TEXTUAL ATTRIBUTES OF INTEGRATED REPORTS AND ECONOMIC BENEFITS

EVIDENCE FROM A VOLUNTARY SETTING

EBBA KJELLBERG

TILDE EMILSSON

Master Thesis

Stockholm School of Economics

2023

Textual Attributes of Integrated Reports and Economic Benefits:
Evidence from a voluntary setting

Abstract

This paper examines the economic benefits associated with textual attributes of Integrated Reports. Integrated Reporting is a corporate disclosure combining financial and non-financial information, and has since 2010 been mandatory for firms listed on the Johannesburg Stock Exchange. Correspondingly, the body of literature on Integrated Reporting Quality primarily examines this mandatory setting. However, due to the Corporate Sustainability Reporting Directive (CSRD) adopted in 2022 and the merger between the International Financial Reporting Standards (IFRS) and International Integrated Reporting Council (IIRC) in 2021, the trend of adopting Integrated Reporting is increasing in Europe. We therefore investigate the association between textual attributes of Integrated Reports and economic benefits in the European voluntary setting between the years 2016 and 2021. We find no statistical relationship between high-quality textual attributes of Integrated Reports and positive market effects. We also extend existing literature by examining if the release of an intertwined Integrated Report has a moderating effect on the above mentioned relationship and find some evidence. The preference for intertwined Integrated Reports consist of high reading difficulty and conciseness, whereas stand-alone Integrated Reports are preferred to be readable and longer.

Keywords:

Integrated Reporting, Textual Attributes, Voluntary Disclosures, IIRF

Authors:

Ebba Kjellberg (24487)
Tilde Emilsson (50593)

Tutor: Irina Gazizova, Associate Professor, Department of Accounting

Master Thesis

Master Program in Accounting, Valuation and Financial Management
Stockholm School of Economics

Acknowledgements

We would like to sincerely thank Irina Gazizova for tutoring and supporting us throughout our thesis work. We are also grateful for the help and support from Antonio Vazquez regarding empirical considerations. The responsibility for all errors is entirely in the hand of the authors.

Stockholm 14th of May 2023

Ebba Kjellberg Tilde Emilsson

Table of contents

Definitions

Integrated Reports: An Integrated Report includes both financial and non-financial information. It explains how these two parameters together create value over time by adhering to the seven principles of the International Integrated Reporting Framework (International Integrated Reporting Council, 2023). An Integrated Report according to this study must also explicitly refer to the International Integrated Reporting Framework (IIRF) or the International Integrated Reporting Council (IIRC). The Integrated Report can be issued either intertwined or stand-alone.

Integrated Reporting Quality: Integrated Reporting Quality can be measured using content analysis, i.e., how well the Integrated Report adheres to the principles relating to content of the IIRF, for example using scoring models. Another approach is to use textual analysis, i.e., how well the Integrated Report aligns with the principles relating to textual attributes of the IIRF (Caglio et al., 2020), which is used in this study.

Intertwined Integrated Report: Firms who release an intertwined Integrated Report are firms who combine the IIRF framework with existing financial reports. Most commonly, firms merge the IIRF with their annual report (International Integrated Reporting Council, n.d.).

Stand-alone Integrated Report: Firms who release a stand-alone Integrated Report are firms who issue a separate report that follow the IIRF (International Integrated Reporting Council, n.d.).

Textual attributes: In this study, textual attributes are used as an umbrella term for syntactic analysis and includes the parameters for readability and length. The parameters compute the cognitive difficulty of reading information and can be seen as proxies for Integrated Reporting Quality (Jones & Shoemaker, 1994; Caglio et al., 2020).

Abbreviations

CSRD: Corporate Sustainability Reporting Directive was adopted in June 2022 by the European Union. The law requires large and listed companies to disclose information on risks and opportunities related to social and environmental factors, as well as how the operations affect these factors (European Commission, 2023).

IIRC: International Integrated Reporting Council. The council was founded to improve the quality of information and the communication of corporate reports related to value creation (International Integrated Reporting Council, 2023).

IIRF: International Integrated Reporting Framework is the foundation for Integrated Reports. It consists of seven guiding principles. Namely the Integrated Reports should be: 1) of strategic focus with a long-term perspective, 2) holistic and connect information that is required for their value creation, 3) present stakeholder relationships, 4) disclose material information that could impact their value creation ability, 5) concise, 6) reliable and complete and 7), show consistency and comparability in the information presented (International Integrated Reporting Framework, 2021).

1. Introduction

Integrated Reporting was formally established in 2010 when the International Reporting Council (IIRC) was founded. The IIRC was created as a response to the 2008/2009 Global Financial Crises with the aim to improve the information environment of corporate disclosures and guide companies on how to clearly communicate value creation over time. This was materialized through the publication of the International Integrated Reporting Framework (IIRF) in 2013. The trend of Integrated Reporting has been growing over time and in the beginning of 2023 the principle-based framework has been adopted by more than 2,500 companies in over 70 countries. In 2021 the IIRC merged with the Sustainability Accounting Standard's Board (SASB) thereby creating the Value Foundation (International Integrated Reporting Council, 2023). This contraction of reporting standards and frameworks was further fueled in 2022 when the Value Foundation consolidated with the International Financial Reporting Standards (IFRS) Foundation. IFRS will by this means incorporate the IIRF into its standards (Guillot, 2023).

Consequently, the authors of this paper argue that the trend of adopting Integrated Reporting in Europe will continue as the IFRS are mandatory for public firms in the European Union (International Financial Reporting Standards, 2022a). Further support of this argument is the increased regulation on the reporting of non-financial data through the adoption of the CSRD in 2022 (KPMG International, 2022). Due to this anticipated expansion of adopting Integrated Reporting it is of importance to investigate how the textual attributes of a firm's Integrated Report is associated with market effects. This paper defines an Integrated Report as a corporate disclosure that includes both financial and non-financial information, and explicitly refers to the IIRF. The Integrated Reports can either be disclosed stand-alone or be intertwined into another report, such as the annual report. More specifically, this paper will examine *if Integrated Reports with high-quality textual attributes have a positive association with market effects and whether this link is affected by the chosen disclosure type, i.e., intertwined or stand-alone.*

Textual analysis in accounting research has gained traction due to the increase of information in corporate disclosures over the years (Loughran & McDonald, 2011). In broad terms, textual analysis investigates the quality of qualitative accounting information by transforming it into a quantitative measure, often related to readability, length, and tone (Caglio et al., 2020). So far, textual analysis has mainly focused on financial reports which has confirmed that readability is positively associated with stock liquidity (Boubaker et al., 2019; Lang & Stice-Lawrence, 2015) and how clarity and completeness have a positive association with firm valuation (Botosan & Plumlee, 2002). The positive association between high-quality textual attributes and market effects have also been confirmed when examining Integrated Reports (Caglio et al., 2020).

Furthermore, Integrating Reporting Quality, defined by how well the Integrated Report aligns with the IIRF's content elements, has also proven to be linked to positive market effects such

as forecast dispersion (Zhou et al., 2017), stock liquidity (Barth et al., 2017), expected future cash flows (Barth et al., 2017) and firm valuation (Lee & Yeo, 2016). However, these positive effects relating to both content- and textual analysis have only been confirmed in South African settings where Integrated Reporting is mandatory (Caglio et al., 2020). In voluntary settings, neither the release of an Integrated Report (Wahl et al., 2020) nor its high-quality (as proxied by its alignment with the framework's content elements) have proven to have a relationship with positive market effects (Leukhardt et al., 2022).

This study examines a European sample of 287 firm-year observations between 2016 and 2021. The research question is investigated through a quantitative research design that utilizes OLS regressions. In line with prior research (Caglio et al., 2020), the proxies for market effects are Tobin's Q and Bid-ask spread, while the main explanatory variables are reading difficulty and length.

Our findings show no significant relationship[1] between high-quality textual attributes and positive market effects when examining a voluntary setting. The results relating to whether an intertwined Integrated Report has a moderating effect on the relationship between high-quality textual attributes and positive market effects are mixed. For some specifications, relating to the examined proxy for textual attributes and market effects, the association holds but is not consistently robust. More specifically, there is only evidence that an intertwined Integrated Report has a moderating effect on the relationship between reading difficulty and Bid-ask spread as well as the association between length and Tobin's Q.

This study contributes to research in several ways. First, we build on the recent stream of studies that examine Integrated Reporting voluntary settings (Leukhardt et al., 2022; Wahl et al. 2020), whereas prior research has mainly examined Integrated Reports in the mandatory setting of South Africa (Caglio et al., 2020; Barth et al., 2017; Zhou et al., 2017; Lee & Yeo, 2016). Second, to our knowledge, we investigate for the first time a sample solely focusing on a voluntary setting in the European market compared to Leukhardt et al. (2022) and Wahl et al. (2020) who investigate international firms. This is of importance due to the recent regulatory shifts and consolidations of standard setting bodies in the European market. Third, while the majority of papers have proxied Integrated Reporting Quality (IRQ) by performing a content analysis (Barth et al., 2017; Zhou et al., 2017; Lee & Yeo, 2016; Leukhardt et al., 2022), we add to the body of literature on textual analysis as proxy for IRQ (Caglio et al., 2020). Last, this study marks the first attempt to study the effect of releasing an intertwined Integrated Report on the association between textual attributes and market effects.

[1] Note that the word "relationship" in this study is used as a synonym for association. While we may observe a correlation between variables, this does not necessarily mean that one variable causes the other. Other factors or variables may be of importance. Hence, we do not imply causality in our findings.

The remainder of this paper is structured as follows. *Section 2* provides institutional background of Integrated Reporting and a literature review on the research domain. *Section 3* develops the hypotheses that are tested for. *Section 4* discusses methodological considerations and *Section 5* presents descriptive statistics. *Section 6* outlines the results which is followed by a discussion in *Section 7*. *Section 8* presents conclusions and implications.

2. Institutional background and literature review

2.1 Definition of an Integrated Report and disclosure methods

Per definition, an Integrated Report is;

"A concise communication about how an organization's strategy, governance, performance and prospects, in the context of its external environment, lead to the creation of value in the short, medium and long term" (International Integrated Reporting Framework, 2021, p.10).

An Integrated Report includes both financial and non-financial information and explains how these parameters together create value over time (International Integrated Reporting Framework, 2021). In this study, an Integrated Report must also refer to the International Integrated Reporting Framework (IIRF) which guides companies on how to produce Integrated Reports. Thus, an Integrated Report is not only a report combining financial information and non-financial information but must also adhere to the principles of the IIRF.

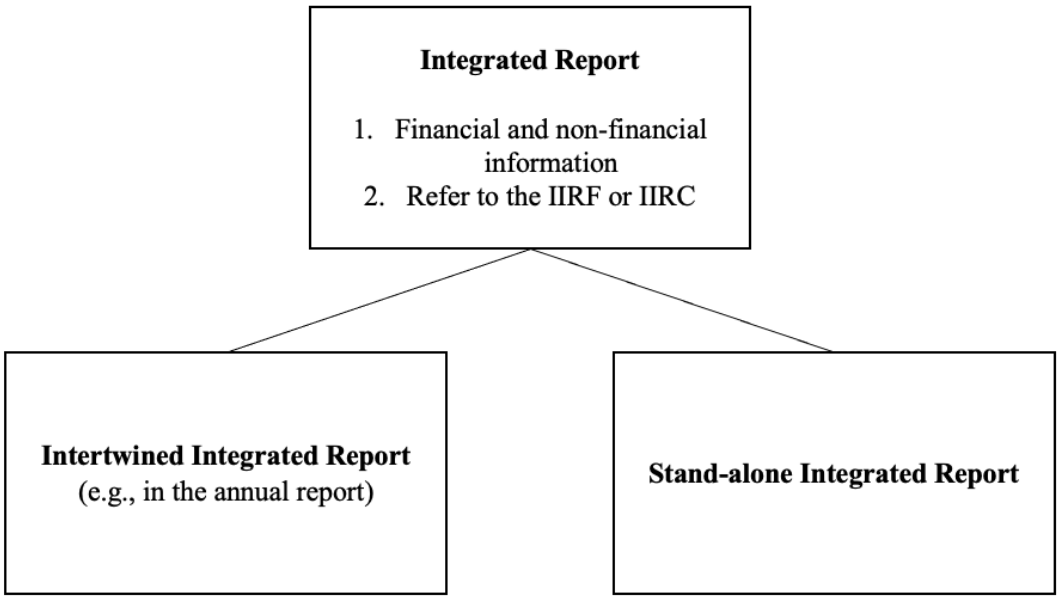
The aims of the IIRF includes enhancing information provision to external stakeholders, and also to improve companies' abilities to utilize integrated thinking for decision making purposes. This implies that the company considers the creation, preservation, or erosion of value in the short, medium, and long term when making decisions. Thus, the framework provides benefits both to external stakeholders and the company per se. To ensure the realization of benefits, the framework builds on principles aimed to guide companies on how to produce concise and cohesive reports that facilitate information sharing with various stakeholders (International Integrated Reporting Council, 2023).

In total, there are seven guiding principles constituting the IIRF: 1) The Integrated Reports should be of strategic focus with a long-term perspective, 2) be holistic and connect information that is required for their value creation, 3) present stakeholder relationships, and 4) disclose material information that could impact their value creation ability. The remaining principles guide the textual attributes related to the report. Explicitly, the information should be 5) concise, 6) reliable and complete, and 7) show consistency and comparability in the information presented (International Integrated Reporting Framework, 2021).

Despite being referred to as a 'report', an Integrated Report is a voluntary disclosure issued by companies. The only exception from this voluntary setting is in South Africa, where firms listed on the Johannesburg Stock Exchange must release an Integrated Report (Caglio et al., 2020). The presence of a voluntary setting implies that companies themselves choose *if* and *how* they want to produce their Integrated Reports. If deciding to release an Integrated Report, this can be done in two ways (see Figure 1). First, the Integrated Report can be intertwined with another report or disclosure, such as the annual report. The second option is to release the Integrated Report on a stand-alone basis (International Integrated Reporting Council, n.d.). The main difference between these different disclosing methods relates to length, where an intertwined

Integrated Report is naturally longer, stemming from including more information, such as financial statements, compared to a stand-alone report.

Figure 1. Disclosure methods of Integrated Reports



2.1 Institutional background on Integrated Reporting

In 2010, the International Integrated Reporting Council (IIRC) was founded as a response to the 2008/2009 Global Financial Crises. Initiated by regulators, investors, businesses, academics, standard setters, accountants, and non-governmental organizations, the IIRC was established to improve the quality of information and the communication of corporate reports related to value creation over time, believing that it would help mitigate another financial collapse. Their mission materialized through the launch of the International Integrated Reporting Framework (IIRF) in December 2013 (International Integrated Reporting Council, 2023).

In the beginning of 2023, the IIRF has been adopted by more than 2,500 companies in over 70 countries, referred to in approximately 40 stock exchanges' guiding documents, as well as inspired corporate governance regulations worldwide (International Integrated Reporting Council, 2023). Out of the 100 largest companies by revenue, from a sample of 58 countries (N100), 55% of all companies in the Middle East and 30% of all companies in Asia Pacific have adopted Integrated Reporting. In Latin America, 28% of N100 firms apply the IIRF, and in Africa, Eastern Europe, and Western Europe the percentage of companies corresponds to approximately 15% each (KPMG International, 2022). South Africa is currently the only country which requires firms listed on the Johannesburg Stock Exchange (JSE) to release Integrated Reports. This was adopted in 2010 and adheres to the rule of 'apply or explain',

meaning that firms can choose not to produce an Integrated Report if the company believes that another type of report will benefit stakeholders (Caglio et al., 2020).

In Europe, it is especially relevant to examine the textual attributes of Integrated Reports due to the European Commission's approval of the Corporate Sustainability Reporting Directive (CSRD) in June 2022. Replacing the Non-Financial Reporting Directive (NFRD), which was adopted in October 2014, the CSRD will extend the scope of companies required to disclose non-financial information from large public-interest companies with more than 500 employees to also include large companies and listed SMEs. In absolute terms, this will increase the number of European companies reporting on non-financial data from approximately 11,700 to circa 50,000. The regulation is stipulated by the European Green Deal, which aims to improve corporate reporting on risk and opportunities arising from social and environmental issues, as well as the entities' impact on people and the environment (European Commission, 2023). Not only is the increased regulation on reporting of non-financial data driving the trend of Integrated Reporting (KPMG International, 2022), but also the current work of incorporating the IIRF into the IFRS Sustainability Disclosure Standards (Guillot, 2023).

In June 2021, the IIRC and the Sustainability Accounting Standards Board (SASB) merged into the Value Reporting Framework as an effort to simplify the corporate reporting landscape (International Integrated Reporting Council, 2023). Further contraction of reporting standard bodies emerged in August 2022 when the Value Reporting Foundation consolidated with the IFRS Foundation. This was performed as an effort to secure the future of the SASB standards and the IIRF, and to simplify and create a baseline for sustainability disclosures for investors. The International Accounting Standards Board (IASB) and the International Sustainability Standards Board (ISSB) will build on and incorporate the IIRF into their standard setting and encourage continued use of it. The IIRF and Integrated Thinking Principles are currently being utilized for the establishment of the IFRS Sustainability Disclosure Standards (Guillot, 2023). Since the IFRS Accounting standards must be followed by all domestic public companies in the European Union (EU), as well as required by foreign companies trading at a European Exchange (International Financial Reporting Standards, 2022a), its incorporation of the IIRC will further drive the trend of intertwined Integrated Reporting in Europe. Even though the IFRS Sustainability Standards are not yet adopted by national jurisdiction in the EU, it can be assumed to follow the path of the IFRS Accounting Standards and become mandatory in European countries once finalized.

2.2 Literature review

2.2.1 Textual analysis

The field of textual analysis in accounting literature has gained importance over the years as the amount of information included in company reports increases, and continues to do so, as the user friendliness of computer software improves (Loughran & McDonald, 2011). Textual analysis relates to the quality of accounting information, which includes a process where qualitative accounting information is converted into a quantitative measure. These quantitative measures commonly explore the level of 'readability', 'length', and 'tone' of the analyzed information (Caglio et al., 2020). There are two main approaches to textual analysis in research: thematic analysis and syntactic analysis.

The thematic analysis identifies attitudes and themes of information, also known as the parameter ‘tone’ (Jones & Shoemaker, 1994). According to Loughran and McDonald (2011), tone can be measured using word categorization. Using external word lists including Harvard’s General Inquirer, the accounting information is tagged with categories such as 'positive', 'negative', and 'weak', which infers the tone of the text. There are critics who argue that word lists are inadequate as there is no dictionary available specifically for corporate information. Hence, some authors argue that tone should be measured with statistical software like Diction to avoid subjectivity (Caglio et al., 2017; Caglio et al., 2020). Diction calculates a quantitative score on dimensions for optimism and certainty, where a higher score indicates a more positive tone (Loughran & McDonald, 2011). Still, there are issues with measuring tone. For example, accounting information can be framed in a way where negative news are described using positive words (e.g., “did not benefit”), which generates an incorrect assessment. Another issue of tone measures appears if the instrument is not case-sensitive. In this case, the word 'may' will be classified as an indicator of uncertainty despite the fact that the text refers to the month of May. These errors will affect the tone both for statistical software and the frequency of word lists (Loughran & McDonald, 2016).

The second approach of textual analysis involves syntactic analysis and the parameters ‘readability’ and ‘length’. Syntactic analysis, which is the focus for this study, computes the cognitive difficulty of reading information (Jones & Shoemaker, 1994). Historically, syntactic analysis has focused on selected topics of information, examined only a small sample, and utilized a single readability metric. For instance, an empirical review performed by Jones and Shoemaker (1994) examine 32 syntactic studies on accounting information and found 26 of the studies to investigate annual report narratives. Only two studies performed the textual analysis on the entire text as opposed to selected topics in the annual report. Two studies sampled more than 100 annual reports, and only 11 studies adopted more than one readability measure. The explanation for these weaknesses generally concerns time constraints since reports and indices are hand-collected by researchers.

Some studies incorporate the separate variable 'length' as a proxy for syntactic analysis. A higher number of sentences, characters or pages, increases the effort and cognitive difficulty as there is more information for the reader to process (Caglio et al., 2017; Caglio et al., 2020; Lang & Stice-Lawrence, 2015; Li, 2008).

'Readability' on the other hand is commonly included in syntactic analysis and is examined using readability formulas like the Gunning Fog Index, the Flesch-Kincaid Index, the Smog Index or the Bog Index. These metrics incorporate for example, the number of words per sentence and syllables per word into a combined metric where a higher grade equals higher reading difficulty. Readability per say has been defined with great variety in the literature; however, the most common definition, also used in this report, follows Chall and Dale (1949):

"Readability is 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. The success is the extent to which they understand it, read it at an optimum speed, and find it interesting" (Chall & Dale, 1949, p. 23)

The Flesch-Kincaid index was used in 26 of the 32 syntactic studies examined in the Jones and Shoemaker (1994) paper. The Flesch-Kincaid Index is based on the Flesch Reading Ease Score developed by the readability expert, Rudolf Flesch, in the 1940s. One had to convert the reading score into a grade level using a conversion table, and henceforth, the Flesch Reading Ease evolved to the Flesch-Kincaid Index in the 1970s, with assistance from scientist J. Peter Kincaid, where the Index's output instantly generated the grade level. The Flesch-Kincaid Index was originally developed for the U.S Navy but its application has progressed also in accounting research.

Nonetheless, the Flesch-Kincaid Index has received a lot of criticism within textual analysis of accounting. Critics argue that the over 50-year-old Index is based on an outdated language and has failed to incorporate linguistic evolution over the years. Additionally, the Flesch-Kincaid Index was not intended for accounting information when constructed (Boubaker et al., 2019). The lack of fit is also evident in the Gunning-Fog Index developed by Robert Gunning in the 1950s. As Loughran and McDonald (2016) point out, frequently used words in corporate documents such as 'financial', 'operations', 'management' and 'customers' contain several syllables which generate a high reading difficulty grade in these indices, despite being highly understood by investors, hence not impairing the actual level of readability.

Subsequently, in the 1970s, psychologist Harry McLaughlin developed a new readability index, the SMOG Index. The intention of the SMOG index was to provide a simpler measure for readability with at least the same level of accuracy as the Flesch-Kincaid and Gunning-Fog indices, thereby subject to the same limitations as the previous indices. The SMOG Index is only based on a selection of sentences, namely, 10 consecutive sentences in the beginning of the text, 10 in the middle and 10 in the end of the text. Henceforth, this measure is inappropriate for shorter texts, though not an issue for financial disclosures (McLaughlin, 1969).

As a response to criticism claiming that these readability indices are inadequate for accounting texts, the Bog Index was developed, particularly intended for corporate information. The Bog Index determines information complexity based on word familiarity using a proprietary word list in addition to counting syllables, used by for example, the Gunning Fog Index (Soliman & Ben-Amar, 2022). However, the Gunning Fog Index is found in most recent articles on textual analysis and financial disclosures and is still more accessible compared to the Bog Index (Li, 2008; Alduais et al., 2022; Boubaker et al., 2019; Lo et al., 2017; Lang & Stice-Lawrence, 2016).

Studies still tend to rely on a single readability measure as a proxy for readability difficulty. This study therefore aims to improve robustness of textual analysis on accounting information by sampling over 280 Integrated Reports in total, while also combining three readability metrics: the Gunning Fog Index, the Flesch-Kincaid Index and the Smog Index, as well as the parameter length proxied by number of characters and number of sentences.

2.2.2 Textual analysis on financial disclosures

Research on textual attributes of annual reports has been conducted since the 1950s and as requirements on firm disclosures increase, the research field of textual analysis on financial disclosures has also grown (Caglio et al., 2017). The intent of increased disclosure requirements is to decrease information asymmetry between the firm and the investors, and in this manner improve transparency. According to the agency theory, the cost of equity should decrease as the risk is lowered, resulting in positive market outcomes (Caglio et al., 2020). On the contrary, an increase in disclosure information can potentially cause an overload of information for investors. This strategy can be utilized by firms who hope to hide their poor performance with low quality disclosures as a smokescreen (Caglio et al., 2017). Therefore, discussions on the quality of disclosure information instead of the quantity of disclosures have gained meaning.

Previous research has found that disclosure quality using textual analysis, including style and length, correlates with economic effects. Boubaker et al. (2019) examine French firms adopting IFRS and the relationship between readability and stock liquidity. They argued that information asymmetry causes friction in the market and liquidity is the most effective way to measure this impact. Theory claims that uninformed investors become less willing to trade at a low cost when informed investors are present, which lowers demand and therefore liquidity. Greater disclosure quality should then reduce information asymmetry and increase liquidity. Using the Gunning Fog Index as proxy for readability and the high-low spread estimator, bid-ask spread, and zero return days for stock liquidity, Boubaker et al. (2019) find low readability to be statistically and economically significantly related to low stock liquidity.

This association is also identified by Lang and Stice-Lawrence (2015) in an exploratory methodology using firm data from 42 countries between the years 1998 and 2011. The authors measure length, readability (using the Gunning Fog Index), as well as more novel approaches like boilerplates. Boilerplates are defined as 4-word phrases that are frequently used in the

home country, and the authors measure how often these occur in the public annual report. It is argued that commonly used phrases across firms are highly uninformative as they do not contain firm-specific information. Even though the study uses solely English-language reports, which may cause translation effects for non-English countries, the authors still find a positive association between readability and stock liquidity in terms of the bid-ask spread and zero return days when restricting the sample to English-speaking countries and reports.

The findings also show that firm valuation improves as disclosure quality increases. For instance, Botosan and Plumlee (2002) examine 668 firms between 1986 and 1996 and find that firms providing annual reports with high disclosure levels benefit in terms of lower cost of equity capital. The disclosure score was provided by the 1985/86-1995/96 Association for Investment Management and Research (AIMR) Reports where industry subcommittees have evaluated the adequacy of firms' reporting practices, where clarity and completeness are considered. By benchmarking firms operating in the same industry and between years, a relative ranking for each firm-year observation is obtained. On average, the cost of equity capital differed by 0.7 percentage points between the highest and lowest scoring firms. This indicates that the annual report is an important source of information for capital markets, where completeness and clarity reduce information asymmetries and risk premiums.

The opposite direction of the link between economic benefits and readability has also been studied. According to the concept 'management by exception', poor performing companies are subject to higher pressure to explain their failure to investors. This implies that poor performing companies are compelled to issue longer annual reports with higher reading difficulty. This link is identified by Li (2008) when examining how earnings impact the reading difficulty and length of annual reports using 50,000 US firm years. Another explanation brought forward to why readability is lower and length is higher for low performing companies is explained by the 'incomplete revelation hypothesis'. The incomplete revelation hypothesis indicates that managers can reduce the market response to negative news by making the information itself more costly to analyze. That is, by hiding negative news in complex and long annual reports, the market does not fully understand the effect and fails to reflect the information in the stock price (Bloomfield, 2008).

In line with findings on financial information readability and economic benefits, studies on CSR readability show that the market reacts positively when reports are more readable (Du & Yu, 2021; Gao et al., 2022). Interestingly, the economic effect is significantly higher when both readability and CSR performance are high. This relationship shows a complex link between CSR performance, readability and financial performance, which potentially requires additional investigation (Gao et al., 2022).

2.2.3 Textual analysis on Integrated Reports

Even though prior literature mainly has performed textual analysis on financial disclosures, there is a growing interest for examining Integrated Reports. The few papers covering the impact of Integrated Reports foremost examine the economic benefits of high-quality

Integrated Reports, i.e., how well the report aligns with the IIRF (Rinaldi et al., 2018). More explicitly, previous research has shown that analyst forecast errors, analyst earnings dispersion, implied cost of equity, stock liquidity, expected future cash flows, and market valuation are positively affected by higher quality Integrated Reports in mandatory settings (Zhou et al., 2017; Barth et al., 2017; Lee & Yeo, 2016). Interestingly, this relationship has not been confirmed for voluntary settings (Leukhardt et al., 2022).

Zhou et al. (2017) investigate the potential benefits of Integrated Reporting Quality (IRQ) from a capital market perspective. The authors examine 443 company-year observations listed on the Johannesburg Stock Exchange (JSE) from 2009 to 2012. IRQ is proxied using 31 disclosure components that cover the IIRF content elements. From this, the authors conclude that analyst forecast errors are reduced when the Integrated Report aligns with the IIRF, thus showing a higher IRQ. Furthermore, the dispersion of financial forecasts, measured as the standard deviation of earnings per share (EPS) forecasts divided by the median forecast, is weakly reduced as companies align with the IIRF. These findings indicate that following the framework facilitates the assessment of future financial performance from the perspective of a financial analyst. The enhanced information transparency facilitated by high-quality Integrated Reports provides additional benefits to the firm in terms of lower implied cost of equity, suggesting that reduced information asymmetry enable investors to reduce their risk premium.

Instead of proxying IRQ by hand, Barth et al. (2017) use proprietary data from Ernst & Young (EY) South Africa to examine whether higher IRQ is associated with higher firm value, captured through financial market effects as well as real effects. The IRQ measure is based on EY Excellence in Integrated Reporting Awards in South Africa, where EY publicly assesses the Integrated Reporting Quality by evaluating how well the company manages to convey information that is relevant, understandable and accessible to the reader. By proxying firm value on stock liquidity, cost of capital, and expected future cash flows, the authors conclude that IRQ is positively associated with liquidity (capital market effects) and expected future cash flows stemming from higher investment efficiency (real effects). However, the relationship between cost of capital and IRQ is not proved. Collectively, these findings support the concept that alignment with the IIRF improves both information provision to investors (capital market effects) and internal decision making (real effects).

Providing further support for IRQ and economic benefits, Lee and Yeo (2015) show a positive association between higher IRQ and Tobin's Q. By constructing an IRQ measure based on an evaluation of five aspects of the IIRF content elements, they perform a cross-sectional analysis on the intensity of Integrated Report disclosures and firm value of South African firms between 2010 and 2013. The evidence is stronger for firms with high operational complexity, such as high intangible assets, multiple business segments, and large firms. They also show that firms in need of capital injection generally score higher IRQ and are associated with higher firm valuation. These results are supported by the agency theory that higher quality Integrated Reports provide better and more transparent information to external investors, ultimately reducing information asymmetry. The reason why complex firms and capital constrained firms

produce more high-quality Integrated Reports is because their inherent nature and needs demand them to do so. Without providing investors with clear information on how they will create value in the future, these complex firms would most likely receive lower valuations and less external capital.

Unlike the mentioned studies on Integrated Reporting which investigate a mandatory setting, Wahl et al. (2020) examine if voluntary Integrated Reporting disclosure affects analyst earnings forecast accuracy and firm value. No significant effect of voluntary Integrated Reporting disclosure is identified by the authors when examining 167 Integrated Reporting companies between 2011 and 2018. Building on this, Leukhardt et al. (2022) continues to extend the body of literature on voluntary Integrated Reporting by examining IRQ and analyst forecast dispersion. The authors examine 101 companies between 2015 and 2019 who voluntarily issue an Integrated Report. IRQ is proxied through a scoring model that assesses the Integrated Report on its background, contents, assurance and reliability, as well as form. No support for an association between IRQ and economic benefits is confirmed.

Instead of proxying IRQ by performing a content analysis, Caglio et al. (2020) utilize textual analysis of Integrated Reports and examine the association between high-quality textual analysis and economic benefits. The examined textual attributes are broken down into 3 different factors: readability (proxied by Flesch-Kincaid, Fog, and Smog readability scores), length (number of words and number of characters), and tone (optimism and certainty). These readability attributes align with the IIRF recommendation of reports being concise, favoring plain language over the use of jargon or technical terms, and have a neutral representation and tone. Economic consequences are represented by firm valuation (proxied by Tobin's Q), stock liquidity (proxied by bid-ask spread), and analysts forecasts dispersion (proxied by the standard deviation of analysts' EPS forecast dispersion divided by the median forecast). Their sample consists of the 160 largest firms listed on the Johannesburg Stock Exchange ranked by market capitalization in 2015. They show that readability is associated with higher market valuation, lengthier reports are associated with lower stock liquidity, and a more biased tone of optimism and certainty provides less dispersed EPS forecasts.

This paper builds on the mentioned papers above and aims to make several contributions. First, we add to the literature of a voluntary setting (Leukhardt et al., 2022; Wahl et al., 2020) compared to previous literature which mainly focus on a mandatory setting (Zhou et al., 2017; Barth et al., 2017; Lee & Yeo, 2016; Caglio et al., 2020). Second, instead of focusing on international firms in a voluntary setting (Leukhardt et al., 2022; Wahl et al., 2020) the European market is chosen due to the new regulatory landscape as well as the recent consolidations of reporting standard bodies. Third, this paper will contribute to prior literature by extending the number of papers that investigate textual analysis (Caglio et al., 2020), rather than content analysis (Zhou et al., 2017; Barth et al., 2017; Lee & Yeo, 2016; Leukhardt et al., 2022) of Integrated Reports. Lastly, this study marks the first attempt to study the effect of releasing an intertwined Integrated Report on the association between textual attributes and market effects.

3. Hypotheses development

3.1 High-quality and low-quality textual attributes of Integrated Reports

Drawing upon the International Integrated Reporting Framework (2021), there are 7 guiding principles relating to the preparation and presentation of the Integrated Report. Of these principles, 2 include recommendations on textual attributes, namely 'Conciseness' and 'Reliability and completeness'. Referring to the former principle, the report should include enough information to ensure understanding of the *"…organization's strategy, governance, performance and prospects, without it being burdened with less relevant information"* (p.33, paragraph 3.37). The framework further states that a concise Integrated Report *"…expresses concepts clearly and in as few words as possible"* (pp.33-34, paragraph 3.38) and *"…favours plain language over the use of jargon or highly technical terminology"* (pp.33-34, paragraph 3.38), amongst other factors. The latter principle, 'Reliability and completeness', highlight the need of presenting a balanced Integrated Report that *"…has no bias in the selection or presentation of information…to change the probability that it will be received either favorably our unfavorably"* (p.35, paragraph 3.44). These 2 guiding principles on textual attributes presented in the IIRF, i.e., 'Conciseness' and 'Reliability and completeness', are the foundation to how this paper defines high-quality and low-quality textual attributes of Integrated Reports.

To summarize, high-quality textual attributes of Integrated Reports relate to (I) conciseness, (II) readability, and (III) balance of information. In comparison, low-quality textual attributes of Integrated Reports relate to them being verbose, difficult to read, and presenting biased information. Hereafter, this paper focuses on textual attributes of Integrated Reports relating to conciseness and readability due to limited access to coding software necessary to analyze balance of information.

3.2 Hypotheses

The theoretical link between textual attributes and economic benefits is mainly represented by the agency theory. The agency theory builds on the principle that investors (principals) delegate operational decisions to management (agents). Management in turn are financially incentivized to act in favor of the investors' interests. This transaction causes management to be more informed than the investors, and thus, investors add a risk-premium and require higher compensation for their increased risk. In general, financial disclosures are utilized to reduce information asymmetry. Henceforth, readable and concise Integrated Reports can reduce the risk-premium by leveling out the playing field between informed and uninformed investors, as well as decreasing investors' cost of capital (Barth et al., 2017). It is also argued that voluntary disclosures can complement the lack of compulsory required information, thus improving the information base and reducing uncertainty further (Wahl et al., 2020).

As presented under *Section 2.2.2 Textual analysis on financial disclosures* and *2.2.3 Textual analysis on Integrated Reports*, the link to agency theory has been proven in the context of

Integrated Reporting as readability and conciseness are associated with improved firm valuation and stock liquidity.

3.2.1 Hypothesis I

Previous literature has shown that corporate disclosures that exhibit high readability are associated with economic benefits such as lower cost of equity capital (Botosan & Plumlee, 2002), and therefore higher market valuation. Extending this evidence to Integrated Reports, Zhou et al. (2017) show that firms that follow the IIRF content elements to a greater extent benefit in terms of lower monitoring costs and implied cost of equity, thereby improving firm valuation. By instead proxying firm value with Tobin's Q, Lee & Yeo (2016) confirm the findings of Zhou et al. (2017). Building on the findings of Zhou et al (2017) and Lee & Yeo (2016) relating to Integrated Reporting Quality proxied by alignment with the IIRF's content elements and its relationship with higher firm valuation, it is reasonable to hypothesize that alignment with the framework's principles on textual attributes also will yield economic benefits. This assumption is supported by Caglio et al. (2020) which show that firms producing more readable reports, i.e., following the IIRF textual attribute principles, are associated with higher firm valuation. Based on these findings, we predict that high-quality textual attributes of Integrated Reports will facilitate the communication of future value creation, reducing information asymmetry and therefore improving firm valuation.

Furthermore, multiple papers (Barth et al., 2017; Plumlee, 2003) have shown that high-quality textual attributes of financial reports, defined as readability, improve stock liquidity. Boubaker et al. (2019) and Lang & Stice-Lawrence (2015) show that financial reports using less length and complexity enable a higher share of market participants to understand and incorporate accounting information, ultimately improving stock liquidity. Extending this evidence to Integrated Reports, Barth et al. (2017) confirm that there is a positive relationship between Integrated Reporting Quality and stock liquidity. Furthermore, the findings by Caglio et al. (2020) state that lengthier Integrated Reports are associated with lower stock liquidity, implying that following the IIRF's principles on textual attributes has positive economic effects for firms. Based on these findings, we predict that high-quality textual attributes of Integrated Reports will facilitate the communication of future value creation and improve stock liquidity due to the lower information asymmetry.

Combining these insights, we hypothesize:

H1 Integrated Reports with high-quality textual attributes, i.e., readability and conciseness, are associated with positive market effects

3.2.2 Hypothesis II

Mentioned in section *2.1 Definition of an Integrated Report and its disclosure methods,* firms can choose to release their Integrated Report either intertwined or stand-alone (see Figure 1). This paper aims to contribute to current literature by examining what effect the release of an

intertwined Integrated Report has on the relationship between textual attributes and economic effects. To the best of our knowledge, this has not been explored before in the literature of textual analysis on Integrated Reports.

Surveying 5,800 investors, Chang et al. (1983) prove that annual reports are important. The authors examine three groups of investors in three countries. Explicitly, individual investors, institutional investors and financial analysts across the United Kingdom, United States and New Zealand. Findings show that individual investors in all countries consider the annual report one of the top three sources of information, while both institutional and financial analysts ranked the annual report as the most or second most important source of information.

This line of reasoning is also discussed in Barber and Odean (2008). The authors argue that investors have cognitive limitations to the amount of information that they can process. For this reason, investors face a search problem as to which stocks to consider when buying a stock. By constructing a choice set of a limited number of stocks in consideration, the buying decision is more manageable. It is in fact easier to choose between ten alternatives compared to a hundred alternatives. One can therefore argue that intertwined Integrated Reports, where the IIRF is combined with the financial report, imposes a higher likelihood of being included in the investor's choice set as the financial report already is a known source to the investor.

Additionally, existing research on financial disclosures shows that there is indeed an effect of simply providing the information within the financial report, despite the fact that the information itself is not new to the market. For example, Christensen et al. (2017) examine SEC-registered mine owners who are required to include mine-safety records in their financial reports with non-SEC-registered mine owners, and whether the inclusion of the mine-safety records in the financial reports have an effect, even though the records already are publicly available to the market through press releases. The authors show that the inclusion of social responsibility information in financial reports has an incremental real effect. Since financial reports are broadly disseminated and have a low acquisition cost, the inclusion of mine-safety records in the financial report increases awareness and attention among investors. This effect is explained by changes in short-window stock returns and mutual fund holdings.

Henceforth, combining the findings presented by Cheng et al. (1983), Barber and Odean (2008), and Christensen et al. (2017) we hypothesize:

H2 The association between textual attributes and market effects is moderated by the effect of issuing an Integrated Report intertwined

4. Method

4.1 Sample and data collection

The sample of this study is based on Integrated Reporting companies acknowledged in the IIRC Database and that are publicly listed on a European exchange. We collect data attributable to two main datasets, one being textual attributes of firms' Integrated Reports, and the other being firms' economic attributes. The Integrated Reports were downloaded from each company's official website, and the financial information was obtained via Capital IQ. These two datasets were then merged into our final sample.

Our original list of firms is based on European Integrated Reporting companies acknowledged in the IIRC Database. Previous studies focusing on Integrated Reporting in voluntary settings have also referred to the IIRC Database as their sourcing strategy for identifying firms adopting the IIRF (Leukhardt et al., 2022; Wahl et al., 2020). According to the IIRC Database, the list includes companies in Europe who:

"… refer to the IIRC, or the Integrated Reporting Framework, or are influenced by the Framework through participation in Integrated Reporting Networks" (International Integrated Reporting Council Database, n.d).

Considering the regulatory advancement of non-financial information within the European Commission, our sample includes companies subject to CSRD (Corporate Sustainability Reporting Directive, see *Section 2.1 Institutional background on Integrated Reporting*). Explicitly, listed EU companies are required to follow the CSRD as well as non-EU companies with branches or subsidiaries in the EU exceeding 150 million euros in revenue (Council of the EU, 2022). Therefore, it is of importance to investigate a European setting and provide insights on Integrated Reporting textual quality due to the anticipated increase in Integrated Reporting amongst firms subject to the CSRD (KPMG International, 2022).

Although the United Kingdom and Switzerland are not member states of the European Union, firms listed in these countries are also admitted to the sample since IFRS are widely applied. The United Kingdom applies the IFRS with minor modifications to all domestic public companies and listings by foreign companies (International Financial Reporting Standards, 2022b). In Switzerland, IFRS is permitted for all domestic public companies and listings by foreign companies. Even though IFRS is not required for Swiss companies, more than 50% of public Swiss companies adopt IFRS (International Financial Reporting Standards, 2021). Therefore, firms in these countries are considered to operate under similar accounting standards as EU-firms and therefore relevant to include. Also, since the IFRS will incorporate the IIRF principles into its standard setting these firms could potentially benefit from this research in terms of policy and managerial implications.

Considering the growth of Integrated Reporting adoption in recent years (KPMG International, 2022), the sampled period is set to 2016-2021 to absorb an increase of reporting companies,

and thereby improving our sample size. In 2017, IIRC announced that the IIRF was entering a breakthrough phase with global adoption gaining momentum. Additionally, the IIRF received recognition as the Chair of IASB announced that the IIRF is compatible with IASB's own conceptual framework in 2018 (International Integrated Reporting Council, 2017b).

After filtering on European companies in the IIRC Database, 161 unique reporting companies and associations were returned. Subsequently, we evaluated each firm by hand to confirm that all criteria were satisfied for our final sample. Companies that did not meet the criteria listed below have been dropped:

1. IIRC or IIRF must be mentioned specifically in the published Integrated Report between 2016 and 2021 to ensure that the Integrated Report analyzed adopts the Integrated Reporting principles.
2. The company must be listed on a European stock exchange between 2016 and 2021. Member states of the European Union are included, as well as firms from the United Kingdom and Switzerland, due to their IFRS adoption[2].
3. The company must be publicly listed and obtain a unique trading ID between 2016 and 2021.

The assessment resulted in 61 unique companies. 44 companies did not mention IIRC or the IIRF in their published Integrated Report, 8 companies were excluded as they were listed outside of the European market, 38 companies were excluded as they are not publicly listed, and 10 firms were duplicates or subsidiaries, and hence, did not obtain a unique trading ID.

[2] For robustness, we perform an additional regression solely focusing on EU firms, i.e., excluding the United Kingdom and Switzerland. See *Appendix* Table A.14-A.16

Table 1. Sample selection and distribution

Panel A: Tobin's Q Sample Selection	
Total number of observations fulfilling the selection criteria	287
Missing observations for Tobin's Q	24
Missing observations for control variables	28
Final number of observations Tobin's Q	*235*
Panel B: Bid-Ask Spread Sample Selection	
Total number of observations fulfilling the selection criteria	287
Missing observations for the Bid-ask spread	26
Missing observations for control variables	26
Final number of observations Bid-ask spread	**235**
Panel C: Integrated Reporting Sample Distribution by Report Type	
Integrated Report Stand-alone	28.92%
Integrated Report Intertwined	71.08%

This identification process generated 287 firm-year observations for the 61 firms between the year 2016 and 2021. We then collected the financial information from Capital IQ for our dependent and control variables. The final observations for each regression model and outcome variable after adjusting for missing data are presented in Table 1. Out of our initial 287 observations, we end up with 235 observations in Panel A, and 235 observations in Panel B, which corresponds to approximately 81%.

Table 2. Distribution of sample on public exchange (country level[3])

Country of exchange	Sample distribution
Netherlands	18.2%
Italy	17.5%
Spain	14.2%
United Kingdom	13.6%
France	7.6%
Germany	7.0%
Switzerland	6.0%
Finland	3.6%
Denmark	3.0%
Belgium	2.0%
Sweden	2.0%
Austria	2.0%
Total	*100.0%*

Sample demographics distribution is presented in Table 2. 18.2% of our sample is listed on the Dutch exchange. Integrated Reporting is well established in the Netherlands. To be precise, more than one third of Dutch listed companies prepared towards adopting IIRF in 2015 after Eumedion announced their encouragement for the framework. Eumedion is a member group of institutional investors in the Netherlands (International Integrated Reporting Council, 2015). It is also noted that 17.5% of our sample is represented by Italian listed firms. According to the IIRC, the Italian market obtains leading practical case studies of Integrated Reporting which drives the adoption of the framework among companies and practitioners (International Integrated Reporting Council, 2017a), which potentially explains the skewness towards Italian firms in our sample.

4.2 Measures of economic effects – dependent variables

The full list of variables and definitions is presented in Table 3. The market effects are measured both using economic effects proxied by Tobin's Q and market reactions proxied by the bid-ask spread.

Tobin's Q reflects firm value by comparing the market valuation of a firm's assets to their book value of total assets. By including market values of assets, Tobin's Q also incorporates the intangible value associated with aspects such as intellectual, human and social capital (Caglio et al., 2020). As in Caglio et al. (2020) and Barth et al. (2017), we calculate the market value

[3] Country fixed effects are controlled for in our main regressions based on "Country of Exchange" in alignment with Lang and Stice-Lawrence (2015) application of country fixed effects.

of equity on the release day of the Integrated Report in the current year to ensure that the information provided in the report is reflected in the firm value.

Asset liquidity cannot be directly observed which makes measuring complicated. Nonetheless, a commonly accepted approach in research identifies liquidity as the ability to buy or sell securities without a substantial impact on stock prices (Boubaker et al., 2019). In research, the bid-ask spread, and zero return days are the most common proxies for stock liquidity. Lang and Stice-Lawrence (2015) argue that zero return days is the preferred measure since it often generates the largest potential sample with less missing data. However, a risk of misclassification of zero-return days is evident, and zero trading activity may not necessarily be explained by stock liquidity, but could rather be due to market frictions. Henceforth, in alignment with Caglio et al. (2020) and Barth et al. (2017) we proxy stock liquidity using the bid-ask spread. The bid-ask spread is an inverse measure explaining information asymmetry. Barth et al. (2017) argue that investors are more willing to trade when information asymmetry is small, therefore lower bid-ask spread. We calculate the bid-ask spread variable as the natural logarithm of the median of the difference between the daily bid and ask prices divided by their midpoint measured the day after the previous year's Integrated Report release date until the current year's Integrated Report release date (Caglio et al., 2020; Barth et al., 2017).

4.3 Measures of Integrated Reporting textual attributes – independent variables

The Integrated Report for each firm-year is downloaded in PDF format from the official company website and thereafter imported to R Studio. The report is cleaned in R Studio and converted to text format by removing graphs, tables, numbers, symbols, URLs, and stop words as presented in Caglio et al. (2020), Li (2008) as well as Loughran and McDonald (2016).

This study considers two main aspects of textual attributes: length and reading difficulty. The research presented confirms that how information is presented and framed is important in the context of Integrated Reporting (Caglio et al., 2020). Within the IIRF a critical component is to present information in a concise manner, thus stressing the importance of expressing concepts in as few words as possible. Accounting research on textual attributes and Integrated Reporting has therefore included measures of length to capture conciseness as they are easily calculated and understood (Caglio et al., 2020; Caglio et al., 2017). The information processing cost is assumed to be higher in longer documents. We measure length as the number of sentences and number of characters in the released Integrated Report. The two variables (number of sentences and number of characters) are calculated in R Studio. A principal component factor analysis is presented in *Section 6.1 Factor Analysis* where the two variables for number of sentences and number of characters are combined into a factor for the purpose of simplifying the representation of length in Integrated Reports. This also mitigates the issue of multicollinearity with high correlation between the explanatory variables (Caglio et al., 2020; Lang & Stice-Lawrence 2015).

In addition, readability is promoted by the IIRF to facilitate understandable and transparent reporting. We measure reading difficulty using three different indices: the Flesch-Kincaid, the

Gunning-Fog Index, and the Smog Index. The construction of the readability indices was completed in R Studio utilizing the Quanteda package. As previously mentioned, the Flesch-Kincaid Index has received criticism due to its lack of fit to today's accounting information. Since the index is both old and based on school texts, recent studies have relied on the Gunning-Fog Index or the accounting specific index "Bog" instead. We chose to include both the Flesch-Kincaid Index and the Gunning-Fog Index, as well as the Smog Index used in Caglio et al. (2020), to avoid bias and skewness of using a single metric. Ideally, we would also have incorporated the Bog Index in our readability proxies, but since the Quanteda package does not include the construction of this index, this was not possible.

The Gunning-Fog Index indicates the number of years of formal education an average reader needs to understand the text with such word-sentence workload. The measure is based on the number of words per sentence and the percentage of complex words, that is, words with at least three syllables, also known as polysyllables. A higher percentage of complex words and longer sentences is associated with higher reading difficulty and thus a higher score (Caglio et al., 2020). The score is calculated as follows:

$$0.4 \cdot \left(\frac{total\ words}{total\ sentences}\right) + 100 \cdot \left(\frac{complex\ words}{total\ words}\right)$$

The Gunning-Fog Index formula

The Smog Index also measures readability. A higher score indicates that more years of education are required to grasp the content of the text. The Smog Index is also based on the number of polysyllables as the Gunning-Fog Index, however, the Smog Index is intended to simplify the readability calculation (Caglio et al., 2020). The Smog Index Formula is presented below:

$$3 + \sqrt{polysyllabic\ count}$$

The Smog Index formula

The Kincaid-Flesch Score Index is included as a readability proxy. The index scores the text according to US high school levels. The measure incorporates the ratio of words per sentence as well as the ratio of syllables per word. A higher score indicates a more difficult text (Caglio et al., 2020). The Kincaid-Flesch Score Index is calculated using the following formula:

$$0.39 \cdot \left(\frac{total\ words}{total\ sentences}\right) + 11.8 \cdot \left(\frac{total\ syllables}{total\ words}\right) - 15.59$$

The Kincaid-Flesch Score Index formula

All our calculations were compared with descriptive statistics presented in Caglio et al. (2020), Caglio et al. (2017) and Li (2008) to confirm the validity of our calculations. Again, the three readability indices are combined through a factor analysis to represent reading difficulty in Integrated Reports (see *Section 6.1 Factor analysis*).

However, one should be aware that readability formulas do not measure understandability. Critique has been addressed to literature using readability and understandability interchangeably, even though these two are not synonyms. Understandability incorporates the reader's knowledge, background, and reading skills. Readability in turn is solely related to the text and does not take personal factors into account (Jones & Shoemaker, 2011).

4.4 Control variables

In addition to our independent variables, there may be firm characteristics that explain the outcome. The selection of control variables follows mainly Caglio et al. (2020). See full variable definitions in Table 3. We control for return on assets, negative profit, accruals, leverage, book to market, firm size, beta, firm complexity and firm age, which will be further elaborated below.

We include two controls for firm performance as in Caglio et al. (2020): *return on assets* and *negative profit*, since firms with higher profitability are presumed to be higher valued by the market (Caglio et al., 2020; Li, 2008; Zhou et al., 2017; Barth et al., 2017; Boubaker et al., 2019).

We control for financial reporting quality (also referred to as earnings quality). Barth et al. (2017) argue that financial reporting quality is negatively associated with cost of capital and thereby positively associated with firm value. Additionally, stock liquidity is higher for firms with less earnings management. *Accruals* is the most used proxy for financial reporting quality. Since accruals are non-cash journal entries, they cause a gap between reported earnings and economic performance. We measure *accruals* as the difference between net income before extraordinary items and preference dividends and net cash flow from operating activities, scaled by total assets (Caglio et al., 2020; Barth et al., 2017; Lee & Yeo, 2016).

Controls for *leverage* and *book to market ratio* are included as these potentially impact investors willingness to trade (Barth et al., 2017). Control variables are also included to account for factors that affect the market value of equity. Hence, *firm size* proxied by the natural logarithm of total assets and market *beta* are controlled for. In particular, Boubaker et al. (2019)

assert that larger firms experience diminished information asymmetry, attributable to the increased availability of public information concerning the firm.

We also add controls for *firm complexity*. Lee and Yeo (2015) discuss that firms with greater complexity face greater information asymmetry as their operations are not fully understood by market participants. Complexity may apply to specific industries, such as environmentally sensitive sectors or sectors with advanced technical assets. Boubaker et al. (2019) additionally emphasize that industry influences may infer the relationship between financial disclosure readability and stock liquidity. Complexity can be proxied as the number of product segments (Caglio et al., 2020), the number of geographical segments (Li, 2008) or share of intangible assets (Lee and Yeo, 2015). Counting the number of segments has been criticized as firms can be active in several markets while still having a low-complex business model, such as the Coca-Cola Company (Loughran & McDonald, 2016). Consequently, we control for firm complexity by proxying the share of intangible assets.

In line with Caglio et al. (2017) and Li (2008) we also control for *firm age*. Research has shown that older firms face less information asymmetry compared to younger firms which potentially explain the variation in the bid-ask spread. Furthermore, it is argued that older firms' Integrated Reports are different from younger firms' Integrated Reports for this reason. The natural logarithm of firm age has been used due to the large values this metric yields, reducing the risk of heteroscedastic error (Newbold et al., 2013).

Table 3. Variable definitions

Variable	Definition
Dependent variables	
TobinsQ	Tobin's Q is the total assets minus total common equity plus common shares outstanding at the year-end multiplied by the share price at the release date of the current year's Integrated Report, divided by total assets (Caglio et al., 2020)
BidAsk	The bid-ask spread is the logarithm of the median of the difference between the daily closing bid and ask prices, divided by their midpoint of the two prices measured from the day after the release date of the prior year's Integrated Report on the release date of the current year's Integrated Report (Caglio et al., 2020)
Independent variables	
FleschKincaid	Flesch-Kincaid is a readability index computed through R Studio. Flesch-Kincaid Score = (11.8*syllables per word) + (0.39 *words per sentence) − 15.59. The Flesch-Kincaid Readability index rates texts by US grade school levels from 0-100, 0 being the easiest, and 100 the most difficult (Caglio et al., 2020).
FOG	Gunning Fog is a readability index computed through R Studio. Fog= (number of words per sentence + the percentage of complex words) * 0.4. The index indicates the number of years of formal education an average reader would need to read and understand the text with such a word-sentence workload. The index ranges from 0 to 20, 0 being the easiest and 20 the most difficult (Caglio et al., 2020).
SMOG	SMOG is a readability index computed through R Studio. Smog is the McLaughlin Readability index, which estimates the years of education needed to understand a text. SMOG grade = 3 + Square Root of Polysyllabic Count. The index ranges from 4-18, with 4 being the easiest, and 18 the most difficult (Caglio et al., 2020).
NoCharacters	The number of characters is computed in R Studio based on the Integrated Report
NoSentences	The number of sentences is computed in R Studio based on the Integrated Report
ReadingDifficulty	Factor computed in STATA combining the readability variables: FleschKincaid, FOG, and SMOG.
Length	Factor computed in STATA combining the length variables: NoCharacters and NoSentences
Control variables	
ROA	Return on assets is the ratio of net income before extraordinary items to prior year’s total assets (Caglio et al., 2020)
NegativeProfit	Negative profit is a dummy variable generating "1" if net income before extraordinary items is negative and “0” if positive (Caglio et al., 2020)

Accruals	Accruals is the difference between net income before extraordinary items and preference dividends and net cash flow from operating activities, scaled by total assets. An inverse proxy for financial reporting quality (Caglio et al., 2020)
DtoCE	Debt to common equity, or leverage, is the ratio of total debt to sum of total debt and book value of common equity (Caglio et al., 2020)
Beta	Beta is the market beta at the end of the fiscal year as a proxy for market sensitivity (Caglio et al., 2020)
BTM	The book-to-market ratio is the ratio of book value of common equity divided by the number of common shares outstanding, multiplied by the year-end share price (Caglio et al., 2020)
LogAssets	The logarithm of total assets as a proxy for firm size (Caglio et al., 2020)
Complexity	Complexity is measured as intangible assets divided by total assets (Lee & Yeo, 2016)
LogFirmAge	The logarithm of the number of years from the date of first incorporation (Caglio et al., 2017)
Interaction term models	
ReadingDifficulty	Factor computed in STATA combining the readability variables: FleschKincaid, FOG, and SMOG. In the interaction term models this variable is only represented by the reading difficulty of stand-alone Integrated Reports
ReadingDifficulty *IR_ITW	An interaction term representing the moderating effect of releasing an intertwined Integrated Report on the relationship between reading difficulty and market effects
Length	Factor computed in STATA combining the length variables: NoCharacters and NoSentences. In the interaction term models this variable is only represented by the reading difficulty of stand-alone Integrated Reports
Length*IR_ITW	An interaction term representing the moderating effect of releasing an intertwined Integrated Report on the relationship between length and market effects
IR_ITW	IR_ITW is a dummy variable that generates "1" if the firm issues an intertwined Integrated Report and “0” if the firm issues a stand-alone Integrated Report

4.5 Principal component factor analysis

As mentioned above, the three readability indices (Flesch-Kincaid Index, Gunning-Fog Index, and Smog Index) and the two length variables (number of characters and number of sentences) are combined into two separate factors to avoid multicollinearity. This is achieved through utilizing the principal component analysis. By doing this, the proxies for reading difficulty and length can be estimated based on multiple input variables rather than limiting our selection to one measure. Therefore, we achieve more robust results (Stock & Watson, 2020). A principal components analysis with k variables (i.e., our three readability indices and two length measures) can be explained as:

"The principal components of the k variables $X_1, \ldots, X_k$ *are the linear combinations of those variables that are mutually uncorrelated, have squared weights that sum to 1, and maximize the variance of the linear combination controlling for the previous principal components"* (Stock & Watson, 2020, p.534).

To ensure that our data set of readability indices and length proxies is adequate to perform principal component factor analysis on, both the Bartlett's test of sphericity and Kaiser-Meyer-Olkin (KMO) is performed (Caglio et al., 2020; Lang & Stice-Lawrence, 2015; Li, 2008). Bartlett's test of sphericity tests whether the variables are uncorrelated. The KMO measures sampling adequacy where the results should be above 0.5. Both criteria were fulfilled. Detailed results for Bartlett's test and KMO are presented in Appendices A.1 and A.2 respectively. Furthermore, factor loadings are rotated to obtain a clearer pattern (Al Amin & Qin, 2023; Caglio et al., 2020). Based on these factors, the new variables $ReadingDifficulty$ and $Length$ could adequately be constructed in Stata (see *Section 6.1 Factor analysis* for factor loadings).

4.7 Models

4.7.1 Model Hypothesis I

To test Hypothesis I, i.e., the relationship between market effects and textual attributes of Integrated Reports (reading difficulty and length), we estimate the following model:

Model 1

$$Market\ Effects_{it} = \beta_0 + \beta_1 ReadingDifficulty_{it} + \beta_2 Length_{it} + \beta_3 ROA + \beta_4 NegativeProfit_{it} + \beta_5 Accruals_{it} + \beta_6 DtoCE_{it} + \beta_7 Beta_{it} + \beta_8 BTM_{it} + \beta_9 LogAssets_{it} + \beta_{10} Complexity_{it} + \beta_{11} \text{Log}FirmAge_{it} + \varepsilon_{it}$$

The market effects are operationalized as Tobin's Q (proxy for firm valuation) and bid-ask spread (proxy for stock liquidity) and are therefore separated into two regressions. From here on, when market effects are proxied by Tobin's Q the Model is referred to as 1a and when instead proxied by bid-ask spread the Model is labeled 1b. Both models include control variables (NegativeProfit, Accruals, DtoCE, Beta, BTM, LogAssets, Complexity, and

LogFirmAge), clustered standards errors by firm (to control for heteroscedasticity and autocorrelation in error terms[4]), various fixed effect structures (no fixed effects structure, year, country, and industry fixed effects[5]). This paper utilizes both year, country, and industry fixed effects as the main structure to account for cross-sectional variation (on a year, country and industry level) that is not captured by the model's explanatory variables. This is important as our sample includes several countries and industries who are exposed to different regulatory settings and market conditions that potentially cause variation in firms' market effects not explained by our explanatory variables.

Reading difficulty and length are included in the same model despite the fact that both variables aim to proxy the quality of textual attributes of Integrated Reports[6]. To ensure that this does not cause multicollinearity, both the Pearson correlation test (see *Section 5.2 Correlation table*) and the Spearman correlation test (Appendix A.5) are performed on each model's independent variables where a threshold of 0.7 is considered as the maximum level of allowed correlation. In addition, a principal component factor analysis is performed (See Table 6) to ensure that the underlying proxies are included in appropriate factors.

4.7.2 Model Hypothesis II

A second model is established to test the Hypothesis II, i.e., whether issuing an intertwined Integrated Report has a moderating effect on the association between textual attributes and market effects. To examine this, an interaction term is utilized which help to investigate whether the association between textual attributes (reading difficulty and length) and market effects (Tobin's Q and bid-ask spread) changes depending on the Integrated Report's disclosure type, i.e., stand-alone or intertwined. If the Integrated Report is released intertwined, the dummy variable IR_ITW receives a "1", and if the report is released stand-alone a "0" is allocated. Therefore, our model categorizes Integrated Reports into two different groups depending on their disclosure method. To simplify the economic interpretation of the interaction terms, we construct two different models that separate reading difficulty and length.

Model 2 includes $ReadingDifficulty$ and Model 3 includes $Length$. The interaction term of Model 2 is presented as $\beta_2 ReadingDifficulty_{it} * IR\_ITW$, which represents the reading difficulty of intertwined Integrated Reports, whilst $\beta_1 ReadingDifficulty_{it}$ represent the reading difficulty of stand-alone Integrated Reports. Similarly, the interaction term of Model 3

[4] See *Appendix* "OLS Assumptions" for the full list of OLS assumptions and corresponding tests for heteroscedasticity, autocorrelation, multicollinearity and non-normality.

[5] The Stata package "Sumdfe" is used to ensure that a sufficient level (minimum 40%) of the model's within variation is kept after applying each reported fixed effect structure.

[6] As robustness, we re-run the models with reading difficulty and length separated in two different regressions. Results are presented in Appendix A.8-A.11.

is labeled as $\beta_2 Length_{it} * IR\_ITW$, which represents the length of an intertwined Integrated Reports and $\beta_1 Length$ represent the length of stand-alone Integrated Reports. See Table 3 for complete variable definitions.

Model 2

$$Market\ Effects_{it} = \beta_0 + \beta_1 ReadingDifficulty_{it} + \beta_2 ReadingDifficulty_{it} * IR_{ITW_{it}} + \beta_3 IR_{ITW_{it}} + \beta_4 ROA + \beta_5 NegativeProfit_{it} + \beta_6 Accruals_{it} + \beta_7 DtoCE_{it} + \beta_8 Beta_{it} + \beta_9 BTM_{it} + \beta_{10} LogAssets_{it} + \beta_{11} Complexity_{it} + \beta_{12} \text{Log} FirmAge_{it} + \varepsilon_{it}$$

Model 3

$$Market\ Effects_{it} = \beta_0 + \beta_1 Length + \beta_2 Length_{it} * IR_{ITW_{it}} + \beta_3 IR_{ITW_{it}} + \beta_4 ROA + \beta_5 NegativeProfit_{it} + \beta_6 Accruals_{it} + \beta_7 DtoCE_{it} + \beta_8 Beta_{it} + \beta_9 BTM_{it} + \beta_{10} LogAssets_{it} + \beta_{11} Complexity_{it} + \beta_{12} \text{Log} FirmAge_{it} + \varepsilon_{it}$$

Once again, the market effects in Model 2 and Model 3 are operationalized as Tobin's Q (proxy for firm valuation) and bid-ask spread (proxy for stock liquidity). From here on, Model 2 is labeled as Model 2a when the tested dependent variable is Tobin's Q and as Model 2b when the tested dependent variable is Bid-ask spread. Using the same logic, Model 3 is referred to as Model 3a when including Tobin's Q, and as Model 3b when including bid-ask spread. The control variables as well as fixed effect structures are the same as in Model 1. Clustered standards errors at firm level are applied to account for heteroscedasticity and autocorrelation in error terms[7]. The main fixed effect structure consists of industry, year, and country[8] (Lang & Stice-Lawrence, 2015; Caglio et al., 2020).

[7] See *Appendix* "OLS Assumptions" for the full list of OLS assumptions and corresponding tests for heteroscedasticity, autocorrelation, multicollinearity and non-normality.

[8] The Stata package "Sumdfe" is used to ensure that a sufficient level (minimum 40%) of the model's within variation is kept after applying each reported fixed effect structure.

5. Descriptive statistics

5.1 Summary statistics

Table 4. Descriptive statistics of explanatory variables

Variable	Obs	Mean	Std.Dev.	Min.	Max.
FleschKincaid	287	26.950	3.667	15.740	42.520
FOG	287	31.840	3.835	18.900	46.650
SMOG	287	25.640	2.735	16.620	33.460
NoCharacters	285	490,775.000	311,911.000	52,535.000	1.741e+06
NoSentences	285	1,597.000	1,187.000	147.000	9623.000
TobinsQ	263	1.418	0.980	0.457	11.920
BidAsk	261	-7.200	1.432	-9.150	-1.298
ROA	275	0.021	0.087	-0.582	0.320
NegativeProfit	287	0.185	0.389	0.000	1.000
Accruals	278	-0.055	0.181	-2.015	0.690
DtoCE	278	0.489	38.290	-509.400	26.800
Beta	287	0.991	0.553	-0.101	4.427
BTM	273	0.810	0.791	-0.713	6.242
LogAssets	287	9.355	2.810	-0.113	14.770
Complexity	254	0.178	0.165	0.000	0.654
LogFirmAge	287	3.970	1.070	0.000	5.922
IR_ITW	287	0.711	0.454	0.000	1.000

The number of observations, means, standard deviations, minimum, and maximum values of our main variables are presented in Table 4. Overall, the presented values are in line with descriptive statistics presented in Caglio et al. (2020), Barth et al. (2017), and Zhou et al. (2017).

The mean values on readability indices are reported as follows: $FleschKincaid$ 26.950, FOG 31.840 and $SMOG$ 25.640, implying that the Integrated Reports are difficult to read in general. Our mean values slightly exceed those reported in South African settings (Caglio et al., 2020; Caglio et al., 2017), though the standard deviation is smaller, indicating lower variation in Integrated Reporting readability for European companies.

The Integrated Reports are lengthy with an average $NoCharacters$ equal to 490,775 (or 1,597 $NoSentences$) which again exceeds the South African setting in Caglio et al. (2020). The standard deviation is higher in our sample (std. dev. 311,911) which may be explained by the variation of stand-alone and intertwined Integrated Reports in the European setting where stand-alone reports on average are shorter than intertwined Integrated Reports.

The outcome variables $TobinsQ$ and $BidAsk$, have mean values of 1.418 and -7.200 respectively, and levels with values presented in Caglio et al. (2020) and Barth et al. (2017).

The same logic is applicable for the control variables: ROA, $Accruals$, $DtoCE$, $Beta$, BTM, $LogAssets$, $Complexity$, $and\ LogFirmAge$.

5.2 Correlation table

The Pearson Correlation Matrix (See Table 5) shows correlation between the explanatory variables. No significant multicollinearity is identified as all correlations between explanatory variables are less than 0.5. Additionally, the two main independent variables, $ReadingDifficulty$ and $Length$, are uncorrelated and therefore presumed to be included in the same regression model as independent variables.

The strongest correlation is identified between book to market and the $LogAssets$ (0.473) at a significance level of <0.001. Strong correlations are also identified between $Length$ and IR_ITW (0.457), suggesting that intertwined Integrated Reports are longer. A strong correlation between $Accruals$ and ROA (-0.457) is also identified, implying that firms with a low return on assets have higher accruals. Similar findings are discovered in the Spearman correlation test (see Appendix A.5). Regardless of these correlations, we base our results on the multivariate regression presented in *Section 6 Results*.

Table 5. Pearson correlation test

	ReadingDifficulty	Length	ROA	NegativeProfit	Accruals	DtoCE	Beta	BTM	LogAssets	Complexity	LogFirm Age	IR_ITW
ReadingDifficulty	1											
Length	0.008	1										
ROA	-0.074	-0.188**	1									
NegativeProfit	0.157*	0.063	-0.422***	1								
Accruals	0.044	-0.061	-0.457***	0.044	1							
DtoCE	-0.043	-0.062	-0.045	0.029	0.008	1						
Beta	0.272***	-0.026	-0.292***	0.293***	0.318***	-0.058	1					
BTM	-0.010	-0.039	-0.307***	0.060	0.268***	0.187**	0.210***	1				
LogAssets	-0.129*	0.012	-0.032	-0.187**	0.083	0.039	-0.157*	0.473***	1			
Complexity	0.086	0.082	0.189**	-0.173**	-0.222***	-0.013	-0.253***	-0.430***	-0.197**	1		
LogFirmAge	-0.334***	-0.041	0.189**	-0.261***	-0.148*	-0.041	-0.306***	-0.016	0.183**	0.126*	1	
IR_ITW	0.035	0.462***	0.123	-0.077	-0.155*	0.132*	-0.104	-0.138*	-0.145*	0.123	-0.0521	1

*** p<0.01, ** p<0.05, * p<0.1

6. Results

6.1 Factor analysis

Table 6. Principal component factor analysis and rotated factor loadings of textual attributes

	Factor pattern			Factor pattern (rotated)		
Variable	Factor 1	Factor 2	Unique-ness	Factor 1	Factor 2	Unique-ness
FleschKincaid	0.978	0.162	0.018	0.994	-0.038	0.010
SMOG	0.966	0.162	0.040	0.989	-0.057	0.018
FOG	0.978	0.182	0.010	0.978	-0.055	0.040
NoCharacters	-0.116	0.983	0.020	0.103	0.985	0.020
NoSentences	-0.424	0.894	0.021	-0.217	0.965	0.021
Eigenvalue	3.040	1.851				

A principal component factor analysis is performed to obtain a parsimonious measure of our textual attributes, as mentioned in the methodology section. The unrotated factor loadings and the rotated factor loadings are presented in Table 6. The unrotated factor loadings can show cross-loadings and include high loadings across factors, we therefore rotate the factor loading structure to minimize the variable loading to each factor. Factor rotations are performed using the varimax method (Caglio et al., 2020). For both loadings, Factor 1 is strongly influenced by readability indices and Factor 2 by length measures.

The unrotated loadings in Factor 1 are mostly represented by the $FleschKincaid$, $SMOG$ and FOG since all variables exceed a loading of 0.96. A lower uniqueness implies that the variables are explaining unique variance; hence, since all variables obtain a uniqueness of less than 0.04, they are all presumed relevant to the factor model. After constructing the factors through the principal component factor analysis, the Eigenvalue of each factor is controlled to ensure that the threshold of minimum 1.0 is achieved. This is critical since a value below 1.0 indicates that the factors do not explain a sufficient level of variance and should be excluded. When rotating the factor loadings, the loadings increased (> 0.97).

The unrotated loadings in Factor 2 are mostly represented by the $NoCharacters$ and $NoSentences$ where the loadings of both variables exceeded 0.89. Since the uniqueness of Factor 2 is small (<0.03) the two variables are highly relevant for the factor loading. The Eigenvalue threshold is fulfilled (>1.0). When rotating the factor loadings, the loadings increased (>0.96).

6.2 Multivariate analysis

6.2.1 Hypothesis I

We separate the two market effects: firm value and stock liquidity. Hence, the results of the main OLS regressions are presented in Table 7 and Table 8 for Tobin's Q and the bid-ask spread respectively. Tobin's Q is intended to reflect the market value of a firm's assets in comparison to its book values, and the bid-ask spread is intended to capture the willingness to trade following information asymmetry (Barth et al., 2017). Therefore, in Table 7 we test Model 1a whether high-quality textual attributes are associated with higher firm value. Correspondingly, in Table 8 we test Model 1b, whether high-quality textual attributes are associated with higher stock liquidity[9]. High-quality textual attributes are measured as Integrated Reports with low reading difficulty and less length i.e., readability and conciseness. In the following paragraphs, the results from applying the main fixed effect structure (industry, country, and year) are foremost discussed.

[9] Three robustness tests are performed for Model 1a and Model 1b, respectively. The first robustness test substitutes factors with their underlying indices/proxies, the second robustness test only includes EU firms and the third replaces missing values with median values. Results are in general aligned with Table 7 and Table 8 where deviations are further discussed under A.8-A.11, A.14 and A.18.

Table 7. Model 1a: OLS regression of Tobin's Q and textual attributes

	(1) TobinsQ	(2) TobinsQ	(3) TobinsQ	(4) TobinsQ	(5) TobinsQ
ReadingDifficulty	0.018	0.020	0.012	0.033	0.024
	(0.44)	(0.44)	(0.33)	(0.66)	(0.54)
Length	0.031	-0.054	0.017	0.002	-0.055
	(0.86)	(-1.08)	(0.63)	(0.04)	(-1.12)
ROA	0.046	0.041	0.052	0.033	0.040
	(2.41)**	(3.56)***	(2.90)***	(1.45)	(3.55)***
NegativeProfit	0.135	0.116	0.187	0.231	0.144
	(2.34)**	(0.82)	(3.03)***	(1.61)	(1.03)
Accruals	0.812	2.012	0.774	1.088	1.841
	(1.24)	(1.96)*	(1.38)	(1.39)	(1.88)*
DtoCE	0.001	0.001	0.001	-0.000	0.001
	(1.84)*	(1.67)	(2.07)**	(-0.36)	(1.67)
Beta	-0.215	-0.338	-0.229	-0.331	-0.361
	(-1.32)	(-2.71)***	(-2.16)**	(-3.35)***	(-2.99)***
BTM	-0.110	-0.084	-0.105	-0.045	-0.073
	(-1.60)	(-1.16)	(-1.72)*	(-0.71)	(-1.01)
LogAssets	-0.135	-0.100	-0.134	-0.075	-0.102
	(-3.85)***	(-2.79)***	(-5.15)***	(-1.86)*	(-2.92)***
Complexity	1.616	1.222	0.797	1.321	1.217
	(2.00)*	(2.55)**	(1.67)	(2.97)***	(2.55)**
LogFirmAge	0.198	0.102	0.146	0.012	0.103
	(2.57)**	(1.30)	(2.60)**	(0.19)	(1.32)
Constant	1.797	2.097	2.126	2.134	2.118
	(3.04)***	(4.84)***	(4.77)***	(4.19)***	(4.99)***
Year FE	Yes	Yes	No	No	No
Industry FE	Yes	No	Yes	No	No
Country FE	Yes	No	No	Yes	No
Observations	232	235	232	235	235
Adj. R-squared	0.849	0.445	0.832	0.619	0.453

Robust t-statistics in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Note: Clustered standard errors at firm level are applied to account for heteroscedasticity and autocorrelation in error terms. See Table 3 for variable definitions.

In Table 7(1) between high-quality textual attributes and firm value, we find a statistically insignificant association between $ReadingDifficulty$ and $TobinsQ$ (0.018, t=0.44) when controlling for both year, industry, and country fixed effects. This result implies that complex Integrated Reports do not affect firm value. The outcome is also statistically insignificant when controlling for only year fixed effects, industry fixed effects, country fixed effects and no fixed effects separately.

Our findings conflict with findings on readability and Tobin's Q presented in Caglio et al. (2020). The authors find a negative and significant association between Tobin's Q and the level of reading difficulty, implying that a lower reading difficulty indeed impacts firm value

positively. The contradictory results may be explained by our voluntary setting where firms who voluntarily issue Integrated Reports potentially are of similar nature, i.e., with less variation compared to the sample investigated in Caglio et al. (2020).

Our results on firm valuation further suggest that investors are indifferent to the conciseness of the Integrated Report. We find a statistically insignificant relationship between $TobinsQ$ and the $Length$ of the Integrated Report when controlling for both year, industry and country fixed effects (0.031, t=0.86). The sign of the coefficient changes when only year and no fixed effects are applied. This suggests there is variation driven by country and industry effects, and when not controlling for these effects cause biased coefficients.

The adjusted r-squared when controlling for both year, industry and country fixed effects (0.849) exceeds both those presented in Caglio et al. (2020) and Barth et al. (2017). To conclude, no evidence for Hypothesis I is identified.

Table 8. Model 1b: OLS regression of bid-ask spread and textual attributes

	(1) BidAsk	(2) BidAsk	(3) BidAsk	(4) BidAsk	(5) BidAsk
ReadingDifficulty	0.110	-0.295	-0.412	-0.212	-0.293
	(1.40)	(-1.53)	(-2.55)**	(-1.70)*	(-1.60)
Length	0.130	0.055	0.080	-0.100	0.051
	(1.66)	(0.49)	(0.70)	(-1.63)	(0.48)
ROA	-0.060	-0.040	-0.064	-0.089	-0.040
	(-2.52)**	(-3.28)***	(-2.98)***	(-3.70)***	(-3.17)***
NegativeProfit	-0.122	-0.165	-0.113	-0.297	-0.160
	(-0.70)	(-0.66)	(-0.70)	(-1.65)	(-0.65)
Accruals	-2.242	0.198	-1.269	-1.052	0.026
	(-2.33)**	(0.10)	(-0.95)	(-0.75)	(0.01)
DtoCE	0.000	0.001	0.001	0.000	0.001
	(0.40)	(0.71)	(0.84)	(0.43)	(0.88)
Beta	-0.019	-0.199	-0.419	-0.297	-0.197
	(-0.11)	(-0.92)	(-1.80)*	(-1.87)*	(-0.90)
BTM	0.106	0.248	0.007	0.296	0.243
	(0.90)	(1.46)	(0.04)	(2.11)**	(1.48)
LogAssets	-0.332	-0.364	-0.301	-0.367	-0.362
	(-3.65)***	(-4.95)***	(-2.60)**	(-6.26)***	(-4.91)***
Complexity	1.888	-1.133	0.272	-0.925	-1.133
	(2.48)**	(-1.78)*	(0.22)	(-1.45)	(-1.76)*
LogFirmAge	0.191	-0.265	-0.353	-0.156	-0.264
	(1.08)	(-1.24)	(-1.31)	(-1.01)	(-1.26)
Constant	-5.249	-2.350	-2.431	-2.618	-2.375
	(-4.91)***	(-1.95)*	(-1.53)	(-2.91)***	(-1.95)*
Year FE	Yes	Yes	No	No	No
Industry FE	Yes	No	Yes	No	No
Country FE	Yes	No	No	Yes	No
Observations	232	235	232	235	235
Adj. R-squared	0.849	0.454	0.656	0.651	0.461

Robust t-statistics in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Note: Clustered standard errors at firm level are applied to account for heteroscedasticity and autocorrelation in error terms. See Table 3 for variable definitions.

In Table 8(1), we find a positive insignificant association between $ReadingDifficulty$ and $BidAsk$ (0.110, t=1.40) when controlling for year, industry, and country fixed effects. This suggests that investors are indifferent about the readability of the Integrated Report. These findings align with those reported in Caglio et al. (2020). When only industry and country fixed effects are applied alone, the coefficient changes direction and turns statistically significant. This could potentially be due to an endogeneity issue where unexplained variation causes bias in the coefficient and statistical level. In turn, we do not consider these coefficients as 'valid'.

We find no statistically significant results in Table 8(1) between the $Length$ of the Integrated Report and $BidAsk$ (0.130, t=1.66). This suggests that investors are indifferent about the conciseness of the Integrated Report. These results are robust also when controlling for all selected fixed effect structures. Caglio et al. (2020) on the other hand statistically significantly prove that investors appreciate concise Integrated Reports. Mixed findings could be explained by variations in examined market, time-period and regulatory environment, see *Section 7.1 Limitations* for extensive discussion.

Reported adjusted R-squared is higher than those presented in Caglio et al. (2020) and Barth et al. (2017). To summarize, no evidence for Hypothesis I is identified.

6.2.2 Hypothesis II

The results from separating Model 2 on $TobinQ$ (Model 2a) and $BidAsk$ (Model 2b) are presented in Table 9 and Table 10. The interaction term of focus is $ReadingDifficulty * IR\_ITW$ which examines if the association between reading difficulty and market effects is moderated if the Integrated Reports is released intertwined[10].

The results from separating Model 3 on $TobinQ$ (Model 3a) and $BidAsk$ (Model 3b) are presented in Table 11 and Table 12. The interaction term of focus is $Length * IR\_ITW$ which examines if the association between length and market effects is moderated if the Integrated Reports is released intertwined[11].

[10] Two robustness tests are performed for Model 2a and Model 2b, respectively. The first robustness test only includes EU firms and the second replaces missing values with median values. Results are in general aligned with Table 9 and Table 10 where deviations are further discussed under A.15 and A.19.

[11] Two robustness tests are performed for Model 3a and Model 3b, respectively. The first robustness test only includes EU firms and the second replaces missing values with median values. Results are in general aligned with Table 11 and Table 12 where deviations are further discussed under A.16 and A.20.

Table 9. Model 2a: OLS regression of Tobin's Q and reading difficulty interaction term

	(1) TobinsQ	(2) TobinsQ	(3) TobinsQ	(4) TobinsQ	(5) TobinsQ
ReadingDifficulty	0.068	-0.032	0.080	-0.003	-0.027
	(0.94)	(-0.58)	(1.30)	(-0.05)	(-0.51)
ReadingDifficulty *IR_ITW	-0.075	0.085	-0.081	0.062	0.082
	(-1.09)	(1.18)	(-1.13)	(0.68)	(1.15)
IR_ITW	0.191	0.062	0.219	0.266	0.053
	(1.13)	(0.52)	(1.42)	(1.80)*	(0.44)
ROA	0.042	0.044	0.047	0.035	0.044
	(2.34)**	(3.71)***	(2.75)***	(1.65)	(3.71)***
NegativeProfit	0.151	0.133	0.213	0.309	0.164
	(2.31)**	(1.03)	(2.86)***	(2.15)**	(1.28)
Accruals	0.803	2.228	0.751	1.161	2.031
	(1.28)	(2.12)**	(1.38)	(1.43)	(2.01)**
DtoCE	0.000	0.001	0.001	-0.001	0.001
	(1.05)	(1.64)	(1.30)	(-1.10)	(1.70)*
Beta	-0.205	-0.345	-0.192	-0.319	-0.366
	(-1.20)	(-3.10)***	(-1.56)	(-3.03)***	(-3.36)***
BTM	-0.118	-0.079	-0.109	-0.034	-0.066
	(-1.69)*	(-1.22)	(-1.90)*	(-0.63)	(-1.03)
LogAssets	-0.114	-0.097	-0.113	-0.062	-0.099
	(-2.78)***	(-2.76)***	(-3.09)***	(-1.70)*	(-2.88)***
Complexity	1.428	1.201	0.673	1.318	1.199
	(2.22)**	(2.56)**	(1.66)	(3.15)***	(2.57)**
LogFirmAge	0.233	0.098	0.189	0.011	0.098
	(2.53)**	(1.16)	(2.72)***	(0.16)	(1.17)
Constant	1.348	2.039	1.583	1.789	2.067
	(1.64)	(4.28)***	(2.25)**	(3.12)***	(4.39)***
Year FE	No	Yes	No	No	No
Industry FE	No	No	Yes	No	No
Country FE	No	No	No	Yes	No
Observations	232	235	232	235	235
Adj. R-squared	0.860	0.442	0.843	0.641	0.449

Robust t-statistics in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Note: Clustered standard errors at firm level are applied to account for heteroscedasticity and autocorrelation in error terms. See Table 3 for variable definitions.

As presented in Table 9(1), the association between $ReadingDifficulty$ and $TobinsQ$ is statistically insignificant (0.068, t=0.94). Additionally, there is no moderating effect of releasing an intertwined Integrated Report on the relationship between reading difficulty and

market valuation since the interaction term $ReadingDifficulty * IR\_ITW$ (-0.075, t=-1.09) is insignificant. The dummy variable IR_ITW (0.191, t=1.13) is also insignificant which indicates that there is no difference in average firm value between firms releasing an intertwined versus firms releasing a stand-alone Integrated Report. The results are robust when controlling for all selected fixed effects structures (see Table 9(2)-9(5)). Isolating the findings in Table 9, there is no support for Hypothesis II.

Table 10. Model 2b: OLS regression of bid-ask spread and reading difficulty interaction term

	(1) BidAsk	(2) BidAsk	(3) BidAsk	(4) BidAsk	(5) BidAsk
ReadingDifficulty	0.263	-0.056	-0.091	-0.045	-0.047
	(2.22)**	(-0.64)	(-0.79)	(-0.51)	(-0.56)
ReadingDifficulty *IR_ITW	-0.415	-0.378	-0.677	-0.258	-0.386
	(-3.55)***	(-1.77)*	(-3.00)***	(-2.02)**	(-1.84)*
IR_ITW	-0.319	0.007	-0.171	-0.299	0.008
	(-2.75)***	(0.03)	(-0.84)	(-1.72)*	(0.03)
ROA	-0.052	-0.047	-0.049	-0.079	-0.047
	(-2.42)**	(-2.96)***	(-2.02)**	(-3.62)***	(-2.91)***
NegativeProfit	-0.218	-0.144	-0.187	-0.311	-0.147
	(-1.12)	(-0.58)	(-1.13)	(-1.87)*	(-0.60)
Accruals	-2.384	0.287	-1.336	-0.494	0.180
	(-2.91)***	(0.14)	(-1.15)	(-0.35)	(0.09)
DtoCE	0.001	0.001	0.001	0.001	0.001
	(1.46)	(0.95)	(1.34)	(1.77)*	(1.11)
Beta	-0.051	-0.160	-0.378	-0.263	-0.161
	(-0.28)	(-0.87)	(-2.09)**	(-1.91)*	(-0.86)
BTM	0.071	0.231	-0.031	0.302	0.226
	(0.65)	(1.36)	(-0.21)	(2.26)**	(1.39)
LogAssets	-0.314	-0.361	-0.290	-0.394	-0.360
	(-4.67)***	(-4.21)***	(-3.26)***	(-6.83)***	(-4.21)***
Complexity	1.614	-1.160	0.545	-0.971	-1.170
	(2.12)**	(-1.80)*	(0.48)	(-1.57)	(-1.81)*
LogFirmAge	0.259	-0.221	-0.228	-0.088	-0.219
	(1.76)*	(-1.27)	(-1.10)	(-0.70)	(-1.28)
Constant	-5.405	-2.546	-3.042	-2.458	-2.566
	(-7.14)***	(-2.24)**	(-2.44)**	(-3.10)***	(-2.21)**
Year FE	Yes	Yes	No	No	No
Industry FE	Yes	No	Yes	No	No
Country FE	Yes	No	No	Yes	No
Observations	232	235	232	235	235
Adj. R-squared	0.867	0.477	0.702	0.670	0.485

Robust t-statistics in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Note: Clustered standard errors at firm level are applied to account for heteroscedasticity and autocorrelation in error terms. See Table 3 for variable definitions.

As presented in Table 10(1), the positive association between $ReadingDifficulty$ and $BidAsk$ is statistically significant (0.263, t=2.22) at a 5% level. Since this explanatory variable is presented by the group of Integrated Reports that are released stand-alone it implies that

investors prefer them to be readable. The interaction term $ReadingDifficulty * IR\_ITW$ (-0.415, t=-3.55), which represents the group of intertwined Integrated Reports, is negatively associated with $BidAsk$ on a 1% significance level. The total association between reading difficulty and bid-ask spread of intertwined Integrated Reports is determined by adding the coefficients of $ReadingDifficulty$ and $ReadingDifficulty * IR\_ITW$. Combining these coefficients yields -0.152 which indicates that a one-unit increase in reading difficulty reduces bid-ask spread for intertwined Integrated Reports by -0.152. Since IR_ITW is negatively associated with $BidAsk$ (-0.319, t=-2.75) on a 1% significance level we show that firms issuing an intertwined Integrated Report on average have a 0.319 lower bid-ask spread than Integrated Reports issued stand-alone.

When controlling for the selected fixed effect structures, as shown in Table 10(2)-10(5), the results vary. This is not surprising as omitting any or all the main fixed effects (year, industry, and country) could potentially lead to endogeneity issues creating biases in signs and significance. Studying Table 10(1) in isolation and combining the insights from the variables $ReadingDifficulty$, $ReadingDifficulty * IR\_ITW$, and IR_ITW, we find support for Hypothesis II. We can conclude that firms releasing an intertwined Integrated Report with high reading difficulty benefits in terms of higher stock liquidity. On the other hand, firms that release stand-alone Integrated Reports with high reading difficulty are punished by investors in terms of lower stock liquidity. Thus, there is a moderating effect of releasing an intertwined Integrated Reports on the relationship between reading difficulty and bid-ask spread.

Table 11. Model 3a: OLS regression of Tobin's Q and length interaction term

	(1) TobinsQ	(2) TobinsQ	(3) TobinsQ	(4) TobinsQ	(5) TobinsQ
Length	0.293	0.129	0.376	0.277	0.133
	(1.99)*	(0.94)	(3.52)***	(1.89)*	(0.99)
Length *IR_ITW	-0.338	-0.255	-0.475	-0.373	-0.256
	(-1.99)*	(-1.63)	(-3.40)***	(-2.11)**	(-1.66)
IR_ITW	0.050	0.027	0.019	0.152	0.014
	(0.40)	(0.17)	(0.14)	(1.18)	(0.09)
ROA	0.048	0.039	0.049	0.043	0.038
	(2.57)**	(3.22)***	(3.06)***	(2.07)**	(3.21)***
NegativeProfit	0.154	0.109	0.175	0.295	0.136
	(2.28)**	(0.80)	(2.66)**	(2.15)**	(1.00)
Accruals	0.963	2.215	0.951	1.232	2.047
	(1.45)	(2.18)**	(1.67)	(1.53)	(2.10)**
DtoCE	0.001	0.001	0.001	0.000	0.001
	(1.66)	(1.57)	(2.31)**	(0.00)	(1.63)
Beta	-0.221	-0.327	-0.277	-0.278	-0.354
	(-1.31)	(-2.72)***	(-2.36)**	(-2.84)***	(-2.95)***
BTM	-0.092	-0.071	-0.079	0.002	-0.058
	(-1.45)	(-1.03)	(-1.52)	(0.03)	(-0.85)
LogAssets	-0.092	-0.091	-0.059	-0.062	-0.094
	(-2.13)**	(-2.47)**	(-1.51)	(-1.60)	(-2.60)**
Complexity	1.489	1.325	0.882	1.413	1.322
	(2.24)**	(2.75)***	(1.99)*	(3.45)***	(2.75)***
LogFirmAge	0.207	0.098	0.182	-0.002	0.096
	(3.00)***	(1.46)	(3.29)***	(-0.03)	(1.44)
Constant	1.375	2.040	1.347	1.886	2.089
	(1.89)*	(4.11)***	(2.04)**	(3.54)***	(4.26)***
Year FE	Yes	Yes	No	No	No
Industry FE	Yes	No	Yes	No	No
Country FE	Yes	No	No	Yes	No
Observations	232	235	232	235	235
Adj. R-squared	0.836	0.461	0.856	0.655	0.467

Robust t-statistics in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Note: Clustered standard errors at firm level are applied to account for heteroscedasticity and autocorrelation in error terms. See Table 3 for variable definitions.

In Table 11(1) we show that there is a statistically significant positive relationship between $Length$ and $TobinsQ$ (0.293, t=1.99) at a 10% significance level which indicates that firms that issue longer stand-alone Integrated Reports benefit in terms of higher firm value. The

interaction term $Length * IR\_ITW$ has a negative association with $TobinsQ$ (-0.338, t=-1.99) on a 10% level. The result from combining the coefficients of $Length$ and $Length * IR\_ITW$ is slightly negative (-0.045) which indicates that longer intertwined Integrated Report is associated with lower firm valuation. Since there is no statistically significant association between IR_ITW and $TobinsQ$ (0.050, t=0.40) we cannot claim inference relating to how market valuation differs on average between firms releasing an intertwined versus stand-alone Integrated Report.

When examining the other selected fixed effect structures as presented in Table 11(2)-11(5) the results are comparable with the main fixed effect structure, but with varying magnitude and significance levels. Once again, there is a risk of omitting explanatory variables when not fixing the effects of year, industry, and country which can cause endogeneity problems and biases in estimations. Isolating the findings in Table 11(1) we find support of Hypothesis II showing that investors prefer intertwined Integrated Reports to be shorter and stand-alone Integrated Reports to be longer.

Table 12. Model 3b: OLS regression of bid-ask spread and length interaction term

	(1) BidAsk	(2) BidAsk	(3) BidAsk	(4) BidAsk	(5) BidAsk
Length	0.264	0.040	0.414	-0.162	0.043
	(1.04)	(0.19)	(1.14)	(-0.61)	(0.21)
Length *IR_ITW	-0.010	0.099	-0.320	0.160	0.068
	(-0.04)	(0.38)	(-0.78)	(0.56)	(0.27)
IR_ITW	-0.562	-0.139	-0.162	-0.250	-0.119
	(-2.56)**	(-0.56)	(-0.40)	(-1.27)	(-0.48)
ROA	-0.039	-0.039	-0.073	-0.092	-0.039
	(-1.92)*	(-2.70)***	(-2.69)***	(-2.89)***	(-2.70)***
NegativeProfit	-0.152	-0.214	-0.094	-0.350	-0.211
	(-0.86)	(-0.86)	(-0.55)	(-2.06)**	(-0.87)
Accruals	-2.003	0.139	-0.736	-1.005	0.032
	(-2.26)**	(0.06)	(-0.39)	(-0.60)	(0.01)
DtoCE	0.001	0.001	0.001	0.001	0.001
	(1.68)*	(1.04)	(0.58)	(0.81)	(1.24)
Beta	-0.052	-0.335	-0.578	-0.377	-0.306
	(-0.29)	(-1.10)	(-1.64)	(-1.78)*	(-1.00)
BTM	0.065	0.238	-0.020	0.296	0.223
	(0.69)	(1.17)	(-0.10)	(2.11)**	(1.11)
LogAssets	-0.385	-0.380	-0.201	-0.379	-0.373
	(-4.50)***	(-4.36)***	(-1.18)	(-6.69)***	(-4.24)***
Complexity	2.546	-1.475	-0.193	-1.065	-1.453
	(3.65)***	(-1.59)	(-0.12)	(-1.55)	(-1.59)
LogFirmAge	0.135	-0.137	-0.146	-0.063	-0.137
	(1.09)	(-0.75)	(-0.53)	(-0.44)	(-0.75)
Constant	-4.227	-2.457	-3.804	-2.636	-2.557
	(-5.55)***	(-2.02)**	(-2.79)***	(-2.93)***	(-2.11)**
Year FE	Yes	Yes	No	No	No
Industry FE	Yes	No	Yes	No	No
Country FE	Yes	No	No	Yes	No
Observations	232	235	232	235	235
Adj. R-squared	0.865	0.399	0.577	0.636	0.404

Robust t-statistics in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Note: Clustered standard errors at firm level are applied to account for heteroscedasticity and autocorrelation in error terms. See Table 3 for variable definitions.

Table 12(1) show no statistically significant results for neither $Length$ $(0.264, t = 1.04)$ nor $Length * IR\_ITW$ (-0.010, t=-0.04) and their respective association with $BidAsk$. Since IR_ITW (-0.562, t=-2.56) is negatively associated with $BidAsk$ on a 5% significance level we

show that firms releasing an intertwined Integrated Reports have on average 0.562 lower bid-ask spread than firms releasing a stand-alone Integrated Report. The results are somewhat consistent when considering the other selected fixed effect structures as presented in Table 12(2)-12(5). Studying Table 12(1) in isolation we find no support of Hypothesis II.

6.2.3 Summary of findings

Table 13. Summary of results Hypothesis I

	TobinsQ	BidAsk
ReadingDifficulty	No evidence	No evidence
Length	No evidence	No evidence

Note: See Table 3 for variable definitions.

Table 13 summarizes the findings related to Hypothesis I. To conclude, there is no supporting evidence for Hypothesis I. Neither the reading difficulty nor the length of the Integrated Report affect the firm value or stock liquidity.

Table 14. Summary of results Hypothesis II

	TobinsQ	BidAsk
ReadingDifficulty*IR_ITW	No evidence	Very strong negative association (1% significance level)
Length*IR_ITW	Weak negative association (10% significance level)	No evidence

Note: See Table 3 for variable definitions.

As seen in table 14, there are mixed findings whether releasing an intertwined Integrated Report has a moderating effect on the relationship between textual attributes of Integrated Reports and market effects. Thus, there is some support for Hypothesis II. From the findings above, we can conclude that there is an indication of a moderating effect of releasing an intertwined Integrated Report when investigating $ReadingDifficulty * IR\_ITW$ and $Length * IR\_ITW$.

7. Discussion

While previous research have established that high Integrated Reporting Quality i.e., the level of alignment with IIRF, have a positive impact on market effects (Zhou et al., 2017; Barth et al., 2017, Lee & Yeo., 2016), high-quality textual attributes as proxy for Integrated Reporting Quality has gained little attention. Nonetheless, Caglio et al. (2020), show that readable Integrated Reports are associated with higher market value, and shorter Integrated Reports associated with lower stock liquidity. We do not find confirmation of these associations and therefore reject Hypothesis I. These contradictory results can stem from multiple different explanations, discussed below.

It is important to consider the different settings of the studies and the type of report investigated. While Caglio et al. (2020) investigate South African firms, where Integrated Reporting is mandatory, our study considers the European market where Integrated Reporting is voluntary for public firms. One potential explanation why our results differ from Caglio et al. (2020) could be that European investors might already have incorporated the information from the Integrated Report into stock pricing as the information environment in Europe exhibits a high levels of transparency, resulting in no additional value add from an Integrated Report (Wahl et al., 2020; Leukhardt et al. 2022).

Moreover, Leukhardt et al. (2022) claim that European firms follow the IIRF inadequately which potentially explains why there is no identified positive relationship between high-quality textual attributes and market effects. One way to fuel the realization of positive market effects from Integrated Reporting in a voluntary setting could therefore be to improve the understanding of the framework both from a firm and investor perspective. Building on this reasoning, we observe that European firms' Integrated Reports are different compared to South African firms' Integrated Reports. When comparing our mean values and standard deviations for the reading indices (Flesch-Kincaid, Gunning-Fog and SMOG) it is clear that European Integrated Reports are on average more complex and lengthier than the South African Integrated Reports. The average reading difficulty for our sample is more aligned with those reported for financial annual reports (Boubaker et al., 2019), which potentially indicate that Integrated Reports of European Firms are more similar to annual reports than Integrated Reports published by South African firms.

Other evident differences between this study and the one presented by Caglio et al. (2020) is the sample size. This study uses a sample of firms listed on an European exchange with 235 firm-year observations for Tobin's Q and 235 firm-year observations for bid-ask spread. On the other hand, Caglio et al. (2020) has a data set of 444 firm-year observations for the Tobin's Q analysis and 435 firm-year observations for the bid-ask analysis only including firms listed on the Johannesburg Stock Exchange. A smaller sample size could lead to a larger margin of error thereby, making it more difficult to find statistically significant results (Newbold, 2013).

Furthermore, this study investigates the time period 2016-2021 whilst Caglio et al. (2020) examines the time period 2011-2016. Different time periods can be associated with different social, economic, environmental and political contexts that affect the population. For example, the behavior of firms or investors may have changed over time due to the Covid-19 pandemic, thereby making them reason differently, which explains why this study finds contradictory results.

When investigating the differences between stand-alone and intertwined Integrated Reports, we find preference for intertwined Integrated Reports to have high reading difficulty but being more concise. On the other hand, investors prefer stand-alone Integrated Reports to be readable but longer. Thus, the preference for high-quality textual attributes depends on the Integrated Report's disclosure type. These mixed findings and lack of robust evidence (see *Section 6.3 Robustness tests*) can stem from fallacies in empirics (see *Section 7.1 Limitations)* and future studies are encouraged to investigate this moderating effect further.

One explanation to why European investors do not appreciate low reading difficulty of intertwined Integrated Reports is that low reading difficulty potentially causes simplification of information, thereby inducing loss of information. As discussed previously, there are critics arguing that neither the Flesch-Kincaid, the Gunning-Fog nor the SMOG index are accurate measures for accounting information (Boubaker et al., 2019). Low reading difficulty in this case may reflect the absence of common accounting phrases with several syllables like "financial", "management" and "operations" (Loughran & McDonald, 2016), and the absence of these phrases concerns investors. An Integrated Report with high reading difficulty in the context of accounting could therefore be an indicator of simplified or missing information rather than complexity.

7.1 Limitations

Our study is sensitive to limitations caused by the voluntary setting where firms who issue Integrated Reports self-select. The self-selection problem imposes endogeneity concerns of the model specification describing the relationship between high-quality textual attributes and market effects. Endogeneity infers that the error term correlates with any of the explanatory variables. This results in biased estimations and inaccurate inferences (Ullah et al., 2018). This issue possibly explains why our findings are highly insignificant and to some extent contradictory to the body of literature on the topic of textual analysis within financial disclosures and Integrated Reporting. We are therefore aware of the validity limitations of our model and are therefore cautious to draw causal inference. Our results should be interpreted as associations between the variables rather than causations.

One underlying issue causing endogeneity relates to omitted variable bias. Omitted variable bias occurs if there are variables not included in the specified model that provide additional, or alternative, explanations to the relationship (Ullah et al., 2018). We acknowledge that it is likely that firm characteristics play a role in the self-selection. Possibly, firms who are considered "good" are choosing to issue an Integrated Report, or issues an Integrated Report with high-

quality textual attributes. Though, the reason for why they experience positive market effects are also due to the fact that they are indeed "good" firms. In this case, the explanatory variables and the dependent variables are related to the omitted variable, thereby causing bias. It is difficult however to identify these omitted variables, and to understand what a "good" firm is defined as. To the best of our knowledge, we have incorporated control variables to prevent the omitted variable bias. For example, we control firm size as firms with more resources have better capacity to produce an Integrated Report, and for firm age as younger firms could be more prone to adopt frameworks for sustainability information. Nonetheless, we cannot be certain that all proper control variables are included in the specified model.

Moreover, our model may have a problem of simultaneity which causes endogeneity. Simultaneity occurs when the independent variables are jointly determined with the dependent variables (Ullah et al., 2018). This means that the direction of causality runs in both directions, i.e., high-quality textual attributes increase market effects due to less information asymmetry, but high market effects also cause high-quality textual attributes as firms with good performance have less pressure to explain their performance (Bloomfield, 2008).

One approach would be to include an exogenous shock affecting the independent variables in a difference-in-difference model. For example, a regulatory introduction making simplification of Integrated Reports mandatory only for a random selection of firms. Typically, this occurs as a pilot study where firms e.g., on S&P 500, Euro Stoxx 50, or another index, are randomly allocated to a treatment group or control group. The randomized treatment as well as the comparison before-and-after the treatment would form a clean setting to examine the effect of the exogenous shock and allow us to make causal inference of the relationship (Ahmed et al., 2019) between textual attributes of Integrated Reports and market effects. This method is valid given that the two groups of firms have similar characteristics and are affected by the same market conditions, in order to isolate the effect of textual attributes on market effects.

To minimize the extent of endogeneity, in addition to adding control variables based on the field of literature, we have explored the robustness of our results, for example by performing additional tests on a sub-sample subject to the same legislative environment (European Union) in order to reduce the extent of unobserved factors affecting both the independent and dependent variables. Likewise, we have controlled unobserved heterogeneity across industries and years by including fixed effect structures.

We also recognize that our fairly small sample size may inhibit the ability to draw inferences of reported results.

8. Concluding remarks

8.1 Conclusion

This paper empirically assesses whether the textual attributes of Integrated Reports are associated with market effects. OLS regressions were performed to test the association on an unbalanced dataset from 2016 to 2021 including 61 unique companies and 287 firm-year observations.

The first hypothesis tested if high-quality textual attributes (readability and conciseness), as proxy for Integrated Reporting Quality, were associated with positive market effects, namely higher firm valuation and stock liquidity. In opposition to previous literature on Integrated Reporting Quality and economic benefits in mandatory settings (Caglio et al., 2020; Barth et al., 2020; Zhou et al., 2017; Lee and Yeo, 2015), we find no evidence of this association when examining a voluntary setting. The results are nonetheless supporting the findings in voluntary settings where neither Wahl et al. (2020) nor Leukhardt et al. (2022) find Integrated Reporting disclosures, or IRQ, to affect economic outcomes. It is evident that the Integrated Reports in our sample differ from those in mandatory settings as both the reading difficulty and length of the reports are higher. It is therefore possible that the IIRF alignment in voluntary settings is lower compared to mandatory settings, which explains why no benefits are identified.

The second hypothesis examined whether there is an effect of issuing the Integrated Report intertwined. The results show that there is some preference for intertwined Integrated Reports to include high reading difficulty and conciseness. The opposite association is identified for stand-alone Integrated Reports, i.e., readable and lengthier. Since this association is only identified for one proxy of market effects independently, and due to the lack of supporting literature in this area, we are cautious when interpreting the findings of Hypothesis II.

Overall, our results contribute to the growing body of literature on disclosure quality by investigating the textual attributes of Integrated Reports. While previous literature within Integrated Reporting research has studied a mandatory setting, we investigate a voluntary setting. Additionally, we discover the implications of issuing an Integrated Report stand-alone versus intertwined, which to our knowledge has not been examined previously.

8.2 Policy and managerial implications

Our findings have implications for policy makers. As an effort to potentially make the benefits of Integrated Reporting (Caglio et al., 2020; Zhou et al., 2017; Barth et al., 2017; Lee & Yeo, 2016) materialize for European firms, regulators are recommended to make it mandatory. The future integration of IIRF into the IFRS accounting standards and the IFRS Sustainability Disclosure Standards (Guillot, 2023) could potentially be sufficient for the relationship between high-quality textual attributes of Integrated Reports and market effects to unfold for European firms. Though, this is most likely affected by how clear and to what extent the IIRF principles are incorporated.

Our findings also have practical managerial implications for firms operating in a voluntary setting of Integrated Reporting. Given that managers are encountered with the decision to improve the textual attributes of their Integrated Report, the lack of market effects can guide resource allocations and help managers evaluate the benefits of such an implementation. Since there are in fact perceived economic benefits with improving quality of textual attributes in financial reporting (Li, 2008; Lang & Stice-Lawrence, 2015; Botosan & Plumlee, 2002; Boubaker et al., 2019) managers are advised to prioritize the quality of these textual attributes.

8.3 Future research

We recommend future research to further investigate the moderating effect of releasing an intertwined versus stand-alone Integrated Report due to our mixed findings. Since this is an unexplored field, additional research is encouraged to build consensus.

The scarcity of studies within the field of benefits on Integrated Reporting Quality in voluntary settings witness a knowledge gap. While no positive market effects were identified in this study, the IIRF still emphasizes a need for integrated thinking among firms. We therefore propose an agenda for future research to investigate potential non-financial benefits with adopting the IIRF. One example of this is to evaluate if the firm experiences a higher CSR performance as a result of applying principle 3) stakeholder relationships.

We also recommend additional studies to include the readability index "Bog" as it is specifically designed for accounting information and thus generating a holistic view of reading difficulty in the corporate context.

9. References

Ahmed, A.S., Li, Y. And XU, N. (2020), Tick Size and Financial Reporting Quality in Small-Cap Firms: Evidence from a Natural Experiment. *Journal of Accounting Research*, 58: 869-914. https://doi.org/10.1111/1475-679X.12331

Al Amin, M., Qin, Y. (2023). Factor Analysis in Stata: Getting started with Factor Analysis. *Princeton University*. Assessed 2023-03-19. https://libguides.princeton.edu/factor

Alduais, F. Ali Almasria, N., Samara, A., & Masadeh, A. (2022). Conciseness, Financial Disclosure, and Market Reaction: A Textual Analysis of Annual Reports in Listed Chinese Companies. *International Journal of Financial Studies*, 10(4), 104–. https://doi.org/10.3390/ijfs10040104

Barber, M. B., & Odean, T. (2008). All That Glitters: The Effect of Attention and News on the Buying Behavior of Individual and Institutional Investors, *The Review of Financial Studies.* Vol. 21. No. 2. pp. 785-818. https://doi.org/10.1093/rfs/hhm079

Barth, M., Cahan, S. F., Chen, L., & Venter, E. R. (2017). The economic consequences associated with Integrated Report quality: Capital market and real effects. *Accounting, Organizations and Society*, *62*, 43–64. https://doi.org/10.1016/j.aos.2017.08.005

Bloomfield, R. (2008). Discussion of "Annual report readability, current earnings, and earnings persistence." *Journal of Accounting & Economics*, *45*(2), 248–252. https://doi.org/10.1016/j.jacceco.2008.04.002

Botosan,C.A., & Plumlee, M. A. (2002). A Re-examination of Disclosure Level and the Expected Cost of Equity Capital. *Journal of Accounting Research*, *40*(1), 21–40. https://doi.org/10.1111/1475-679X.00037

Boubaker, S. Gounopoulos, D., & Rjiba, H. (2019). Annual report readability and stock liquidity. *Financial Markets, Institutions & Instruments*, 28(2), 159–186. https://doi.org/10.1111/fmii.12110

Caglio, A., Melloni, G., & Perego, P. (2020). Informational Content and Assurance of Textual Disclosures: Evidence on Integrated Reporting. *The European Accounting Review*, *29*(1), 55–83. https://doi.org/10.1080/09638180.2019.1677486

Caglio, A. Melloni, G., & Perego, P. (2017). Saying more with less? Disclosure conciseness, completeness and balance in Integrated Reports. *Journal of Accounting and Public Policy*, 36(3), 220–238. https://doi.org/10.1016/j.jaccpubpol.2017.03.001

Chang, S, L., Most, S, K., Brain, W, C. (1983). The utility of annual reports: an international study. *Journal of International Business Studies*, 14(1). 64-84 https://doi.org/10.1057/palgrave.jibs.8490507

Christensen, H, B., Floyd, E., Yao Liu, L., & Maffett, M. (2017). The real effects of mandated informatoin on social responsibility in financial reports: Evidence from mine safety records. *Journal of Accounting and Economics.* 64, 284-304. https://doi.org/10.1016/j.jacceco.2017.08.001

Council of the EU. (2022, November 28). Council gives final green light to corporate sustainability reporting directive. *European Council of European Union.* Retrieved March 26, 2023, from https://www.consilium.europa.eu/en/press/press-releases/2022/11/28/council-gives-final-green-light-to-corporate-sustainability-reporting-directive/

Chall, J.C & Dale, E. (1949). The Concept of Readability. *Elementary English*, 26(1), 19–26. Print.

Du, S & Yu, K. (2021). Do Corporate Social Responsibility Reports Convey Value Relevant Information? Evidence from Report Readability and Tone. *Journal of Business Ethics*, 172(2), 253–274. https://doi.org/10.1007/s10551-020-04496-3

European Commission (2023) Corporate sustainability reporting (2023) *Finance. European Commission.* Retrieved May 2, 2023.Available at: https://finance.ec.europa.eu/capital-markets-union-and-financial-markets/company-reporting-and-auditing/company-reporting/corporate-sustainability-reporting_en

Gao, J. Hartmann, F. G.., Zhang, M., & Chen, Y. (2022). The impact of CSR performance and CSR disclosure readability on investors' earnings estimates. *Accounting and Finance (Parkville)*. https://doi.org/10.1111/acfi.12938

Guillot, J. (2023). An update on progress post consolidation. *Integrated Reporting.* Retrieved March 22, 2023, from https://www.integratedreporting.org/news/an-update-on-progress-post-consolidation/

International Financial Reporting Standards (2022a). European Union. Who uses IFRS Accounting standards? Retrieved March 26, 2023, from https://www.ifrs.org/use-around-the-world/use-of-ifrs-standards-by-jurisdiction/view-jurisdiction/european-union/

International Financial Reporting Standards. (2022b). United Kingdom. Who uses IFRS Accounting Standards? . Retrieved March 26, 2023, from https://www.ifrs.org/use-around-the-world/use-of-ifrs-standards-by-jurisdiction/view-jurisdiction/united-kingdom/

International Financial Reporting Standards. (2021). Switzerland. Who uses IFRS Accounting Standards? Retrieved March 26, 2023, from https://www.ifrs.org/use-around-the-world/use-of-ifrs-standards-by-jurisdiction/view-jurisdiction/switzerland/

International Integrated Reporting Council Database. (n.d.). Reporters. *Integrated Reporting Examples Database.* Retrieved March 21, 2023, from https://examples.integratedreporting.org/ir-reporters/

International Integrated Reporting Council. (2023). Celebrating 10 years. *10 Years of the IIRC.* Retrieved March 10, 2023, from https://www.integratedreporting.org/10-years/10-years-summary/

International Integrated Reporting Council. (2022). IFRS Foundation completes consolidation with Value Reporting Foundation. *Integrated Reporting*. Retrieved March 10, 2023, from https://www.integratedreporting.org/news/ifrs-foundation-completes-consolidation-with-value-reporting-foundation/

International Integrated Reporting Council. (2015, July 15). More than a third of Dutch listed companies are working towards IR. *Integrated Reporting*. Retrieved March 17, 2023, from https://www.integratedreporting.org/news/more-than-a-third-of-dutch-listed-companies-are-working-towards/

International Integrated Reporting Council. (2017a). What is the future of corporate reporting? *Integrated Reporting.* Retrieved March 17, 2023, from https://www.integratedreporting.org/news/what-is-the-future-of-corporate-reporting/

International Integrated Reporting Council. (2017b). "Breakthrough moment for Integrated Reporting" – international Integrated Reporting council publishes OWN annual Integrated Report. *Integrated Reporting.* Retrieved March 16, 2023, from https://www.integratedreporting.org/news/breakthrough-moment-for-integrated-reporting-international-integrated-reporting-council-publishes-own-annual-integrated-report/

International Integrated Reporting Council. (n.d.). Frequently Asked Questions. *Integrated Reporting*. Retrieved March 15, 2023, from https://www.integratedreporting.org/FAQS/

International Integrated Reporting Framework. (2021). The international <IR> framework. *International Financial Reporting Standards Foundation*. Retrieved March 16, 2023, from https://www.integratedreporting.org/resource/international-ir-framework/

Jones, M.A. & Shoemaker, P. A. (1994). Accounting Narratives: A Review of Empirical Studies of Content and Readability. *Journal of Accounting Literature*, 13, 142–.

KPMG International. (2022). *Big shifts, small steps- Survey of Sustainability Reporting 2022.* https://assets.kpmg.com/content/dam/kpmg/se/pdf/komm/2022/Global-Survey-of-Sustainability-Reporting-2022.pdf

Lang, M & Stice-Lawrence, L. (2015). Textual analysis and international financial reporting: large sample evidence. *Journal of Accounting & Economics*, 60(2-3), 110–135. https://doi.org/10.1016/j.jacceco.2015.09.002

Lee, K.-W., & Yeo, G. H.-H. (2016). The association between Integrated Reporting and firm valuation. *Review of Quantitative Finance and Accounting*, *47*(4), 1221–1250. https://doi.org/10.1007/s11156-015-0536-y

Leukhardt, L., Charifzadeh, M., & Diefenbach, F. (2022). Does integrated reporting quality matter to capital markets? Empirical evidence from voluntary adopters. *Corporate Social-Responsibility and Environmental Management*, *29*(5), 1482–1494. https://doi.org/10.1002/csr.2285

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

Lo, K. Ramos, F., & Rogo, R. (2017). Earnings management and annual report readability. *Journal of Accounting & Economics*, 63(1), 1–25. https://doi.org/10.1016/j.jacceco.2016.09.002

Loughran, T & Mcdonald, B. (2016). Textual Analysis in Accounting and Finance: A Survey. *Journal Of Accounting Research*, 54(4), 1187–1230. Https://Doi.Org/10.1111/1475-679x.12123

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

Mc Laughlin, H. (1969). SMOG Grading – a New Readability Formula. *Journal of Reading 12(8)*, 639-646 https://www.jstor.org/stable/40011226

Newbold, P., Carlson, W.L. and Thorne, B. (2013) *Statistics for Business and Economics. Boston,* MA: Pearson.

Plumlee, M. A. (2003), 'The Effect of Information Complexity on Analysts' Use of That Information', *The Accounting Review*, Vol. 78, No. 1, pp. 275–96. Web.

Rinaldi, L., Unerman, J., & de Villiers, C. (2018). Evaluating the Integrated Reporting journey: insights, gaps and agendas for future research. *Accounting, Auditing & Accountability Journal*, *31*(5), 1294–1318. https://doi.org/10.1108/AAAJ-04-2018-3446

Soliman, M & Ben-Amar, W. (2022). Corporate social responsibility orientation and textual features of financial disclosures. *International Review of Financial Analysis*, 84, 102400–. https://doi.org/10.1016/j.irfa.2022.102400

Stock, J.H. and Watson, M.W. (2020) *Introduction to econometrics 4th Edition*. New York, NY: Pearson.

Ullah, S., Akhtar, P., & Zaefarian, G. (2018). Dealing with endogeneity bias: The generalized method of moments (GMM) for panel data. *Industrial Marketing Management*, 71, 69–78. https://doi.org/10.1016/j.indmarman.2017.11.010

Wahl, A., Charifzadeh, M., & Diefenbach, F. (2020). Voluntary Adopters of Integrated Reporting – Evidence on Forecast Accuracy and Firm Value. *Business Strategy and the Environment*, *29*(6), 2542–2556. https://doi.org/10.1002/bse.2519

Zhou, S., Simnett, R., & Green, W. (2017). Does Integrated Reporting Matter to the Capital Market? *Abacus (Sydney)*, *53*(1), 94–132. https://doi.org/10.1111/abac.12104

Appendix

Factor analysis

Table A.1. Bartlett's test of sphericity and Kaiser-Meyer-Olkin measure of sampling adequacy

Barlett's test of sphericity	
Chi-square	2662.994
Degrees of freedom	10
p-value	0

Bartlett's test of sphericity tests the null hypothesis that variables are unrelated and not ideal for performing a factor analysis. The test compares the matrix of correlations with an identity matrix filled with zero correlations. The p-value presented is less than 0.05 (Caglio et al., 2020), hence, implying that the variables $FleschKincaid$, $SMOG$, FOG, $NoCharacters$, and $NoSentences$ are indeed related.

Table A.2. Kaiser-Meyer-Olkin measure of sampling adequacy

Kaiser-Meyer-Olkin measure of sampling adequacy	
KMO	0.648

The KMO tests examine whether the data is suited for a factor analysis by investigating partial correlation, i.e., the proportion of variance among the variables that are common variance. The presented value is larger than 0.5 (Caglio et al., 2020) and hence, a factor analysis is appropriate for variables: $FleschKincaid$, $SMOG$, FOG, $NoCharacters$, and $NoSentences$.

OLS assumptions

To test our hypotheses, multiple Ordinary Least Squares (OLS) regressions are performed. Applying this method allows us to test for several independent variables simultaneously affecting a dependent variable. The coefficients are obtained by performing the least squares procedure, meaning that the estimated coefficients minimize the sum of the residuals squared. By utilizing this method, a linear equation that best explains the observed data is achieved (Newbold et al., 2013). The general OLS model is formulated as:

$$y_{it} = \beta_0 + \beta_1 x_{1,it} + \ \dots \ + \beta_k x_{k,it} + \varepsilon_{it} \qquad (i = 1, \ \dots, I; t = 1, \ \dots, T)$$

Where y_{it} is an outcome variable that shows both within and between variation, $(x_{1,it}, \dots, x_{k,it})$ is a vector of explanatory variables with both within and between variation, ε_{it} is the unobservable random error term, and $(\beta_0, \dots, \beta_k)$ is a vector of constant parameters to be estimated (Newbold et al., 2013).

To achieve the best linear unbiased estimator, the following assumptions need to hold (Newbold et al., 2013). Violation of these assumptions may impair the validity of presented findings. Therefore, assumptions are tested for, and appropriate responses are incorporated to assure robustness of our findings.

1. The x_{it} terms are fixed numbers, or they are realization of random variables, X_{it} , that are independent of the error terms, ε_{it}
2. The expected value of the random variable Y is a linear function of the X_{it} independent variables[12]
3. The error terms are normally distributed random variables with a mean of 0 and the same variance
4. The random error terms, ε_{it} , are not correlated with one another
5. It is not possible to find a set of non-zero numbers, $c_1, \dots, c_K$, such that $c_1 x_{1,it} + c_2 x_{2,it} \dots + c_k x_{k,it}$

Table A.3. Breusch-Pagan/Cook-Weisberg

Variable	H0	chi2(1)	Prob>chi2
TobinsQ	Constant variance	80.41	0
BidAsk	Constant variance	73.89	0

Heteroskedasticity implies that error terms have unequal variance. If this is present, the least squares procedure is not the best method to utilize for estimating coefficients, and thus the OLS regression is not appropriate for hypothesis testing. The Breusch-Pagan tests the null hypothesis for homoscedasticity, i.e., that the error terms have constant variance. For both Tobin's Q and the Bid-Ask spread, the null hypothesis is rejected, hence, the error terms do not have constant variance. This implies that there is heteroscedasticity present. To account for this, robust standard errors at firm level are included in the OLS regression.

Table A.4 Wooldridge test

Variable	H0	F	Prob>F
TobinsQ	No first-order autocorrelation	37.617	0
BidAsk	No first-order autocorrelation	98.127	0

The underlying data set used for the OLS regressions is classified as unbalanced panel data, which means that it is a combination of cross-sectional (multiple firms) and time series (multiple years per firm). Due to the attributes of time-series data, we acknowledge that

[12] Our untabulated scatterplots of our main explanatory variables against our dependent variables show a weak linear relationship.

autocorrelation between error terms might be present. This stems from the fact that error terms in close time periods often exhibit correlation, thus often being the case when using time-series data (Newbold et al., 2013). Therefore, the Woolridge test for autocorrelation in panel data is performed. For both Tobin's Q and the bid-ask spread, the null hypothesis is rejected, hence, autocorrelation in error terms is present. To adjust for autocorrelation, standard errors are clustered at firm level.

Table A.5. Spearman correlation test

	ReadingDifficulty	Length	ROA	NegativeProfit	Accruals	DtoCE	Beta	BTM	LogAssets	Complexity	Log FirmAge	IR_ITW
ReadingDifficulty	1											
Length	0.115*	1										
ROA	0.042	-0.133**	1									
NegativeProfit	0.160**	0.104*	-0.469***	1								
Accruals	0.050	-0.093	-0.429***	0.019	1							
DtoCE	-0.211***	0.086	-0.433***	-0.088	0.183***	1						
Beta	0.172***	-0.010	-0.273***	0.181***	0.351***	-0.113*	1					
BTM	0.042	-0.136**	-0.621***	0.113*	0.394***	0.340***	0.432***	1				
LogAssets	-0.076	-0.006	-0.365***	-0.099	0.207***	0.430***	-0.050	0.505***	1			
Complexity	0.053	0.089	0.521***	-0.158**	-0.335***	-0.362***	-0.258***	-0.548***	-0.305***	1		
LogFirmAge	-0.243***	-0.063	0.111*	-0.229***	-0.118*	0.030	-0.137**	-0.025	0.181***	0.136**	1	
IR_ITW	0.077	0.522***	0.162**	-0.077	-0.184***	0.024	-0.152**	-0.247***	-0.157**	0.078	-0.056	1

*** p<0.01, ** p<0.05, * p<0.1

Assuming no multicollinearity, i.e., no perfect correlation between independent variables, is important for the purpose of assessing the separate effects of the explanatory variables. If moving in union, this will not be possible (Newbold et al., 2013). To test for multicollinearity, the parametric Pearson correlation test (see *Section 5.2 Correlation table*), the non-parametric Spearman correlation test and the VIF test (See Appendix A.6) have been performed.

The Spearman Correlation is a non-parametric test which investigates the strength and direction of a relationship between two variables. The test does not require normality. No correlations are exceeding 0.7 which presumes that no multicollinearity is present.

Table A.6. Variance Inflation Factor (VIF)

	TobinsQ		BidAsk	
Variable	VIF	1/VIF	VIF	1/VIF
BTM	1.84	0.544	1.88	0.531
ROA	1.77	0.566	1.83	0.546
LogAssets	1.52	0.656	1.55	0.646
Beta	1.49	0.669	1.50	0.667
Accruals	1.49	0.673	1.48	0.674
Negativeprofit	1.46	0.685	1.42	0.702
LogFirmAge	1.39	0.720	1.38	0.723
Complexity	1.28	0.779	1.35	0.739
ReadingDifficulty	1.24	0.805	1.30	0.770
Length	1.10	0.912	1.11	0.904
DtoCE	1.08	0.930	1.08	0.930
Mean VIF		1.42		1.44

The VIF tests the magnitude of multicollinearity, i.e., if there is correlation between the explanatory variables in the specified model. The mean VIF for both models are less than 2 which implies that multicollinearity is not an issue.

Table A.7. Shapiro Wilk test

Variable	Obs	W	V	z	Prob>z
res_TobinsQ	235	0.932	11.600	5.686	0
res_BidAsk	235	0.946	9.267	5.164	0

Shapiro Wilk tests the null hypothesis of normal distributed residuals for a continuous variable. The reported results show that the null hypothesis for normal distributed variables can be rejected for Tobin's Q and the Bid-ask spread.

Robustness tests

We perform additional robustness tests to validate our main analysis. For Hypothesis I, we re-run Model 1 with the independent variables $ReadingDifficulty$ and $Length$ in two separate regressions. In addition, for both Hypothesis I and II, we perform our regressions on a sample covering only EU-firms, regressions substituting factors with their underlying indices/proxies, and regressions with no missing values i.e., replacing missing data with median values to increase the number of observations. The reasons why we choose to perform these additional robustness tests are outlined below.

First, the two independent variables for Model 1 are separated in two different regressions to confirm our findings. Given that both factor $ReadingDifficulty$ and $Length$ are proxies for textual attributes, it could be argued that they are explaining similar variance and should be separated to avoid issues of multicollinearity.

Second, as discussed in *Section 2.1 Institutional Background on Integrated Reporting*, the approval of CSRD in June 2022 will require firms within the European Union to disclose non-financial information. Hence, the number of firms who report on non-financial information is assumed to increase from 11,700 to approximately 50,000 which is presumed to also drive the trend of adopting IIRF (KPMG International, 2020). The CSRD applies to all large public firms within the European Union. Even though non-EU firms located in Europe such as Switzerland and the United Kingdom are assumed to be highly impacted by initiatives within the Union, it is of interest to validate the main analysis by solely focusing on EU-firms who operate in a similar environment and are required to follow EU legislation.

Third, we acknowledge that the construction of our factors for reading difficulty and length imposes a risk of losing variable data for the individual measures as these are grouped into one combined variable. This occurs if the individual measures are explaining slightly different variances. Additionally, it is moreover interesting to investigate if one measure is more strongly associated with the market effects than others. Therefore, we also perform our main regression with single measures for our readability indices and length proxies instead of the factors. That is, reading difficulty is substituted with the Flesch-Kincaid index, the Gunning Fog index, and the Smog index. Length is substituted with number of characters and number of words. The natural logarithm is performed on number the of characters and the number of words in order to account for the large data points.

Forth, a common approach to increase the sample size as a robustness test is to replace the missing values with the median value for each variable. This is performed in Caglio et al. (2017) and we therefore follow their approach.

Table A.8. Hypothesis 1. Model 1a with independent variable reading difficulty in separate regressions

	(1) TobinsQ	(2) TobinsQ
ReadingDifficulty	0.018	0.026
	(0.43)	(0.57)
ROA	0.045	0.043
	(2.44)**	(3.71)***
NegativeProfit	0.124	0.156
	(2.45)**	(1.18)
Accruals	0.795	2.053
	(1.24)	(2.09)**
DtoCE	0.001	0.001
	(1.83)*	(1.94)*
Beta	-0.215	-0.358
	(-1.30)	(-3.14)***
BTM	-0.101	-0.068
	(-1.60)	(-1.04)
LogAssets	-0.133	-0.103
	(-3.74)***	(-3.03)***
Complexity	1.537	1.192
	(2.03)**	(2.52)**
LogFirmAge	0.191	0.107
	(2.62)**	(1.34)
Constant	1.817	2.100
	(3.16)***	(5.10)***
Year FE	Yes	No
Industry FE	Yes	No
Country FE	Yes	No
Observations	232	235
Adj. R-squared	0.853	0.450
Robust t-statistics in parentheses		
*** p<0.01, ** p<0.05, * p<0.1		

Note: Clustered standard errors at firm level are applied to account for heteroscedasticity and autocorrelation in error terms. See Table 3 for variable definitions.

Table A.8(1) show no evidence of a relationship between $ReadingDifficulty$ and $TobinsQ$. This effect is also robust when no fixed effect structures are applied (column 2). Findings in Table A.8(1) are in line with main results presented in Table 7.

Table A.9. Hypothesis 1. Model 1a with independent variable length in separate regressions

	(1) TobinsQ	(2) TobinsQ
Length	0.031	-0.055
	(0.86)	(-1.13)
ROA	0.047	0.040
	(2.40)**	(3.54)***
NegativeProfit	0.135	0.148
	(2.35)**	(1.06)
Accruals	0.835	1.837
	(1.27)	(1.87)*
DtoCE	0.001	0.001
	(1.84)*	(1.66)
Beta	-0.220	-0.352
	(-1.35)	(-2.95)***
BTM	-0.110	-0.071
	(-1.64)	(-0.99)
LogAssets	-0.136	-0.102
	(-3.97)***	(-2.89)***
Complexity	1.641	1.248
	(1.93)*	(2.60)**
LogFirmAge	0.190	0.095
	(2.98)***	(1.33)
Constant	1.838	2.138
	(3.51)***	(4.98)***
Year FE	Yes	No
Industry FE	Yes	No
Country FE	Yes	No
Observations	232	235
Adj. R-squared	0.854	0.455
	Robust t-statistics in parentheses	
	*** p<0.01, ** p<0.05, * p<0.1	

Note: Clustered standard errors at firm level are applied to account for heteroscedasticity and autocorrelation in error terms. See Table 3 for variable definitions.

Table A.9(1) show no evidence of a relationship between $Length$ and $TobinsQ$. This effect is also robust when no fixed effect structures are applied (column 2). Findings in Table A.9(1) are in line with main results presented in Table 7.

Table A.10. Hypothesis 1. Model 1b with independent variable reading difficulty in separate regressions

	(1)	(2)
	BidAsk	BidAsk
ReadingDifficulty	0.116	-0.294
	(1.41)	(-1.57)
ROA	-0.067	-0.044
	(-2.76)***	(-3.09)***
NegativeProfit	-0.160	-0.162
	(-0.84)	(-0.67)
Accruals	-2.428	-0.166
	(-2.59)**	(-0.09)
DtoCE	0.000	0.001
	(0.47)	(0.85)
Beta	-0.019	-0.202
	(-0.11)	(-0.92)
BTM	0.153	0.238
	(1.12)	(1.43)
LogAssets	-0.322	-0.359
	(-3.36)***	(-4.69)***
Complexity	1.532	-1.114
	(1.89)*	(-1.71)*
LogFirmage	0.181	-0.271
	(1.01)	(-1.25)
Constant	-5.259	-2.363
	(-4.70)***	(-1.92)*
Year FE	Yes	No
Industry FE	Yes	No
Country FE	Yes	No
Observations	232	235
Adj. R-squared	0.845	0.461
	Robust t-statistics in parentheses	
	*** p<0.01, ** p<0.05, * p<0.1	

Note: Clustered standard errors at firm level are applied to account for heteroscedasticity and autocorrelation in error terms. See Table 3 for variable definitions.

Table A.10(1) show no evidence of a relationship between $ReadingDifficulty$ and $BidAsk$ This effect is also robust when no fixed effect structures are applied (column 2). Findings in Table A.10(1) are in line with main results presented in Table 8.

Table A.11. Hypothesis 1. Model 1b with independent variable length in separate regressions

	(1)	(2)
	BidAsk	BidAsk
Length	0.133	0.063
	(1.65)	(0.57)
ROA	-0.058	-0.043
	(-2.41)**	(-3.04)***
NegativeProfit	-0.112	-0.205
	(-0.64)	(-0.79)
Accruals	-2.164	0.008
	(-2.29)**	(0.00)
DtoCE	0.000	0.001
	(0.43)	(0.91)
Beta	-0.019	-0.306
	(-0.11)	(-1.01)
BTM	0.083	0.221
	(0.70)	(1.12)
LogAssets	-0.345	-0.362
	(-3.63)***	(-4.66)***
Complexity	2.087	-1.453
	(2.93)***	(-1.62)
LogFirmage	0.092	-0.142
	(0.51)	(-0.78)
Constant	-4.733	-2.697
	(-4.23)***	(-2.16)**
Year FE	Yes	No
Industry FE	Yes	No
Country FE	Yes	No
Observations	232	235
Adj. R-squared	0.848	0.406
	Robust t-statistics in parentheses	
	*** p<0.01, ** p<0.05, * p<0.1	

Note: Clustered standard errors at firm level are applied to account for heteroscedasticity and autocorrelation in error terms. See Table 3 for variable definitions.

Table A.11(1) show no evidence of a relationship between $Length$ and $BidAsk$. This effect is also robust when no fixed effect structures are applied (column 2). Findings in Table A.11(1) are in line with main results presented in Table 8.

Table A.12. Pearson correlation table with single items (no factors)

	Flesch Kincaid	SMOG	FOG	No Characters	No Sentences	ROA	Negative Profit	Accruals	DtoCE	Beta	BTM	LogAssets	Complex-ity	LogFirm Age	IR_ITW
Flesch Kincaid	1.														
SMOG	0.942***	1													
FOG	0.992***	0.955***	1												
LogNo Characters	0.048	0.052	0.070	1											
No Sentences	-0.263***	-0.259***	-0.242***	0.906***	1										
ROA	-0.042	-0.072	-0.076	-0.195**	-0.163*	1									
Negative Profit	0.139*	0.163*	0.151*	0.082	0.023	-0.422***	1								
Accruals	0.037	0.056	0.045	-0.044	-0.080	-0.457***	0.044	1							
DtoCE	-0.040	-0.040	-0.032	-0.078	-0.037	-0.045	0.029	0.008	1						
Beta	0.268***	0.274***	0.262***	0.017	-0.100	-0.292***	0.293***	0.318***	-0.058	1					
BTM	-0.027	0.006	-0.000	-0.044	-0.032	-0.307***	0.060	0.268***	0.187**	0.210***	1				
LogAssets	-0.120	-0.133*	-0.131*	-0.003	0.041	-0.032	-0.187**	0.083	0.039	-0.157*	0.473***	1			
Complex-ity	0.070	0.083	0.091	0.084	0.067	0.189**	-0.173**	-0.222***	-0.013	-0.253***	-0.430***	-0.197**	1		
LogFirm Age	-0.350***	-0.315***	-0.322***	-0.058	0.015	0.189**	-0.261***	-0.148*	-0.041	-0.306***	-0.0158	0.183**	0.126*	1	
IR_ITW	-0.006	0.012	0.020	0.472***	0.424***	0.123	-0.077	-0.155*	0.132*	-0.104	-0.138*	-0.145*	0.123	-0.0521	1

*** p<0.01, ** p<0.05, * p<0.1

Table A.12 shows the correlations between all explanatory variables. Significant (1% significance level) and very strong correlations (>0.9) are identified between reading indices: $FleschKincaid$, $SMOG$ and FOG, implying that these variables explain similar variance. $NoSentences$ is also highly correlated (>0.9) with $NoCharacters$ at a 1% significance level. Not surprisingly, correlations are also identified between all reading indices and $NoSentences$ as the formulas calculating the indices are based on the sentences within the analyzed text.

Table A.13. Spearman correlation table with single items (no factors)

	Flesch Kincaid	SMOG	FOG	No Characters	No Sentences	ROA	Negative Profit	Accruals	DtoCE	Beta	BTM	LogAssets	Complex-ity	Log FirmAge	IR_ITW
Flesch Kincaid	1														
SMOG	0.977***	1													
FOG	0.990***	0.983***	1												
No Characters	0.103	0.118*	0.131**	1											
No Sentences	-0.164***	-0.127**	-0.130**	0.945***	1										
ROA	0.073	0.036	0.037	-0.128**	-0.146**	1									
Negative Profit	0.146**	0.167***	0.166***	0.106*	0.092	-0.469***	1								
Accruals	0.045	0.061	0.052	-0.097	-0.095	-0.429***	0.019	1							
DtoCE	-0.235***	-0.206***	-0.201***	0.069	0.118*	-0.433***	-0.088	0.183***	1						
Beta	0.172***	0.168***	0.161**	0.007	-0.049	-0.273***	0.181***	0.351***	-0.113*	1					
BTM	0.037	0.051	0.054	-0.125*	-0.147**	-0.621***	0.113*	0.394***	0.340***	0.432***	1				
LogAssets	-0.074	-0.083	-0.081	-0.002	-0.008	-0.365***	-0.099	0.207***	0.430***	-0.050	0.505***	1			
Complex-ity	0.043	0.043	0.056	0.084	0.081	0.521***	-0.158**	-0.335***	-0.362***	-0.258***	-0.548***	-0.305***	1		
LogFirmAge	-0.248***	-0.230***	-0.234***	-0.067	-0.002	0.111*	-0.229***	-0.118*	0.030	-0.137**	-0.025***	0.181	0.136	1	
IR_ITW	0.041	0.054	0.061	0.517***	0.512***	0.162**	-0.077	-0.184***	0.024	-0.152**	-0.247**	-0.157	0.078	-0.056***	1

*** p<0.01, ** p<0.05, * p<0.1

Table A.13 shows the correlations between all explanatory variables. Significant (1% significance level) and very strong correlations (>0.9) are identified between reading indices: $FleschKincaid$, $SMOG$ and FOG, implying that these variables explain similar variance. $NoSentences$ is also highly correlated (>0.9) with $NoCharacters$ at a 1% significance level. Not surprisingly, correlations are also identified between all reading indices and $NoSentences$ as the formulas calculating the indices are based on the sentences within the analyzed text.

Table A.14. OLS regression of market effects and textual attributes with EU Sample firms

	(1) TobinsQ	(2) TobinsQ	(3) BidAsk	(4) BidAsk
ReadingDifficulty	-0.040	0.042	0.190	-0.241
	(-1.59)	(0.92)	(1.81)*	(-1.49)
Length	-0.017	-0.037	0.152	0.023
	(-0.61)	(-0.97)	(1.45)	(0.22)
ROA	0.036	0.038	-0.044	-0.030
	(1.94)*	(4.92)***	(-1.69)	(-2.21)**
Negativeprofit	0.046	0.138	-0.171	-0.437
	(0.99)	(1.08)	(-1.18)	(-1.28)
Accruals	0.881	1.051	-2.008	-0.212
	(1.54)	(1.34)	(-2.01)*	(-0.12)
DtoCE	0.009	0.026	0.002	0.080
	(1.00)	(2.75)***	(0.03)	(3.60)***
Beta	-0.207	-0.272	-0.004	-0.082
	(-1.01)	(-2.55)**	(-0.02)	(-0.33)
BTM	-0.106	-0.099	0.155	0.041
	(-1.64)	(-1.69)*	(1.09)	(0.23)
LogAssets	-0.156	-0.097	-0.329	-0.457
	(-4.26)***	(-2.66)**	(-3.05)***	(-4.84)***
Complexity	0.668	1.194	2.284	-1.054
	(1.85)*	(2.57)**	(2.54)**	(-1.23)
LogFirmAge	0.088	0.011	0.335	-0.206
	(3.39)***	(0.22)	(1.54)	(-0.80)
Constant	2.578	2.195	-5.988	-1.970
	(11.33)***	(4.57)***	(-5.40)***	(-1.42)
Year FE	Yes	No	Yes	No
Industry FE	Yes	No	Yes	No
Country FE	Yes	No	Yes	No
Observations	189	185	185	186
Adj. R-squared	0.841	0.534	0.858	0.520

Robust t-statistics in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Note: Clustered standard errors at firm level are applied to account for heteroscedasticity and autocorrelation in error terms. See Table 3 for variable definitions.

As shown in Table A.14(1), there is no statistically significant association between $ReadingDifficulty$ and $TobinsQ$ nor between $Length$ and $TobinsQ$, which is in line with results presented in the main analysis. It supports the finding that firm value is unaffected by

both Readability and the Length of the Integrated Report. The EU-sample shows no difference to our main analysis in regard to significance and magnitude of coefficients.

Presented in Table A.14(3), there is a positive association between $ReadingDifficulty$ and the $BidAsk$ (0.190, t=1.81). This association was not identified in our main sample which suggests that for EU-firms, investors indeed prefer lower reading difficulty as it decreases the bid-ask spread. However, the statistical insignificance between $Length$ and $BidAsk$ is corresponding with those presented in the main analysis suggesting that investors are indifferent about the conciseness of the Integrated Report.

Table A.15. Model 2: OLS regression of market effects and reading difficulty interaction term with EU firms

	(1) TobinsQ	(2) TobinsQ	(3) BidAsk	(4) BidAsk
ReadingDifficulty	-0.076	-0.082	0.329	-0.017
	(-3.32)***	(-1.81)*	(1.84)*	(-0.15)
ReadingDifficulty*IR_ITW	0.032	0.210	-0.557	-0.362
	(0.66)	(2.84)***	(-3.40)***	(-1.85)*
IR_ITW	-0.157	-0.042	-0.288	-0.198
	(-2.99)***	(-0.37)	(-1.95)*	(-0.67)
ROA	0.036	0.041	-0.035	-0.032
	(1.86)*	(5.11)***	(-1.41)	(-1.88)*
NegativeProfit	0.016	0.102	-0.310	-0.375
	(0.31)	(0.86)	(-1.85)*	(-1.20)
Accruals	0.852	0.906	-2.209	-0.130
	(1.51)	(1.25)	(-2.36)**	(-0.07)
DtoCE	0.012	0.027	0.014	0.085
	(1.21)	(2.56)**	(0.27)	(3.36)***
Beta	-0.217	-0.316	0.031	-0.017
	(-1.08)	(-2.92)***	(0.16)	(-0.08)
BTM	-0.115	-0.101	0.064	0.032
	(-1.79)*	(-1.86)*	(0.52)	(0.20)
LogAssets	-0.173	-0.099	-0.332	-0.473
	(-5.29)***	(-2.52)**	(-3.96)***	(-4.36)***
Complexity	0.725	1.140	1.777	-0.897
	(1.93)*	(2.62)**	(1.82)*	(-1.06)
LogFirmAge	0.052	-0.035	0.390	-0.118
	(2.25)**	(-0.66)	(1.92)*	(-0.52)
Constant	3.008	2.472	-5.933	-2.160
	(15.17)***	(4.63)***	(-5.29)***	(-1.62)
Year FE	Yes	No	Yes	No
Industry FE	Yes	No	Yes	No
Country FE	Yes	No	Yes	No
Observations	189	189	185	186
Adj. R-squared	0.846	0.564	0.879	0.545

Robust t-statistics in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Note: Clustered standard errors at firm level are applied to account for heteroscedasticity and autocorrelation in error terms. See Table 3 for variable definitions.

As seen in Table A.15(1), both $ReadingDifficulty$ (-0.076, t= -3.32) and IR_ITW (-0.157, t= 2.99) show statistically significant associations with $Tobin's\ Q$ on a 1% level. These findings are inconsistent with the results from the main test presented in Table 9(1) which show

no indication of significant results. The interaction term $ReadingDifficulty * IR\_ITW$ (0.032, t= 0.66) in Table A.15(1) supports the findings in Table 9(1) of no moderating effect of releasing an intertwined Integrated Reports on the relationship between reading difficulty and firm value.

To conclude, when only including EU firms we show that reading difficulty is negatively associated with firm valuation for stand-alone Integrated Reports and there is no moderating effect of choosing to disclose the Integrated Report intertwined. Also, firms releasing an intertwined Integrated Report has on average 0.157 lower firm valuation than firms issuing a stand-alone Integrated Report. When examining Table A.15(1) in isolation we find no support for Hypothesis II which is consistent with the main test presented in Table 9(1).

When testing the relationship between reading difficulty and bid-ask spread on EU firms the results in Table A.15(3) are highly consistent with the main test presented in Table 10(1). More specifically, we find a positive relationship between $ReadingDifficulty$ (0.329, t= 1.84) and $BidAsk$ which indicate that investors prefer stand-alone Integrated Reports to have low reading difficulty. Examining the negative relationship between the interaction term $ReadingDifficulty * IR\_ITW$ (-0.557, t= -3.40) and $BidAsk$, we can conclude that releasing an intertwined Integrated Report has a moderating effect on the relationship between reading difficulty and bid-ask spread. When adding the coefficients of $ReadingDifficulty$ and $ReadingDifficulty * IR\_ITW$ we get -0.228 which indicates that investors prefer intertwined Integrated Reports to have high reading difficulty as this reduces bid-ask spread and improves stock liquidity. The dummy variable IR_ITW (-0.198, t= -0.67) indicates that firms that release an intertwined Integrated Report have on average 0.198 lower bid-ask spread than firms releasing a stand-alone Integrated Report.

In line with the main test presented in Table 10(1) we find support of Hypothesis II and that releasing an intertwined Integrated Report has a moderating effect on the relationship between reading difficulty and bid-ask spread.

Table A.16. Model 3: OLS regression of market effects and length interaction term with EU firms

	(1) TobinsQ	(2) TobinsQ	(3) BidAsk	(4) BidAsk
Length	0.026	-0.086	0.088	-0.358
	(0.24)	(-0.63)	(0.20)	(-1.33)
Length*IR_ITW	-0.023	0.045	0.209	0.539
	(-0.20)	(0.30)	(0.45)	(1.67)
IR_ITW	-0.119	0.071	-0.577	-0.064
	(-1.70)*	(0.61)	(-1.88)*	(-0.23)
ROA	0.036	0.037	-0.038	-0.017
	(1.90)*	(4.49)***	(-1.59)	(-1.12)
NegativeProfit	0.049	0.152	-0.269	-0.490
	(1.01)	(1.18)	(-1.83)*	(-1.45)
Accruals	0.798	1.065	-1.963	-0.327
	(1.42)	(1.33)	(-2.13)**	(-0.16)
DtoCE	0.006	0.022	0.009	0.117
	(0.75)	(1.71)*	(0.17)	(4.01)***
Beta	-0.207	-0.253	-0.041	-0.104
	(-1.02)	(-2.48)**	(-0.19)	(-0.33)
BTM	-0.116	-0.084	0.048	-0.064
	(-1.55)	(-1.48)	(0.46)	(-0.28)
LogAssets	-0.156	-0.093	-0.419	-0.531
	(-4.05)***	(-2.09)**	(-3.77)***	(-5.33)***
Complexity	0.676	1.232	2.671	-1.175
	(1.82)*	(2.69)**	(3.44)***	(-1.16)
LogFirmAge	0.095	-0.002	0.123	-0.063
	(4.70)***	(-0.05)	(0.87)	(-0.27)
Constant	2.661	2.124	-3.828	-2.025
	(12.04)***	(4.11)***	(-4.51)***	(-1.42)
Year FE	Yes	No	Yes	No
Industry FE	Yes	No	Yes	No
Country FE	Yes	No	Yes	No
Observations	189	189	185	186
Adj. R-squared	0.841	0.527	0.871	0.512
Robust t-statistics in parentheses *** p<0.01, ** p<0.05, * p<0.1				

Note: Clustered standard errors at firm level are applied to account for heteroscedasticity and autocorrelation in error terms. See Table 3 for variable definitions.

The findings in Table A.16(1) show somewhat contradictory results from the main test presented in Table 11(1) since neither $Length$ (0.026, t= 0.24) nor $Length * IR\_ITW$ show significant results when controlling for EU firms. Also, in Table A.16(1) the dummy variable

IR_ITW(-0.119, t=-1.7) shows a negative association with $TobinsQ$ on a 10% significance level, which is not found in the main test presented in Table 11(1). Studying Table A.16(1) in isolation we find no support of Hypothesis II, which contradicts the conclusion from our main test.

Studying Table A.16(3) we find similar results as in the main test presented in Table 12(1) since both $Length$ (0.088, t=0.2) and $Length * IR\_ITW$ (0.209, t=0.45) show no significant association with $BidAsk$. The dummy variable IR_ITW (-0.577, t=1.88) shows a negative relationship with $BidAsk$ on a 10% significance level, which is aligned with the main test. From the findings in Table A.16(3) we find no support of Hypothesis II which is consistent with the conclusion drawn from the main test as presented in Table 12(1).

Table A.17. Model 1: OLS regression of market effects and textual attributes without factors

	(1)	(2)	(3)	(4)
	TobinsQ	TobinsQ	BidAsk	BidAsk
FleschKincaid	-0.299	-0.125	0.635	0.304
	(-1.96)*	(-0.96)	(2.97)***	(1.10)
FOG	0.336	0.125	-0.585	-0.281
	(2.01)*	(0.93)	(-2.77)***	(-0.98)
SMOG	-0.037	0.091	0.114	-0.077
	(-2.52)**	(3.07)***	(2.86)***	(-1.50)
LogNoCharacters	-0.430	-1.239	-1.087	-0.461
	(-1.00)	(-2.28)**	(-1.44)	(-0.51)
LogNoSentences	0.394	1.136	1.223	0.518
	(0.95)	(2.24)**	(1.74)*	(0.56)
ROA	0.046	0.045	-0.049	-0.048
	(2.37)**	(3.75)***	(-2.09)**	(-3.02)***
NegativeProfit	0.130	0.130	-0.091	-0.131
	(2.29)**	(0.94)	(-0.51)	(-0.56)
Accruals	0.940	1.790	-1.959	0.032
	(1.46)	(1.85)*	(-2.08)**	(0.02)
DtoCE	0.001	0.001	0.001	0.001
	(1.73)*	(1.67)	(0.89)	(0.81)
Beta	-0.171	-0.333	-0.088	-0.234
	(-0.99)	(-2.96)***	(-0.53)	(-1.04)
BTM	-0.124	-0.146	0.149	0.310
	(-1.89)*	(-1.51)	(1.21)	(1.73)*
LogAssets	-0.124	-0.080	-0.301	-0.377
	(-3.04)***	(-2.37)**	(-4.08)***	(-4.62)***
Complexity	1.382	1.005	2.252	-0.934
	(2.06)**	(2.11)**	(3.25)***	(-1.43)
LogFirmAge	0.120	0.092	0.430	-0.215
	(2.48)**	(1.20)	(2.51)**	(-1.25)
Constant	3.060	6.908	-2.742	2.481
	(1.80)*	(2.68)***	(-0.73)	(0.45)
Year FE	Yes	No	Yes	No
Industry FE	Yes	No	Yes	No
Country FE	Yes	No	Yes	No
Observations	232	235	232	235
Adj. R-squared	0.858	0.486	0.857	0.469
	Robust t-statistics in parentheses			
	*** p<0.01, ** p<0.05, * p<0.1			

Note: Clustered standard errors at firm level are applied to account for heteroscedasticity and autocorrelation in error terms. See Table 3 for variable definitions. The natural logarithm is performed on NoCharacters and NoSentences to account for large datapoints.

In Table A.17(1) significant associations between the reading indices and $TobinsQ$ are identified. $TobinsQ$ is significant and negatively associated with $TobinsQ$ (-0.299, t=-1.96) at a 10% statistical significance level. FOG on the other hand, is significant and positively associated with $TobinsQ$ at a 10% significance level (0.336, t=2.01) and $SMOG$ negatively associated with $TobinsQ$ at a 5% significance level. $NoCharacters$ is negatively associated with $TobinsQ$. The result is statistically insignificant (-0.430, t=-1.00), and $NoSentences$ is positively associated with $TobinsQ$. Again, the effect is statistically insignificant (0.394, t=0.95).

Table A.17 (3) shows the relationship between the three readability proxies ($FleschKincaid$, FOG, and $SMOG$) and $BidAsk$ as well as the relationship between the length proxies ($LogNoCharacters$ and $LogNoSentences$) and $BidAsk$. Again, mixed signs are identified for Readability proxies. $FleschKincaid$ is positively associated with $BidAsk$ (0.635, t=2.97), FOG negatively associated with $BidAsk$ (-0.585, t=-2.77) and $SMOG$ positively associated with $BidAsk$ (0.114, t=2.86). All these associations show statistical significance at a 1% level. $NoCharacters$ (-1.087, t=-1.44) and $NoSentences$ (1.223, t=1.74) have opposite signs, and $NoSentences$ also show statistical significance at a 10% level.

When performing correlation matrices (see Appendix A.12 and A.13) as well as the factor analysis (see Table 6), it is evident that there are high correlations between the readability indices and between the length proxies (>0.9), which implies presence of multicollinearity. Hence, the inconsistency and instability in coefficients and signs compared to the main analysis could be due to the presence of correlation between explanatory variables.

Table A.18. Model 1: OLS regression of market effects and textual attributes with median values

	(1) TobinsQ	(2) TobinsQ	(3) BidAsk	(4) BidAsk
ReadingDifficulty	0.045	0.033	0.106	-0.173
	(1.09)	(0.71)	(1.18)	(-1.16)
Length	-0.034	-0.112	-0.042	0.103
	(-0.98)	(-1.89)*	(-0.38)	(0.89)
ROA	-0.048	-0.013	-0.024	-0.006
	(-1.04)	(-0.68)	(-1.85)*	(-0.53)
NegativeProfit	-0.269	-0.139	-0.064	0.003
	(-1.06)	(-0.69)	(-0.31)	(0.02)
Accruals	-1.670	-2.397	1.391	1.136
	(-6.93)***	(-11.82)***	(2.52)**	(3.87)***
DtoCE	0.001	0.001	0.001	0.001
	(0.39)	(1.72)*	(1.76)*	(0.97)
Beta	-0.300	-0.250	-0.013	-0.188
	(-1.65)	(-2.22)**	(-0.09)	(-0.92)
BTM	-0.130	-0.204	0.216	0.326
	(-1.79)*	(-2.01)**	(1.77)*	(2.27)**
LogAssets	-0.145	-0.054	-0.371	-0.368
	(-3.00)***	(-2.25)**	(-4.05)***	(-8.12)***
Complexity	1.264	0.980	-0.273	-1.102
	(1.74)*	(2.09)**	(-0.18)	(-1.78)*
LogFirmAge	0.125	0.116	-0.017	-0.128
	(2.55)**	(1.69)*	(-0.14)	(-0.99)
Constant	2.500	1.603	-3.666	-3.089
	(7.60)***	(5.98)***	(-4.19)***	(-4.46)***
Year FE	Yes	No	Yes	No
Industry FE	Yes	No	Yes	No
Country FE	Yes	No	Yes	No
Observations	286	290	286	290
Adj. R-squared	0.647	0.389	0.798	0.603
Robust t-statistics in parentheses *** p<0.01, ** p<0.05, * p<0.1				

Note: Clustered standard errors at firm level are applied to account for heteroscedasticity and autocorrelation in error terms. See Table 3 for variable definitions. The natural logarithm is performed on NoCharacters and NoSentences to account for large data points.

When replacing the missing values for each variable with the corresponding median value, our regression includes 286 observations in Table A.18(1). The reported results are highly in line with those presented in the main analysis. An insignificant positive association between

$ReadingDifficulty$ and $TobinsQ$ is identified, suggesting that firm value is unaffected by the length of the Integrated Report.

The robustness analysis in Table A.18(3) also supports the main analysis for stock liquidity. We find statistically insignificant associations between $ReadingDifficulty$ and $BidAsk$ as well as between $Length$ and $BidAsk$ which supports the argument that investors are indifferent about the complexity and conciseness of Integrated Reports.

Table A.19. Model 1: OLS regression of market effects and reading difficulty interaction term with median values

	(1) TobinsQ	(2) TobinsQ	(3) BidAsk	(4) BidAsk
ReadingDifficulty	0.030	0.023	0.243	0.023
	(0.48)	(0.42)	(1.76)*	(0.42)
ReadingDifficulty*IR_ITW	0.061	0.013	-0.354	0.013
	(0.62)	(0.16)	(-2.35)**	(0.16)
IR_ITW	0.249	0.020	-0.484	0.020
	(1.36)	(0.19)	(-2.79)***	(0.19)
ROA	-0.050	-0.012	-0.018	-0.012
	(-1.07)	(-0.61)	(-1.54)	(-0.61)
NegativeProfit	-0.222	-0.157	-0.170	-0.157
	(-0.94)	(-0.81)	(-0.86)	(-0.81)
Accruals	-1.627	-2.413	1.321	-2.413
	(-6.40)***	(-10.63)***	(2.45)**	(-10.63)***
DtoCE	0.000	0.001	0.001	0.001
	(0.14)	(1.98)*	(2.78)***	(1.98)*
Beta	-0.284	-0.233	-0.025	-0.233
	(-1.54)	(-2.23)**	(-0.18)	(-2.23)**
BTM	-0.117	-0.203	0.144	-0.203
	(-1.72)*	(-2.13)**	(1.27)	(-2.13)**
LogAssets	-0.142	-0.057	-0.391	-0.057
	(-2.99)***	(-2.31)**	(-4.98)***	(-2.31)**
Complexity	1.132	0.911	-0.020	0.911
	(1.78)*	(1.89)*	(-0.01)	(1.89)*
LogFirmAge	0.118	0.112	0.021	0.112
	(2.21)**	(1.58)	(0.19)	(1.58)
Constant	2.319	1.623	-3.248	1.623
	(5.50)***	(5.34)***	(-3.87)***	(5.34)***
Year FE	Yes	No	Yes	No
Industry FE	Yes	No	Yes	No
Country FE	Yes	No	Yes	No
Observations	286	290	286	290
Adj. R-squared	0.652	0.374	0.814	0.374

Robust t-statistics in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Note: Clustered standard errors at firm level are applied to account for heteroscedasticity and autocorrelation in error terms. See Table 3 for variable definitions.

Presented in Table A.19(1), insignificant associations are identified for $ReadingDifficulty$ (0.030, t=0.48), the dummy variable IR_ITW (0.249, t=1.36), and the interaction term $ReadingDifficulty * IR\_ITW$ (0.061, t=0.62) with $TobinsQ$. These

findings are consistent with those presented in the main analysis (Table 9(1)). These results are also robust when no fixed effects are applied in Table A.19(2).

Results are also highly aligned in Table A.19(3) with those results presented in the main analysis (Table 10(3)) between $ReadingDifficulty$ and $BidAsk$. To be precise, we find a positive relationship between $ReadingDifficulty$ (0.243, t=1.76) and $BidAsk$. This implies that investors prefer stand-alone Integrated Reports to be readable. There is a negative relationship between the interaction term $ReadingDifficulty * IR\_ITW$ and $BidAsk$ (-0.354, t=-2.35) inferring that there is a moderating effect of issuing an intertwined Integrated Report. The total effect after adding the coefficients for $ReadingDifficulty$ and $ReadingDifficulty * IR\_ITW$ is –0.111 which indicates that investors prefer Integrated Reports to have high reading difficulty as this reduces bid-ask spread and improves stock liquidity. The dummy variable IR_ITW (-0.484, t=-2.79) shows that firms issuing intertwined Integrated Reports on average have a 0.484 lower $BidAsk$ compared to firms issuing stand-alone Integrated Reports. A slightly lower significance level is identified when replacing missing values with median values, but overall, the results are highly consistent with the main analysis.

Table A.20. OLS regression of market effects and length interaction term with median values

	(1)	(2)	(3)	(4)
	TobinsQ	TobinsQ	BidAsk	BidAsk
Length	-0.293	0.003	-0.222	0.003
	(-1.81)*	(0.02)	(-0.45)	(0.02)
Length*IR_ITW	0.231	-0.177	0.333	-0.177
	(1.41)	(-1.06)	(0.71)	(-1.06)
IR_ITW	0.410	0.078	-0.360	0.078
	(1.79)*	(0.53)	(-1.01)	(0.53)
ROA	-0.051	-0.013	-0.020	-0.013
	(-1.09)	(-0.70)	(-1.50)	(-0.70)
NegativeProfit	-0.227	-0.127	-0.141	-0.127
	(-0.97)	(-0.66)	(-0.68)	(-0.66)
Accruals	-1.609	-2.355	1.298	-2.355
	(-6.25)***	(-12.08)***	(2.33)**	(-12.08)***
DtoCE	-0.000	0.001	0.001	0.001
	(-0.16)	(1.21)	(0.98)	(1.21)
Beta	-0.281	-0.237	-0.039	-0.237
	(-1.51)	(-2.16)**	(-0.27)	(-2.16)**
BTM	-0.118	-0.181	0.118	-0.181
	(-1.71)*	(-1.93)*	(1.10)	(-1.93)*
LogAssets	-0.135	-0.050	-0.401	-0.050
	(-2.79)***	(-2.15)**	(-4.44)***	(-2.15)**
Complexity	1.108	1.061	0.406	1.061
	(1.71)*	(2.26)**	(0.28)	(2.26)**
LogFirmAge	0.122	0.110	-0.035	0.110
	(2.10)**	(1.76)*	(-0.32)	(1.76)*
Constant	2.087	1.522	-3.135	1.522
	(4.38)***	(4.82)***	(-3.67)***	(4.82)***
Year FE	Yes	No	Yes	No
Industry FE	Yes	No	Yes	No
Country FE	Yes	No	Yes	No
Observations	286	290	286	290
Adj. R-squared	0.655	0.394	0.806	0.394

Robust t-statistics in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Note: Clustered standard errors at firm level are applied to account for heteroscedasticity and autocorrelation in error terms. See Table 3 for variable definitions.

Presented in Table A.20(1), a significant association for $Length$ (-0.293, t=-1.81) of stand-alone Integrated Reports is identified with $TobinsQ$ at a significance level of 10%. This result stipulates that the firm value of firms issuing stand-alone Integrated Reports is negatively affected by the length of the report. This is inconsistent with results from the main analysis

presented in Table 11(1) where the opposite direction for stand-alone Integrated Reports was reported. Findings on no missing values are also inconsistent with the main analysis when considering the interaction term $Length * IR\_ITW$ and the dummy variable for IR_ITW. Whereas an insignificant association is presented in Table A.20(1) for $Length * IR\_ITW$ (0.231, t=1.41), a negative significant association is reported in the main analysis, implying that firm value is negatively affected by the length of intertwined Integrated Reports. We can neither confirm the robustness of the effect of IR_ITW (0.410, t=1.79) as this is significant in Table A.20(1) but insignificant in the main analysis. Importantly, the results in the main analysis (Table 11(1)) and the robustness results in Table A.20(1) are only statistically significant at a 10% level which increases the risk of type I errors compared to if the regression showed significance levels at 5% or 1%. Additionally, the statistical significance ceases when no fixed effect structures are controlled for.

In Table A.20(3) the moderating effect of releasing an intertwined Integrated Reports on the relationship between length and bid-ask spread is examined. The results are highly consistent with findings reported in the main analysis Table 12(1). No significant associations are identified for $Length$ (-0.222, t=-0.45) or the interaction term $Length * IR\_ITW$ (0.333, t=0.71). While the main analysis indeed found IR_ITW to have a negative effect on $BidAsk$, this association is not identified in the robustness analysis.

To summarize, the robustness results where the missing values are replaced with median values are to some extent inconsistent with results presented in the main analysis. We are therefore careful with claiming inference of our results. Considering the findings of Table A.20 and Table 11 in isolation, we cannot conclude support for Hypothesis II due to contradictory and unstable coefficients and significance levels.