Erasmus University Rotterdam

Erasmus School of Economics

Bsc Thesis- International Bachelor Economics and Business Economics

The Impact of Obscuring Firm Performance Through the Manipulation of Annual Report Readability

Abstract

This paper studies how the readability of annual reports relates to firm performance over the two years following their filing date. Furthermore, this paper looks into whether analysts' earnings forecasts are affected by low readability of annual reports. To understand these relationships, I look at 48,642 unique annual reports and quantify their reading ease. The impact of readability on firm performance and subsequently analyst forecasts is then examined. I find that firms whose annual reports score low on readability have on average lower earnings than firms whose annual reports score well. Further, I find that analysts' forecasts are impacted by the readability of annual reports. Analysts forecasting error increases and consensus amongst analysts decreases as annual reports become more difficult to read. Lastly, I find that low readability of annual reports results in analyst forecasts that are systematically lower than actual earnings, supporting the need for firms to produce clear disclosures.

Student: Youri Lempers

Student number: 445886

Supervisor: Sjoerd Van den Hauwe

Keywords: Readability, analyst forecasts, financial disclosures, earnings per share, forecasting error

Table of Contents

1. Introduction

If analysts see annual reports as important tools through which they can base their earnings forecasts, and if annual report readability impacts an analyst's ability to make accurate forecasts, then poorly performing firms are incentivized to manipulate their report readability. The purpose of annual reports is to lower the asymmetric information between investors and the firm. If firms wish to avoid being mis-valued by investors they can facilitate information transfer (Healy and Krishna, 2001). However, if firms expect poor performance, they may be less willing to provide analysts with the information they require (Lang & Lundholm, 1996). If written in such a way as to mask poor future expectations, then analysts may naturally provide more favorable forecasts.

In 2012 Jack Abramoff an American political lobbyist appeared on the tv show 60 minutes. Here he explained how he manipulated language to get laws passed in congress.

"What we did was we crafted language that was so obscure so confusing so uninformative but so precise to change the US code."

Jack Abramoff

This practice of obscure and uninformative language also lends itself to business jargon and annual reports. Below is a letter from the CEO of Enron, Jeffrey Schilling, as a foreword to Enron's 2000 annual report, just 9 months before Enron filed for bankruptcy.

"We have robust networks of strategic assets that we own or have contractual access to, which give us greater flexibility and speed to reliably deliver widespread logistical solutions….We have metamorphosed from an asset-based pipeline and power generating company to a marketing and logistic company whose biggest assets are its well-established business approach and its innovative people"

Jeffrey Schilling

This excerpt was presented by Fugere, Hardaway and Warshawsky (2005) as the introduction to the first chapter in their book "*Why Business People Speak Like Idiots: A Bullfighter's Guide"*. It is their first example of "indigestible" language with a lack of substance. Indeed, when read over it seems this text could have been greatly simplified to help communicate the intended meaning of the text to its audience. However, the intention may not have been to be as clear as possible, instead difficult to read text may serve its own purpose.

From 1990 till 1998 Enron saw only modest growth compared to the average S&P500 firm. However, Enron saw its stock share spike 56% in 1999 and 87% in 2000. Enron's market cap was 70 times earnings and 6 times its book value. However, Enron filed for bankruptcy just one year later. Unethical practices resulted in misrepresenting earnings and manipulating the balance sheet to indicate good performance (Healy & Palepu, 2003). Enron's financial statements were furthermore confusing for investors and analysts (Bratton, 2001). Language as used by Jeffrey Schilling in his letter to investors may thus have also been used to mask Enron's illicit behavior. In particular, low readability allows for positive language to have a greater impact on investor sentiment (Tan, Wang & Zhou, 2014).

Improved firm disclosure facilitates the roles of analysts (Lang & Lundholm, 1996). Obscuring the information content of annual reports could thus make analysts jobs more difficult. Analysts role is to bridge the gap in information between the firm and its investors. Hindering readability may thus impact the ability of analysts to fulfill this role. If the firm can paint a picture more positive than actual expectations, then analysts might have too rosy forecasts. This paper therefore asks the question whether analyst forecasts can be altered by strategically manipulating the readability of a firm's annual reports.

This paper extends current research in the following ways. Although there is past research on the relationship between readability and performance, these were conducted using readability measures that may not be accurate for financial documents (Loughran & McDonald, 2016). This paper aims to improve and use measurements of readability better suited to financial disclosure than past research. Past research has used readability measures that employ the length of a word and the length of a sentence to measure readability. However, there may be many more factors involved. Therefore, this paper uses guidelines published by the SEC and the US Department of Education to determine the factors that most hinder readability in business reports. Once identified I combine these measures into a single readability index, similar to the ones popular in past research. Furthermore, this paper focuses on some aspects of analyst forecasts which are not sufficiently researched. Past research has focused on readability's impact on analyst following (Lehavy, Li, & Merkley, 2011), or its impact on analyst dispersion (Alblanque, 2018). This paper however focuses on the accuracy of analyst forecasts and particularly whether obscuring information in annual reports leads to an overestimation of earnings per share by analysts. Lastly, this paper is one of view studies using a large sample size when assessing readability of 10-ks and firm performance, this allows for more generalizability of the study results. Small scale studies such as Leisurs (2018) find no relationship between readability and performance within very specific criteria, but it is not known whether this holds with regards to the larger market. The large sample may thus create new insight into whether readability is an indicator of future performance.

This paper is organized in the following way. First in the literature review I discuss relevant literature regarding financial disclosures and asymmetric information. I then go into readability, its measures, and its manipulation by public firms. I then discuss the purpose of analysts and how analyst's forecasts are affected by annual reports. After the literature review section, I present my sample selection method. Then in the methodology section I discuss which independent and control variables are used as well as their sources. In particular I go into detail on how I conceptualized readability. I then proceed to discuss the models I use to determine the impact of readability on firm performance and analyst forecasts. In the results section I present the results of the research into improving readability for business reports. Secondly, I present the findings of this paper regarding readability, performance, and analyst behavior. Finally, In the discussion I discuss my conclusions, limitations of the study, and suggestions for further research

2. Literature Review

The literature this paper will focus on stems from multiple areas of focus. Particularly asymmetric information between firms and their stockholders, the role of readability on 10-ks ability to transfer information, and the role of readability on analyst forecasts.

2.1 Asymmetric Information

Standard finance theory assumes that outside investors and firm managers have the same information about the firm (Modigliani and Miller, 1961). However, a more realistic assumption is that there is a fundamental asymmetry in the information possessed by the managers of a firm and the firm's shareholders (Saorín and Lopez, 2013; Miller and Rock, 1985). One of the important pillars of a well-functioning financial market is that information is readily available. The view in the 1990s and early 2000s was that the solution to the asymmetry problem was annual reports (Farvaque, Refait-Alexandre & Saidane, 2011).

To facilitate financial dissemination, firms publish financial disclosures such as annual reports. In the United States public companies are required to publish yearly 10-k disclosures which are closely regulated by the Securities and Exchange Comission (SEC). One of the reasons for strict regulation of financial disclosures is that firms might weigh the costs and benefits of adequate financial reporting to maximize profit (Beaver, 1998). Firms that are performing well would want to avoid misvaluation and are thus incentivized to publish clear financial disclosures (Healy and Krishna, 2001). However poorly performing firms may have an incentive to obscure the information content they are required to disclose. In doing so, managers hope to remove or delay this information from being considered in the stock price (Bloomfield, 2002; Merkl-Davies & Brennan, 2007; Adelberg, 1979). This ability of managers to manipulate the communication process prompted Courtis (1998) to develop the 'obfuscation hypothesis'. Simply stated this hypothesis believes that management is not a neutral agent in creating firm disclosures. Some managers would obfuscate their failures and highlight successes.

The annual report serves to paint a picture of the corporation and its financials of the previous year. If managers are not neutral agents, this picture might be a distortion of reality. As Hines (1988) says "we create a picture of an organization… and on the basis of that picture, people think and act, and by responding to that picture of reality they make it so". It can therefore work in a manager's favor to influence investors perception of the firm through firm disclosures if firm performance is poor (Lang & Lundholm, 1996).

2.2 Readability of Financial Disclosures.

In 1967 the SEC conducted a study dubbed the "Wheat Report" named after SEC Commissioner Frances M. Wheat. One of the conclusions from this study was that prospectuses were too complex and wordy for the average investor (Wheat, 1967). In response the SEC has taken multiple steps to encourage clear disclosure. In 1982 the SEC revised its current rules through a new integrated disclosure system that encouraged firms to submit clear and concise text to the SEC. In 1998 the SEC developed its *Plain English Handbook* encouraging firms to adhere to a list of factors which could affect the readability of the disclosure.

It is no surprise then that research into readability has increased. Until recently most research into readability and firm performance has been conducted on small samples of firms (Li, 2008). Scolaro (2014) used the annual reports of 40 European banks to determine whether disclosure readability affects stock returns and found no relation. Healy (1977) analyzed the reading ease of 50 New Zealand firms' footnotes to financial statements. On the other hand, few studies have looked at a larger overall sample to draw conclusions to the wider effect of readability. Ablanque (2018) used 7139 US firms in her study on equity compensation, readability, and analysts' forecasts. This paper aims to contribute to the research using a large sample size of annual reports.

The concept of readability itself is abstract, and the process of quantifying it varies. Dale and Chall (1949) found that definitions of readability were confusing, a text could be defined as readable if it was legible, interesting, or through ease of reading. However, these terms are not mutually exclusive. Dale and Chall define readability thus as "the sum total (including the interactions) of all those elements within a given piece of printed material that affects the success that a group of readers have with it". One of the most common measures of readability was developed by Robert Gunning in 1944. His Gunning-Fog index considers the average sentence length of a text and the average word length. McKee (1967) conducted a survey of businesses and found that the most common yardstick regarding readability was the Fog Index.

Current research continues to apply the fog index regarding financial disclosure (Li, 2008; Lehavy, Li and Markly, 2011; Lawrence, 2013). The fog index is a measure of sentence length and word length/complexity. However, these two factors are likely not the only determinant of readability. The fog index is therefore not the only proxy for readability that has been used to analyze 10-ks. Document length may also be a hindrance to readability (Li, 2008; You and Zhang, 2009; Leisurs, 2014). However, in regard to finance these measures are not enough to quantify readability (Loughran and McDonald, 2016). For example, according to the fog index words such as "management", "liabilities" and "depreciation" are considered complex and decrease the readability score of a text. These words, although long, are expected to be understood by the average reader. Therefore, the Fog index might not always be accurate. Furthermore, there are many other factors that can affect readability such as word choice, active or passive voice, and paragraph length.

2.2.1 Readability of Financial Disclosures: Current Performance

Much research has been done with regards to obscuring current firm performance. The benefit of obscuring firm performance would be that the market reaction to negative news would be delayed or even avoided (Merkl-Davies & Brennan, 2007). However, Benston (1973) finds that the number of cases of abuse are minimal. Furthermore, annual reports would provide no new information to the market and therefore not affect market price. On the other hand, Singhvi and Desai (1971) found that inadequate financial disclosure likely leads to high fluctuations in the market price of a security.
Several studies find that firms do take advantage of their ability to manipulate information disclosure. Scrand and Walther (2000) find that managers emphasize improvements in comparison with their own strategically created targets whilst obscuring other performance measures. Jonest et al (2017) Finds that European banks navigated the financial crises by omitting stock market performance graphs or placing them in non-prominent places. This omission led to a more positive impression of the banks. Rutherford (2003) found that poorly performing firms would release reports that were more complex. Rutherford states that this supports the ‘obfuscation hypothesis, that managers are not neutral when writing reports. However, there may be a directionality problem. If performing poorly is more difficult to explain in text, then firm disclosures will naturally be more complex without being an indication of manager obfuscation. In that regard the impact of readability on future performance avoids this problem. In addition, Li (2008) states several factors which would make the relationship between current performance and annual report readability dubious at best. Annual reports contain a lot of historical and financial information and data, therefore attempts at hiding current performance would be weak. In addition, if good current performance numbers are partly due to strategic manipulation there is less of an incentive to make annual reports more readable.

2.2.2 Readability of Financial Disclosures: Future Performance

Researching the impact on future performance also avoids the directionality problem encountered by Rutherford (2003) If a relationship exists between readability and future performance this can only stem from readability. If managers believe their current poor earnings will be persistent or that the firm will perform poorly, they could manipulate readability to give investors the opinion that the firm has good opportunities in the future. Li (2008) found that firms with readable disclosures had higher earnings as well as greater earnings persistence. Bloomfield (2008) further corroborated this idea. In his case study he found that obfuscation was likely not to hide poor current performance but instead the nature of said negative performance.

For these reasons I have come to the following hypothesis for my research.

H1: Annual reports that are complex to read are indicators of poor future performance.

2.3 Analyst Response to Annual Reports

Due to the large processing costs of financial information and the asymmetry of information between managers and investors there exists a market for information. Analysts can fill this gap by providing their services (Schipper, 1991). Analysts ensure quick dissemination of information from the firm to investors(Barth, Beaver, & Landsman, 1998) . As the readability of firm disclosures decreases there is an increasing demand for analysts (Lehavy, Li & Merkley ,2011) The abilities of financial analysts ensure that markets are efficient and improve investor confidence (Beyer et al, 2010;Schipper, 1991). Tan, Wang and Zhou (2014) Furthermore, the abilities of analysts are complimented by readable disclosure documents. More readable documents would mean a lower information processing cost for analysts; implicitly this would lead to higher stock returns (Ablanque, 2018; Hsieh, Hui and Zhang, 2015)

Prior literature shows analysts behavior and consequently their forecasts are affected by firm disclosure policies (Land & Lundholm, 1996; Previts, 1994). Increased disclosure readability facilitates analysts roles and reduces analyst needs for additional private information for their forecasts (Li, 2008; Asay, Elliott & Rennekamp, 2017). Tan, Wang and Zhou (2014) find that as readability decreases, language sentiment influences analyst judgements. However as analyst experience increases so does their forecasting accuracy (Clement, 1999). Tan, Wang and Zhou believe that as analysts gain experience they would be less prone to sentiment manipulation in financial disclosures.

Ablanque (2018) does find that forecast dispersion is negatively corelated with annual report readability. This implies that analysts are affected by the readability of the reports they use to build forecasts. However, less readable texts might increase the demand for analysts, and a greater collective could result in a lower forecast dispersion. Despite this, increasing processing cost would incentivize analysts to use other sources of information and potentially cause analysts to interpret the same information differently. Therefore, I come to the following hypotheses

H2: The reading ease of an annual report is negatively correlated with the variance amongst analyst forecasts.

H3: The reading ease of an annual report is negatively correlated with analyst forecast accuracy.

Therefore, when reading difficulty increases I expect to see lower consensus amongst analysts forecasts as well as a decrease in forecast accuracy.

As managers would only be incentivized to obscure negative information, firms with positive expectations will be more likely to have readable disclosures. Thus, accuracy and precision of forecasts would be worse for firms with poor expectations. As managers are not neutral agents, they may underscore their success while breezing through failures (Adelberg, 1979). As a result negative information is obscured and it may be logical that analysts tend to overpredict future earnings for these firms. This would mean that as the complexity of annual reports increases, analysts tend to provide forecasts that are more favorable than actual future performance. Tang, Wang and Zhou (2014) do find that amongst MBA students this is the case. As readability is lower, students were more likely to be influenced by framing effects. As a result, students overpredicted future earnings. For these reasons I come to my last hypothesis.

H4: The readability of an annual report is negatively correlated with the directionality of the analyst forecast error.

I thus expect that if annual report readability is low that the average error amongst forecasts has a higher probability to be positive than if readability was high.

3. Data

From WRDS SEC Analytics Suite I download the original sample size of all 10-K publications between 2002 and 2017. Here data imported includes number of words per document, sentences per document, paragraphs per document as well as sentiment measures provided by Loughran and McDonald. Those being weak and strong modal verbs as a proportion of the total text; as well as the proportion of litigious words.

In Table 1 I present the process by which firms were selected for my sample.

Reason for removal/ retrieval H1	**Firm Years Remaining**
All disclosures considered as a "10-k" on the WRDS SEC Analytics Suite database between the years of 2002 and 2017.	134884
Removing all 10-ks with less than 1000 words	133011
Removing all 10-ks where no Cusip code is provided	132859
Removing all 10-ks without a matching Cusip and Cik number and GVkey.	82724
Removing firms with no data on Debt/Equity or Book/Market	59271
Removing firms where 10-ks have an abnormally high Gunning Fog Index above 30	59253
Removing firms with paragraph numbers under 10 or words per paragraph over 2000	59076
Removing firms with no earnings data	48,642

Reasons for further removal H2,H3,H4	**Firm Years remaining**
Removing firms with no forecasts	14504
Removing firms with analyst following under 3	12250
Removing firms with no age data	6653

Table 1: Firm Selection. The left column indicates for what reason annual reports were omitted from the final sample. The right column indicates how many annual reports remain after the removal of the annual reports as per the corresponding reason in the left column.

As presented in Table 1, I start off with 134,884 firm years. I remove firms where data is missing. Further. I also remove some anomalies from my sample.

3.1 Readability Measures

The usage of the Gunning Fog Index has become the standard for analyzing financial disclosure readability and I therefore incorporate it in this study as one of my variables of interest. The Fog index relays the grade level a reader must have achieved to be able to comfortably read a text. It is calculated as follows

$$0.4[\left(\frac{Words}{Sentences}\right) + 100\left(\frac{Complex\,Words}{Words}\right)] \tag{1}$$

This is a function of the average number of words per sentence, and the percentage of words that are deemed to be complex. Complex words are words with three or more syllables. The presence of longer sentences and more complex words are considered variables that determine the complexity of an overall text. As sentence length and word complexity increases the Fog score would increase, indicating a higher-grade level requirement to be able to understand a text. As readability decreases the fog index increases.

A concept such as readability may be difficult to define using only one variable, or even two such as in the Fog Index. Readability is difficult to quantify and is likely determined through a mix of numerous variables unknown applied in weights unknown. Therefore, I look to literature published by the SEC and the Plain Writing act of 2010 to find the variables that are most likely to impact investors' ability to read and interpret financial disclosures. I follow Loughran and McDonald in looking at the SEC's *1998 Plain English Guidebook* for what the SEC determines to be most important to make disclosures readable. In particular, I look at the rules set out by the SEC as well as the provided examples of poor disclosure. Rule 421(b) is especially helpful as it provides a summary of the overall guidebook.

The selection of variables will primarily be based on the SEC's advice regarding written text although the guidebook also includes rules for charts, tables and lists. I contrast the *Plain English Guidebook* with the other SEC publication *Disclosure to Investors - A Reappraisal of Federal Administrative Practices Under the 1933 and 1934 Acts* as well as the *Plain Writing Act of 2010* published by the US department for Education. Furthermore, I look to past research on readability of financial documents such as Leisurs (2014) and You and Zhang (2009). Not all variables regarding text that reoccur in these documents were analyzed, in particular factors with infrequent reoccurrences such as the use of multiple negatives, and capitalization of common terms are omitted. Variables that are more difficult to quantify such as font choice are also omitted.

Definite, concrete everyday words: Both the SEC guidebook and the federal guidelines advocate for simple or concrete everyday words to be used throughout the text. As Robert Gunning (1969) did I take the length of a word to be an indication of its complexity, despite there being many words that may be long yet commonly used in a financial setting such as "management" and "consolidated". I accept that in general smaller words are more likely to be understood accurately. I use the number of syllables per word as a proxy for everyday words

Legalese: Rule 421(b) advocates for no legal jargon or highly technical business terms. The term “legal jargon” is especially prominent in the examples of poor writing presented at the end of the guidebook. The federal guidelines likewise dedicate a section to the removal of jargon. Loughran and McDonald (2016) developed a list of legal jargon specifically for financial disclosures. To measure the extent of legal jargon in texts I take the number of occurrences of these words as a proportion of the total number of words in the 10-k file. This data is obtained from WRDS SEC Analytics Suite.

Short Sentences: Rule 421(b) further asks for “short sentences whenever possible”. Likewise, the federal guidelines advocate for short sentences to avoid dense and wordy text. Shorter sentences promote readability by breaking up text into bite sized pieces. This variable is computed as the number of words of the 10-k document divided by the number of sentences to find the average sentence length.

Paragraph Length: Page 47 of the *Plain English Guidebook* is dedicated to the usage of short paragraphs, and the issue of “blocky” and “lengthy” paragraphs again seems prominent in the guidebook’s examples. The federal guidelines likewise promote short paragraphs stating that lengthy paragraphs discourage readers from attempting to understand the material. I express this variable as the natural logarithm of the average number of words per paragraph.

Active Voice: The Plain English Guidebook promotes usage of the active voice with usage of strong verbs. Weak verbs usually come with the passive voice as well as hidden verbs and can add confusion to a sentence. Modal verbs play an important role in writing in the active voice As a proxy for active voice I therefore use the proportion of strong and weak modal verbs present in a text as obtained through WRDS SEC Analytics Suite, which uses Loughran and Mcdonald(2016)’s word lists. Said lists contain a list of weak modal verbs or strong modal verbs respectively. The active voice component is calculated as follows

$$\ln\left(\frac{1 + WeakModal}{1 + StrModal}\right) \tag{5}$$

Where *WeakModal* is the number of weak modal verbs and *StrModal* is the number of strong modal verbs as a proportion of total word count in the 10-k files.

Average Filing Length: The SEC publication *Disclosure to Investors - A Reappraisal of Federal Administrative Practices Under the 1933 and 1934 Acts* recommends avoiding lengthy text to facilitate the average investor’s understanding of the filing. I follow Leisurs (2014) and You and Zhang (2009) In using word count as an estimate of the text size. An alternative option would be to use file size as it is less prone to calculation error (Ablanque, 2018). However as the SEC specifically mentions lengthy text I use word count as a variable instead, as file size may be distorted by images and tables.

3.2 Creating a Readability Index

The purpose then is to build one measure to indicate readability using the above-mentioned variables. Having one measure makes it easier to quantify what readability is and thus present a clear conclusion to the research questions. Loughran and McDonald(2016) calculated their *Plain English* variable as the sum of the standardized variables which they found to impact readability. This leads to each variable having the same weight which is unlikely to be very realistic as most of the variation in readability may come from a single variable as opposed to another.

To avoid linearly correlated variables I use principal component analysis (PCA) to determine the variable weights and create one readability measure. Principal component analysis is used to reduce the number of potentially correlated variables to explain most of the variation in the original data set (Abeyasekera, 2005). PCA analysis forms new variables which are uncorrelated and linear composites of the original variables, called principal components (Sharma, 1996). The first principal component explains the maximum amount of variation possible from the variables. The second principal component then explains the maximum variation of that which remains, and so on until all variation is explained. One of the requirements of PCA is that the variables are mean-corrected or standardized, as each original variable operates under a different scale. If this was not accounted for the weights given to the principal components would be affected by their respective distributions. Therefore, I standardize each variable ensuring that their mean is 0 with a variance of 1; after which I compute the principal components.

There is precedence of using the first principal component as an index for what the original variables are intended to show (Primpas et al, 2010). However in his applied multivariate techniques book , Sharma (1996) recommends selecting principal components with eigenvalues greater than 1. I thus use only principal components who adhere to this criteriaThe results of the PCA are presented at the start of the results section.

3.3 Earnings and Earnings Forecasts

For the dependent variable earnings, I use earnings data retrieved from Compustat. I retrieve the earnings from the year following the 10-k filing as well as for the year two years after the filing. Earnings are a measure of firm performance and are used by a wide range of users. Furthermore, earnings are universally used to determine executive bonus compensation, giving it credit as a performance measure (Dechow, 1994).

I collect earnings figures from Compustat. Here I collect earnings before interest and taxes and I scale it to book value to account for firm size. I collect earnings per share (diluted) from the same source. I divide the scaled earnings numbers into two variables EARNINGS+1 and EARNINGS+2. EARNINGS+1 represents the firm earnings one year after the annual report filing, and EARNINGS+2 represents the firm earnings two years after the annual report filing.

The most widely available forecast provided by analysts is earnings per share (EPS). Earnings per share forecasts are collected from the I/B/E/S database. Forecasts are only considered if they are published within one quarter following the filing date of the annual report. Furthermore, forecasts are only used if at least three analysts provided forecasts for the firm. This is done as no forecast dispersion is present with only one forecast and a minimum of three forecasts may give greater accuracy of dispersion. I use the forecasts for the date one year from the annual report. I call the earnings forecasts EPSFORECASTS and the variance of the forecasts EPSVARIANCE. I remove several outliers as they heavily skew the data. To compute the accuracy of the forecasts I take the absolute value of the difference between the average forecast for each filing and the actual earnings per share as reported by the firm one year after the date of the forecast.

3.4 Control Variables

In the following section I discuss control variables that can impact both firm performance and readability and in the second section I discuss control variables that can impact both analyst forecasts and readability. I have included the variable names in parenthesis next to the topics they are meant to represent.

3.4.1 Control variables: Firm Performance & Readability

Market-to-Book Value (M2B): Bloomfield (2008) suggested further research look into market to book values, stating that market to book ratios have robust associations with returns. Growth stocks may have riskier investments, and this may require more complex and longer disclosure. The market-to-book value is collected from the Financial Ratios Suite by WRDS.

Firm Size (SIZE): It is possible that larger firms will need to explain more and have more complex operations. Therefore, their readability could be affected. Furthermore, firm size is also a determinant of disclosure quality and quantity. Chow and Wong-Boren (1987) find that there is a positive relationship between firm size and voluntary disclosure quantity. Firm size is measured as the natural logarithm of total assets. Data on firm size comes from Compustat Fundamentals

Financial Leverage (D2E): In their study on readability and stock returns Hsieh et al (2016) use financial leverage as a control variable. This is to account for any difference in financial disclosure as a result of a different financing structure. I collect the debt to equity ratio for each firm from the WRDS Financial Ratios Suite.

Industry (INDUSTRY)*:* Certain industries may require more complex disclosure. Greenspan (2001) for example reported that bank disclosures require more effort to be understood. Ablanque (2018) purposely omits firm data for pharmaceutical and bioinformatics companies as they require more complex disclosures. I therefore include dummy variables to represent industries. This is done through the SIC codes which are four-digit industry classification codes. I use the divisions structure of the SIC as a categorical variable. Only the Public Administration division is split into two for the purpose of this paper, Public Administration and Non-classifiable Establishments. The SIC codes are retrieved from Compustat Fundamentals.

Firm Year (YEAR): Lastly, I control for the year using yearly dummies. This is to account for any changes in firm performance due to the general market as well as any potential changes in reporting practices. The year is the year in which the annual report filing date occurs. This information is retrieved from WRDS SEC Analytics Suite.

3.4.2 Control Variables: Analyst Forecasts and Readability.

Most variables that were used to control for earnings and readability are also relevant to analyst forecasts with some changes and additions. Below I will note all additions as well as additional reasoning for some variables where needed.

*Directionality of Net Income(D*NI): It may take longer and naturally be more complex to explain losses than to explain positive earnings. Furthermore Ablanque (2018) states that positive earnings disclosures tend to result in lower analyst forecast dispersion than negative disclosures. I therefore use a dummy variable for whether net income is positive or negative e to account for the directionality of income.

Analyst Following(FOLLOWING): A large analyst following may lead to herding behavior, whereby individuals are likely to be influenced by the group. This is especially true for inexperienced analysts who tend to deviate less from consensus earnings forecasts (Hong, Kubik & Solomon, 2000). The variable is equal to the number of analysts in the quarter following the annual report release. The data required to compute analyst following is retrieved from I/B/E/S.

Firm Size (SIZE)*:* Large firms tend to have greater analyst following as well as providing a friendlier environment for analysts (Lang & Lundholm, 1996). Firm size is measured as the natural logarithm of the total assets reported by Compustat.

Market-to-Book Value (M2B) : Li (2008) states that firms with a high market-to-book value have significantly different investment opportunities and options for growth. Analysts may find firms with considerable investment opportunities more attractive. Analysts may also find it more difficult to value high growth stocks (Lehavy, Li & Merkley, 2011). I use market-to-book value as a proxy for growth. The data is retrieved from the Financial Ratios Suite by WRDS.

Firm Age (AGE): Information for older firms may be more available requiring less time and effort to understand. Analysts would thus have an easier time developing forecasts for established companies. Furthermore, there may also be a relationship between firm age and performance. Loderer and Waelchli (2010) find that as firms age profitability tends to decrease. I use time since the firm's IPO date as a proxy for age. This data is gathered from Compustat.

Industry and debt-to-equity ratio are also control variables used for the analyst forecast and readability regressions. After collecting all the data I use the provided GVkeys, Cusips, and Central Index Keys combined with the date from which the data originates to match each variable to the corresponding annual report.

4. Methodology

4.1 Regressions

To answer hypothesis 1 I run the following regressions:

$$EARNINGS+1=\alpha_0+\beta_1 Readability+\ \beta_2 M2B+\beta_3 SIZE+\beta_4 D2E+\beta_5 INDUSTRY+\beta_6 YEAR+\varepsilon \tag{3}$$

$$EARNINGS+2=\alpha_0+\beta_1 Readability+\ \beta_2 M2B+\beta_3 SIZE+\beta_4 D2E+\beta_5 INDUSTRY+\beta_6 YEAR+\varepsilon \tag{4}$$

In the first regression, EARNINGS+1 represents the earnings of the firm one year after the 10-k filing. In the second regression EARNINGS+2 represents the earnings of the firm one year after the 10-k filing. The Readability variable in each regression represents the readability of the annual report. I use two measures for this, for each regression I run it once using the established FOG method. I use this to check whether the results remain the same regardless of how I define readability. In the follow up I instead use the readability index that I calculated in section 3.2. These regressions should provide an answer on whether readability provides information on whether readability is an indicator of future performance. In both the Fog and the readability index an increase implies a more difficult text. Therefore, hypothesis one assumes that an increase in the readability measure would result in lower earnings.

To answer hypothesis 2 I run the following regressions

$$\begin{aligned}VARIANCE=\alpha_0+\beta_1 Readability+\beta_2 DNI+\beta_3 FOLLOWING+\ \beta_4 M2B+\beta_5 SIZE+\beta_6 D2E+\beta_7 Age\\+\beta_8 Readability*Age+\beta_9 INDUSTRY+\beta_{10} YEAR+\varepsilon\end{aligned} \tag{5}$$

In this regressions VARIANCE corresponds to the variance in the analyst forecasts. This is a measure of the consensus amongst analysts. As consensus decreases analyst forecasts will be more dispersed so the variance amongst forecasts increases. Again, I use the standard Fog as well as the Readability Index as independent variables in two separate regressions. I include an interaction effect between readability and firm age. Clarke and Shastri (2000) state that there is less information asymmetry between investors and firms when the firm is older. Therefore there may be less of a benefit to obscure information. I thus expect to see that a higher readability has less of an impact on forecast variance as the firm ages. Hypothesis 2 predicts that as readability becomes more difficult, that the variance amongst analysts would increase. Thus, an increase in the readability measure would result in an increase in the variance.

To answer hypothesis 3 I run the following regressions

$$|EPS - EPSFORECASTS| = \alpha_0 + \beta_1 Readability + \beta_2 DNI + \beta_3 FOLLOWING + \beta_4 M2B + \beta_5 SIZE + \beta_6 D2E + \beta_7 INDUSTRY + \beta_8 Age + \beta_9 YEAR + \varepsilon \quad (6)$$

Here |EPS-EPSFORECASTS| is the average distance between analyst forecasts and the actual earnings per share reported by the firm. This would thus be the average forecast error as a result of the readability. Again, this regression will be run twice, once using Fog and once using my Readability Index. Hypotheses 3 would predict that as the difficulty in readability increases, so would the difference between actual earnings per share and the forecasts.

For hypothesis 4 The dependent variable, the direction of forecast error, can only take two values, a positive error or a negative error. A positive error would imply that the average analyst forecast would be higher than actual earnings per share. A negative error would imply that the average forecast would be lower than actual earnings per share. For this reason, I utilize a probit model. A probit model is used to determine with what probability an observation will fall into a specific category. In this case whether the forecast has a positive or negative error. The regression is placed below.

$$P_i = \Phi(\alpha_0 + \beta_1 Readability + \beta_2 DNI + \beta_3 FOLLOWING + \beta_4 M2B + \beta_5 SIZE + \beta_6 D2E + \beta_7 INDUSTRY + \beta_8 Age + \beta_9 YEAR + \varepsilon) \quad (7)$$

Here P_i is the probability that the forecasting error is positive. In this regression a one unit increase in a regressor results in a β_i change in the z-score of the dependent variable. Φ is the Cumulative Distribution Function. This regression would test how an increase in the complexity of an annual report would impact the probability of the average analyst forecast being higher than the actual earnings per share. The model includes year dummies to control for year effects, as well as industry dummies to control for unique industry factors. ε is the error term and assumed to be normally distributed.

With regards to the error term in these regressions certain assumptions are required of an ordinary least squares regression. Firstly, the mean of the residuals should be zero. An analysis of the errors concludes that the means of the errors are indeed zero. Further to prevent potential heteroskedasticity I use robust standard errors.

4.2 Descriptive statistics

In table 2 I provide the descriptive statistics for each variable.

Table 2: Descriptive Statistics

Variable	*Mean*	*1st*	*25th*	*75th*	*99th*	*Standard Deviation*
FOG	20.06	17.87	19.41	20.62	23.41	1.06
ReadabilityINdex	3.92	2.11	3.45	7.39	6.02	0.78
EPSFORECASTS Error	1.25	.01	0.16	1.14	13.91	3.47
EPSVARIANCE	0.58	0.11	0.19	0.62	2.28	2.86
EPS	1.25	-6.87	-0.15	1.71	8.67	3.22
EARNINGS+1	0.12	-2.69	0.00	0.28	2.40	0.96
EARNINGS+2	0.14	-3.08	0.00	0.31	2.72	1.38
M2B	1.08	0.30	1.18	1.37	3.45	0.78
SIZE	6.52	2.06	5.00	7.87	11.81	2.14
D2E	2.79	-9.26	0.43	2.79	19.97	6.13
DNI	0.75	0.00	0.00	1.00	1.00	0.45
FOLLOWING	16.65	3..00	6	22	77	4.05

This table provides the descriptive statistics for the variables measured. The mean value as well as the value at the indicated percentile is presented. Lastly standard deviation is also included.

The FOG readability index shows that all annual reports are at a high difficulty reading level. The bottom percent of FOG would be considered university level, whilst the average annual report requires 20.07 years of formal schooling to be understood. From the Readability Index measure it becomes clear that a majority of the data is centered on the lower end of the scale with a few texts being significantly more difficult to read than the average.

EPS forecasts are on average equal to actual EPS. .However the EPS at percentiles 1st and 25th are lower than forecasts whilst the EPS for percentiles 75th and 99th are higher than the forecasts. Analysts may thus be wary of extreme evaluations. There is also a large difference in number of analysts following a firm, indicating that some firms are more attractive as shown by Lehavy, Li and Merkley (2011). The earnings in the year following the report and two years after the report are very similar, as expected. This would be the case as many of the years would overlap, only the years 2003 and 2019 would be mutually exclusive to EARNINGS+1 and EARNINGS+2 respectively. From DNI it is clear that the average firm has positive earnings.

5. Results

5.1 Principal Component Analysis

The eigenvalues, and the cumulative percentage of variance explained by each component are listed in Table 3.

Table 3: Principal Components and their proportion of variance explained

PRINCIPAL COMPONENT	EIGENVALUE	PROPORTION	CUMULATIVE PROPORTION
COMP1	2.016	0.336	0.336
COMP2	1.403	0.234	0.570
COMP3	0.937	0.156	0.726
COMP4	0.668	0.111	0.837
COMP5	0.573	0.096	0.933
COMP6	0.402	0.067	1.000

. Column one presents the principal components. Column two provides the eigenvalue of the component. Column three presents the proportion of variation amongst input variables that is explained by the principal component. Column four presents the cumulative proportion of variance explained if you account for all previous components.

The choice of which components to retain for the final measure is dependent on how much information one wants to retain and is thus subjective. I follow the eigenvalue-greater-than-one rule as suggested by Sharma (1996). I therefore select all principal components with an eigenvalue greater than one. The reasoning behind this rule is that for standardized data the amount of variance accounted for by each principal component should be at least equal to that of one of the original variables. Because components 1 and 2 have eigenvalues greater than one I combine both factors to create my readability measure. This should explain 57% of the variation between the variables.

Table 3 presents the composition of the principal components. This aids in understanding what each component represents, as they are linear functions of the input variables. Furthermore, the makeup of the components determines the composition of the final readability measure.

Table 4: Composition of the Principal Components

Variable	Comp1	Comp2	Comp3	Comp4	Comp5	Comp6
Word Length	-.106	0.686	0.067	0.616	0.2.16	-.296
Legalese	0.461	-0.035	-0.611	0.068	0.587	0.257
Sentence Length	0.593	-0.051	0.057	-0.203	-0.045	-.774
Paragraph Length	0.383	-0.166	0.766	0.186	0.347	0.291
Active Voice	0.079	0.671	0.154	-0.686	0.093	0.202
Filing Length	0.522	0.221	-0.096	0.264	-0.692	0.349

Column 1 provides the input variables for the PCA. The following columns correspond to the principal components. Each column provides the weight attributed to each input variable in the composition of the principal component.

Although principal component analysis avoids multicollinearity another problem that occurs is the difficulty in interpreting the exact meaning of each principal component, and hence the final readability measure. Therefore, I look to Table 4 to draw meaning. Component 1 is primarily made up off factors that hinder reading by making it tedious through length. In particular Sentence Length and Filing Length have the largest impact on component 1, 0.522 and 0.593 respectively. Component 2 focuses more on the syntax and word choice, hindering readability not by tiring out the reader, but my complicating direct understanding of the text. This is seen as Word Length and Active Voice account for the largest makeup of the component.

The makeup of the final readability component unscaled is an addition of the two retained principal components. The measure is provided in equation (8) below.

$$Readability\ Index = 0.580WordLength + 0.426Legalese + 0.542SentenceLength + 0.217ParagraphLength + 0.750ActiveVoice + 0.743FilingLength \tag{8}$$

As in the FOG index, word length and sentence length are large factors in determining readability. However, the length of the report and the utilization of active voice provide a greater impact. To allow for easier interpretation I rescale the readability measure to be between 0 and 10. A 10 would be the most difficult report to read and a 0 the easiest. This would coincide with the FOG index to the extent that higher numbers imply greater difficulty in reading comprehension.

5.2 Regressions

The first hypothesis tested in this paper states that annual reports that are complex to read are indicators of poor future performance. Table 5 below tests this relationship between readability and the earnings one year ahead and two years ahead.

Table 5: Regressions of Readability on Earnings

Dependent Variable	*Earnings 1 Year Ahead*	*Earnings 1 Year Ahead*	*Earnings 2 Years Ahead*	*Earnings 2 Years Ahead*
Variable of Interest	[1] Readability Index	[2] FOG	[3] Readability Index	[4] FOG
Independent Variables				
Readability Index	-0.0761(-11.08)***		-0.0735(-8.28)***	
FOG		-0.0379(-8.36)***		-0.0391(-7.39)***
M2B	-0.0361(-5.49)***	-0.034(-5.24)***	-0.0189(-1.93)**	-0.0169(-1.72)*
Size	0.0754(21.27)***	0.7071(20.95)***	0.0823(15.01)***	0.0780(14.87)***
D2E	0.0384(11.32)***	0.0387(11.44)***	0.0408(6.57)***	0.0411(6.63)***
Industry dummies	Yes	Yes	Yes	Yes
Year dummies	Yes	Yes	Yes	Yes
Cons	-.3803(.99)	0.4737(5.55)***	-0.077(-1.7)*	0.4646(4.78)***
R^2	0.0916	0.0903	0.0508	0.0503
Observations	48642	48642	48642	48642

This table shows the result for the regressions where Earnings are the dependent variable and readability is the variable of interest. Regressions [1] and [3] use the Readability Index as a measure of readability and regressions [2] and [4] use FOG.
The t-statistics are provided within brackets. The significance is indicated as follows. * Significant at the 0.1 level, ** Significant at the 0.05 level, *** Significant at the 0.01 level

Regression [1] shows that the readability index I calculated is statistically significant at the 0.01 level. Firms with more complex to read annual reports tend to have lower earnings in the following year. This provides evidence that annual report readability indicates future firm performance. Interesting to note is that when comparing the impact of the Fog in regression [2] and [4], the impact of readability is greater two years after the publication of the annual report as opposed to one year after. However, with the Readability Index the impact on earnings decreases slightly. In both cases the t-score drops from the first to the second year, implying a lower significance of readability on firm performance in the second year as opposed to the first. In all 4 of the regressions the readability measure is significant at the 0.01 level. Although the t-score drops when using FOG compared to the Readability Index (from -11.08 to -8.36; from -8.28 to -7.39), both are still significant at the 0.01 level. Therefore, regardless of how readability has been designed a more difficult to read annual report tends to prelude poor firm performance. Therefore, I cannot reject hypothesis 1, that annual reports that are complex to read are indicators of future poor performance. However due to the very high t-scores, I conduct a robustness check at the end of this section to confirm these results.

In regressions [1] and [2] all control variables are significant at the 0.01 level. However, one year later the impact of these variables tends to decrease. Regression [3] shows that the market to book value goes from significant at the 0.01 level to the 0.05 level. This drop is more extreme in regression [4], where the market to book value drops from significance at the 0.01 level to the 0.1 level. All other control variables despite dropping their significance remain significant at the 0.01 for both regressions [3] and [4].

Hypothesis 2 states that the readability of an annual report is negatively correlated with analyst forecast accuracy. This would imply that as annual reports become more complex, the variance in analyst forecasts increases. In Table 6, regressions [5] and [6] are directed towards this hypothesis.

Table 6: Readability and Analyst Forecasts

Dependent Variable	*Variance in Forecasts*	*Variance in Forecasts*	*Forecast Error*	*Forecast Error*	*Direction of Forecast Error*	*Direction of Forecast Error*
Variable of Interest	[5] Readability Index	[6] FOG	[7] Readability Index	[8] FOG	[9] Readability Index	[10] FOG
Independent Variables						
Readability Index	0.9501(1.05)		0.1342(2.39)**		-0.1041(-3.74)***	
FOG		0.9249(1.07)		0.0994(3.17)***		-0.0329(-2.01)***
Following	0.0057(2.16)**	0.0068(2.16)**	0.0133(2.91)***	0.1352(2.97)***	-0.0002(-0.16)	-0.0004(-0.31)
DNI	-0.1096(-1.98)**	-0.1102(-1.94)*	-1.291(-13.49)***	-1.2864(-13.71)***	-0.1998(-4.85)***	-0.1801(-4.42)***
M2B	0.5147(1.02)	.5250(1.02)	0.1177(1.58)	0.1193(1.6)	0.1659(3.98)***	0.1693(4.05)***
Size	-.1048(-1.67)*	-0.0985(-1.70)*	0.1614(4.14)***	0.1652(4.26)***	0.1123(8.46)***	0.1076(8.16)***
D2E	-0.0020(0.79)	0.0021(0.76)	-0.0002(-0.04)	-0.0002(-0.03)	0.0520(2.05)**	0.0054(2.11)**
Firm Age	0.187(0.97)	.8940(1.03)	0.0090(1.73)*	0.0077(1.49)	-0.0018(-0.86)	-0.0003(-0.16)
ReadabilityIndex Firm Age*	-0.4457(-1.00)					
FOG Firm Age*		-0.4435(-1.03)				
Industry Dummies	Yes	Yes	Yes	Yes	Yes	Yes
Year Dummies	Yes	Yes	Yes	Yes	Yes	Yes
Cons	-2.6608 (-0.68)	-17.3474 (-0.99)	1.08(0.234)	-0.3461(-0.32)	0.5756(1.39)	0.7918(0.128)
R^2	0.0043	0.0062	0.0576	0.0580	0.0329	0.0317
Observations	6653	6653	6653	6653	6653	6653

This table shows the results for the regressions where the dependent variables are forecast characteristics and the variable of interest is readability. Regressions [5],[7], and [9] use the Readability Index as a measure of readability, regressions [6],[8], and [10] use FOG. Regressions [5] and [6] test the relationship between readability and the variance between analyst forecasts. Regressions [7] and [8] test the relationship between readability and the magnitude of the forecasting error. Regressions [9] and [10] test the relationship between readability and the directionality of the forecasting error. A positive(negative) value here indicated that the forecast is greater(less) than actual earnings per share.

The t-statistics are provided within brackets. The significance is indicated as follows. * Significant at the 0.1 level, ** Significant at the 0.05 level, *** Significant at the 0.01 level

Regression [5] uses the Readability Index as a measure of readability. As expected, the direction of the effect of readability on variance is positive. However, no conclusions can be drawn as its coefficient is not significant. Regression [6] likewise finds that FOG is insignificant. Therefore, I don’t reject the null that the variance of forecasts is not affected by readability.

Contrary to what was expected larger analysts' followings result in higher variances. This is significant at the 0.05 level in regression [5] and [6]. Herding behavior would imply that analysts are more likely to seek similar estimations. The directionality of income is also significant at the 0.05 level. When annual reports are published in years that the net income was positive, the variance of forecasts tend to be lower. This may imply that forecasting becomes more difficult when firm income is negative.. The size of the firm is significant in both regression [5] and [6]. As the size of the firm increases variance of forecasts tend to decrease. This supports Lang and Lundholm (1996)'s notion that large firms provide friendlier environments for analysts.

Hypothesis 3 states that the readability of an annual report is negatively correlated with the analyst forecasting error. This means that the easier an annual report is to read the lower the forecasting error. Better readability would facilitate analysts in making their forecasts. Table 6 regressions [7] and [8] are used to test this hypothesis.

Regression [7] finds that as the readability index increases so does the forecast error. Therefore, more complex annual reports do tend to result in analysts being less accurate in their forecasts. The coefficient for the readability index is significant at the 0.05 level. An increase in the readability index of one-point leads to an increase in the error by 13 cents. Likewise regression [8] finds that more complex reports lead to greater forecasting error. Here the coefficient of the fog index is significant at the 0.01 level, an increase of 1 in the fog index leads to an increase in the error by 10 cents. This implies that a text that requires one extra year of schooling to understand results in an error of 10 cents. I therefore reject the null that more complex reports do not increase the forecast error.

The number of analysts is significant at the 0.01 level for regression [7] and [8]. As the number of analysts increases the difference between actual earnings per share and the forecasted earnings per share increases. Furthermore, the direction of net income in the year the annual report was released is also significant in both regressions at the 0.01 level. Again, negative earnings result in more difficult to forecast earnings per share, as the forecasting error increases. Size of the firm is also significant at the 0.01 level in both regressions. As the size of the firm increases the forecasting error decreases, further supporting Lang and Lundholm (1996).

The final hypothesis states that the readability of an annual report is negatively correlated with the directionality of the analyst forecast error. So if an annual report was complex to read the likelihood of the forecast being greater than the actual earnings per share would be greater. Table 6 has regressions [9] and [10] to test this hypothesis. These regressions are probit models, and positive coefficients represents an increased likelihood that the forecast is greater than the actual earnings per share. Regression [9] uses the Readability Index as its measure of readability. Contrary to the hypothesis, this regression finds that more complex reports tend to result in forecasts lower than actual earnings per share significant at the 0.01 level. Regression [10] makes the same conclusions while using FOG as a measure of readability. Therefore, I do not reject the null hypothesis, in fact these regressions provide evidence that greater readability increases the likelihood that analysts provide favorable forecasts. Contrary to past regressions the number of analysts is not significant in either of the two regressions. The direction of net income is significant. If net income was positive, then forecasts for next year tend to be lower than actual earnings per share in that year. Furthermore, the market-to-book ratio is significant at the 0.01 level in both regressions. If the market-to-book ratio increases the likelihood of an overestimation of eps increases. Significant at the 0.05 level, debt-to-equity increases also increase the likelihood of an overestimation.

5.3 Robustness Testing

5.4

Using two measures of readability aids in determining whether readability is impactful regardless of how its measured. However, this does not mean that the effects of readability are robust to changes in the model. Drawing inferences from a single precisely defined model ignores the many other accurate representations that can alternatively be used as a basis for data analysis. Leamer and Leonard (1983) believe the task of an econometrician should be to analyze a range of inferences generated by varying models, instead of the unique inferences drawn by a single model.

Therefore, I adapt the model specifications of my first model which tested the relationship between readability and performance and found exceptionally high t-values for readability. Firstly I challenge some of the assumptions I made in the process of creating my baseline model. In the theoretical framework I dived into how current performance was likely not a cause for obscuring annual reports, and therefore did not see it relevant enough to put into my regression model, as it should thus not impact how readability influences firm performance. However, there may be an alternative reason for why current performance is relevant. Negative information can be more difficult to disclose. If a firm performs poorly their annual reports may be more difficult to read regardless of whether this is done on purpose to obscure information or not. As a proxy for whether a firm performs poorly, I use the directionality of earnings DNI. This variable takes on a 1 if earnings were positive or a 0 if current earnings were negative in the filing year.

Next I also account for firm age. Older firms may find less success obscuring information, this is as there is more information available from an established company (Clarke & Shastri, 2000) Therefore, as a firm gets older there should be less of an incentive for firms to obscure negative information.
The results are presented in Table 7 (Appendix A). The coefficients of readability in both regressions decrease. In regression [11] the coefficient of readability is -0.0265. This is a large decrease in

comparison with regression [1] where the coefficient was -0.0761. However it is still significant at the 0.05 level. The coefficient for FOG in regression [12] is also lower having dropped from -0.0379 in regression [2] to -0.0211. It remains significant at the 0.01 level. The t-values drop in both regressions compared to regressions [1] and [2]. In regression [11] the Readability Index's t-value is -2.51, in regression [12[the Fog's t-value is -3.03.

Next I check for multicollinearity in regressions [1] till [12] using VIFs (Appendix B). I find that the interaction terms in regressions [5] and [6] have VIFs higher than 5. As this is an interaction term it is likely that there is multicollinearity between the interaction terms and their components, therefore I do not take corrective measures. All other variables report VIFs lower than 5 and so no corrective measures are taken.

6. Discussion

6.1 Discussion of Results

The aim of this paper was to find whether the readability of annual reports is an indicator of future performance and analyst forecasts. With regards to the impact of readability on future performance I find that readability is a significant indicator of future performance. More complex annual reports indicate poor performance in at least the two years following the annual report. This is a similar conclusion as was drawn by Li (2008) who used the Gunning-Fog Index as a measure of readability. He found that earnings are more persistent the easier annual reports are to read. This also supports Bloomfield (2008) who stated that annual report obfuscation was likely sought after to obscure negative information regarding future performance of the firm.

My paper also addresses whether annual report readability can be used to manipulate analyst forecasts. In particular, I look at three key characteristics of analyst forecasts. These characteristics are the variance in analyst forecasts, the accuracy of forecasts, and the direction of the forecast error. A higher variance in forecasts would indicate that there is less consensus between analysts. This could be due to a more difficult time interpreting information and thus differences in forecasts. A higher accuracy in forecasts despite difficult readability would indicate that analyst forecasts accuracy is not affected by the difficulty in reading annual reports. Either annual reports would have little impact on analyst forecast, or readability itself is a non-factor. The direction of the forecast error is the third characteristic that is tested. Manager obfuscation of annual reports is first and foremost a way to mitigate negative information. Therefore, it would be expected that the greater the difficulty to read an annual report the likelier an analyst would overestimate future performance.

Although the coefficients of readability imply that variance amongst analyst forecasts increases as reading difficulty increases, the coefficient are not significant. Therefore I cannot conclude that analyst consensus is affected by readability.

Forecast accuracy also tends to decline as annual report readability becomes more complex. This indicates that analysts face greater difficulty accurately creating their forecasts as a result of difficulty reading annual reports. However, I do not find that these forecasts to be systematically higher than actual performance. In fact, the worse the readability the more likely that a forecast is lower than actual earnings per share in the following year. Therefore, despite forecasts being less accurate, mitigation of negative information may not be accomplished by annual report obfuscation. Instead analysts are more likely to provide forecasts that are lower than actual performance.

One possibility is that analysts are not fooled by the difficult readability. Although obscuring information makes forecasting more difficult, analysts' confidence of firm performance decreases with reading difficulty. Analysts may question the credibility of annual reports when positive expectations are not credible and poor performance is obscured through low readability. To analysts, low readability may therefore be an indicator that firm future expectations are low. Therefore, analysts would attach a penalty to the firm for its low readability.

In Tan, Wang and Zhou (2014)'s study on MBA students they find that overall, low readability results in students being more prone to positive sentiment in text and therefore overestimate firm performance. However, they also found that the more experienced MBA students were less likely to be fooled by positive sentiment in combination with low readability. I then find similar results as Tan Wang and Zhou. As on average analysts can be expected to be more experienced than students, they would be more likely to see through annual report obfuscation, and therefore not overestimate earnings.

The question then remains, why if analysts can see through obfuscation would the penalty overcompensate for poor performance, why do we still see analyst forecasts being predominantly lower than actual firm performance? If the penalty was equal to the amount indicative of actual firm performance due to obscuring information, then we would not see analysts systematically under predicting firm performance. Analysts must then believe that the extent to which firms expect poor performance is greater than it actually is.

Analysts may believe that certain firms with poor expectations would not obscure their disclosures as the penalty attached to doing so would lower forecasts below actual expectations. However, analysts may have a wrong view of which firms obscure information. If they expect firms to take this penalty into account when deciding to obscure information they would expect only firms with negative expectations better than the size of the penalty to obscure information. As all firms who do not meet this criterion would not be incentivized to obscure information only firms with negative expectations greater than the forecasting penalty would remain, thereby increasing the penalty further. Eventually only firms with the worst expectations of future performance would remain and obscure information.

If firms do not consider this forecasting penalty or believe they can fool analysts and avoid such a penalty then firms will obscure information even when the penalty attached to doing so would leave forecasts lower than that without obscuring information. As a result, we would see analyst forecasts being systematically lower than actual performance. It would therefore imply that it would be in a firm's best interest to relay information as accurately as possible regardless of its expectations of future performance.

This paper does provide evidence that managers may choose to manipulate readability when their expectations of future performance are poor. This coincides with Bloomfield (2002)'s notion that managers might try and delay or stop the impact of negative information. However, with regards to forecasts, despite resulting in greater dispersion and less accurate forecasts, obscuring negative information does not result in forecasts that are greater than actual performance. This paper finds that analysts are more likely to provide forecasts underestimating actual performance in the case of information obfuscation.

6.2 Discussion of Contributions and Limitations

This study addresses many limitations other studies faced. The first is the small sample size identified by Li(2008). A majority of research on annual report readability was conducted with a small sample of firms. This study uses a large number of firms to create greater generalizability. Another limitation identified in past literature was the lack of a readability measure designed specifically for business reports (Loughran & McDonald, 2006). Past papers have used only general readability formulas such as the Gunning-fog index, or overall length of the paper. By analyzing legibility guidelines published by the SEC as well as the US department of education, I was able to identify some of the key features that would make business reports more difficult to read. I then combined them into one readability index. Using this readability index tailored to business reports I was able to corroborate past research on the impact of readability on future performance.

Furthermore, there was limited research regarding the impact of readability on analyst forecasts. Ablanque (2018) found that analyst dispersion would tend to increase as annual reports got more difficult to read. This paper has similar findings but also found that manipulating readability did not result in more favorable forecasts, instead the worse the readability the greater the likelihood of forecasts being below actual performance.

This paper however does come with its own limitations. A large sample size may aid in generalizability, however all firms analyzed in this paper are firms from the United States. Firms abroad may face different reporting requirements and economic environments. The applicability of readability thus still needs to be tested. A large sample size also does not consider the nuances that separate sectors face. With regards to the methods of defining readability, there are many ways to do so. This paper utilizes two, but better readability methods may be available or can still be developed to be better applicable to business reports. In particular the SEC guidelines mention several more factors that hinder readability that were not tested in this paper such as double negatives, font choice, and the usage of personal pronouns. Furthermore, the SEC guidelines recommend using graphs and tables to improve readability.

As of yet, readability measures only regard text in their analysis. Therefore, characteristics such as the use of graphs and tables, as well as formatting and font choice and size are not represented.

This paper also found that annual report readability is a significant indicator for firm performance in the two years following the annual report. However, the length of such an effect has not been determined as this study only collected data for two years following the 10-k publication

6.3 Suggestions for further research

As certain industries face differences in reporting standards (Scolaro, 2018) it may be beneficial to study certain industries, and particular the differences between industries. Although this paper uses industry dummies, it does not focus on the nuances between these industries and how they impact the link between readability, performance, and forecasts.

There are several improvements possible to the methods of testing readability. The methods provided in this paper are only a basis for further research, it is only the start for a tailored readability index for business reporting. Other variables that are not tested but mentioned by the SEC may be invaluable components in determining the readability of an annual report. Further analysis may utilize these variables and create a better measure for the readability of business reports.

Lastly, this paper found that annual report readability's impact on firm performance is relevant for the two years following the annual report, however, says nothing about the years following. Further research could measure the persistence of negative firm performance as a function of the readability represented in the annual report.

7. Conclusion

This paper aims to find the relationship between annual report readability, future firm performance as well as analyst forecasts. With regards to firm performance this paper suggests that higher reading difficulty is an indicator of future poor performance. This could be due to managers purposely obscuring negative information to mitigate the transfusion of negative information to the market and analysts. As a result, complex annual reports are more likely to lead to negative firm performance in the two years following the annual report.

This paper also found that more complex annual reports hinder analysts' ability to create accurate forecasts. Lastly, contrary to expectations, this paper finds that complex annual reports result in analysts being more likely to provide earnings per share forecasts that underestimate actual earnings per share. It would therefore be beneficial for firms to clearly convey all information that could impact analyst forecasts regardless of future performance expectations.

References

Ablanque, P. (2018, November 8). *The net effect of disclosure readability on analyst forecast dispersion*. *Economics*. Retrieved from http://hdl.handle.net/2105/44598

Adelberg, A. H. (1979). Narrative disclosures contained in financial reports: means of communication or manipulation?. *Accounting and Business Research*, *9*(35), 179-190.

Asay, H. S., Elliott, W. B., & Rennekamp, K. (2017). Disclosure readability and the sensitivity of investors´ valuation judgements to outside information. The Accounting Review , 1-25.

Barth, M. E., Beaver, W. H., & Landsman, W. R. (1998). Relative valuation roles of equity book value and net income as a function of financial health . Journal of Accounting and Economics, 25(1), 1-34

Beaver, W. H. (1981). *Financial reporting: an accounting revolution*. Prentice Hall.

Benston, G. J. (1973). Required disclosure and the stock market: An evaluation of the Securities Exchange Act of 1934. *The American Economic Review*, *63*(1), 132-155.

Beyer , A., Cohen , D. A., Lys, T. Z., & Walther , B. R. (2010). The financial reporting environment: Review of the recent literature. Journal of Accounting and Economics, 50(23), 269-343.

Bloomfield, R.J.,2002. The "incomplete revelation hypothesis" and financial reporting. Account Horions 16, 233-243

Bratton, W. W. (2001). Enron and the dark side of shareholder value. *Tul. L. Rev.*, *76*, 1275.

Chow, C. W., & Wong-Boren, A. (1987). Voluntary financial disclosure by Mexican corporations. *Accounting review*, 533-541.

Clarke, J., & Shastri, K. (2000). On information asymmetry metrics. *Available at SSRN 251938*.

Clement, M. B. (1999). Analyst forecast accuracy: Do ability, resources, and portfolio complexity matter?. *Journal of Accounting and Economics*, *27*(3), 285-303.

Courtis, J. K. (1998). Annual report readability variability: tests of the obfuscation hypothesis. *Accounting, Auditing & Accountability Journal*, *11*(4), 459-472.

Dale, E., & Chall, J. S. (1949). The concept of readability. *Elementary English*, *26*(1), 19-26.

Dechow, P. M. (1994). Accounting earnings and cash flows as measures of firm performance: The role of accounting accruals. *Journal of accounting and economics*, *18*(1), 3-42.

Farvaque, E., Refait-Alexandre, C., & Saïdane, D. (2011). Corporate disclosure: a review of its (direct and indirect) benefits and costs. *Economie internationale*, (4), 5-31.

Fugere, B., Hardaway, C., & Warshawsky, J. (2005). *Why Business People Speak Like Idiots: A Bullfighter's Guide*. Simon and Schuster.

Gunning, R. (1969). The fog index after twenty years. *Journal of Business Communication*, *6*(2), 3-13.

Healy Paul M., Krishna G. P. (2001). Information Asymmetry, Corporate Disclosure, and the Capital Markets: A Review of the Empirical Disclosure Literature. Journal of Accounting and

Economics 31, 405-440

Healy, P. (1977). Can you understand the footnotes to financial statements. *Accountants Journal, 56*(July), 219-222.

Healy, P. M., & Palepu, K. G. (2003). The fall of Enron. *Journal of economic perspectives*, *17*(2), 3-26.

Hong, H., Kubik, J. D., & Solomon, A. (2000). Security analysts' career concerns and herding of earnings forecasts. *The Rand journal of economics*, 121-144.

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

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

Lang, M. H., & Lundholm, R. J. (1996). Corporate disclosure policy and analyst behavior. *Accounting review*, 467-492.

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

Leamer, E., & Leonard, H. (1983). Reporting the fragility of regression estimates. *The Review of Economics and Statistics*, 306-317.

Leisurs, M. (2014, July 24). *Disclosure readability regulation and firm-specific economic benefits. Accounting & Control*. Retrieved from http://hdl.handle.net/2105/20322

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

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

McKee, B. K. (1967). Readability formulas can aid good writing. *Public Relations Journal*, *23*(7), 8-11.

Merkl-Davies D. M., Brennan N. M. (2007). Discretionary Disclosure Strategies in Corporate Narratives: Incremental Information or Impression Management?. Journal of Accounting Literature 26, 116-19

Miller, M., & Modigliani, F. (1961). Dividend policy, growth, and the valuation of shares

Previts, G. J., Bricker, R. J., Robinson, T. R., & Young, S. J. (1994). A content analysis of sell-side financial analyst company reports. *Accounting Horizons*, *8*(2), 55.

Primpas, I., Tsirtsis, G., Karydis, M., & Kokkoris, G. D. (2010). Principal component analysis: Development of a multivariate index for assessing eutrophication according to the European water framework directive. *Ecological Indicators*, *10*(2), 178-183.

Ross, S. A. (1973). The economic theory of agency: The principal's problem. *The American economic review*, *63*(2), 134-139.

Rutherford, B. A. (2003). Obfuscation, textual complexity and the role of regulated narrative accounting disclosure in corporate governance. *Journal of management and governance*, *7*(2), 187-210.

Saorín, E., & López, F. J. (2013). Corporate financial communication and the internet: manipulating investor audiences?. *Online information review*, *37*(4), 518-537.

Schipper, K. (1991). Analysts' forecasts. Accounting horizons, 5(4), 105.

Schrand, C. M., & Walther, B. R. (2000). Strategic benchmarks in earnings announcements: The selective disclosure of prior-period earnings components. The Accounting Review, 75(2), 151-177.

Scolaro, Federico. (2018, July 13). *The association between disclosure readability and stock returns: evidence from the 40 largest listed European banks*. *Accounting & Financial Management*. Retrieved from http://hdl.handle.net/2105/44456

Sharma, Subhash, and Subhash Sharma. "Applied multivariate techniques." (1996).

Singhvi, S. S., & Desai, H. B. (1971). An empirical analysis of the quality of corporate financial disclosure.

The Accounting Review, *46*(1), 129.

Tan, H. T., Ying Wang, E., & Zhou, B. O. (2014). When the use of positive language backfires: The joint

effect of tone, readability, and investor sophistication on earnings judgments. *Journal of Accounting Research*, *52*(1), 273-302.

United States. The Securities and Exchange Commission, & Wheat, F. M. (1967). *Disclosure to Investors: A Reappraisal of Federal Administrative Policies Under the'33 and'34 Acts*. Commerce Clearing House.

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

Appendix A

Table 7: Regressions of Readability on Earnings

Dependent Variable	*Earnings 1 Year Ahead*	*Earnings 1 Year Ahead*
Variable of Interest	[11] Readability Index	[12] FOG
Independent Variables		
Readability Index	-0.0265(-2.51)**	
FOG		-0.0211(-3.03)***
M2B	-0.0130(-1.81)*	-0.0122(-1.69)*
Size	0.0567(10.71)***	0.5543(10.91)***
D2E	0.0346(7.05)***	0.0347(7.08)***
Age	0.0026(2.55)**	0.0029(2.90)***
DNI	0.3662(25.72)***	0.0141(26.15)***
Industry dummies	Yes	Yes
Year dummies	Yes	Yes
Cons	-.291(-4.08)***	0.4737(5.55)***
R^2	0.1024	0.1025
Observations	26015	26015

This table shows the results for the regressions where the dependent variable is earnings one year after the annual report filing date and the variable of interest is readability, defined in regression [1] as the readability index and in [2] as the Gunning-Fog Index .

The t-statistics are provided within brackets. The significance is indicated as follows.

* Significant at the 0.1 level

** Significant at the 0.05 level

*** Significant at the 0.01 level

Appendix B

VIFS

Regression no.	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]	[9]	[10]	[11]	[12]
Readability Index	1.31		1.31		4.55		1.21		1.21		1.52	
FOG		1.12		1.12		4.38		1.21		1.13		1.20
M2B	1.08	1.08	1.08	1.08	1.15	1.15	1.15	1.15	1.08	1.15	1.12	1.11
Size	1.28	1.21	1.28	1.21	2.08	2.05	2.05	2.05	2.07	2.05	1.51	1.41
D2E	1.21	1.21	1.21	1.21	1.07	1.07	1.07	1.07	1.07	1.07	1.15	1.14
DNI					1.44	1.41	1.43	1.40	1.43	1.40	1.35	1.32
Firm Age					56.46	412.45	1.15	1.15	1.21	1.15	1.22	1.17
Readability Index* Firm Age					53.46							
FOG*Firm Age						408.00						
Following					1.63	1.63	1.63	1.63	1.63	1.63		

The variance inflation factors (VIFS) for each variable are presented for each regression. The regression numbers correspond to the number mentioned in the text.