

Efretuei, E, Usoro, A and Koutra, C

Complex information and accounting standards: Evidence from UK narrative reporting

<https://researchonline.ljmu.ac.uk/id/eprint/17709/>

Article

Citation (please note it is advisable to refer to the publisher's version if you intend to cite from this work)

Efretuei, E ORCID logoORCID: <https://orcid.org/0000-0002-4671-9841>, Usoro, A and Koutra, C (2021) Complex information and accounting standards: Evidence from UK narrative reporting. South African Journal of Accounting Research. pp. 1-24. ISSN 1029-1954

LJMU has developed [LJMU Research Online](#) for users to access the research output of the University more effectively. Copyright © and Moral Rights for the papers on this site are retained by the individual authors and/or other copyright owners. Users may download and/or print one copy of any article(s) in LJMU Research Online to facilitate their private study or for non-commercial research. You may not engage in further distribution of the material or use it for any profit-making activities or any commercial gain.

The version presented here may differ from the published version or from the version of the record. Please see the repository URL above for details on accessing the published version and note that access may require a subscription.

For more information please contact researchonline@ljmu.ac.uk

Complex information and accounting standards: Evidence from UK narrative reporting

Ekaete Efretuei, Abel Usoro & Christina Koutra

To cite this article: Ekaete Efretuei, Abel Usoro & Christina Koutra (2021): Complex information and accounting standards: Evidence from UK narrative reporting, South African Journal of Accounting Research, DOI: [10.1080/10291954.2021.1970450](https://doi.org/10.1080/10291954.2021.1970450)

To link to this article: <https://doi.org/10.1080/10291954.2021.1970450>



© 2021 The Author(s). Co-published by NISC Pty (Ltd) and Informa UK Limited, trading as Taylor & Francis Group



Published online: 02 Dec 2021.



Submit your article to this journal [↗](#)



Article views: 954



View related articles [↗](#)



View Crossmark data [↗](#)

Complex information and accounting standards: Evidence from UK narrative reporting

Ekaete Efretuei *

Newcastle University Business School, Barrack Road, NE1 4SE Newcastle University, Newcastle Upon Tyne, United Kingdom

Abel Usoro

Department of Mathematics and Computer Science, Ritman University, Nigeria

Christina Koutra

College of Business, Abu Dhabi University, Abu Dhabi, United Arab Emirates

(Received 1 November 2020; accepted 11 August 2021)

The application of International Financial Reporting Standards (IFRS) has been introduced in many countries to enhance efficiencies in financial markets and improve communication in financial reporting. However, extant studies have suggested that the introduction of IFRS has increased narrative complexity, owing to the demand for more reporting. Considering that accounting complexity can be either informative (enhancing understanding) or non-informative, thereby causing obfuscation, this study performs an empirical analysis to highlight which of the two types of complexities may be affected by IFRS application. Using the setting of IFRS adoption in the UK and a word list-adjusted component of the fog index, this study decomposes complexity into two components: information (common complexity) and obfuscation (uncommon complexity). The results reveal that IFRS adoption has increased the common complexity of accounting narratives (information) but does not necessarily increase obfuscation. The study's contribution is twofold: methodological through the decomposition of complexity using the term weighting concept and policy-related by identifying areas of increased narrative comparability in IFRS reports. Moreover, the study's application of complexity decomposition to IFRS is novel. Future studies may apply this by using the identified information and obfuscation components to investigate the economic consequences of IFRS-associated complexity.

Keywords: complexity; International Financial Reporting Standards (IFRS); information; accounting narratives; readability; obfuscation

Data availability

Data are available from the public sources cited in the text.

*Corresponding Author. Email: ekaete.efretuei@newcastle.ac.uk

© 2021 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

South African Journal of Accounting Research is co-published by NISC Pty (Ltd) and Informa Limited (trading as Taylor & Francis Group)

1. Introduction

This study investigates changes in the narrative complexity of corporate annual reports associated with the mandatory adoption of accounting standards. There are two motivations for this study. The first is the increase in complex financial narrative disclosures¹ (Lo et al., 2017), and the associated economic consequences (Asay et al., 2017; Callen et al., 2013; De Franco et al., 2015; Lee, 2012). The second is the role of accounting standards in increasing the complexity of narrative disclosures (Dyer et al., 2017). Here, the complexity of narrative disclosures refers to report readability, specifically the use of complex words, while accounting standards refer to the International Financial Reporting Standards (IFRS). Extant research on narrative complexity in IFRS compliant annual reports shows that IFRS increase complexity (Lang & Stice-Lawrence, 2015) and that this increase can be driven by increased disclosures commingled with complex narratives. In this study, we address the gap in the literature in relation to whether this increased complexity is informative or obfuscating. This study adopts the view that reporting incentives define reporting choices, implying that IFRS, despite being aimed at report comparability, allow discretion in reporting (Leuz & Wysocki, 2016). Firms exercise reporting discretion according to their expectation of positive outcomes. For example, Li (2008) shows that obfuscation of narratives increases with negative changes in earnings.

Within this incentives view, firms have the choice to provide either informative or obfuscated disclosures (Guay et al., 2016), moderated by the application of IFRS (Neel, 2017). This study is based on the premise that narrative complexity commingles the two latent components of information and obfuscation (Bushee et al., 2018). The information component of narrative complexity represents the technical nature of accounting narratives – for example, the use of complex words such as “financial” or “incorporated”. The obfuscation component represents obscurity or perplexity and is intended to reduce the informativeness of narrative disclosure – for example, the excessive use of non-technical complex words. This study decomposes narrative complexity into i) an informative component measured as “common complexity”, that is, complex words that are commonly used in annual reports, and ii) an obfuscation component measured as “uncommon complexity”, or complex words that are not regularly used in annual reports. This study uses both components to examine whether the adoption of IFRS increases the narrative complexity of annual reports.

The findings show that narrative complexity associated with the adoption of IFRS appears to increase in terms of informative complexity. Further analysis shows that this increase in narrative complexity is significantly associated with IFRS adoption after controlling for the expected increase over the years, that is, the changes in complexity with IFRS adoption and the changes in reported economic effects with IFRS adoption. However, there is an increase in narrative obfuscation² that tends to disappear after controlling for time effects. The observed increase in obfuscation appears to be consistent with a general increase in obfuscation over time.

This study contributes to the existing literature in two respects. First, in terms of methodology, this study applies a measure of complexity based on the reporting incentives view to analyse the two components of narrative complexity: incentive to obfuscate and incentive to inform. It does this using the term weighting concept (Jegadeesh & Wu, 2013). Second, with regard to policy, this study contributes to the International Accounting Standard Board’s (IASB, 2017) Disclosure Initiative project and supports the European Financial Reporting Advisory Group’s (EFRAG) recommendation to address complexity in the conceptual framework (EFRAG, 2014) by indicating which

components of complexity are influenced by IFRS adoption. The results show that the application of IFRS increases the informative component of disclosure. The IASB's Disclosure Initiative project is focused on addressing three aspects of the disclosure problem: (i) not enough relevant information, (ii) irrelevant information, and (iii) ineffective communication of information.³ This study finds that the application of IFRS does not complement the disclosure problem. In fact, the observed increase in "common complexity" suggests an increase in the informativeness and, to an extent, comparability of narratives post-IFRS adoption.

The remainder of the paper is structured into six sections. Section 2 reviews the relevant literature. Section 3 presents the narrative complexity classification. Section 4 develops the hypothesis. Section 5 discusses the research design. Section 6 presents the empirical results, and Section 7 concludes the paper.

2. Literature review

Studies show that the reading difficulty of annual reports has increased over the years (Dyer et al., 2017). Li (2008) finds that firms whose annual reports are less readable or difficult to read have lower earnings persistence, while firms whose annual reports are easier to read have more persistent positive earnings. Even if companies apply the same reporting standards, differences arise in reporting outcomes owing to internal and external factors that shape incentives (Leuz & Wysocki, 2016). Supporting the view that reporting incentives influence readability, Lo et al. (2017) find that increased reading difficulty is associated with firms that have managed to surpass the prior year's earnings, concluding that obfuscation incentives increase the reading difficulty of narratives. Similarly, Dempsey et al. (2012) conclude that narrative complexity is higher in the reports of poorly performing firms.

Both internal and external factors shape reporting incentives. Internal factors include the company's operating characteristics, while external factors include legal institutions and enforcement regimes or capital market forces (Leuz & Wysocki, 2016). In applying IFRS, firms with incentive to provide informative disclosures tend to increase the information components of their narrative disclosures, while obfuscation incentives are more likely to increase the non-informative components of disclosures. A number of studies find that bigger, more complex, and more volatile firms tend to have more complex disclosures (Li, 2008). However, application of IFRS can moderate this complexity through disclosure requirements. Dyer et al. (2017) analyse the narrative trends in annual reports over the years and report poorer readability. They find that regulatory requirements explain most of the changes and topics; for example, fair value and risk factor disclosures.

Studying IFRS and readability, Cheung and Lau (2016) examine the notes appended to financial statements and report that they have become lengthier following the adoption of IFRS. Interestingly, they observe that narrative disclosures have increased in length, while readability scores have decreased. A lower readability score implies improved readability of the text. Thus, in addition to improving disclosure quantity, IFRS may increase the informative components of disclosure. Accounting regulations such as IFRS are viewed as governing extraction cost (Bloomfield, 2002). The incomplete revelation hypothesis (IRH) is based on a "noisy rational expectations" model, which assumes that the cost of extracting information from publicly available data increases with information complexity (Bloomfield, 2002). However, by introducing IFRS, information becomes standardised, comparable, and accessible to users, thereby fulfilling the standards' role of increasing the informative components of disclosure.

Even if IFRS were to be introduced with the intention of reducing extraction cost,⁴ a firm's reporting incentives would affect the application of the standards; moreover, firms' discretion in application provides obfuscation opportunities. Firms increase the complexity of their narrative disclosures when there are benefits to be gained by doing so. Numerous studies show that narrative complexity is associated with market incentives (Lawrence, 2013; Miller, 2010; Rennekamp, 2012; Tan et al., 2015; You & Zhang, 2009) and investor disclosure choices (Asay et al., 2017; Filzen & Peterson, 2015). According to You and Zhang (2009), investors' under-reaction to 10-K disclosures appears sluggish and tends to be more prominent for firms with complex disclosures. Filzen and Peterson (2015) reveal that analysts' reliance on narratives increases with the complexity of financial disclosures. Lawrence (2013) and Miller (2010) show that the complexity of disclosures affects retail and small investors' reactions and disclosure choices.

IFRS also increase the narrative complexity of disclosures. Richards and Staden (2015) hypothesise that the complexity of accounting standards tends to increase the associated narrative complexity in annual reports, contrary to Cheung and Lau (2016), who find that IFRS adoption improves the readability of narrative disclosures. Most of these studies use the "fog index" readability measure, which is widely used in the accounting literature to measure the reading difficulty of accounting narratives (e.g., Li, 2008; Lo et al., 2017). The fog index uses a combination of complex words and complex sentences based on the assumption that, on average, more words per sentence and more syllables per word make a document more difficult to read.

The impact of IFRS on narrative complexity appears to be mixed, but there is consensus that the accounting standard increases disclosure length. Lang and Stice-Lawrence (2015) find that the "fog index" of annual reports has increased over the years and attribute this trend to the disclosure of complex topics, such as financial instruments, currencies, and taxes. The authors also study the benefits of IFRS and show that firms with a higher fog index experience fewer benefits from IFRS.⁵ If IFRS are associated with increased complexities characterised by the fog index, and a higher fog index is associated with poor economic outcomes – such as a reduction in IFRS benefits, earnings management, and lower earnings – then IFRS application would likely provide discretionary opportunities for increasing reporting outcomes and variation in narrative complexity.

Christensen et al. (2013) find increased liquidity benefits with enforcement during the IFRS application period. In the United Kingdom (UK), enforcement was characterised by the change from reactive to proactive reviews. A reactive approach is a limited review of companies brought to the regulator's notice, while a proactive approach involves identification of a broad sample of companies for review (Christensen et al., 2013).⁶ Proactive reviews evaluate 15% of interim and annual financial reports, which creates new incentives for narrative obfuscation. This is because in a proactive regime, consistent with the expectation of the management obfuscation hypothesis (Li, 2008), firms that have incentives to manage earnings through poor application of IFRS tend to obfuscate reported disclosures to circumvent selection. Meanwhile, firms that are more likely to enhance reporting outcomes with IFRS adoption are also more likely to report clear and informative disclosures.

Isolated benefits of IFRS are observed in countries with smaller differences between local accounting standards used pre-IFRS adoption and IFRS. For instance, using a sample of UK firms, Brochet et al. (2013) observe that IFRS increase overall comparability. Lang and Stice-Lawrence (2015) observe an increase in the narrative comparability of reports following IFRS adoption and attribute this to an increase in relative word similarities across reports. While the UK has had limited differences in standards pre- and post-IFRS adoption (Bae et al., 2008), the exogenous effect of increased comparability

influences disclosure behaviour (Brochet et al., 2013). Bae et al. (2008) show that the pre-IFRS UK accounting standards are similar to IFRS, indicating limited changes in disclosure. The motivation for a change in disclosure post-IFRS adoption implies increased uniformity in standards across multiple countries (Brochet et al., 2013). Firms tend to signal private information using narratives because standards have become more accessible and comparable. The annual reports of UK firms have both regulated and unregulated narrative components, observed as a potential mechanism for discretionary reporting (private information).

IFRS have focused less on wider narrative reporting, with no separate standard for it (only a practice statement, “The IFRS Management Commentary Practice Statement”). This creates the expectation that IFRS regulation will increase the complexity of narrative reporting because it is not standardised and, thus, allow obfuscation. Conversely, the application of standards can increase narrative complexities through improved discussions on complex financial disclosures, such as fair value disclosures (Dyer et al., 2017), and an increase in narrative comparability (Lang & Stice-Lawrence, 2015). In the next section, we discuss the decomposition of complexity in narrative reporting and how it is applied in this study.

3. Narrative complexity classification

In this study, we utilise the concept of “word complexity” (derived from the fog index)⁷ for classifying narrative complexity. In particular, this study is based on the premise that the theoretical construct of narrative complexity is composed of two latent components, information and obfuscation (Bushee et al., 2018), and that the target audience tends to understand only one component of complex words in annual reports (Loughran & McDonald, 2014). This complex component is labelled the informative component. Thus, this study decomposes the observed complex words in narratives into two components: i) information labelled as common complexity and ii) obfuscation labelled as uncommon complexity.

To measure these common and uncommon complexity components, the theoretical concept of “term weighting” is applied. The core idea underlying the term weighting technique is that the effectiveness or impact of a word in any text is inversely related to its occurrence; that is, if a word occurs with higher frequency, it should be weighted with less impact than a word that occurs less frequently in the text (Manning & Schutze, 2003). For example, the word “and” occurs frequently in most texts, but its high frequency does not equate with its impact on those who read the text because the more readers see the word “and”, the more likely they are comfortable with the meaning of the word.

The process of term weighting involves the calculation of the TF.IDF weight of each word, which considers the occurrence frequency of words both in the defined text and in the defined corpus. In this context, the term “defined text” is a company’s annual report, whereas “defined corpus” is the universe of all company annual reports in the sample. By calculating the frequency of a word in the corpus, we can consider whether that word is concentrated in the defined text or in the defined corpus. The purpose is to account for the likelihood that the readers of similar texts are “comfortable” with the word. The occurrence of words with higher TF.IDF weights is less frequent in the defined corpus than words with lower TF.IDF weights.

Based on the concept of term weighting, more “familiar” words in the annual report corpus are expected to have lower weights, and less “familiar” words are expected to have higher weights. The term TF.IDF has been applied in accounting narrative studies to

moderate narrative measures. Specifically, Loughran and McDonald (2011) apply it to moderate the impact of high-frequency words in the repetition of tonal words in 10-K disclosures. TF.IDF originates from the concept of “term weighting” in the literature on information retrieval.⁸ It is composed of two terms: TF and IDF. The first term, TF, attenuates the impact of high-frequency complex words in a document. It is a normalised frequency of the word in the document applied. The second term, IDF, modifies the impact of a complex word in a document based on its commonality within a defined corpus (Loughran & McDonald, 2011). It indicates whether the word is common or rare in the corpus of documents. The assumption is that the impact of a complex word reduces as its frequency increases within the corpus. Thus, a high-IDF term decreases the impact of the first term, TF.

Manning and Schütze (2003) denote TF.IDF as (i) highest when a term (complex word) occurs repeatedly in a few documents in the defined corpus,⁹ (ii) lower when the occurrence is less frequent in a document or occurs in multiple documents, and (iii) lowest when the term occurs in virtually all documents.

This study applies the theoretical assumptions of “term weighting” to identify the two categories of complexity. It measures uncommon complexity based on the assumption that complex words with high TF.IDF weights tend to increase obfuscation. This is because high TF.IDF words are less frequent in the defined corpus and, thus, are less familiar to the users of the corpus (Manning & Schütze, 2003). Following this assumption, we expect that words with high frequency are more likely to be familiar to readers, consistent with prior literature assumptions (e.g., Loughran & McDonald, 2014). If a complex word occurs in a firm’s annual report but is not common in the identified corpus, it is assumed that its persistent use is associated with obfuscation. Kim et al. (2019) find that an adjusted word complexity component of the fog index, which is adjusted based on the concept of “familiar words”, is significantly related to the post-filing date return volatility as opposed to the traditional fog index.

Conversely, words with low TF.IDF scores have greater narrative complexity,¹⁰ although they are more informative. Complex words in this category occur more frequently in the annual report corpus and are considered familiar to investors. Based on Manning and Schütze’s (2003) definition, TF.IDF terms are lowest when they occur in virtually all documents. In this study, the implication is that words identified as low TF.IDF complex words tend to occur in all documents in the defined corpus. Thus, we identify words with low TF.IDF scores to test the association between the application of IFRS and common or uncommon complexity.¹¹ The concept of TF.IDF is an appropriate measure to decompose words theoretically into informative and obfuscating components of complexity, given its extensive application in the information retrieval literature to moderate the impact of words in a corpus.¹²

Based on the theoretical observation that narrative complexity measures two latent components (information and obfuscation), this study deems that words with low TF.IDF weights indicate information. These words are either informative (necessary technical details) or common (persistent across reports).¹³ To identify low TF.IDF words, this study draws from the identification of 52 words that account for a quarter of the total count of complex words in annual report narratives¹⁴ (Loughran & McDonald, 2014). Loughran and McDonald (2014) identify them as “familiar words” that users of annual reports are accustomed to, and their complexity is informative rather than obfuscating.

4. Hypothesis development

The study aims to answer the following research question: what is the role of accounting standards in mitigating obfuscation and enhancing the informative component of complex

annual report narratives? The hypothesis is developed based on the theoretical views of information and obfuscation (Asay et al., 2018; Bushee et al., 2018; Lo et al., 2017; Stone & Parker, 2016). Information is interpreted as driven by the increased complexity of technical disclosures, while obfuscation is interpreted as driven by increased complexity not attributable to informative disclosures (Bloomfield, 2008; Bushee et al., 2018; Merkl-Davies & Brennan, 2007). The analysis of complexity is at the word level; this implies that increased disclosure of uninformative complex words indicates increased complexity (obfuscation), while disclosure of informative technical words represents information (Courtis, 2004).

Several studies report that IFRS adoption increases the fog index (Lang & Stice-Lawrence, 2015; Richards & Staden, 2015). However, it is unclear whether this increase is driven by the increased disclosure of informative complex words or the increased disclosure of obfuscating complex words. With IFRS application, discretionary opportunities in disclosure can lead to either informative reporting outcomes or uninformative outcomes depending on the reporting incentive (Christensen et al., 2015; Gao & Sidhu, 2017; Leuz & Wysocki, 2016). If IFRS application increases the complexity of reports, and firms have lower IFRS benefits such as reduced liquidity and analyst following (Lang and Stice-Lawrence, 2015), it is more likely that IFRS application increases the obfuscation component of complexity. Moreover, lower liquidity and analyst following indicate that a firm's information environment is obscure. However, if IFRS foster improved communication (Tarca, 2004), clear disclosure relative to local requirements, and increased technical disclosure (Lang & Stice-Lawrence, 2015), then IFRS adoption would be more likely to increase the disclosure of complex informative words.

The complexity index that is used to measure information and obfuscation is a function of weighted word occurrence and syllabification. The first input is a component of the fog index studied by Lo et al. (2017) in the context of financial performance. Li (2008) finds a negative relationship between firm profitability and the fog index. Bloomfield (2008) provides several alternative explanations for Li's (2008) observed results, including obfuscation and information. Li's (2008) findings are consistent with the observation that management uses linguistic obfuscation to make textual disclosures more complex when performance is poor. Bushee et al. (2018) provide a theoretical view by decomposing the two components using the fog index, where information is aimed at explaining disclosure complexity as a necessary detail. This study uses a component measure of the fog index to estimate the complexity index.

Specifically, this study uses the complex-words component of the fog index to calculate "common" and "uncommon" complexity. Loughran and McDonald (2014) analyse the readability of accounting narratives and find that the component of complex words in the fog index does not explain the market's reaction to disclosures. The authors explain that investors easily understand words such as "management" and "corporation" despite that they are termed "complex". While these words have two or more syllables, they are familiar to users of such reports. The authors show that these complex words account for the first quartile of the most frequently occurring complex words in annual reports. Bonsall et al. (2017) identify the components of plain English attributes that capture narrative complexity, namely, the bog index. The word "bog" is designed to capture "word difficulty", which "bogs" readers down. It is calculated using a general vocabulary, identifying words that are consistent with the notion of "familiar" words in annual reports (Bonsall et al., 2017). Both studies find that certain words that are often identified as complex are familiar to investors, and the use of such words in annual reports is informative rather than obfuscating.

Following this notion, if IFRS adoption increases the disclosure of “familiar” words in annual reports, it would increase the information component of narrative complexity, and this is identified as common complexity (information). Alternatively, if the adoption of IFRS increases the disclosure of complex words that are not identified as “familiar” in annual reports, it would increase obfuscation, and this is identified as uncommon complexity. Given that IFRS application is expected to improve informativeness through better quality technical disclosures, it is likely to persist after first-time adoption. Therefore, we propose the following testable hypothesis:¹⁵

Hypothesis 1 (H₁): Informative disclosures in annual reports (common complexity) are positively correlated with IFRS application for 2 years after first-time IFRS adoption.

5. Research design

The empirical analysis tests whether increased narrative complexity following IFRS application is associated with the disclosure of complex words that are more informative in annual reports. The general equation used to test H₁ is:

$$Complexity_{it} = Intercept + \beta_1 year_ifrs_{it} + \beta_2 Controls + \varepsilon_{it} \quad (1a)$$

The sampling process and variables are discussed in the below subsections.

5.1 Sample selection and data preparation

To test H₁, this study uses the annual reports of a sample of companies listed on the FTSE All-Share (FTALSH) index. This study assesses the changes in disclosure complexity between the pre- and post-IFRS adoption periods. It begins with firms listed on the FTSE All-Share index between 2003 and 2007. To capture this, the study uses the annual reports of firms for the years ended 2003 and 2004 for the pre-adoption period and for 2006 and 2007 for the post-adoption period. Similar to Brochet et al. (2013), the sample allows for a balanced set of years and excludes the consequences of narrative reporting during the 2008 financial crisis. The adoption year is excluded to eliminate the effects of narrative reporting from the results for first-time IFRS adoption. The narrative data for each firm must include 4 years of data to enable a matched comparison of the pre- and post-adoption years.

To retrieve the annual report narratives for each firm in the FTALSH index, the reports were downloaded from the Thomson ONE database and the respective company websites of firms included in the sample. In line with the methods used in the readability literature, the text extraction process excludes sections with more than 50% numerical content as opposed to textual content (Li, 2008; Miller, 2010). The data collection process excludes interim reports presented as annual reports. For firm-year observations, the study eliminates annual reports with fewer than 2000 words from the sample and firm-year observations with missing or extreme values. Furthermore, the study eliminates firm years with no matching financial statement data (see Table 1 for details of the data collection and sampling process). The final study sample comprises 939 firm years following incorporation of financial statement data results. The sample size varies across the control variables owing to data availability. The sample sizes are reported for the regression analysis.

Table 1. Data collection.

Event	Firm Years
Initial annual report textual output for 2003, 2004, 2006, and 2007	1,505
Less: firm years that have data missing for the above 4 years	(497)
Less: firm years with missing/extreme values for <i>lowtfidf</i> scores (close checking indicated poor word conversion of pdf report)	(56)
Final narrative firm-year observations =	952
Less: firm years with missing financial statement data	(13)
Final firm-year observations (used in hypothesis test)	939

5.2 Variable definitions

Table 2 presents the definitions of the key variables and control variables applied in the regression. The identified control variables are as applied in similar studies on the readability of annual reports (e.g., Li, 2008). The data collection steps indicating the estimation of both common and uncommon complexity are discussed in Appendix I. The list of words identified as low TF.IDF is provided in Appendix II.

H₁ involves the influence of mandatory IFRS adoption regulation regarding textual word complexity (common and uncommon complexity). Therefore, it requires measures of word complexity decomposed into two components. This study does not measure readability (fog index) or length but, rather, focuses on word complexity in evaluating the informative and obfuscating components driving complexity following IFRS application. To provide a benchmark for the increase in information during the sample years, the logarithm of the number of words is captured as a control variable.

6. Empirical results

6.1 Descriptive statistics

Table 3 provides summary statistics of the narrative and control variables. Panel A provides descriptive statistics for the complete sample. Panel B shows descriptive statistics for the pre- and post-IFRS group. Reports, on average, contain 36,000 words. From panel B, it can be observed that the length of reports is smaller in the pre-IFRS period, compared to the post-IFRS period. The *numcompwords* variable shows an increase in IFRS years, which may indicate the disclosure of complex topics. Deconstructing this further, the variables *hightfidf comp words* and *lowtfidf comp words* both increase in the post-IFRS years. This indicates that both the informative and obfuscating components of complexity, as measured in this study, increase during the IFRS application years. The variable *comp words* is a component of the fog index and captures average word complexity. As seen in Table 3, the average fog index of the sample of reports is 22, with a minimum of 16 and maximum of 31. This is consistent with prior literature, which observes similar levels of fog index scores.¹⁶ *Profit_loss* is an indicator variable where 1 = profit firms and 0 = loss firms. The mean of 0.92 for *profit_loss* indicates that the sample is skewed toward profit-making firms, comprising 92% of the sample.

Table 4 presents correlation results of the variables included in the testing of H₁ in the regressions. The variables *hightfidf* and *lowtfidf* are positively correlated, although the correlation is below 10%, indicating that they are not interchangeable. The correlations between the narrative variables and determinant variables are consistent with the

Table 2. Variable definitions.

Variable	Full Name	Definitions
<i>numrepwords</i>	Report Length	Number of words in the annual report
<i>numcompwds</i>	Report Complexity	Number of complex words in the annual report
<i>comp words</i>	Percent Report Complexity	[(Number of Complex Words in the Annual Report / Total Number of Words in the Same Report) *100]. Complex words are words with three or more syllables. This is the component of the fog index of interest in this study. The fog index is defined as: $Fog = 0.4(words\ per\ sentence + percent\ of\ complex\ words)$.
<i>hightfidf comp words</i>	High TF.IDF Words	Number of high TF.IDF complex words in the annual report
<i>hightfidf</i>	Percent High TF.IDF	Percentage of total words in the annual report that are high TF.IDF complex words
<i>lowtfidf comp words</i>	Low TF.IDF Words	Number of low TF.IDF complex words in the annual report
<i>lowtfidf</i>	Percent Low TF.IDF	Percentage of total words in the annual report that are low TF.IDF complex words
<i>length</i>	Length	Log of the number of words in the annual report
<i>year_ifrs</i>	Year IFRS	Year IFRS is an indicator variable equal to 1 for post-adoption years (2006 and 2007) and 0 for pre-adoption years (2003 and 2004)
<i>Control Variables</i>		
<i>firm_size</i>	Firm Size	Natural logarithm of a firm's total assets
<i>firm_age</i>	Firm Age	Date in 2012 (02/07/2012) less date of incorporation
<i>earn_vol</i>	Earnings Volatility	Standard deviation of annual earnings (operating earnings after depreciation) of the 5 years prior to the year of analysis
<i>price_vol</i>	Price Volatility	Average annual price movement to a high and low price from a mean price for each year
<i>bus_comp</i>	Business Complexity	Natural logarithm of the number of business segments
<i>geo_comp</i>	Geographical Complexity	Natural logarithm of the number of geographical segments
<i>earn</i>	Earnings	Firm's operating income after depreciation scaled by the total assets of the firm
<i>time1</i>	Time 1	Measures the increase in word complexity for the pre- and post-IFRS adoption years; increases by 1 for each year
<i>time2</i>	Time 2	Measures the increase in word complexity for the post-IFRS adoption years; takes the value of 0 for the pre-adoption years and increases by 1 for each post-IFRS adoption year
<i>profit_loss</i>	profit loss	An indicator variable where 1 = profit firms and 0 = loss firms

expectations of this study. Consistent with prior studies, earnings are negatively associated with the word complexity measures (Li, 2008). However, the correlation is not significant for the *hightfidf* measure, indicating the broad nature of the complexity measure. The primary variable of interest *year_ifrs* is positively correlated with the narrative variable *comp words* (21%), which is consistent with the view that accounting standards increase narrative disclosures in annual reports (Dyer et al., 2017). *hightfidf* and *lowtfidf* are positively correlated with IFRS, consistent with the proposed hypothesis, H₁, that narrative

Table 3. Descriptive statistics.

Panel A: Whole sample					
Variable	Obs	Mean	Std. Dev.	Min	Max
<i>Narrative Variables</i>					
<i>numrepwords</i>	952	36012.27	24887.97	3323.00	224890.00
<i>numcompwds</i>	952	9104.21	6497.61	808.00	56550.00
<i>comp words</i>	952	25.12	1.42	4.50	29.99
<i>hightfidf comp words</i>	952	7314.40	5217.72	654.00	44381.00
<i>hightfidf</i>	952	20.16	1.20	3.68	23.83
<i>lowtfidf comp words</i>	952	1795.12	1317.44	154.00	12235.00
<i>lowtfidf</i>	952	4.98	0.66	0.82	7.58
<i>fog</i>	952	21.94	1.35	16.50	30.59
<i>Control Variables</i>					
<i>firm_size</i>	931	6.80	2.07	1.70	14.46
<i>firm_age</i>	856	40.80	28.83	2.55	111.98
<i>earn_vol</i>	913	146.68	581.57	0.00	7090.75
<i>price_vol</i>	856	28.76	10.78	11.82	75.41
<i>bus_comp</i>	806	0.88	0.62	0.00	2.30
<i>geo_comp</i>	793	1.00	0.59	0.00	2.30
<i>earnings</i>	929	0.08	0.09	-0.46	0.74
<i>profit_loss</i>	929	0.92	0.27	0	1
Panel B: Pre- and post-IFRS adoption periods					
<i>year_ifrs=0</i>	Obs	Mean	Std. Dev.	Min	Max
<i>numrepwords</i>	476	28452.95	19615.83	3323.00	162584.00
<i>numcompwds</i>	476	7100.15	5039.08	808.00	40534.00
<i>comp words</i>	476	24.82	1.30	16.89	29.53
<i>hightfidf comp words</i>	476	5748.38	4051.95	654.00	31708.00
<i>hightfidf</i>	476	20.08	1.13	13.09	23.56
<i>lowtfidf comp words</i>	476	1355.90	1015.73	154.00	8883.00
<i>lowtfidf</i>	476	4.76	0.61	2.94	6.99
<i>year_ifrs=1</i>	Obs	Mean	Std. Dev.	Min	Max
<i>numrepwords</i>	476	43571.58	27217.97	3634.00	224890.00
<i>numcompwds</i>	476	11108.26	7147.37	1090.00	56550.00
<i>comp words</i>	476	25.42	1.48	4.50	29.99
<i>hightfidf comp words</i>	476	8880.42	5759.60	866.00	44381.00
<i>hightfidf</i>	476	20.24	1.26	3.68	23.83
<i>lowtfidf comp words</i>	476	2234.34	1434.10	205.00	12235.00
<i>lowtfidf</i>	476	5.19	0.64	0.82	7.58

complexity increases with the adoption of IFRS. The variables in the correlation matrix are included in the regression model with other controls and firm- and year-fixed effects to test H_1 .

6.2 Primary results

6.2.1 Determinants

Table 5 presents the results of the determinants of the dependent variables included as controls in the hypothesis testing in the regressions. Consistent with prior studies (Hasan, 2018; Li, 2008), earnings and earnings volatility are significantly associated with complexity variables, including *compwords* and *lowtfidf*. However, they are not significantly

Table 4. Correlation matrix – selected variables.

		1	2	3	4	5	6	7	8
<i>comp words</i>	1	1.0000							
<i>hightfidf</i>	2	0.8871*** (0.0000)							
<i>lowtfidf</i>	3	0.5461*** (0.0000)	0.0980** (0.0025)	1.0000					
<i>firm_size</i>	4	0.0974** (0.0029)	0.1236*** (0.0002)	−0.0118 (0.7184)	1.0000				
<i>price_vol</i>	5	−0.0478 (0.1620)	−0.0199 (0.5605)	−0.0684* (0.0454)	−0.2913*** (0.0000)	1.0000			
<i>bus_comp</i>	6	0.0552 (0.1174)	0.0625 (0.0764)	0.0067 (0.8494)	0.3518*** (0.0000)	−0.1760*** (0.0000)	1.0000		
<i>earns</i>	7	−0.1156*** (0.0004)	−0.0630 (0.0548)	−0.1354*** (0.0000)	−0.0355 (0.2794)	−0.2248*** (0.0000)	0.0008 (0.9829)	1.0000	
<i>year_ifrs</i>	8	0.2105*** (0.0000)	0.0705* (0.0296)	0.3258*** (0.0000)	0.0915** (0.0052)	−0.1393*** (0.0000)	0.0459 (0.1926)	0.1104*** (0.0000)	1.0000

Parentheses denote significance: * < 0.05, ** < 0.01, *** < 0.001.

Table 5. Determinants of complexity.

	Predicted Sign	(1) Length	(2) <i>comp words</i>	(3) <i>hightfidf</i>	(4) <i>lowtfidf</i>
<i>firm_size</i>	+	0.1665*** (19.42)	0.0070 (0.23)	0.0356 (1.36)	−0.0278 (−1.90)
<i>earn_vol</i>	+	0.0001* (2.04)	0.0003** (3.05)	0.0001 (1.49)	0.0002*** (3.92)
<i>price_vol</i>	+	0.0024 (1.55)	0.0008 (0.15)	0.0025 (0.52)	−0.0015 (−0.57)
<i>bus_comp</i>	+	0.1124*** (4.60)	−0.0632 (−0.71)	−0.0734 (−0.98)	0.0108 (0.26)
<i>geo_comp</i>	+	0.1596*** (6.44)	0.5785*** (6.41)	0.5806*** (7.68)	−0.0029 (−0.07)
<i>earns</i>	+/−	0.5156** (3.11)	−1.6804** (−2.79)	−0.5319 (−1.05)	−1.1438*** (−4.04)
<i>_cons</i>		8.8472*** (100.64)	24.5686*** (76.69)	19.3519*** (72.16)	5.2208*** (34.76)
N		709	709	709	709
adj. R^2		0.580	0.092	0.102	0.037
F-statistic		169.90 ***	17.74***	16.19 ***	5.54 ***

t statistics in parentheses.

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Table 5 presents results from estimating the determinants of the key dependent variables. See Table 2 for variable definitions. The regression analysis is performed with robust standard errors to obtain unbiased estimates under heteroskedasticity. We first present results of a model without firm fixed year effects in Table 5. Doing so permits us to investigate the impact of variables that exhibit little within-firm variation over our sample period. However, given that our controls variables have limited variables that exhibit these characteristics, we perform the remaining analysis with firm fixed effects.

associated with the occurrence of *hightfidf* words. This is expected because the *lowtfidf* word complexity variable is more likely to capture financial words related to earnings reporting. Prior literature reports mixed results on the relationship between earnings and narrative length. Li (2008) finds a negative relationship between earnings and report length, while Lang and Stice-Lawrence (2015) show that profitable companies have lengthy reports. The results in Table 5 are consistent with a positive association between earnings and length, indicating that firms with positive earnings tend to have more informative (longer) annual reports than firms reporting lower earnings. This finding is consistent with the results of Lang and Stice-Lawrence (2015),¹⁷ who studied an international sample. In addition, this difference is expected because the content of annual reports differs significantly between corporates in the United States (US) and the UK, as reports in the UK are unstructured in comparison to the US 10-K filings (El-Haj et al., 2019).

Geographical complexity tends to increase the use of complex words. However, *lowtfidf* words show a negative association with geographical complexity, indicating that firms with more geographic segments have less “common” complex words. The significant association between geographical complexity and *hightfidf* shows that firms with more geographic segments are likely to have higher variation in narrative reporting. They are more likely to report words that are less common within the corpus of all reports and more likely unique to the firm. This is consistent with the expectations of TF.IDF assumptions. Based on the TF.IDF literature, words with higher TF.IDF occur less frequently in the universe of the corpus considered but could have a higher occurrence in one or a small number of documents in the sample. We also note the positive association between *bus_comp* and the length of

reports, which supports the view that firms with more complex operations have longer reports. This supports the ontological explanation that ‘*some aspect of reality ... that requires more dictionary entries*’ in its description tends to be more complex (Bloomfield, 2002). This finding may be caused by complex firms needing more text to explain their context. Most of the determinant variables appear to be consistent with the predicted signs, which are based on prior literature (e.g., Li, 2008). Most of the variables in the regression are significantly associated at the 5% level with the length or complexity variables, making them suitable as controls for testing H_1 in the regression.¹⁸

6.2.2 Narrative complexity and mandatory IFRS adoption

Table 6 presents the test results for H_1 . The interpretation of the variable *year_ifrs* in this analysis is that it increases after mandatory adoption of IFRS. This means that it is an indicator variable identified as 1 for the years 2006 and 2007 and 0 for the years 2003 and 2004. Four dependent variables are used in this analysis. The first two (i.e., *length* and *comp words*) act as a benchmark for complexity, incorporating both information and obfuscation components. The results show that textual disclosures in annual reports increase in length following IFRS adoption, which is consistent with other studies (e.g., Lang & Stice-Lawrence, 2015).

The dependent variables’ word complexity throughout the entire report (*comp words*) increases with IFRS adoption. Specifically, the decomposed variables of interest are significantly and positively associated with the IFRS adoption years. This indicates that IFRS adoption increases both common and uncommon complexity, as measured using

Table 6. Narrative complexity and mandatory IFRS adoption.

	(1) Length	(2) <i>comp words</i>	(3) <i>highfidf</i>	(4) <i>lowtfidf</i>
<i>year_ifrs</i>	0.424*** (15.64)	0.718*** (7.18)	0.247** (3.17)	0.470*** (10.06)
<i>firm_size</i>	0.044 (1.44)	−0.138 (−1.23)	−0.078 (−0.85)	−0.062 (−1.26)
<i>earn_vol</i>	0.000 (0.56)	−0.000* (−2.20)	−0.000 (−0.79)	−0.000 (−1.36)
<i>price_vol</i>	0.00202 (0.69)	0.000 (0.02)	0.000 (0.08)	−0.000 (−0.14)
<i>bus_comp</i>	0.0657 (1.27)	−0.083 (−0.45)	−0.043 (−0.31)	−0.038 (−0.51)
<i>geo_comp</i>	0.084* (2.16)	0.348* (2.33)	0.269* (2.04)	0.079 (1.20)
<i>earns</i>	0.116 (0.64)	−1.047 (−1.49)	−0.369 (−0.62)	−0.672* (−2.23)
Firm Fixed Effects	Yes	Yes	Yes	Yes
Year Fixed Effects	Yes	Yes	Yes	Yes
N	709	709	709	709
adj. R^2	0.608	0.127	0.021	0.308
F-statistic	84.88***	20.50***	4.09***	21.70***

t statistics in parentheses.

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Table 6 presents the results from estimating the regression of complexity variables and IFRS adoption. *year_ifrs* is an indicator variable equal to 1 for the post-IFRS adoption years and 0 for pre-adoption years. All control variables are included in the regression model. See Table 2 for variable definitions. The regression analysis is performed with robust standard errors to obtain unbiased estimates under heteroskedasticity.

the word list in Appendix II. The coefficients of both variables, *lowtfidf* and *hightfidf*, are positive and significant. The results are deemed significant at the 1% level. The significant coefficient of *lowtfidf* is consistent with H₁, namely, that informative disclosures in annual reports (common complexity) are positively correlated with IFRS application in the 2 years following first-time IFRS adoption. While we do not make predictions regarding *hightfidf*, the results indicate a significant positive association between IFRS and obfuscating disclosures in annual reports (uncommon complexity).

6.2.3 Narrative complexity and IFRS adoption with economic and time effects

From the findings, a concern arises as to whether the results capture an increase in informative and obfuscating components of annual reports, given the economy-wide effects that are external to the firm and occur over time. Thus, the regression analysis is estimated on the assumption that complexity increases over time, as recorded in previous studies (Lang & Stice-Lawrence, 2015; Li, 2008). To exclude these external effects, we introduce two “time” variables in Equation 1a: *time1* measures the increase in word complexity for the pre- and post-IFRS adoption years and takes the value of 1 for 2003 and 2 for 2004, increasing in value by 1 every year to 5 in 2007. The variable *time2* measures the increase in word complexity for the post-IFRS adoption years, taking the value of 0 for the pre-adoption years and increasing in value by 1 for each post-IFRS adoption year. The variable *time1* captures the underlying assumption that the complexity of annual report increases over time. The variable *time2* captures an increasing trend post-IFRS adoption, which models the expected increase in economic activities inherent in the introduction of a new mandatory regulation¹⁹. The purpose of including the two “time” variables is to capture the time effects not captured by time fixed effects.

Table 7 shows that the coefficients of *year_ifrs* remain significant for all of the dependent variables except for the *hightfidf* complex measure. The increase in *hightfidf* word complexity appears to be driven primarily by *time1*, indicating that after IFRS adoption, there is no significant increase in uncommon complexity. The variable *hightfidf* indicates uncommon complexity, while *lowtfidf* indicates common complexity. The variable *lowtfidf* remains positive and significant, supporting the hypothesis that IFRS application increases the disclosure of informative complex words (common complexity) in the 2 years post-IFRS adoption. There is a 0.6% increase in the disclosure of informative complex words in the 2 years post-IFRS adoption. The results of the variable *hightfidf* add to those of other studies; for example, Kim et al. (2019) show that complex words that investors are not “familiar” with add noise to investors’ information processing and are not associated with financial information. The significant association between *hightfidf* and *time1* indicates that this is also driven by wider time trends that have allowed an increase in the disclosure of these words. It can also be associated with new words that have been incorporated into the financial corpus not captured in the *hightfidf* word list.

6.3 Further analysis – The impact of firm performance and firm complexity

To extend the analysis of obfuscation and information, the study reports additional regression results in Table 8; interaction terms are included in the regression model in Equation 1a. This aids in better understanding the relationship between the variables driving increased word complexity and the impact following IFRS adoption. The regression model used to test H₁ models the level of word complexity based on the identified firm characteristics and whether the reported narratives are in the pre- or post-IFRS

Table 7. Narrative complexity and IFRS adoption with economic and time effects.

	(1) Length	(2) <i>comp words</i>	(3) <i>highfidf</i>	(4) <i>lowtfidf</i>
<i>year_ifrs</i>	0.265*** (4.90)	0.804*** (3.47)	0.247 (1.26)	0.560*** (6.22)
<i>time1</i>	0.107*** (8.71)	0.087 (1.88)	0.133** (3.29)	-0.046 (-2.04)
<i>time2</i>	-0.0583* (-2.54)	-0.137* (-1.16)	-0.155 (-1.58)	0.016 (0.41)
<i>firm_size</i>	0.023 (0.85)	-0.135 (-1.42)	-0.085 (-1.07)	-0.051 (-1.06)
<i>earn_vol</i>	0.000 (0.49)	-0.000* (-2.03)	-0.000 (-0.68)	-0.000 (-1.30)
<i>price_vol</i>	0.004 (1.21)	0.001 (0.09)	0.002 (0.22)	-0.001 (-0.25)
<i>bus_comp</i>	0.070 (1.38)	-0.081 (-0.43)	-0.039 (-0.27)	-0.040 (-0.54)
<i>geo_comp</i>	0.087* (2.29)	0.352* (2.36)	0.274* (2.09)	0.078 (1.18)
<i>earns</i>	0.016 (0.09)	-1.066 (-1.46)	-0.436 (-0.70)	-0.623* (-2.02)
Firm Fixed Effects	Yes	Yes	Yes	Yes
Year Fixed Effects	Yes	Yes	Yes	Yes
N	709	709	709	709
adj. R^2	0.628	0.126	0.023	0.308
F-statistic	72.24 ***	20.26 ***	4.39 ***	22.70 ***

t statistics in parentheses.

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Table 7 presents results from estimating the regression of complexity variables, including *time1* and *time2* variables as independent variables. *time1* indicates the increasing complexity of annual reports over time, and *time2* indicates increasing complexity in the post-IFRS adoption years. All control variables are included in the regression model. See Table 2 for variable definitions. The regression analysis is performed with robust standard errors to obtain unbiased estimates under heteroskedasticity.

adoption period. The regression equation is estimated as follows:

$$Complexity_{it} = Intercept + \beta_1 year_ifrs_{it} + \beta_2 firmcharacteristics_{it} + \varepsilon_{it} \quad (1b)$$

firmcharacteristics are labelled Controls in Equation 1a.

Li (2008) finds primary results consistent with the IRH that management makes bad news more costly to extract by writing excessively long and complex reports. This is in line with the management obfuscation concept that where earnings are poor, reports are likely to be more complex. Above, we find that IFRS reports are more complex; however, we do not know whether this effect is increased by firms reporting poorer earnings upon application of IFRS. If firms with poor performance use more complex words in their disclosures post-IFRS and these increase with IFRS adoption, then it supports the view that IFRS adoption itself does not change observed reporting obfuscation.

Bloomfield (2002) offers an ontological explanation that firms use excessively complex language because the context to be described requires more dictionary categories, for example, good news being inherently easier to communicate than bad news (Lo et al., 2017). Likewise, in the post-IFRS period, it is possible that firms with complex operations

Table 8. Earnings, business complexity, and IFRS adoption.

	(1) Length	(2) <i>comp words</i>	(3) <i>hightfidf</i>	(4) <i>lowtfidf</i>
<i>year_ifrs</i>	0.416*** (10.42)	0.413 (2.40)	-0.056 (-0.43)	0.469** (5.90)
<i>bus_obf</i>	0.013 (0.35)	0.100 (0.99)	0.108 (1.23)	-0.007 (-0.06)
<i>earns_obf</i>	-0.045 (-0.20)	2.257* (1.98)	2.169** (2.45)	0.082 (0.44)
<i>firm_size</i>	0.043 (1.42)	-0.123 (-1.14)	-0.064 (-0.72)	-0.061 (-1.27)
<i>earn_vol</i>	0.000 (0.54)	-0.000* (-2.29)	-0.000 (-1.17)	-0.000 (-1.37)
<i>price_vol</i>	0.002 (0.74)	-0.005 (-0.33)	-0.004 (-0.36)	-0.001 (-0.17)
<i>bus_comp</i>	0.059 (1.08)	-0.123 (-0.71)	-0.088 (-0.66)	-0.034 (-0.45)
<i>geo_comp</i>	0.086* (2.10)	0.284* (2.30)	0.208 (1.93)	0.076 (1.13)
<i>earns</i>	0.153 (0.57)	-2.502** (-2.81)	-1.761* (-2.20)	-0.730 (-1.85)
Firm Fixed Effects	Yes	Yes	Yes	Yes
Time Fixed Effects	Yes	Yes	Yes	Yes
N	709	709	709	709
adj. R^2	0.608	0.133	0.032	0.306
F-statistic	90.23 ***	18.84 ***	5.46***	17.76 ***

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 8 presents results from estimating the regression of the complexity variables and the independent variables *bus_obf* and *earns_obf*, which are the interaction of *year_ifrs* with business complexity and earnings, respectively. All control variables are included in the regression model. See Table 2 for variable definitions. The regression analysis is performed with robust standard errors to obtain unbiased estimates under heteroskedasticity.

tend to use a higher proportion of complex words owing to increased disclosure classes, whereas firms with less complex operations tend to use fewer complex words. Another likelihood is that operational complexity affects the level of word complexity during both the pre- and post-IFRS adoption periods, but it may increase with the adoption of IFRS given, for example, the accounting standards for specific business segment reporting. This supports the perspective that the level of word complexity in annual reports post-IFRS adoption is associated with increased informative disclosures.

We test this by adding an interaction term to Equation 1b in which two firm characteristics, business complexity (*bus_comp*) and earnings (*earns*) variables, are multiplied by the variable *year_ifrs*:

$$\begin{aligned} \text{Complexity}_{it} = & \text{intercept} + \beta_1 \text{year_ifrs}_{it} \\ & + \beta_2 \text{firmcharacteristics}_{it} + \beta_3 \text{firmcharacteristics}_{it} * \text{year_ifrs}_{it} + \varepsilon_{it} \end{aligned} \quad (2)$$

A significant coefficient for the interaction term indicates that the effect of *bus_comp*/*earns* on the level of word complexity varies with the values of the *year_ifrs* variable. To test the business complexity and earnings effect, both the interaction term and the individual variables *bus_comp* and *earns*, respectively, without interaction (in this case, β_2), are

relevant. Adding an interaction term to the model changes the interpretation of the coefficient of β_2 . In Equation 1b, that is, without an interaction term, and consistent with prior literature that tests obfuscation using performance, the interpretation of this variable shows the unique impact of earnings on the level of complexity. Including an interaction term in the model shows that the effect of earnings on word complexity at different levels of *year_ifrs* is not limited to β_2 but also depends on the values of β_3 . This implies that β_2 excludes the effect of β_3 , as it is held constant. β_2 is interpreted as the unique effect of earnings on complexity only when *year_ifrs* = 0. The significance of β_2 indicates that earnings change complexity during the pre-IFRS period, while the coefficient of β_3 indicates that earnings change complexity during the post-IFRS period. The interpretation of β_1 remains the main variable of interest; that is, *year_ifrs* increases word complexity after controlling for the interaction of *year_ifrs* and firm characteristics.

Table 8 reports the results for the regression analysis of equation 2. To explain specific IFRS associations, Table 8's results include firm fixed effects, time fixed effects, and time measured variables used to test H_1 to control for firm and time factors affecting the complexity of annual report disclosures during the IFRS adoption period. Table 8 presents the regression tests for H_1 after controlling for the interaction of firm variables and IFRS adoption. This is to ensure that the observed results are not driven by increased word complexity attributable to reporting disclosures on firm complexity and firm performance post-IFRS application.

Row 3 in Table 8 indicates that the interaction of earnings and the IFRS variable is not associated with increased complexity of *lowtfidf*. When we control for firm-specific effects and include the time variable, it appears from the results that firm performance does not drive the disclosure reporting with more complex and informative words. This supports the IFRS objective of transparency, showing that despite the strong association between earnings and complexity, as reported in Table 4, firm performance and IFRS adoption do not drive common complexity. This provides evidence of achieving the IFRS objectives of improved report transparency and comparability, with respect to the disclosure of "common" complex words in annual report narrative disclosures notwithstanding firm earnings.

The results support H_1 , that post-IFRS application, informative complexity increases after controlling for earnings-driven complexity, as shown by the significant positive coefficient of *year_ifrs* and *lowtfidf*. However, there is an association between the earnings IFRS interaction term and *hightfidf* words, indicating that the interaction of earnings and IFRS adoption is positively associated with disclosures of uncommon complexity. This denotes that uncommon complexity is largely driven by earnings reporting, supporting the concept of uncommon complexity as an extraction cost. Accordingly, complex words that are not "familiar" to investors are more likely to be positively associated with earnings narrative management post-IFRS adoption. The results suggest that with IFRS adoption, obfuscation is driven by earnings; however, the findings do not make specific predictions regarding the obfuscation component of disclosure, as the research design extracts informative words and categorises words that are not considered informative as obfuscation.

Row 2 shows that the interaction of business complexity and IFRS adoption (*bus_obf*) is not associated with word complexity in annual reports. In this study, business complexity is measured as the number of business segments, implying that complexity associated with IFRS adoption is not driven by business complexity. This further strengthens the results of IFRS-specific effects on complex words. This result also includes the time variable, showing that with the increase in complexity over time – that is, increased disclosures for larger firms and changes in disclosure needs – IFRS increase word complexity.

7. Conclusions

This study assesses the impact of IFRS adoption on narrative complexity by decomposing complexity into two components: information and obfuscation. The study's findings provide evidence that IFRS adoption increases the use of complex words in disclosures, and this increased word complexity grows with the informativeness of reports, identified as common complexity. In other words, narrative complexity, when modelled as common complexity, increases with IFRS adoption. The findings also show that increasing narrative obfuscation is identified as uncommon complexity; however, this effect reduces when the assumption that time increases word complexity is introduced into the model. The study shows that the effect of IFRS on disclosure complexity is not driven by obfuscation with respect to earnings and business complexity. While we observe an association between earnings and uncommon complexity, this association is not related to IFRS adoption.

The increase in word complexity is persistent with IFRS adoption, indicating increased complexity of disclosures for users of annual reports. However, this increased word complexity is driven by an increase in the common complexity component of narrative complexity. This contribution is novel and complements prior studies by identifying a component of complexity using an adjusted word complexity component of the widely studied fog index. The study further shows that this component is associated with the IFRS disclosure complexity observed in prior studies. This study does not establish direct causality owing to the existence of mandatory IFRS requirements, the introduction of the new standards, or firms' reporting incentives around this period. However, it reports a significant increase in the complexity and informativeness of reports based on the observed significant increase in common complexity. The study also attempts to control for these effects in the proposed research design.

Future studies can extend this research by enhancing the *lowtfidf* words (common complexity words). This can be done by incorporating words that are not present in Loughran and McDonald's (2014) list of informative words. For example, studies can apply the list adjustment in Kim et al. (2019). In addition, IFRS was adopted in the UK in 2005 and has evolved over the years with amendments to the standards, such as IFRS 9 and IAS 39 Financial Instruments,²⁰ which are perceived as inherently complex (Lang & Stice-Lawrence, 2015). Future studies could complement this work by analysing the changes in the informative and obfuscating components of complexity, focusing on the change in IFRS over time.

This study is relevant for the "Better Communication" component of the IASB's Disclosure Initiative project, which recognises that valuable information is lost in voluminous disclosures. While there is agreement that narrative complexity is a significant concern that reduces accessibility to accounting narratives, it is unclear how accounting standards drive complexity. Further research into the determinants of this increase in common complexity with the adoption of IFRS can guide standard setters on the specific standards driving narrative complexity. The results indicate that IFRS annual reports tend to become more comparable with the increase in common complexity in the post-IFRS adoption period. The IASB recommendation is to actively supply information to reduce complexity, for example, by providing references and, in some instances, further explanation. However, the use of words is a factor worth considering in the IASB's recommendations. Specifically, this study shows that IFRS themselves drive increased use of complex words, although these words are informative. Further debate and discussion are required within the IASB to determine whether the findings of an increase of complex words with IFRS

adoption is consistent with its “Better Communication” initiative and IFRS comparability objective. Although the findings are limited to the setting of 2 years post-IFRS adoption, they lay the basis for future studies to apply this novel identification of complexity decomposition and accounting standards.

Notes

1. We use the term “narrative disclosure”, consistent with prior literature (e.g., Hasan, 2018). This study also uses “narrative” and “disclosure” interchangeably to refer to the textual parts of company reports. These refer to all parts of the reports, excluding financial numbers. According to Beattie (2014), the difference between these two terms is ideological, with reference to traditional philosophical identities. In this study, we use the terms from the perspective of economic information asymmetry for identifying texts as objective economic facts (Beattie, 2014).
2. This study uses the terms “narrative obfuscation” and “uncommon complexity” interchangeably as the obfuscation component is measured as “uncommon complexity”.
3. This disclosure problem is attributable to a lack of guidance, particularly for narrative sections and sample notes for accounts (IASB, 2017).
4. The IRH proposes that the extraction cost can be modelled using market reaction to disclosures and investor disclosure choices (Bloomfield, 2002). An example of the steps taken by the IASB to reduce extraction cost is the “Better Communication” initiative. <https://www.ifrs.org/projects/better-communication/>
5. Lang and Stice-Lawrence (2015) show that firms with the greatest increase in “desirable textual attributes” relating to IFRS experienced IFRS benefits, such as greater increases in liquidity, analyst following, and institutional investment.
6. The reactive review is post-event, while the proactive review is pre-event. A sample of firms is identified for assessment in the proactive regime. In the reactive regime, only firms found to be inconsistent with the guidelines are reviewed.
7. This study does not focus on the fog index/readability. For more details on the fog index, its computation, and its components, see Efretuei (2020a).
8. See Manning and Schutze (2003) for a discussion of the term weighting concept and Jegadeesh and Wu (2013) for a discussion of its application in accounting and finance.
9. Based on the TF.IDF literature, words with higher TF.IDF occur less frequently in the universe of a considered corpus but could occur more often in one or a small number of documents within the sample.
10. Studies show that increased narrative complexity is associated with smaller investor trading, analyst dispersion, and obfuscation (e.g., Lehavy et al., 2011).
11. TF (without IDF) schemes exhibit effectiveness as a TF.IDF scheme, and prior evidence finds common complexity and uncommon complexity as the relevant extraction cost (e.g., Miller, 2010). In the accounting literature, it is assumed that the aggregated measure evaluates extraction cost, and this study attempts to identify the separate components.
12. Manning and Schutze (2003) describe its use, while Loughran and McDonald (2011) also apply it to the accounting literature.
13. Stop words are excluded only from the low TF.IDF word list but not from the high TF.IDF word list. From a complexity estimation, about 8% of words in the stop word list are complex words. Based on the information retrieval concept and the aforementioned latent component assumptions, this study also considers complex stop words to be less informative and includes them in the high TF.IDF word list, primarily because stop words tend to be non-technical in detail and increase the obscurity of useful information through frequent occurrence.
14. The 52 words identified are commonly used and appear in corporate annual reports; the persistent use of these words does not increase obfuscation.
15. The authors do not make specific predictions for the obfuscation component of disclosures, as the research design extracts informative words and categorises words that are not informative as obfuscation. However, based on prior report fog index studies (e.g., Li, 2008), we expect that this component of disclosure complexity changes with IFRS application. The authors also do not make predictions for the period beyond 2 years after the adoption year, given the scope of the data used in this study. However, as Lang and Stice-Lawrence (2015) find increased

complexity persisting for years post-IFRS, the predictions are likely to persist for more than 2 years after the adoption year.

16. Studies have reported average fog index scores from approximately 20 (Lang & Stice-Lawrence, 2015; Li, 2008) up to 22 (Dyer et al., 2017). The type of corpus (management report or annual report), the sample period of the study (later years have higher fog index scores), and the institutional context (US or UK) tend to affect the average fog index reported.
17. Lang and Stice-Lawrence (2015) use an indicator variable coded 1 if the firm reports losses.
18. Variables that are not significantly associated are included to maintain consistency with prior research on readability and accounting variables.
19. For example, in the UK, there was a change in the enforcement regime aligned to the regulation of IFRS adoption (Christensen, 2013). This may impact the form and content of narrative disclosures not due to the application of IFRS but, rather, to align to new enforcement regime requirements.
20. See section 20[4] of Gore-Browne on EU Company Law, which details the changes in standards since the adoption of IFRS in 2005 (Efretuei, 2020b).
21. This is a Linux-based platform that converts a pdf file into a text file. It has an additional check using OCR whereby the pdf files cannot be converted due to inclusion of embedded text.
22. This module, developed by James (2018), is used to split text into words, paragraphs, segments, and tiles. The Perl module is part of a Perl programme and is used to analyse the annual reports and produce a file of split words for each annual report.
23. Lingua::EN::Syllable::syllable estimates the number of syllables in the word passed to it. Further information on the code can be found in Bowers and Fast (1999).

Acknowledgments

We acknowledge the valuable comments of the chair, co-chair, and workshop participants of the Global Emerging Scholars Research Workshop at the American Accounting Association Conference 2018, the attendees of the Financial Reporting and Business Communication Conference 2019, and the attendees of the Newcastle University London Research Workshop series.

Funding

The author(s) reported there is no funding associated with the work featured in this article.

Declaration of interest

The authors declare that they have no competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

- Asay, H. S., Elliott, W. B., & Rennekamp, K. (2017). Disclosure readability and the sensitivity of investors' valuation judgments to outside information. *Accounting Review*, 92(4), 1–25. <https://doi.org/0.2308/accr-51570>
- Asay, H. S., Libby, R., & Rennekamp, K. (2018). Firm performance, reporting goals, and language choices in narrative disclosures. *Journal of Accounting and Economics*, 65(2–3), 380–398.
- Bae, K.-H., Tan, H., & Welker, M. (2008). International GAAP differences: The impact on foreign analysts. *The Accounting Review*, 83(3), 593–628.
- Beattie, V. (2014). Accounting narratives and the narrative turn in accounting research: Issues, theory, methodology, methods and a research framework. *British Accounting Review*, 46(2), 111–134. <https://doi.org/10.1016/j.bar.2014.05.001>
- Bloomfield, R. (2002). The incomplete revelation hypothesis and financial reporting. *Accounting Horizons*, 16(3), 233–243.
- Bloomfield, R. (2008). Discussion of annual report readability, current earnings, and earnings persistence. *Journal of Accounting and Economics*, 45(2–3), 248–252.

- Bonsall, S. B., Leone, A. J., Miller, B. P., & Rennekamp, K. (2017). A plain English measure of financial reporting readability. *Journal of Accounting and Economics*, 63(2–3), 329–357.
- Bowers, N., & Fast, G. (1999). *Lingua::EN::Syllable Perl Module*. Meta CPAN, Module version 0.30 <https://metacpan.org/pod/Lingua::EN::Syllable>
- Brochet, F., Jagolinzer, A. D., & Riedl, E. J. (2013). Mandatory IFRS adoption and financial statement comparability. *Contemporary Accounting Research*, 30(4), 1373–1400.
- Bushee, B. J., Gow, I. D., & Taylor, D. J. (2018). Linguistic complexity in firm disclosures: Obfuscation or information? *Journal of Accounting Research*, 56(1), 85–121.
- Callen, J. L., Khan, M., & Lu, H. (2013). Accounting quality, stock price delay, and future stock returns. *Contemporary Accounting Research*, 30, 269–295. <https://doi.org/10.1111/j.1911-3846.2011.01154.x>
- Cheung, E., & Lau, J. (2016). Readability of notes to the financial statements and the adoption of IFRS. *Australian Accounting Review*, 26(2), 162–176.
- Christensen, H. B., Hail, L., & Leuz, C. (2013). Mandatory IFRS reporting and changes in enforcement. *Journal of Accounting and Economics*, 56(2, Suppl. 1), 147–177.
- Christensen, H. B., Lee, E., Walker, M., & Zeng, C. (2015). Incentives or standards: What determines accounting quality changes around IFRS adoption? *European Accounting Review*, 24(1), 3–61.
- Courtis, J. K. (2004). Corporate report obfuscation: Artefact or phenomenon? *The British Accounting Review*, 36(3), 291–312.
- De Franco, G., Hope, O., Vyas, D., & Zhou, Y. (2015). Analyst report readability. *Contemporary Accounting Research*, 32, 76–104. <https://doi.org/10.1111/1911-3846.12062>
- Dempsey, S. J., Harrison, D. M., Luchtenberg, K. F., & Seiler, M. J. (2012). Financial opacity and firm performance: The readability of REIT annual reports. *Journal of Real Estate Finance and Economics*, 45(2), 450–470.
- Dyer, T., Lang, M., & Stice-Lawrence, L. (2017). The evolution of 10-K textual disclosure: Evidence from latent Dirichlet allocation. *Journal of Accounting and Economics*, 64(2), 221–245.
- EFRAG. (2014). *Getting a better framework*. COMPLEXITY Bulletin, European Financial Reporting Advisory Group. <https://www.efrag.org/Activities/1604011051315205/Conceptual-Framework—Bulletin-Complexity>.
- Efreteui, E. (2020a). Year and industry-level accounting narrative analysis: Readability and tone variation. *Journal of Emerging Technologies in Accounting*. <https://doi.org/10.2308/JETA-18-12-21-26>
- Efreteui, E. (2020b). Financial and narrative reporting in EU company and securities law. In G.-B. Carsten (Ed.), *Gore-Browne on EU Company Law* (Issue 79, Ch. 20). Jordan Publishing.
- El-Haj, M., Alves, P., Rayson, P., Walker, M., & Young, S. (2019). Retrieving, classifying and analysing narrative commentary in unstructured (glossy) annual reports published as PDF files. *Accounting and Business Research*, 50(1), 1–29. <https://doi.org/10.1080/00014788.2019.1609346>
- Filzen, J. J., & Peterson, K. (2015). Financial statement complexity and meeting analysts' expectations. *Contemporary Accounting Research*, 32(4), 1560–1594.
- Gao, R., & Sidhu, B. K. (2017). The impact of mandatory IFRS adoption on investment efficiency: Standards, enforcement and reporting incentives. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2956781
- Guay, W., Samuels, D., & Taylor, D. (2016). Guiding through the fog: Financial statement complexity and voluntary disclosure. *Journal of Accounting and Economics*, 62(2–3), 234–269. <https://doi.org/10.1016/j.jacceco.2016.09.001>
- Hasan, M. M. (2018). Readability of narrative disclosures in 10-K reports: Does managerial ability matter? *European Accounting Review*, 29(1), 1–22. <https://doi.org/10.1080/09638180.2018.1528169>
- IASB. (2017). Disclosure initiative – Principles of disclosure. Discussion Paper DP/2017/1, International Accounting Standards Board. www.ifrs.org
- James, D. (2018). *Lingua::EN::Splitter module*. Module version 0.1 [cited 29/09/2018]. <https://metacpan.org/pod/Lingua::EN::Splitter>
- Jegadeesh, N., & Wu, D. (2013). Word power: A new approach for content analysis. *Journal of Financial Economics*, 110(3), 712–729. <https://doi.org/10.1016/j.jfineco.2013.08.018>
- Kim, C., Wang, K., & Zhang, L. (2019). Readability of 10-K reports and stock price crash risk. *Contemporary Account Research*, 36, 1184–1216. <https://doi.org/10.1111/1911-3846.12452>

- Lang, M., & Stice-Lawrence, L. (2015). Textual analysis and international financial reporting: Large sample evidence. *Journal of Accounting and Economics*, 60(2–3), 110–135.
- Lawrence, A. (2013). Individual investors and financial disclosure. *Journal of Accounting and Economics*, 56(1), 130–147.
- Lee, Y. (2012). The effect of quarterly report readability on information efficiency of stock prices. *Contemporary Accounting Research*, 29, 1137–1170. <https://doi.org/10.1111/j.1911-3846.2011.01152.x>
- Lehavy, R., Li, F., & Merkley, K. (2011). The effect of annual report readability on analyst following and the properties of their earnings forecasts. *The Accounting Review*, 86(3), 1087–1115.
- Leuz, C., & Wysocki, P. D. (2016). The economics of disclosure and financial reporting regulation: Evidence and suggestions for future research. *Journal of Accounting Research*, 54(2), 525–622.
- Li, F. (2008). Annual report readability, current earnings, and earnings persistence. *Journal of Accounting and Economics*, 45(2–3), 221–247.
- Lo, K., Ramos, F., & Rogo, R. (2017). Earnings management and annual report readability. *Journal of Accounting and Economics*, 63(1), 1–25.
- Loughran, T., & McDonald, B. (2011). When is a liability not a liability? Textual analysis, dictionaries, and 10-Ks. *The Journal of Finance*, 66(1), 35–65.
- Loughran, T., & McDonald, B. (2014). Measuring readability in financial disclosures. *The Journal of Finance*, 69(4), 1643–1671.
- Manning, C. D., & Schutze, H. (2003). *Foundations of statistical natural language processing*. The MIT Press.
- Merkel-Davies, D. M., & Brennan, N. M. (2007). Discretionary disclosure strategies in corporate narratives: Incremental information or impression management? *Journal of Accounting Literature*, 26, 116–194.
- Miller, B. P. (2010). The effects of reporting complexity on small and large investor trading. *The Accounting Review*, 85(6), 36.
- Neel, M. (2017). Accounting comparability and economic outcomes of mandatory IFRS adoption. *Contemporary Accounting Research*, 34, 658–690. <https://doi.org/10.1111/1911-3846.12229>
- Rennekamp, K. (2012). Processing fluency and investors' reactions to disclosure readability. *Journal of Accounting Research*, 50(5), 1319–1354. <https://doi.org/10.1111/j.1475-679X.2012.00460.x>
- Richards, G., & Staden, C. V. (2015). The readability impact of International Financial Reporting Standards. *Pacific Accounting Review*, 27(3), 282–303.
- Stone, G. W., & Parker, L. (2016). A comment on “The struggle to fabricate accounting narrative obfuscation: an actor-network-theoretic analysis of a failing project.” *Qualitative Research in Accounting & Management*, 13(1), 86–89.
- Tan, H.-T., Wang, E. Y., & Zhou, B. (2015). How does readability influence investors' judgments? Consistency of benchmark performance matters. *The Accounting Review*, 90(1), 371–393. <https://doi.org/10.2308/accr-50857>
- Tarca, A. (2004). International convergence of accounting practices: Choosing between IAS and US GAAP. *Journal of International Financial Management & Accounting*, 15(1), 60–91.
- You, H., & X.-J. Zhang. (2009). Financial reporting complexity and investor underreaction to 10-K information. *Review of Accounting Studies*, 14(4), 559–586.

Appendices

Appendix I

Data collection process for textual variables

To analyse the text for narrative variables, we first downloaded the PDF files for the FTALSH companies. Then, all text from the PDFs was extracted by converting the PDF files to text format using “pdftotext” with an optical character recognition (OCR)²¹ Linux-based command. After file conversion, each file is input into the Perl “Lingua::EN::Splitter” module,²² which reads the text files and outputs a text file with words split into new lines. The criterion for identifying words is that a word must have letters and at least one vowel sound.

The next step is to input the converted text files into a computer programme that separates the words in the file into two groups: group A is a text file with common complexity words, identified using the high-frequency financial words (Loughran & McDonald 2014) listed in Appendix II; group

B is a text file excluding the common complexity words – uncommon complexity words defined as complex words with three or more syllables and not present in group A. Following this, the text files of groups A and B are input into a Perl programme, “Lingua::EN::Syllable”, which estimates the number of complex words for each group.²³ The data steps are as follows:

- Step 1:
 - Download PDF annual reports
 - PDF-to-text conversion using OCR for text preservation
- Step 2:
 - Split text files into words using <https://metacpan.org/pod/Lingua::EN::Splitter>
- Step 3:
 - Extract common complexity words for group A text files and exclude common complexity words from group B text files using author-written Perl code.
- Step 4:
 - Compute word complexity weighting for *lowtfidf* complex words and *hightfidf* complex words using Perl programme <https://metacpan.org/pod/Lingua::EN::Syllable>
 - This produces two measures of word complexity:
 - *lowtfidf* – common complexity (information)
 - *hightfidf* – uncommon complexity (identified as obfuscation)
- Step 5:
 - Run regression analysis using word complexity weighting as the dependent variable and an IFRS adoption dummy as the independent variable, including relevant controls.

Appendix II

lowtfidf words

financial|company|interest|agreement|including|operations|period|related|management|consolidated|
information|services|provided|pursuant|following|securities|approximately|reference|operating|
material|capital|expenses|corporation|outstanding|additional|effective|accounting|incorporated|
included|compensation|applicable|primarily|accordance|significant|subsidiaries|customers|respect-
ively|registrant|obligations|provisions|liabilities|addition|otherwise|property|employees|benefit|
reporting|principal|development|revenue|equity|insurance