studies typically introduce a lag between the fiscal year-end and the price date. For example, Collins et al. (1997) measure share prices three months after year-end to ensure that annual reports have been published and processed by the market.

2.4 Readability Measures

Several indexes have been developed to assess textual readability, with two of the most prominent being the Fog Index and the Bog Index. The Fog Index captures complexity through sentence length and word structure and has been widely applied in financial text research, including in studies by Li (2008), Biddle et al. (2009) and Bifulco (2025). The Bog Index offers a broader assessment by incorporating elements of writing style, and is similarly well-established in the literature, with applications in Gounopoulos et al. (2025), Weisskopf (2025) and Rjiba et al. (2021).

Readability measures such as the Fog Index have some imitations in financial reporting. Fog treats all multisyllabic words as complex, even though many such terms, such as *"corporation"* and *"management"* are common and easily understood in a financial context (Loughran and McDonald, 2014). It also may not entirely capture actual comprehension difficulty, partly as the metric was not originally developed for annual reports (Jones and Shoemaker, 1994).

Li (2008) similarly argues that high Fog scores in annual reports often reflect industry-specific terminology rather than genuine opacity and should therefore be interpreted as a relative indicator of disclosure complexity across firms, not an absolute measure of readability. These limitations have motivated the development of alternative metrics such as the Bog Index

(Bonsall et al., 2017), which captures stylistic features, including sentence structure, passive voice, weak verbs, jargon, and evaluates word familiarity rather than syllable count.

It is also important to consider firm characteristics, as some industries have more complex business models that naturally require more technical disclosures. Prior evidence shows that industries such as finance, telecommunications and utilities tend to exhibit the highest Fog scores, reflecting the technical nature of their reporting (Garel et al., 2018). Li (2008) similarly hypothesizes that firms with more complex operations or financial structures should produce more complex annual reports. In the study, operational complexity is measured using the number of business and geographic segments, and financial complexity using the number of non-missing financial statement items. However, the empirical results contradict these predictions. This suggests that inherent firm complexity may contribute to disclosure complexity, but the evidence is weaker and less consistent than theory would suggest.

2.5 Tone and Linguistic Attributes

Beyond readability, research shows that the tone of corporate disclosures also provides valuable signals to investors. An important contribution is Loughran and McDonald's (2011) word lists, developed because general dictionaries often misclassify common financial terms (such as *liability* or *capital*) as negative, even though these words are not negative in a financial context. Their dictionaries include both domain-specific terms and general vocabulary commonly used in corporate reporting, which improves the accuracy of tone measurement in narrative disclosures. The word lists capture tone categories including negative, positive, uncertain, litigious, modal and constraining words, each reflecting different interpretive elements of financial text. These lists have become tools in financial text analysis, allowing researchers to quantify disclosure tone.

2.6 Determinants of the Cost of Equity

The cost of equity capital represents the return that investors require to compensate for the risk of providing equity financing to the firm (Botosan, 2006). In traditional asset pricing models such as the Capital Asset Pricing Model (CAPM) (Sharpe, 1964; Lintner, 1965), it is determined by the risk-free rate, the market risk premium and the firm's beta. However, later research shows that this framework explains only part of the variation in expected returns. Fama and French (1993) demonstrate that variables outside traditional asset-pricing theory, particularly firm-specific characteristics such as size and book-to-market, help explain why some firms earn higher or lower returns. Their findings suggest that the cost of equity is influenced by more than just systematic market risk.

Because the cost of equity is forward-looking and cannot be observed directly, it must be estimated (Botosan, 2006). A common approach in the literature is to use ex ante, or implied, cost of equity measures, which infer the discount rate embedded in current stock prices (Botosan, 2006; Reverte, 2012). These models combine the current share price with expectations about future performance, often using analysts' earnings forecasts as a proxy for future cash flows (Gode and Mohanram, 2003). Unlike CAPM-based measures, which restrict the cost of equity to a predetermined set of priced risk factors and therefore cannot capture effects related to disclosure, implied estimates incorporate market expectations and reflect the information environment surrounding the firm. For this reason, implied cost of equity is considered more suitable for research examining how disclosure influences firms' cost of equity (Botosan, 2006).

2.7 Readability and External Financing

Previous research shows that disclosure readability is an important determinant of firms' financing costs. Using a large sample of U.S. firms, Rjiba et al. (2021) find that less readable annual reports are associated with higher implied costs of equity. Evidence from Brazil shows a similar pattern, as the quality of financial information, of which readability is a key component, also influences firms' cost of equity (Telles and Moreira, 2024). Related work demonstrates that lower readability increases firms' borrowing costs, indicating that the effects extend beyond equity markets (Bonsall and Miller, 2017). Furthermore, Diamond and Verrecchia (1991) shows that enhanced disclosure lowers firms' financing costs, and Francis et al. (2008) report similar effects for voluntary reporting. Recent studies also highlight the role of narrative and non-financial information, showing that higher integrated reporting quality reduces the cost of capital by improving the overall information environment (Iqbal et al., 2025).

Another important dimension of readability is the amount of narrative information provided. Athanasakou et al. (2020) document a U-shaped relationship between disclosure volume and the cost of equity. They find that for lower volumes, additional text reduces information asymmetry and lowers the required return. In contrast, once narrative sections become excessively long, the added complexity raises the cost of equity. This notion aligns with Li

(2008), who argues that longer narrative reports generally impose higher information-processing costs, making documents more deterring and more difficult to read.

Finally, research shows that tone can influence how readability affects financing outcomes. Rjiba et al. (2021) find that negative, uncertain and weak modal wording, measured using the Loughran and McDonald (2011) word lists, is associated with a higher implied cost of equity, suggesting that linguistic negativity and ambiguity increase investors' perceived risk. Ertugrul et al. (2017) also apply the Loughran and McDonald (2011) uncertain and weak-modal word categories in their analysis and, using a large U.S. loan sample, show that firms with less readable and more ambiguous annual reports face higher external financing costs and tighter loan terms.

3. Method

This section presents the empirical method used in the study, including the regression models and the construction of the variables employed in the analysis.

3.1 Hypotheses

Based on the previous literature, research and theories within the field, we have constructed two hypotheses:

H_1: Firms with less readable annual reports exhibit a higher implied cost of equity.

H_2: The link between less readable annual reports and a higher implied cost of equity is stronger when the report's tone is more negative or more ambiguous.

3.2 Regression Model for Hypothesis 1

To test our first hypothesis, we estimate whether annual report readability and disclosure complexity are associated with firms' implied cost of equity.

Our baseline regression model is specified as:

$$r_{e,it} = \beta_0 + \beta_1 FOG_{it} + \beta_2 lnWords_{it} + \beta_3 BETA_{it} + \beta_4 BTM_{it} + \beta_5 SIZE_{it} + \beta_6 LEV_{it} +$$

$$+ \beta_7 g_{it} + IndustryFE + YearFE + \varepsilon_{it}$$

Where:

- $r_{e,it}$ measures the implied cost of equity.
- FOG_{it} measures annual-report readability (Fog Index).
- $lnWords_{it}$ captures the report length (number of words) of an annual report.
- $BETA_{it}$ is the market beta.
- BTM_{it} is the book-to-market ratio.
- $SIZE_{it}$ denotes firm size.
- LEV_{it} is the leverage ratio.
- g_{it} is the long-term growth earnings forecast proxy.
- $IndustryFE$ and $YearFE$ represent industry and year fixed effects.

All continuous variables are winsorized at the 1st and 99th percentiles to limit the influence of outliers, and all regressions are estimated with firm-clustered standard errors to account for within-firm correlation.

3.2.1 Dependent Variable

To estimate the implied cost of equity, we calculate firm-level estimates for cost of equity ($r_{e,it}$) using two commonly applied models: the Easton (2004) PEG model and the Ohlson and Juettner-Nauroth (2005) model. As shown in Table 1, the implied cost of equity for each firm *i*, is computed as the arithmetic mean of the two model estimates. Following Rjiba et al. (2021), averaging the two estimates prevents the results from leaning toward a single model and reduces sensitivity to any one input. We use two well-known models rather than the four employed in Rjiba et al. (2021) due to lack of data.

Table 1. Calculation of the implied cost of equity (PEG and OJN models)

Easton (PEG)	**Ohlson-Juettner-Nauroth (OJN)**	**Arithmetic mean**
$r_{e,it}^{PEG} = \sqrt{\frac{EPS_{i\,t+2} - EPS_{i\,t+1}}{P_{0,it}}}$	$r_{e,it}^{OJN} = \frac{D_{i\,t+1}}{P_{0,it}} + \sqrt{\frac{EPS_{i\,t+1} \times g_{it}}{P_{0,it}}}$	$r_{e,it} = \frac{r_{e,it}^{PEG} + r_{e,it}^{OJN}}{2}$

Where:

- $r_{e,it}$ is the implied cost of equity.
- $P_{0,it}$ is the share price used in the calculation
- $EPS_{i\ t+1}$ and $EPS_{i\ t+2}$ are analyst earnings forecasts.
- $D_{i\ t+1}$ is the analyst forecasted dividend for the next year
- g_{it} is the long-term earnings growth rate
- $r_{e,it}^{PEG}$ and $r_{e,it}^{OJN}$ are the PEG and OJN implied cost of equity estimates.

Valuation Date and Model Notation

All variables used to compute the implied cost of equity for year *t* are collected on the same date, 29 April of year *t+1*. This choice is driven by two reasons. First, this timing allows us to capture the value relevance: the date falls after all firms have published their annual reports. This means that market prices, analyst forecasts and expected dividends all reflect the value relevance released with the annual reports. Second, it is set before the Q2 results are announced, which ensures that our inputs are not affected by new information later in the year. Using the same date for all firms avoids timing differences between prices and forecasts, and helps us obtain more consistent implied cost of equity estimates.

To maintain consistency with the PEG and OJN model notation, we use the subscript 0 to denote the valuation date. In our setting, this valuation date corresponds to 29 April of year t+1, after the annual report for fiscal year *t* has been released. Accordingly, $P_{0,it}$ refers to the stock price of firm *i* observed on this date. All remaining inputs are collected on the same valuation date and indexed according to their respective forecast horizons.

Variable Definitions

Implied cost of equity ($r_{e,it}$) represents investors' required rate of return on equity, inferred from market prices and earnings expectations rather than realized returns. It reflects the compensation investors demand for bearing the firm's risk and serves as a forward-looking measure of expected return.

Earnings per share ($EPS_{i\ t+1}$ and $EPS_{i\ t+2}$) represent analysts' expectations of the firms' earnings performance over the next two fiscal years. Ex-ante measures are preferred over ex-post realized returns, as they provide a better reflection of expected returns and are less distorted by noise from external events, factors unrelated to the annual report or the availability of information.

Share price, closing balance ($P_{0,it}$) represents the market's valuation of the firm at that point in time and is used as the price component in both implied cost of equity models.

Forecasted dividend ($D_{i\ t+1}$) represents the expected dividend to be paid in the next period, based on analyst forecasts. It reflects the portion of expected earnings anticipated to be distributed to shareholders and forms part of the expected cash flows used to determine the cost of equity.

Long-term growth in earnings per share (g_{it}) reflects the firm's expected longer-term growth in earnings beyond the short-term forecast horizon. Following the approach of Rjiba et al. (2021) we use the three-year-ahead earnings forecast (EPS_{t+3}) as a proxy because long-term growth estimates were not available for enough firms in our sample. This captures a similar idea by incorporating expectations about performance further into the future. Accordingly, g_{it} is defined as the expected growth rate between year t+2 and t+3, calculated as:

$$g_{it} = \frac{EPS_{i\ t+3}}{EPS_{i\ t+2}} - 1$$

3.2.2 Independent Variables

Our main independent variables capture the readability of firms' annual reports. To measure these aspects, we employ two complementary textual measures: the *Fog Index* and *annual report length*. By including both measures in our regressions, we capture both linguistic and quantitative aspects of textual complexity.

Fog Index

The Fog index is used to measure the readability of annual reports, a widely used metric in accounting and finance research. While the measure has limitations, we use the Fog Index for cross-sectional comparisons, which reduces many of the concerns raised in prior literature. Although more recent studies have used the Bog Index, it requires proprietary software or relies on pre-processed datasets available only for U.S. firms. Because no equivalent Bog data exist for Swedish listed companies, its application is not feasible in our setting. We therefore adopt the Fog Index as a practical and well-established proxy for textual complexity.

The Fog index for firm $i's$ annual report in year *t* is calculated as:

$$FOG_{it} = 0.4 \times (\frac{Words_{it}}{Sentences_{it}} + 100 \times \frac{ComplexWords_{it}}{Words_{it}})$$

where:

- FOG_{it} = the calculated Fog Index in the annual report
- $Words_{it}$= total number of words in the annual report
- $Sentences_{it}$= total number of sentences in the annual report
- $ComplexWords_{it}$= number of words containing three or more syllables in the annual report.

To provide context for interpreting the measure, Fog scores above 18 are typically classified as highly unreadable, scores between 14 and 18 indicate difficult language, and values between 8 and 10 are considered overly simplistic (Li, 2008). Based on the findings in the literature review, we expect the Fog Index to be positively associated with the implied cost of equity.

Report Length

The report length is measured as the natural logarithm of total words in the annual report for firm i in year t ($lnWords_{it}$). This measure is used as an additional proxy for disclosure complexity. This complements the linguistic focus of the Fog Index, while also reducing skewness in word count distributions.

Previous research has commonly used the length of annual reports as a proxy for disclosure complexity. For instance, Bloomfield (2008) argues that firms experiencing poor performance tend to issue lengthier reports to justify their results to investors and to distract investors from bad news. Loughran and McDonald (2014) and Dyer et al. (2017) use the total number of words or file size as indicators of informational complexity. Given this prior research linking longer reports to higher informational complexity, we expect the variable to be positively associated with the cost of equity.

Table 2. Independent Variables

Independent Variable	Construction	Expected sign
Readability measured by the Fog Index (FOG)	0.4 × (average words per sentence + % complex words).	+
Report length (lnWords)	The natural logarithm of total words in an annual report	+

3.3 Regression Model for Hypothesis 2

To test our second hypothesis, we examine whether the association between readability and the implied cost of equity differs depending on the tone of the annual report. Tone is measured using the Loughran-McDonald (2011) word lists, covering three tone dimensions: negative, uncertain, and weak-modal words. In all models, lnWords is included only as a control variable, as the moderation effect is captured exclusively through the interaction between FOG and HighTone.

For each of the three tone variables, we classify firms into HighTone and LowTone groups using an annual median split. Importantly, this classification is based on the share of tone-related words in the report, not the total number of words in each category.

This classification results in a dummy variable, HighTone, constructed as follows:

1. HighTone = 1 for firms with tone values at or above the yearly median.
2. HighTone = 0 for firms with tone values below the yearly median.

This procedure creates six subsamples in total (High vs. Low for each of the three tone dimensions), which allows us to examine whether the effect of readability on the cost of equity differs depending on whether the tone of the annual report is high or low. This results in six separate regressions (High/Low × three tone types). Comparing $\beta_1^{Low,k}$ and $\beta_1^{High,k}$ across these regressions provides initial evidence on whether tone moderates the effect of readability.

To formally test whether the difference between HighTone and LowTone is statistically significant, we estimate a pooled interaction model for each tone dimension.

We estimate the following regressions for each tone category $k \in \{negative, uncertain, weak\_modal\}$

Model	Regression
LowTone	$r_{e,it}^{Low,k} = \beta_0^{Low,k} + \beta_1^{Low,k} FOG_{it} + \beta_2^{Low,k} lnWords_{it} + \beta_3^{Low,k} BETA_{it} + + \beta_4^{Low,k} BTM_{it} + \beta_5^{Low,k} SIZE_{it} + \beta_6^{Low,k} LEV_{it} + \beta_7^{Low,k} g_{it} + + IndustryFE + YearFE + \varepsilon_{it}$
HighTone	$r_{e,it}^{High,k} = \beta_0^{High,k} + \beta_1^{High,k} FOG_{it} + \beta_2^{High,k} lnWords_{it} + \beta_3^{High,k} BETA_{it} + + \beta_4^{High,k} BTM_{it} + \beta_5^{High,k} SIZE_{it} + \beta_6^{High,k} LEV_{it} + \beta_7^{High,k} g_{it} + + IndustryFE + YearFE + \varepsilon_{it}$
Pooled Interaction	$r_{e,it} = \beta_0 + \beta_1 FOG_{it} + \beta_2 HighTone_{it} + \beta_3 (Fog_{it} \times HighTone_{it}) + + \beta_4 lnWords_{it} + \beta_5 BETA_{it} + \beta_6 BTM_{it} + \beta_7 SIZE_{it} + \beta_8 LEV_{it} + + \beta_9 g_{it} + IndustryFE + YearFE + \varepsilon_{it}$

The interaction coefficient β_3 captures the incremental effect of low readability (high Fog) on the cost of equity when tone is high (negative, uncertain, or weak-modal). Based on the findings of Rjiba et al. (2021), a positive sign on β_3 is expected.

We also use a chi-square test to formally assess whether the coefficient on Fog differs significantly between the HighTone and LowTone subsamples, ensuring that any observed difference is not simply driven by random variation in the estimates.

All continuous variables are winsorized at the 1st and 99th percentiles to limit the influence of outliers, and all regressions are estimated with firm-clustered standard errors to account for within-firm correlation.

3.3.1 Moderating Variables

To test Hypothesis 2, the moderating variables are used to examine whether negative or ambiguous tone influences the relationship between readability and the implied cost of equity.

Negative tone is measured using the Loughran-McDonald (2011) negative word list, while ambiguity is measured using two groups of words: uncertain and weak-modal terms. Each tone variable is calculated as the share of category-specific words relative to the total number of words in the annual report.

- Negative words (NEG_share) include terms such as *loss*, *impairment*, and *default*.
- Uncertain words (UNC_share) include words like *may*, *risk*, and *possible*.
- Weak-modal words (WM_share) include words such as *could*, *might*, and *depending*.

Additional examples of words included in each tone category are presented in Appendix 1.

Table 3. Moderating Variables

Moderating Variable	Construction	Interpretation
Negative Tone (NEG_share)	Loughran-McDonald (2011) Negative Words / Total Words	Negative Tone
Uncertain Tone (UNC_share)	Loughran-McDonald (2011) Uncertain Words / Total Words	Ambiguous Tone
Weak-Modal Tone (WM_share)	Loughran-McDonald (2011) Weak-Modal Words / Total Words	Ambiguous Tone

3.4 Control Variables

Market beta (BETA) is included in the regression to control for a firm's systematic risk. Without this control, the readability effect could be confounded with underlying market risk. We use the three-year average beta for all five years. As firms' market risk is generally stable over time, this provides a consistent risk control, and any small variations not captured should not materially affect our regression results. Previous research shows that beta is positively associated with the cost of equity capital (Hail & Leuz, 2006; Athanasakou et al., 2019; Rjiba et al., 2021). In line with this, we expect beta to be positively associated with the cost of equity.

Book-to-market (BTM) is included to control for differences in firms' growth opportunities. For each firm *i* in year *t*, we proxy growth opportunities using the ratio of book value of equity to market value of equity. Consistent with prior literature, we expect BTM to be positively associated with the cost of equity (Gebhardt et al. 2001; Hail and Leuz, 2006; Rjiba et al. 2021).

Firm size (SIZE) is defined as the natural logarithm of total assets for firm *i* in year *t*. Larger and more mature firms tend to exhibit lower costs of equity, as they are perceived as less risky and more transparent (Gebhardt et al., 2001; Gode and Mohanram, 2003). Diamond and

Verrecchia (1991) show that greater disclosure raises stock prices and reduces the required risk premium, an effect that is stronger for larger firms. Therefore, we expect size to be negatively associated with the cost of equity.

Financial leverage (LEV) is defined as long-term debt divided by total assets for firm *i* in year *t*. Gode and Mohanram (2003) and Botosan and Plumlee (2005) contend that higher leverage increases exposure to financial distress and, therefore, risk. Consistent with this view, both studies document a positive relation between leverage and the cost of equity. Thus, we expect leverage to be positively associated with the cost of equity.

Long-term growth (g) is defined as the expected long-term growth in earnings, which we proxy using the earnings-per-share forecast for year t+3. Our sample contains limited formal long-term growth estimates, which requires an alternative measure. Rjiba et al. (2021) use long-term growth when available and replace missing values with the earnings-per-share forecast for year t+3. To ensure consistency across all firms, we adopt a uniform approach and use the EPS_{t+3} estimate as our proxy for long-term growth for the entire sample. Prior research finds that higher long-term growth rates are associated with higher required returns (Gebhardt et al., 2001; Gode and Mohanram, 2003; Rjiba et al., 2021). Therefore, we expect growth to be positively associated with the cost of equity.

Table 4. Dependent Variables

Dependent Variable	Construction	Expected sign
Firm Beta (BETA)	Average 3-year Beta	+
Book-to-market ratio (BTM)	Book value of equity / Market value of equity	+
Firm Size (SIZE)	The natural logarithm of total assets	-
Financial Leverage Ratio	Long-term Debt / Total Assets	+
Long-term growth proxy (g)	Earnings per Share (EPS) t+3 Analyst Forecast	+

3.5 Fixed Effects

Industry Fixed Effects (IndustryFE) controls for unobserved industry-level characteristics that may systematically influence firms' disclosure practices or risk profiles. The fixed effects are constructed by assigning firms to Fama-French industry groups based on their Standard Industrial Classification) SIC codes. Including these dummies ensures that the estimated coefficients reflect within-industry variation rather than differences across sectors.

Year Fixed Effects (YearFE) are used to control for macroeconomic conditions that affect all firms in a given year. Factors such as interest rates, economic crises and regulatory changes create common shocks that could otherwise influence our results. Year fixed effects remove this shared variation and ensure that we do not mistakenly attribute year-specific effects to differences in readability.

4. Empirical Data

In this section, we explain how we selected our sample, collected the data, chose the time period and the industry classification used in the study.

4.1 Sample Selection

We use the same sample of Swedish listed firms for all empirical tests of the relationship between annual report readability and the implied cost of equity, including the tone-based tests in H2. The sample covers the period 2019-2024. Focusing on a single institutional environment reduces cross-country variation in accounting regulation and disclosure practices, improving the comparability of firms' annual reports. The decision to limit the sample to a Swedish context and a six-year window was also driven by practical considerations. The data collection required substantial manual work, as all annual reports had to be downloaded and processed individually, making a longer time horizon infeasible.

The initial sample consists of all publicly listed Swedish firms during the period, resulting in 6,312 firm-year observations.

First, we require firms to have December fiscal year-ends to ensure comparability across reporting periods. Firms with December fiscal year-ends typically publish their annual reports within the same reporting cycle, which is essential for our cost of equity calculation. We use analyst forecasts as of 29 April to capture market expectations after the disclosure of the annual report and firms reporting outside this cycle would not meet this requirement. After applying this restriction, the sample is reduced to 5,430 firm-year observations.

Second, we exclude firms with missing financial data, as these observations do not allow the computation of the implied cost of equity or the required control variables. After this step, 734 firm-year observations remain.

Third, we exclude firms for which an English-language annual report was not available, as the textual analysis requires English disclosures. We also remove the few observations where the report could not be processed for text extraction. After this step, the sample is reduced to 623 firm-year observations.

Finally, we exclude firms for which the implied cost of equity cannot be computed. In the PEG model, certain firms produce negative terms under the square root, which makes the implied cost of equity undefined. A practical alternative is to drop the square-root restriction, but this leads to negative cost of equity values, which are not economically meaningful. We therefore remove these observations. The final sample consists of 547 firm-year observations. Given the remaining sample size, we do not expect this exclusion to affect our ability to examine the relationship of interest. Table 5 summarizes the sample selection process and the resulting reductions at each step.

Table 5. Sample Size

Selection Criteria	**Adjustments**	**Number of companies**
All publicly listed Swedish firms		6312
Firms with non-December fiscal year-ends	-882	5430
Firms with missing required financial data	-4696	734
Firms with missing or unreadable annual reports	-111	623
Firms with undefined implied cost of equity	-92	547
Final Sample		**547**

4.2 Data Collection

The annual reports were manually collected from each company's website. After downloading the reports, we used Python to extract the narrative text and subsequently process it to compute our textual readability metrics. This procedure included calculating the Fog Index, the natural logarithm of total words and the proportions of negative, uncertain, and weak-modal words for each firm-year observation.

All financial data used in this study, including the inputs required to compute the implied cost of equity and all control variables, were obtained from Capital IQ Pro. This includes earnings-per-share forecasts, share prices, dividends, and the financial statement items used to construct our control variables.

4.3 Industry Classification

Table 6 presents the final industry composition of the sample. Industry groups are assigned using firms' Standard Industrial Classification (SIC) codes and mapped into the Fama-French 48 industry classification, which groups firms with similar economic activities into broader industry categories. In our sample, Manufacturing is the largest sector, followed by Services and Finance, while smaller sectors such as Construction, Wholesale, Transportation, Retail, and Other are also represented.

Table 6. Number of Firms per Industry

Industry	**Frequency**
Manufacturing	276
Services	93
Financial	65
Construction	43
Wholesale	37
Transportation, Communications, Utilities	15
Other	11
Retail	7
Total	**547**

5. Results

This section presents the empirical results of the study. We begin with descriptive statistics and a correlation matrix, followed by tests for multicollinearity and heteroskedasticity. We then report the fixed effects regression results for Hypothesis 1 together with its robustness tests, before turning to Hypothesis 2 and its corresponding robustness test.

5.1 Descriptive Statistics

Table 7 reports the descriptive statistics for the main variables used in the analysis. The table provides an overview of the sample distribution, including central tendencies and dispersion measures.

Table 7. Descriptive Statistics

Variable	N	Mean	SD	Min	P1	P25	Median	P75	P99	Max
r_e	547	0.093	0.085	0.017	0.025	0.049	0.069	0.104	0.520	0.759
FOG	547	18.143	1.474	13.591	14.713	17.303	18.239	18.967	21.762	30.841
BETA	547	0.870	0.331	-0.083	0.144	0.611	0.869	1.137	1.552	1.569
BTM	547	0.610	0.623	-0.206	0.038	0.222	0.444	0.814	2.834	6.558
LEV	547	0.174	0.118	0.000	0.000	0.087	0.156	0.238	0.563	0.627
SIZE	547	16.343	1.599	12.108	12.904	15.270	16.429	17.583	19.672	20.387
g	547	0.205	0.424	-0.775	0.010	0.073	0.116	0.193	2.665	4.176
lnWords	547	11.015	0.334	9.193	10.186	10.804	11.039	11.243	11.678	12.080

The cost of equity displays a notable right-skewed distribution, with a mean of 9.3% compared to a median of 6.9%. This reflects a small number of firms with high implied costs of equity, as indicated by the 99th percentile reaching 52%.

FOG exhibits a relatively symmetric distribution around a high level of textual complexity. The mean 18.14 and median 18.24 indicate that firms' annual reports generally have high scores. The dispersion is limited (standard deviation = 1.47), suggesting that readability is fairly similar across firms. A small number of observations take on considerably higher values, with the maximum reaching 30.84.

Beta is close to symmetric, with a mean and median around 0.87. Some firms exhibit higher systematic risk, as indicated by the 99th percentile value of 1.55. The Book-to-market ratio is right-skewed, with a mean of 0.61 and a median of 0.44. This pattern indicates that a large share of the firms in the sample have relatively low book-to-market ratios, while a few firms exhibit higher book-to-market ratios, as seen in the max value of 6.56, which drive the mean upward.

A bias toward larger firms is evident in the distribution of firm size, which shows generally high values with a mean of 16.34 and a median of 16.43. This is consistent with our sample selection and reflects that smaller firms are less likely to have analyst coverage and are therefore commonly excluded from the final sample. The variable also exhibits a substantial variation (standard deviation = 1.60). The leverage ratio is generally low across firms, with a mean of 0.17 and a median of 0.16, indicating that most firms rely relatively little on long-term debt.

The long-term growth proxy is right-skewed, with a mean of 0.21 exceeding the median of 0.12 and a pronounced upper tail (max = 4.18). This pattern is consistent with our use of earnings per share forecasts three years ahead as a proxy for long-term growth, which tend to be more optimistic and more volatile than formal long-term growth estimates. The report length shows relatively limited dispersion, with a mean of 11.02 and a standard deviation of 0.33, indicating that most firms produce annual reports of broadly comparable length.

5.2 Correlation Matrix

Table 8 is presented below and reports the pairwise correlations between the variables.

Table 8. Pearson Correlation Matrix

Variables	r_e	FOG	BETA	BTM	LEV	SIZE	g	lnWords
r_e	1.000							
FOG	0.125***	1.000						
BETA	-0.080*	-0.010	1.000					
BTM	0.547***	0.050	-0.150***	1.000				
LEV	-0.000	0.008	0.049	0.214***	1.000			
SIZE	-0.117***	0.091***	0.431***	0.090**	0.181***	1.000		
g	0.518***	0.053	-0.034	0.097**	-0.089**	0.019	1.000	
lnWords	-0.071*	0.213***	0.377***	0.091**	0.115***	0.715***	0.049	1.000

*** *p < .01*, ** *p < .05*, * *p < .1*

Overall, the correlation coefficients do not indicate any clear signs of multicollinearity, as none of the pairwise correlations are particularly high. The implied cost of equity shows a positive and significant correlation with FOG, indicating that firms with more complex annual reports tend to face higher required returns, consistent with our main hypothesis.

There is a strong and positive correlation between the implied cost of equity and book-to-market, indicating that firms with higher book-to-market ratios tend to have higher implied required returns. Firm size is negatively associated with cost of equity, which is expected given that larger firms typically exhibit lower information asymmetry. The long-term growth proxy shows a positive and significant correlation with the cost of equity, indicating that firms with higher expected earnings growth tend to exhibit higher implied required returns.

In contrast, the report length shows a negative correlation with cost of equity, suggesting that volume alone does not increase perceived risk. The leverage ratio shows no significant correlation with the cost of equity. Beta shows a negative correlation with the cost of equity.

The correlation between FOG and report length is statistically significant and positive, indicating that longer reports are more complex. Among the control variables, there is a positive correlation between report length and firm size, reflecting that larger firms produce longer reports.

5.1.1 VIF Test

To formally evaluate whether multicollinearity is present among the explanatory variables, we conduct a Variance Inflation Factor (VIF) test. The results are presented in Table 9.

Table 9. VIF Results for Multicollinearity

Variable	VIF
FOG	1.06
BETA	1.31
BTM	1.12
LEV	1.09
SIZE	2.25
g	1.03
lnWords	2.19

The Variance Inflation Factors (VIF) results indicate no presence of multicollinearity in our regression models. All VIF results values lie well below the commonly used thresholds of 5 and 10, with the highest value being 2.25 for firm size. This suggests that the explanatory variables are not excessively correlated, and that multicollinearity is unlikely to bias the estimated coefficients or inflate their standard errors.

5.1.2 Breusch-Pagan Test

We perform a Breusch-Pagan test to examine whether the regression models exhibit heteroskedasticity. The results are presented in Table 10.

Table 10. Breusch-Pagan Results for Heteroskedasticity

N	547
Lagrange multiplier	161.4149
p-value	0.0000
f-value	32.2340
f p-value	0.0000

The Breusch-Pagan test indicates heteroskedasticity ($p < 0.001$). This does not affect our inference, as all regressions use heteroskedasticity-robust standard errors clustered at the firm level, which account for both heteroskedasticity and within-firm correlation.

5.3 Regression Results for Hypothesis 1

To test Hypothesis 1, which predicts that lower annual-report readability is associated with a higher implied cost of equity, we estimate an OLS regression model with industry and year fixed effects. Table 11 reports the results, where the implied cost of equity serves as the dependent variable and FOG enters as the main explanatory variable. The remaining firm characteristics are included as controls. Standard errors are clustered at the firm level, which improves the reliability of the coefficient estimates.

Table 11. The Relationship Between Readability and the Implied Cost of Equity

Variable	**Coef.**	**Std. Err**	**t-value**	**p-value**	**95% Conf. Lower**	**95% Conf. Higher**	**Sig.**
FOG	0.006	0.002	2.302	0.021	0.001	0.010	**
BETA	0.026	0.014	1.893	0.058	-0.001	0.010	*
BTM	0.091	0.011	8.552	0.000	0.070	0.112	***
LEV	0.013	0.032	0.420	0.675	-0.049	0.075	
SIZE	-0.007	0.003	-2.562	0.010	-0.013	-0.002	***
g	0.088	0.013	6.532	0.000	0.061	0.114	***
lnWords	-0.026	0.011	-2.385	0.017	-0.047	0.005	**

Mean dependent var	0.092	**SD dependent var**	0.079
Adj. R-squared	0.624	**N obs**	547
F-test	26.277	**Prob. > F**	0.000
Year FE	Yes	**Industry FE**	Yes

**** p < .01, ** p < .05, * p < .1*

The coefficient for FOG is positive (0.006) and statistically significant (p-value = 0.021), indicating that lower readability is associated with a higher implied cost of equity. A one-unit increase in FOG is associated with a 0.6 percentage point increase in the implied cost of equity.

Beta exhibits a positive and significant coefficient. The change in sign relative to the correlation matrix is likely due to omitted-variable bias and the fact that beta is measured as a three-year average, leaving little within-firm variation. In the bivariate correlation, year-to-year changes in the cost of equity are influenced by other firm characteristics, but once these are controlled for and fixed effects are applied, a positive relationship becomes evident.

Book-to-market remains positive and significant. Firm size is negatively associated with the cost of equity, consistent with the view that larger firms are generally more transparent and exhibit lower levels of information asymmetry. The long-term growth proxy also shows a strong and positive relationship with the cost of equity, reflecting that firms with higher expected earnings growth tend to be more volatile and therefore require higher returns. In contrast, the leverage ratio is not statistically significant. Interestingly, the coefficient for report length is negative and significant.

The model demonstrates relatively strong explanatory power, with an adjusted R-squared of 62,4% indicating that the included variables capture a large share of the variation in firms' implied cost of equity. The F-test confirms that the model is jointly significant.

5.3.1 Separate Implied Cost of Equity Models

To ensure that our findings are not driven by the averaging of the two implied cost of equity measures, we re-estimate all baseline regressions using the PEG and OJN models separately. This allows us to assess whether the results are consistent across both model specifications. The robustness tests present the estimated coefficients for readability (FOG) along with the full set of control variables (see Appendix 2). For the PEG model (Panel A), the coefficient on FOG remains positive and statistically significant, and its magnitude increases slightly compared to the baseline estimate. Among the control variables, book-to-market, firm size, and the long-term growth proxy remain significant, while Beta and report length lose statistical significance in this specification. For the OJN model (Panel B), the coefficient on FOG remains positive and statistically significant and the control variables show the same signs and significance patterns as in the baseline model.

5.3.2 Alternative Implied Cost of Equity Measurement Dates

To control that our findings are not driven by the specific date used in the implied cost of equity models, we conduct a robustness test based on alternative reference dates. Results using alternative input dates (See Appendix 3) are based on model inputs collected from 15 April and 15 May, rather than the original reference date of 29 April. This allows us to assess whether the estimated relationship between readability and the cost of equity remains stable when the input date is shifted. We apply a two-week window in each direction because moving the date too far back risks including firms that had not yet published their annual reports, while shifting it too far forward risks incorporating information from Q2 announcements. Across both alternative windows, FOG remains positive and statistically significant. In contrast, Beta and report length lose significance, while firm size and the book-to-market ratio remain significant with coefficients that retain their original signs.

5.3.3 Financial Firms Robustness Test

In a further robustness test for H1, we examine whether financial firms influence the estimated relationship between readability and the cost of equity. Since financial firms make up a relatively large share of our sample (65 observations), and prior evidence shows that these firms tend to produce more complex and technical annual reports (Garel et al., 2018), we investigate whether they affect our main results. Table 12 presents a single robustness table with two panels. Panel A reports the results when we re-estimate the baseline specification after excluding all financial firms, and Panel B shows the results when the analysis is repeated using only financial firms. This setup allows us to assess whether the relationship between readability and the cost of equity is driven by industry-specific reporting characteristics.

Table 12. Robustness Test: Excluding and Including Financial Firms

Panel A: Sample Excluding Financial Firms

Variable	Coef.	Std. Err	t-value	p-value	95% Conf. Lower	95% Conf. Higher	Sig.
FOG	0.004	0.002	1.626	0.104	-0.001	0.008	
BETA	0.020	0.014	1.418	0.156	-0.008	0.048	
BTM	0.100	0.012	8.310	0.000	0.074	0.120	***
LEV	0.062	0.034	1.824	0.068	-0.005	0.130	*
SIZE	-0.008	0.003	-2.801	0.005	-0.014	-0.002	***
g	0.100	0.015	6.430	0.000	0.067	0.123	***
lnWords	-0.021	0.012	-1.830	0.068	-0.044	0.002	*

Mean dependent var	0.094	**SD dependent var**	0.080
Adj R-squared	0.671	**N obs**	482
F-test	23.000	**Prob. > F**	0.000
Year FE	Yes	**Industry FE**	Yes

Panel B: Sample Restricted to Financial Firms

Variable	Coef.	Std. Err	t-value	p-value	95% Conf. Lower	95% Conf. Higher	Sig.
FOG	0.011	0.006	1.823	0.068	-0.001	0.024	*
BETA	0.101	0.043	2.358	0.018	0.017	0.184	**
BTM	0.046	0.022	2.102	0.036	0.003	0.088	**
LEV	-0.165	0.067	-2.468	0.014	-0.296	-0.034	**
SIZE	-0.018	0.011	-1.651	0.099	-0.039	0.003	*
g	0.061	0.022	2.795	0.005	0.018	0.104	***
lnWords	-0.017	0.019	-0.883	0.377	-0.054	0.021	

Mean dependent var	0.078	**SD dependent var**	0.068
Adj R-squared	0.516	**N obs**	65
F-test	13.974	**Prob. > F**	0.000
Year FE	Yes	**Industry FE**	Yes

**** p < .01, ** p < .05, * p < .1*

In Panel A, after removing financial firms, the coefficient on FOG remains positive but is no longer statistically significant. Among the control variables, Beta loses significance, whereas the leverage ratio becomes statistically significant with a positive sign. The remaining controls retain both the expected signs and their significance. This shows that the readability effect documented in our main specification does not remain robust when financial firms are excluded, which indicates that our results in H1 might be driven by the financial firms in our sample.

Panel B reports the results when the regression is estimated only for firms in the financial industry. The FOG coefficient is positive and statistically significant, and all control variables retain their signs and remain statistically significant. The leverage ratio becomes negative and statistically significant in this specification.

Taken together, these results indicate that the readability effect may differ across industries in our sample, and that financial firms could be contributing more strongly to the observed relationship. However, since this robustness test is performed only on the financial industry, we cannot conclusively determine whether similar patterns hold in other sectors.

5.4 Regression Results for Hypothesis 2

Hypothesis 2 predicts that the relationship between annual report complexity and the implied cost of equity is stronger when the tone of the report is more negative or more ambiguous. To assess this, we analyze three tone dimensions: negative tone, uncertain tone, and weak modal tone. We estimate separate OLS models for firms with low and high tone levels and include an interaction term between the Fog Index and the HighTone indicator to capture potential differences in the readability effect. Table 13 reports the results for all three tone categories in Panels A, B and C.

Table 13. Tone as a Moderator of the Readability-Implied Cost of Equity Relationship

Panel A: Negative tone

	LowTone	**HighTone**	**Interaction** **FOG × HighTone**
FOG	**-0.000** (0.004) Chi-square = 11.070 *p-value = 0.001*	**0.023***** (0.006) Chi-square = 11.070 *p-value = 0.001*	**0.010**** (0.004) Chi-square = 11.070 *p-value = 0.001*
Obs.LowTone	272	R-squared LowTone	0.756
Obs. HighTone	275	R-squared HighTone	0.778
Year FE	Yes	industry FE	Yes

Panel B: Uncertain Tone

	LowTone	**HighTone**	**Interaction** **FOG × HighTone**
FOG	**0.001** (0.003) Chi-square = 1.460 *p-value = 0.228*	**0.009** (0.006) Chi-square = 1.460 *p-value = 0.228*	**0.001** (0.003) Chi-square = 1.460 *p-value = 0.228*
Obs.LowTone	272	R-squared LowTone	0.802
Obs. HighTone	275	R-squared HighTone	0.804
Year FE	Yes	industry FE	Yes

Panel C: Weak-modal tone

	LowTone	**HighTone**	**Interaction** **FOG × HighTone**
FOG	**0.006** (0.004) Chi-square = 0.020 *p-value = 0.901*	**0.006** (0.005) Chi-square = 0.020 *p-value = 0.901*	**0.001** (0.004) Chi-square = 0.020 *p-value = 0.901*
Obs.LowTone	272	R-squared LowTone	0.781
Obs. HighTone	275	R-squared HighTone	0.823
Year FE	Yes	industry FE	Yes

*Firm-clustered standard errors are reported in parentheses. *** p<.01, ** p<.05, * p<.1*

Panel A examines whether negative tone moderates the relationship between readability and the cost of equity. The results indicate clear differences across the two tone groups above and below median. For firms with low negative tone, the coefficient on FOG is statistically insignificant, indicating that the baseline effect of readability is weaker within this subsample. In contrast, for firms with high negative tone, the coefficient is positive and significant. The interaction term (FOG × HighTone) is positive and significant, indicating that negative tone functions as a moderator that strengthens the effect of readability on the implied cost of equity. The chi-square tests show that the difference between the two tone groups is statistically significant. The R-squared values are relatively high in both subsamples, with the high-tone specification explaining slightly more of the variation in the cost of equity.

Panels B and C examine whether ambiguous tone moderates the relationship between readability and the cost of equity. Panel B tests this using the uncertainty dimension, while Panel C tests the weak-modal dimension. Across both tone dimensions, we find no evidence of a moderating effect. The coefficients on FOG are statistically insignificant in the LowTone and HighTone subsamples, and the interaction terms likewise show no meaningful effect. The chi-square tests confirm that the differences between tone groups are not statistically significant. Although the models exhibit high explanatory power (adjusted R-squared between 0.78 and 0.82), this appears to reflect the contribution of the control variables rather than any moderating role of tone in either tone category.

5.4.1 Separate Implied Cost of Equity Models

To ensure that the moderating effect of tone is not driven by the averaging of the two implied cost of equity measures, we conduct a robustness test for Hypothesis 2 (see Appendix 4) in which we re-estimate the moderating effect of tone using each implied cost of equity model separately. Across both models, the results for negative tone (Panel A) closely mirror the baseline regression. The coefficient on FOG is positive and statistically significant for firms with high-negative tone group, and the interaction term remains significant. For uncertain tone (Panel B) and weak-modal tone (Panel C), the coefficients on FOG remain statistically insignificant across both models, confirming that these tone categories do not moderate the relationship between readability and the implied cost of equity.

6. Discussion and Analysis

This section discusses our empirical results by interpreting them in relation to prior research and considering potential explanations for the observed patterns. We begin with the findings for Hypothesis 1, followed by Hypothesis 2 and lastly to the results for our control variables before addressing the explanatory power of our models and the broader economic implications of the analysis.

6.1 Analysis and Discussion of Hypothesis 1

Our first hypothesis tested for whether firms with less readable annual reports exhibit a higher implied cost of equity. The *Fog variable* captures the level of complexity for the narrative sections, and the generally high Fog scores in our sample are expected given the technical and standardized language typical of financial reporting. The results provide support for the hypothesis and show that Fog has a positive and significant effect on the implied cost of equity. This reflects that annual report complexity has economic relevance, and that investors demand higher returns when the narrative sections of annual reports are more complex.

Our findings are consistent with prior research showing that disclosure complexity is associated with higher financing costs (Bonsall and Miller, 2017). Rjiba et al. (2021) also document that less readable annual reports are associated with higher implied costs of equity. In practice, this implies that the way firms formulate the discretionary parts of their reports matters, which is especially relevant in an IFRS setting where the numerical statements are standardized and narrative explanations play an important role in explaining those numbers (Efretuei et al., 2022).

To interpret these findings, we draw on the theoretical concept of information asymmetry, which holds that corporate disclosures reduce uncertainty for investors (Diamond and Verrecchia, 1991; Healy and Palepu, 2001). In this context, our results may reflect that when narrative disclosures are harder to read, they do not effectively increase transparency and the reduction in information asymmetry becomes weaker. As a result, investors face greater difficulty in forming precise expectations, which increases perceived risk and leads them to demand higher returns. Another possible explanation is that lower readability increases investors' processing costs. Li (2008) similarly argues that linguistic difficulty raises the effort required to extract relevant information, which makes the processing less efficient. This offers a plausible reason why more complex reports may be associated with higher implied costs of equity, as investors demand compensation for the additional effort.

A third explanation for this association relates to credibility. You and Zhang (2009) notes that when disclosures are difficult to understand, investors may place less weight on the information provided, which reduces confidence and increases the return they require. In this sense, lower readability may result in lower trust in the narrative explanations, which could help explain its association with a higher implied cost of equity. Taken together, these perspectives suggest that the relationship between complexity and the implied cost of equity is reinforced by several factors. Lower readability can weaken the reduction in information asymmetry, increase the effort required to process the information, and reduce investors' confidence in the credibility of the narrative explanations. In our context, it is reasonable to assume that these elements operate simultaneously.

The robustness tests add further nuance to our main results. As firms in the financial industry constitutes a large part of our sample, we excluded these firms to examine whether our main result for H1 depends on their inclusion. The results show that Fog no longer remains significant, indicating that the documented relationship between complexity and the implied cost of equity is sensitive to the presence of financial firms in the sample. When we run the regression on financial firms only, Fog becomes positive and significant, indicating that the association between narrative complexity and the implied cost of equity is strong within this industry.

An intuitive explanation for these results is that some industries naturally produce more complex disclosures due to the underlying complexity of their business models, a notion supported by prior research. Garel et al. (2018) show that firms in more complex industries tend to exhibit higher Fog scores. Our findings are consistent with this theoretical expectation, and the results imply that, even after controlling for industry fixed effects, there appear to be meaningful differences across industries in how narrative complexity relates to the implied cost of equity. Some industries may display levels of narrative complexity that are more strongly linked to variation in the implied cost of equity than others, which could explain why the effect disappears once these firms are excluded. In our analysis, we only test this indirectly by excluding financial firms, and while the results suggests that industry differences may exist, it does not allow us to draw firm conclusions. Rather, it indicates that this may be an area where future research is needed.

We also include the *logarithm of total words* as an additional proxy for readability that reflects the overall length of the narrative sections in the annual report. Our regression results show that report length has a negative and significant effect on the implied cost of equity, indicating that longer reports are associated with a lower implied cost of equity. Even though this appears counterintuitive, it could suggest that the firms in our sample are positioned at a level of disclosure where additional length reduces information asymmetry and consequently lowers the implied cost of equity. This is in line with Athanasakou et al. (2020), who show that when disclosure is relatively limited, providing additional information can enhance transparency and reduce information asymmetry, which in turn lowers the cost of equity.

This finding is particularly interesting in the context of the current reporting environment in Sweden, where IFRS regulate corporate reporting. As discussed earlier, IFRS increases the amount of mandated numerical disclosures, which in turn makes the narrative sections more important for complementing and explaining these figures (Efretuei et al., 2022). One possible interpretation is that adding more content in the narrative sections may help investors interpret the underlying numbers more easily, thereby reducing uncertainty. Our results may reflect this effect.

Furthermore, the robustness tests show that the effect of report length is sensitive to model specification. When estimating the OJN and PEG models separately, the coefficient for our length proxy only remains significant in the OJN model. The two models rely on different types of inputs and reflect different valuation horizons. The PEG model is mainly driven by short-term earnings forecasts, while the OJN model captures longer-term earnings and dividend expectations. Given this, our findings could reflect that disclosure volume plays a more meaningful role when investors form longer-term expectations, possibly having limited influence on the shorter-term earnings focus embedded in the PEG model.

When adjusting the timing of the inputs used in the implied cost of equity calculations, the length proxy becomes insignificant across all specifications, indicating that its effect is not consistently reflected in the implied cost of equity. A similar pattern appears when restricting the sample to financial firms. However, the loss of significance may in this case simply reflect the reduced sample size rather than a substantive change in the relationship. Overall, one possible interpretation is that report length provides a weaker or less stable signal than linguistic complexity. The lack of consistent significance suggests that the market may respond more systematically to narrative complexity than to disclosure volume.

Moreover, the results indicate that Fog and the length of the narrative sections capture different dimensions of disclosure quality. While Fog seems to reflect linguistic difficulty, report length in our sample appears to be a proxy for disclosure transparency. Longer reports may provide more comprehensive context, more detailed explanations and more extensive discussion of firm activities. The opposite signs of the two variables suggest that complexity and volume represent distinct aspects of disclosure, both of which seem to shape how investors assess firms' risk and form expectations about the cost of equity.

Overall, the results indicate that Hypothesis 1 is supported: firms with more textually complex annual reports tend to face a higher implied cost of equity. However, the robustness tests suggest that the pricing relevance of narrative complexity may differ across industries, as the effect disappears once financial firms are removed. Another important finding is the contrasting signs of Fog and report length, which highlight that linguistic complexity and disclosure volume capture distinct aspects of reporting quality. Our results are consistent with previous research showing that annual report readability has economic relevance, and that investors demand higher returns when narrative disclosures are more complex. The findings suggest that readability is value relevant for investors, underscoring the importance of readability in narrative reporting.

6.2 Analysis and Discussion of Hypothesis 2

Hypothesis 2 tests whether negative and ambiguous tone moderate the relationship between annual report readability, now measured solely as textual complexity, and the implied cost of equity. The results show that only negative tone exhibits a statistically significant moderating effect. The interaction term is positive and significant, suggesting that narrative complexity is priced more heavily when the report contains a larger share of negative words. This implies that when the overall message of the report is more negative, investors become particularly attentive to additional complexity. This suggests that in such cases, negative wording makes the situation appear more uncertain or difficult to evaluate, possibly leading investors to interpret it as a signal of heightened risk and therefore require a higher return.

Ambiguous tone is measured using two tone dimensions: uncertain and weak-modal wording. Contrary to our expectations, none of the interaction terms between these tone measures and the Fog Index are statistically significant. Our results therefore provide no evidence that these tone categories moderate the relationship between readability and the implied cost of equity. This suggests that linguistic ambiguity does not materially alter how investors respond to narrative complexity.

While earlier evidence in this area is limited, our results are partly consistent with studies that have examined the relationship between tone and the cost of equity. Rjiba et al. (2021) document moderating effects across all three tone categories, indicating that negative and ambiguous tone can shape how investors interpret disclosure complexity. In contrast, our findings point to a more restricted pattern, as only negative tone strengthens the association between readability and the implied cost of equity in our setting. A useful point of comparison comes from the debt literature. Ertugrul et al. (2017) show that ambiguous language can tighten loan terms, suggesting that ambiguity may have financial implications in other contexts. However, we do not observe a corresponding effect in equity markets. Overall, our results imply that the moderating role of tone appears narrower in our sample than what earlier studies suggest, with negative tone standing out as the only tone dimension that consistently strengthens the pricing effect of narrative complexity.

Our findings are particularly interesting given the uncertain macroeconomic environment that prevailed during our sample period. Between 2019 and 2024, the Swedish market was affected by both the COVID-19 pandemic and rising interest rates, creating a generally uncertain environment for firms (Sveriges Riksbank, 2020; Sveriges Riksbank, 2022). In such a context, investors may have anticipated a higher degree of ambiguity in corporate disclosures, meaning that ambiguous wording conveyed little firm-specific information. As a result, this type of language may not have raised additional concern among investors. Negative tone, by contrast, possibly convey clearer downside risks. Investors may therefore have reacted more strongly when complexity was paired with explicit negative signals, while ambiguity alone might have been too common in this environment to stand out. In this context, the discrepancy between our findings and those of Rjiba et al. (2021), could reflect the macroeconomic environment of our sample period, where ambiguity might have been widely expected and therefore offered limited firm-specific information.

Another possible explanation for our results concerns the characteristics of our dataset. Our sample is considerably smaller than the U.S. samples used in Rjiba et al. (2021) and Ertugrul et al. (2017) which may reduce the statistical power to detect moderating effects of similar magnitude. This issue is particularly relevant for weak-modal tone, as seen in Appendix 5, which contains descriptive statistics for the tone variables and demonstrates that weak-modal wording is both infrequent and displays very limited variation. Such limited dispersion leaves little scope for identifying a meaningful moderating effect. A further consideration is that Rjiba et al. (2021) use the Bog Index, while we use the Fog Index, and these measures capture different aspects of textual complexity. This difference in measurement may also contribute to the contrasting findings, as it could suggest that the Bog Index captures the relevant effects more effectively.

We do not conduct a financial-firm exclusion test for this hypothesis. Financial firms account for 65 observations in our dataset, and once the sample is further split into low-tone and high-tone groups, the resulting subsamples become too small to support a meaningful moderation analysis. However, since the robustness tests for Hypothesis 1 showed that financial firms can influence the results, it is possible that industry differences also matter for this second hypothesis.

In summary, we find partial support for Hypothesis 2. Negative tone strengthens the relationship between complexity and a higher implied cost of equity, suggesting that when the overall message of the report is more negative, investors become particularly attentive to additional complexity. In contrast, we find no corresponding evidence for ambiguous tone. Compared with earlier studies, the moderating role of tone appears narrower in our setting, which may reflect both the macroeconomic uncertainty of the sample period, the limited size of our dataset or differences in readability measurements. Overall, our findings indicate that tone can amplify the pricing effects of narrative complexity, but primarily when disclosures contain negative signals, suggesting that tone may also carry elements of value relevance.

6.3 Analysis and Discussion of Control Variables

Regarding the control variables, some behave as expected, while others deviate from our predictions or display inconsistent patterns across specifications.

Book-to-market has a positive and significant effect on implied cost of equity, which indicates that firms with a higher book-to-market exhibits higher implied cost of equity. This may appear counterintuitive, but it is consistent with the findings of previous research (Gebhardt et al., 2001; Hail and Leuz, 2006; Rjiba et al., 2021). Although value firms are typically viewed as mature and stable companies with lower risk, our results show a positive association between book-to-market and the implied cost of equity. This pattern suggests that the book-to-market variable in our sample does not exclusively capture 'true' value firms, but also firms facing financial distress. Such distressed firms are companies whose market values have declined because investors perceive limited future prospects, a signal of market mistrust. As market value falls for distressed firms, their book-to-market ratio mechanically increases, while their risk, and therefore their cost of equity, rises. Thus, the positive coefficient could possibly reflect that the book-to-market ratio in our model partly proxies for distress rather than pure value characteristics.

Firm size is negative and significant, suggesting that larger firms tend to have lower implied cost of equity, which is consistent with our predictions and previous research (Gebhardt et al., 2001; Gode and Mohanram, 2003). This reflects that larger firms are generally viewed as more stable and transparent, which reduces perceived risk and therefore lowers the required return. However, it is worth noting that our sample might be skewed toward larger firms due to the requirement of analyst coverage. This should be kept in mind, as the relationship we observe may not necessarily generalize to smaller firms.

Long-term growth has a positive and significant effect, implying that higher long-term growth rates are associated with a higher implied cost of equity, which is consistent with previous research (Gebhardt et al. 2001; Gode and Mohanram 2003; Rjiba et al., 2021). This implies that firms with higher expected long-term growth are viewed as riskier because their future earnings are more uncertain, which leads investors to require a higher return.

However, the coefficient on long-term growth becomes insignificant in the robustness test where we apply alternative forecast dates. This variability is likely driven by how the variable is constructed. We use the EPS forecast three years ahead as a proxy for long-term growth and such forecasts tend to be higher and more volatile than traditional long-term growth estimates. As a result, *g* in this context is better interpreted as a cross-sectional comparison rather than a precise long-term measure, which may explain why its significance changes across specifications. Importantly, the main results remain intact when estimating the PEG model, in which *g* is not included, supporting that this proxy has not distorted our overall findings.

Financial leverage is surprisingly not statistically significant in our model, which contrasts with prior research showing that higher leverage increases firms' exposure to financial distress and should therefore raise the implied cost of equity (Gode and Mohanram, 2003; Botosan and Plumlee, 2005). In our sample, however, we do not observe such a relationship.

Beta is positive and significant in our main model, suggesting that firms with higher market risk tend to face higher required returns. This is consistent with previous studies (Hail & Leuz, 2006; Athanasakou et al., 2019; Rjiba et al., 2021). However, its significance fluctuates considerably across the robustness tests. The lack of a consistent pattern suggests that beta does not exhibit a stable relationship with the implied cost of equity in our setting, and its explanatory power appears highly sensitive to how the sample is defined. One possible explanation for this is that beta is defined as a three-year average and applied uniformly across all fiscal years. This construction limits the variation in the variable and may weaken its ability to explain differences in the cost of equity, particularly in smaller subsamples.

The inconsistent behavior of beta and leverage likely reflects characteristics of our sample rather than meaningful economic differences. The relatively small sample size reduces statistical power and makes these controls more sensitive to how the data are partitioned. This helps explain why their significance shifts across specifications, and why leverage only becomes significant in narrowly defined subsamples. Overall, the instability of these coefficients is more indicative of limited variation and sample size constraints than of underlying economic effects.

6.4 Explanatory Power

Our models exhibit relatively high explanatory power. For H1, the adjusted R-squared is approximately 0.60, and for H2 it increases to about 0.75. This indicates that the models explain a substantial share of the variation in the implied cost of equity. These values are considerably higher than those reported in Rjiba et al. (2021), who document an adjusted R-squared of 0.25 in the model corresponding to our H1.

Several factors may explain the high explanatory power in our models. Our sample spans a shorter time period and includes fewer observations, which creates a more homogeneous dataset, while Rjiba et al. (2021) analyze more than 39,000 U.S. firm-year observations from 1995 to 2017, a long period during which reporting standards and market conditions likely changed. In our sample, all firms report under IFRS throughout the entire period, and the Swedish market is likely more uniform than the broader U.S. setting examined in prior work. Such consistency may reduce variation in both reporting practices and market conditions. In addition, the descriptive statistics show that many of our variables display lower dispersion, which generally makes relationships easier to capture within a regression framework. Country-level differences in reporting environments may also contribute, as accounting practices and disclosure styles vary internationally (Nobes, 1998).

7. Conclusion

7.1 Concluding Remarks

This study examines whether the readability of annual reports influences the implied cost of equity for Swedish publicly listed firms during the 2019-2024 period. Specifically, we investigate whether reports with lower readability are associated with a higher implied cost of equity, and whether linguistic tone moderates this relationship. We contribute to the existing literature by providing evidence from a Swedish IFRS-based reporting environment, and according to our knowledge, no prior study has examined this relationship in this context, as earlier studies have predominantly focused on U.S. firms.

Our findings indicate that firms with more textually complex annual reports exhibit higher implied cost of equity, suggesting that readability has economic relevance. This relationship is robust across alternative model specifications and timing adjustments. However, the association weakens when financial firms are excluded, indicating that the pricing of narrative complexity may vary across industries. We also find indications that report length influence the implied cost of equity. In our sample, longer reports are associated with lower implied costs of equity, which could reflect reduced information asymmetry when firms provide more extensive explanations. At the same time, the varying statistical significance across robustness tests suggests that the effect of length is less consistent than the effect of linguistic complexity captured by the Fog Index.

Our results also show that negative tone moderates the relationship between readability and the implied cost of equity. This implies that textual complexity in annual reports is priced more strongly when they contain a higher share of negative wording. In contrast, we find no evidence that ambiguous tone influences this relationship.

Overall, this suggest that narrative reporting is not merely a compliance exercise but a meaningful component of corporate disclosure, as differences in readability are associated with economically relevant outcomes. Our results support the notion that the readability of annual reports has value relevance.

7.2 Limitations of the Study

This study also has limitations. The Fog Index may classify certain technical terms as complex even when they reflect normal financial reporting vocabulary, which should be kept in mind when interpreting the results. This means that direct comparisons with previous studies that rely on the Bog Index should be made with caution. Differences in how the two measures classify and capture linguistic difficulty imply that they are not perfectly interchangeable, and this should be considered when relating our findings to earlier research. Furthermore, our sample is likely skewed toward larger firms, since analyst coverage is required for inclusion and such coverage is more common among larger companies. This represents a limitation of our study, as the findings may not fully generalize to smaller firms that typically do not have analyst coverage. Finally, while the study aims to provide insights into readability effects in Sweden, the findings may not directly generalize to other capital markets or to annual reports written in languages other than English.

7.3 Suggestions for Future research

There are several directions for future research. Given the known limitations of the Fog Index, it would be valuable to examine whether our findings hold when applying alternative readability measures such as the Bog Index, as different metrics capture linguistic difficulty in different ways. Our results also point to potential differences across industries, suggesting that future studies could examine whether the relationship between narrative complexity and the implied cost of equity varies across sectors. Future research could also broaden the linguistic analysis by incorporating additional tone categories beyond the two used in this study. Analyzing a wider set of tone dimensions may reveal moderating effects that are not captured when focusing only on ambiguous and negative wording. Finally, future studies could investigate which parts of the annual report drive the observed relationships. Narrative sections such as the CEO letter, MD&A or risk disclosures may influence the implied cost of equity differently than other parts of the report. Examining these components separately would help clarify where readability matters most.

8. Reference List

Akerlof, G. A. (1970). The market for "lemons": Quality uncertainty and the market mechanism. *The Quarterly Journal of Economics, 84*(3), 488-500

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

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

Bifulco, G. M., Caserio, C., Di Donato, F., & Trucco, S. (2025). Does sustainable performance matter for nonfinancial disclosure readability? A Fog Index analysis on Italian-listed companies. *Business Strategy and the Environment, 34*(5), 5601–5623.

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

Bonsall, S. B., & Miller, B. P. (2017). The impact of narrative disclosure readability on bond ratings and the cost of debt. *Review of Accounting Studies, 22,* 608-643

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. (2006). Disclosure and the cost of capital: What do we know? *Accounting and Business Research, 36*(sup1), 31-40

Collins, D. W., Maydew, E. L., & Weiss, I. S. (1997). Changes in the value-relevance of earnings and book values over the past forty years. *Journal of Accounting and Economics, 24*(1), 39-67

Davis, A. K., Piger, J. M., & Sedor, L. M. (2012). Beyond the numbers: Measuring the information content of earnings press release language. *Contemporary Accounting Research, 29*(3), 845-868

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

Dyer, T., Lang, M., & Stice-Lawrence, L. (2017). The evolution of 10-K textual disclosure: Evidence from Latent Dirichlet Allocation. *Journal of Accounting and Economics, 64*(2-3), 221-245

Easley, D., Hvidkjaer, S., & O'Hara, M. (2002). Is information risk a determinant of asset returns? *The Journal of Finance, 57*(5), 2185-2221

Easton, P. D. (2004). PE ratios, PEG ratios, and estimating the implied expected rate of return on equity capital. *The Accounting Review, 79*(1), 73-95

Efretuei, E., Usoro, A., & Koutra, C. (2022). Complex information and accounting standards: Evidence from UK narrative reporting. *South African Journal of Accounting Research, 36*(3), 171-194

Ertugrul, M., Lei, J., Qiu, J. & Wan, C. (2017). Annual report readability, tone ambiguity, and the cost of borrowing. *Journal of Financial and Quantitative Analysis*, *52*(2), 811-836

Fama, E. F. (1970). Efficient capital markets: A review of theory and empirical work. *The Journal of Finance, 25*(2), 383-417

Fama, E. F., & French, K. R. (1993). Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics, 33*(1), 3-56

Francis, J., Nanda, D., & Olsson, P. (2008). Voluntary Disclosure, Earnings Quality, and Cost of Capital. *Journal of Accounting Research, 46*(1), 53-99

Garel, A., Gilbert, A. B., & Scott, A. (2019). Linguistic complexity and cost of equity capital. *SSRN Working Paper*. https://shorturl.at/EE5lX

Gebhardt, W. R., Lee, C. M. C., & Swaminathan, B. (2001). Toward an implied cost of capital. *Journal of Accounting Research, 39*(1), 135-176

Gode, D., & Mohanram, P. (2003). Inferring the cost of capital using the Ohlson-Juettner model. *Review of Accounting Studies, 8*(4), 399-431

Gounopoulos, D., Huang, W., & Yang, M. (2025). Innovation and annual report readability. *Research Policy, 54*(8), 103601

Hail, L., & Leuz, C. (2006). International differences in the cost of equity capital: Do legal institutions and securities regulation matter? *Journal of Accounting Research, 44*(3), 485-531

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*, 405-440

Hsieh, C.-C., Hui, K. W., & Zhang, Y. (2016). Analyst report readability and stock returns. *Journal of Business Finance & Accounting, 43*(1-2), 98-130

Iqbal, M.A., Sarkar, M.A.R., Alharithi, M., Jiale, M. & Rahman, M.N. (2025). Effect of integrated reporting quality disclosure on cost of equity capital in developed markets: Exploring the moderating role of corporate governance quality. PLOS ONE, 20(6), e0324147.

Jones, M. J., & Shoemaker, P. A. (1994). Accounting narratives: A review of empirical studies of content and readability. *Journal of Accounting Literature, 13*, 142-184

Kothari, S. P. (2001). Capital markets research in accounting. *Journal of Accounting and Economics, 31*(1-3), 105-231

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

Lintner, J. (1965). The valuation of risk assets and the selection of risky investments in stock portfolios and capital budgets. *The Review of Economics and Statistics, 47*(1), 13-37

Loughran, T., & McDonald, B. (2011). When is a liability not a liability? Textual analysis, dictionaries, and 10-Ks. *The Journal of Finance, 66*(1), 35-65

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

Loughran, T., & McDonald, B. (2016). Textual analysis in accounting and finance: A survey. *Journal of Accounting Research, 54*(4), 1187-1230.

Muslu, V., Radhakrishnan, S., Subramanyam, K. R., & Lim, D. (2008). *Causes and consequences of forward-looking disclosures in the MD&A* (Working paper). University of Texas at Dallas & University of Southern California. https://shorturl.at/TeWvv

Neel, M. (2017). Accounting comparability and economic outcomes of mandatory IFRS adoption. *Contemporary Accounting Research, 34*(1), 658-690

Nobes, C. (1998). Towards a general model of the reasons for international differences in financial reporting. *Abacus, 34*(2), 162-187

Ohlson, J. A., & Juettner-Nauroth, B. E. (2005). Expected EPS and EPS growth as determinants of value. *Review of Accounting Studies, 10*(2-3), 349-365

Reverte, C. (2012). The impact of better corporate social responsibility disclosure on the cost of equity capital. *Corporate Social Responsibility and Environmental Management, 19*(5), 253-272

Rjiba, H., Saadi, S., Boubaker, S., & Ding, X. (2021). Annual report readability and the cost of equity capital. *Journal of Corporate Finance, 67*, 101902

Sharpe, W. F. (1964). Capital asset prices: A theory of market equilibrium under conditions of risk. *The Journal of Finance, 19*(3), 425-442

Sveriges Riksbank. (2020). *Långsiktiga effekter av pandemin på svensk ekonomi: Fördjupning i redogörelse för penningpolitiken 2020.* https://shorturl.at/ouzSj

Sveriges Riksbank. (2022) *Monetary policy report: September 2022*. https://shorturl.at/1FQwn

Telles, S. V., & Moreira, J. C. C. (2024). Quality or quantity? The impact on the cost of equity capital. *Revista Contabilidade & Finanças, 35*(95), 1-16

Verrecchia, R. E. (2001). Essays on disclosure. *Journal of Accounting and Economics, 32*(1-3), 97-180

Weisskopf, J.-P. (2024). Simplifying the complex: 10-K readability and asset structure. *International Journal of Hospitality Management, 119*, 103795

You, H., & Zhang, X. (2009). Financial reporting complexity and investor underreaction to 10-K information. *Review of Accounting Studies, 14*(4), 559-586

9. Appendix

Appendix 1. Examples of Dictionary Words by Tone Category (Loughran-McDonald)

Tone Category	Example of Words
Negative	*loss, termination, impairment, default, failure, restated, litigation, limitation, damages, bankruptcy.*
Uncertain	*approximately, risk, believe, assumptions, intangible, estimated, undetermined, potential, anticipate, estimate.*
Weak-Modal	*may, could, possible, might, uncertain, appears, nearly, sometimes, possibly, perhaps.*

Appendix 2. Robustness Results for H1: Separate PEG and OJN Model Estimates

Panel A: PEG Model

Variable	Coef.	Std. Err	t-value	p-value	95% Conf. Lower	95% Conf. Higher	Sig.
FOG	0.004	0.002	1.976	0.048	0.000	0.008	**
BETA	0.010	0.009	1.063	0.288	-0.001	0.029	
BTM	0.060	0.005	11.834	0.000	0.050	0.070	***
LEV	0.019	0.033	0.573	0.567	-0.046	0.083	
SIZE	-0.007	0.003	-2.499	0.013	-0.011	-0.001	**
g	0.105	0.007	15.153	0.000	0.091	0.118	***
lnWords	-0.013	0.010	-1.283	0.200	-0.032	0.007	

Mean dependent var	0.107	SD dependent var	0.067
Adj R-squared	0.658	N obs	547
F-test	81.531	Prob. > F	0.000
Year FE	Yes	Industry FE	Yes

Panel B: OJN Model

Variable	Coef.	Std. Err	t-value	p-value	95% Conf. Lower	95% Conf. Higher	Sig.
FOG	0.008	0.004	2.080	0.040	0.000	0.015	**
BETA	0.038	0.020	1.947	0.052	-0.000	0.076	*
BTM	0.102	0.023	4.416	0.000	0.051	0.147	***
LEV	-0.007	0.040	-0.174	0.862	-0.085	0.071	
SIZE	-0.007	0.004	-1.730	0.084	-0.014	-0.001	*
g	0.061	0.022	2.740	0.006	0.017	0.104	***
lnWords	-0.037	0.013	-2.822	0.005	-0.062	0.011	***

Mean dependent var	0.078	SD dependent var	0.103
Adj R-squared	0.508	N obs	547
F-test	9.515	Prob. > F	0.000
Year FE	Yes	Industry FE	Yes

*** $p < .01$, ** $p < .05$, * $p < .1$

Appendix 3. Robustness Results for H1: Alternative Cost of Equity Measurement Dates

Panel A: Implied Cost of Equity Measured on 15 April

Variable	Coef.	Std. Err	t-value	p-value	95% Conf. Lower	95% Conf. Higher	Sig.
FOG	0.006	0.003	2.095	0.036	0.000	0.012	**
BETA	0.020	0.015	1.342	0.180	-0.009	0.049	
BTM	0.093	0.012	8.045	0.000	0.070	0.115	***
LEV	-0.013	0.035	-0.361	0.718	-0.082	0.056	
SIZE	-0.001	0.003	-2.990	0.003	-0.016	-0.003	***
g	0.001	0.001	1.193	0.233	0.001	0.002	
lnWords	-0.018	0.012	-1.468	0.142	-0.042	0.006	

Mean dependent var	0.092	**SD dependent var**	0.078
Adj R-squared	0.446	**N obs**	547
F-test	16.012	**Prob. > F**	0.000
Year FE	Yes	**Industry FE**	Yes

Panel B: Implied Cost of Equity Measured on 15 May

Variable	Coef.	Std. Err	t-value	p-value	95% Conf. Lower	95% Conf. Higher	Sig.
FOG	0.006	0.003	1.915	0.056	-0.000	0.011	*
BETA	0.023	0.015	1.543	0.123	-0.006	0.053	
BTM	0.095	0.012	8.271	0.000	0.072	0.117	***
LEV	-0.032	0.035	-0.926	0.354	-0.101	0.036	
SIZE	-0.009	0.003	-2.903	0.004	-0.015	-0.003	***
g	0.001	0.001	1.354	0.176	-0.000	0.002	
lnWords	-0.018	0.012	-1.471	0.141	-0.041	0.006	

Mean dependent var	0.092	**SD dependent var**	0.079
Adj R-squared	0.455	**N obs**	547
F-test	14.54	**Prob. > F**	0.000
Year FE	Yes	**Industry FE**	Yes

**** p < .01, ** p < .05, * p < .1*

Appendix 4. Robustness Results for H2: Separate PEG and OJN Model Estimates

Panel A: PEG Model

Tone	Below median	Above median	Interaction	Chi-square	p-value	R^2 low	R^2 high
Panel A: Negative tone	0.001 (0.003)	0.017*** (0.003)	0.008** (0.003)	13.144***	0.000***	0.814	0.889
Panel B: Uncertain tone	0.002 (0.003)	0.004 (0.004)	-0.001 (0.003)	0.090	0.764	0.842	0.906
Panel C: Weak-modal tone	0.006 (0.004)	0.004* (0.002)	0.000 (0.003)	0.179	0.6722	0.846	0.909
Year FE	Yes						
Industry FE	Yes						

Panel B: OJN Model

Tone	LowTone	HighTone	Interaction	Chi-square	p-value	R^2 low	R^2 high
Panel A: Negative tone	-0.001 (0.004)	0.029*** (0.011)	0.012** (0.006)	7.010***	0.008***	0.701	0.682
Panel B: Uncertain tone	-0.002 (0.003)	0.013 (0.009)	0.002 (0.005)	2.224	0.136	0.760	0.704
Panel C: Weak-modal tone	0.005 (0.005)	0.009* (0.008)	0.001 (0.005)	0.148	0.701	0.695	0.743
Year FE	Yes						
Industry FE	Yes						

*Firm-clustered standard errors are reported in parentheses. *** p < .01, ** p < .05, * p < .1*

Appendix 5. Descriptive Statistics of the Moderating Tone Variables

Variable	N	Mean	SD	Min	P1	P25	Median	P75	P99	Max
NEG_share	547	0.0095	0.0019	0.0003	0.0052	0.0083	0.0096	0.0107	0.0135	0.0153
UNC_share	547	0.0093	0.0017	0.0006	0.0048	0.0084	0.0093	0.0104	0.0129	0.0150
WM_share	547	0.0019	0.0005	0.0000	0.0008	0.0016	0.0019	0.0022	0.0032	0.0037

Appendix 6. The Use of AI

For this thesis, we used generative artificial intelligence (AI) through ChatGPT (version 5.1), developed by OpenAI. The tool supported our Python coding, including data preparation and troubleshooting, and was also used to confirm our understanding of certain sources. In addition, ChatGPT assisted with grammar, spelling, and general text refinement by offering suggestions to improve clarity and academic tone. Overall, the use of AI enhanced the efficiency of the thesis process. However, using AI involved risks such as inaccurate suggestions or outdated information. To mitigate these risks, all outputs were critically reviewed, coding suggestions were tested by us, and all factual content was verified against credible academic sources.