UCLouvain

LOUVAIN
School of Management

Louvain School of Management

Structural readability and informativeness of bullet points in EPRs

A topic modeling approach

Author(s): BASSEM Oscar
Supervisor(s): THEWISSEN James
Academic year 2023-2024
Dissertation for the master of Management (GEST2MS/GE)
Corporate finance
Daytime schedule

Abstract :

Information asymmetry in earning press releases (EPRs) have been studied for years, making the readability of these documents a decisive factor in apprehending the link between a company's communication and its investors' understanding. However, the impact of structural readability on investors is little investigated. In this thesis, we study the informativeness of bullet points in EPRs. Using a sample of EPRs from 1,725 companies from 2004 to 2021, we find that bullet points have an impact on investor reaction in the medium term (3 to 5 days before and after the announcement). Moreover, we show that this relationship was even more significant when the company was small. Using an LDA topic modeling method, we identify five topics present in the bullet points and show that the performance-related topic has the greatest impact on investor reaction. Finally, our results suggest that investors' information processing follow a conceptual fluency when reading bullet points.

UNIVERSITÉ CATHOLIQUE DE LOUVAIN
Louvain School of Management

Place des Doyens, 1 bte L2.01.01, 1348 Louvain-la-Neuve
Boulevard Emile Devreux 6, 6000 Charleroi, Belgique
Chaussée de Binche 151, 7000 Mons, Belgique

www.uclouvain.be/lsm

Acknowledgements

First of all, I would like to thank Professor James Thewissen for guiding me in my research. I would also like to thank him for his lectures in the corporate finance option I attended, which greatly helped me to define the field of research and to conduct the appropriate methodology. The various projects undertaken during these courses were invaluable in developing the rigor, critical thinking and structure of this work.

Contents

1 Introduction

Earnings Press Releases (EPRs) have been studied for several years to address problems of information asymmetry between companies and stakeholders (Henry, 2006; Boudt et al., 2018). Studies have already shown that information asymmetries mainly originate in textual readability and processing fluency defined as "how easy it is to process information" (Loughran & McDonald, 2014; Rennekamp, 2012). However, other studies have shown that presentation structures in texts are increasingly present, making structural readability an angle of study to be addressed (Christensen et al., 2024).

While previous studies have focused on the effectiveness of textual readability (Loughran & McDonald, 2014; Boudt et al., 2018), no studies have investigated structural readability - which represents the textual organisation of a document - and its impact on readers' processing fluency. Research by Cheng et al., (2018) has evaluated the structural readability of financial reports and its effects on investor reaction, but a definitive link with information processing has yet to be established. On the other hand, Ho (2023) shows that structures such as bullet points have a cognitive influence on the reader. Furthermore, he demonstrates that bullet points effectively highlight important information, a notion supported by several studies showcasing the significant impact of content in reports (Graas, 2021; Thewissen et al., 2021). These assertions in the literature reveal an unexplored research gap: the informativeness of bullet points. Trough this information, we believe that the use and content of bullet points should provide an advantage in processing EPR information and improve the efficiency of the stock market by reducing information asymmetry.

In this thesis, we investigate the structural readability of Earnings Press Releases (EPRs) to determine whether the use and content of bullet points influence investor reactions. This study addresses the informativeness of bullet points in EPRs through three research questions. First, we explore the impact of the number of bullet points on investors to understand the benefits of using this textual structure. Next, we examine the topics mentioned in the bullet points. Finally, we assess whether these topics affect investor reactions.

Using a sample containing all the EPR bullet points of 1,725 listed companies from 2004 to 2021, we run a regression measuring whether a variation in the number of bullet points is significant with a variation of the Cumulative Abnormal Return (CAR) which represents the investor reaction. This first step enables us to observe whether the latter are read by investors. In the second stage, we use the same database to carry out a targeted

search on the topics present in the bullet points. Textual data processing throught a Latent Dirichlet Allocation (LDA) topic modeling method enabled us to identify the topics frequently used in bullet points. By measuring the frequency of these topics in each bullet point of each EPRs in our sample, we perform a second investigation to understand which content has the greatest impact on investors.

The study shows that the number of bullet points had a positive impact on investor reaction, which confirms the interest of bullet points in EPRs and suggests the existence of a gain in information processing by the reader. Furthermore, it corroborate the literature on the effectiveness of bullet points (Ho, 2023). However, when we look at bullet points efficiency, we found that this impact is influenced by the company's size, aligning with existing literature that demonstrates this effect for textual readability (Dodoo & Gu, 2020). Our research into the content of bullet points also revealed five frequently used topics: governance, finance, performance, business operations and regulations. Our investigation of the impact of these topics showed that performance topics is the most influential with investors. However, it was shown that the impact of the regulation topic is significant but only in the very short term.

By this way, this work contributes to the research as Cheng et al. (2018) by further investigating the structural importance of a text and the quality of its content. By identifying the topics frequently used in bullet points, our study helps to underline the importance of content as an indicator of the quality of information in reports. (Graas, 2021; Thewissen et al., 2021). In addition, our investigations also highlight the effectiveness and usefulness of bullet point in the special case of EPRs which confirms the general statement by Ho (2023). Furthermore, our study provides answers of how to avoid information asymmetries between stakeholders by improving the understanding of processing fluency in financial reports (Rennekamp, 2012). Exploring the content also gives companies information on the most important topics to communicate. Finally, this research provides information on the quality and the credibility of EPRs for investors.

In this report, we begin by taking stock of existing research through a literature review. We then describe the data used and the method we followed. In the third section, dedicated to analysis, we report on the results obtained and put them into perspective with the elements found in the literature review. Then, an additional analysis is provided on the content of bullet points in EPRs. Finally, the conclusion will mark the end of this report with a summary of the findings.

2 Literature review, research question and hypothesis development

An Earnings Press Release (EPR) is primarily a means of communicating information about a company's business and results to its shareholders. It provides an overview of the company's health by transmitting both quantitative data (such as financial ratios) and qualitative data (such as strategic initiatives). The SEC (Securities and Exchange Commission) considers all the information in these reports to be important for making investment decisions. This includes financial, regulatory, and management data, which together give a comprehensive view of the company's performance and activities (SEC). The quality of information is therefore an important attribute for investors when making their choices. Indeed, it has been shown that report quality leads to higher investment volumes due to low information asymmetry (Brown & Hillegeist, 2007; Nwaobia et al., 2013). Furthermore, it has been shown that content takes part of quality in reports (Thewissen et al., 2021).

Given this, it is clear that the stock market serves as a valuable indicator for understanding how investors value information in EPRs and measuring information asymmetry in the market. Several studies have identified the sources of information asymmetry in EPRs by observing general market behavior (Boudt et al., 2018; Davis et al., 2012; Fiske & Taylor, 2013; Cheng et al., 2018). Firstly, linguistic readability was investigated through word complexity and sentence length using the fog index (Loughran & McDonald, 2014). Studies have also looked at the tone (Boudt et al., 2018) and the optimistic/pessimistic aspect of the words used (Davis et al., 2012). Each of these studies revealed the presence of information asymmetries in the market. Additionally, structural readability has been studied in recent years, revealing the impact of different text structures on readers (Fiske & Taylor, 2013; Cheng et al., 2018). For example, it has been shown that the first elements of a list are more salient, with readers assigning less value to subsequent items (Fiske & Taylor, 2016). In the specific context of EPRs, Cheng et al. (2018) studied the order of information in reports and found that this variable influences investor reactions. Finally, readability is affected by other attributes such as company size. Dodoo and Gu (2020) attempted to establish a link between readability and company performance through the analysis of annual reports. They concluded that corporate performance does not affect readability. However, their study shows convincing results regarding company size: annual reports are less readable in larger companies due to the complexity of their operations (Dodoo & Gu, 2020).

Structural readability of documents also includes bullet points. Ho (2023) has demonstrated the cognitive benefits and studied in depth the reader's experience through the use of bullet points. He first demonstrates that the effectiveness of bullet points is varied, improving reading on several levels: organizing information, highlighting information, improving readability, summarizing content and increasing retention. However, the results of his study also show that bullets are most effective when they are simple and organized. What's more, the cognitive effects of bullet points are significant, requiring less cognitive effort and enabling readers to focus directly on the essentials. However, overuse of bullet points makes them less readable (Ho, 2023).

Over and above we can see that information processing is the main attribute that make readers react. Studies talk about "processing fluency" which is defined as the degree of ease with which information is processed (Rennekamp, 2012). Processing fluency is described as an essential aspect of our judgment (Alter & Oppenheimer, 2009) and has multiple facets. These include perceptual fluency, which refers to the ease of recognizing something, and conceptual fluency, which refers to the ease of making quick associations with a stimulus (Graf et al., 2017). Although these two categories of information processing are based on different cognitive systems, it has been shown that they do not affect final judgment (Winkielman et al., 2003). Finally, processing fluency is most often measured by the action following processing. This method provides a partial understanding of the phenomenon, but does not allow us to generalize about it, given the subjective bias of our preferences when faced with information (Novemsky et al., 2007).

2.1 Research questions

The aim of this study is to investigate the readability of EPRs through the effectiveness and content of bullet points. Previous studies have shown that structural readability impacts investors (Fiske & Taylor, 2013; Cheng et al., 2018). Additionally, the effectiveness of bullet points has been confirmed at the cognitive level, thanks to their ability to improve reading (Ho, 2023). In this study, we will begin by answering whether the use of bullets influence investor reaction (RQ1). The question aims to confirm the effectiveness and readability of bullet points specifically in the context of EPRs. Moreover, understanding this will allow us to make a probable link between the general efficiency of text structure and its involvement in information processing. This question is the starting point of our research to ensure the validity of further investigation. Secondly, the literature indicates that bullet points highlight important information (Ho, 2023) and some studies

have shown that content impacts information quality (Thewissen et al., 2021). Therefore, we are going to assess what is the textual content of bullet points (RQ2). Finally, investigating the content of the bullet points enables to target the topics mentioned and to see whether this content affects investors. The third research question aims to know what topics are investors receptive to (RQ3). Studying the content and its impact on investor reactions will help us better target the type of processing fluency that investors respond to.

2.2 Hypothesis development

Initially, studies of structural readability have shown direct results on investor reaction, suggesting that these reactions are consistent and not depend on individual preferences (Fiske & Taylor, 2013 ; Cheng et al., 2018). Furthermore, the literature showed that bullet points had a cognitive impact on readers (Ho, 2023). This latter finding directly link to the way information is processed (Graf et al., 2017), even if this processing may be subjective.

Our first hypothesis investigates the impact of the number of bullets on investors reaction (H1a). The aim of this hypothesis is to ascertain the general effect of a visual structure on investors. Additionally, the results of the hypothesis help to suggest whether bullet points affect investors' information processing. In this way, the results of the hypothesis contribute directly to the literature exploring the structural readability of documents, and provide avenues of exploration into the possible impact of bullet points on information processing.

H1a: *The more bullet points there are in EPRs, the greater the reaction of investors.*

Literature has also shown that readability changes with company size (Gu & Dodoo, 2020). However, since bullet points are associated to greater organization of information and greater reader retention in documents (Ho, 2023), it is reasonable to assume that the quality of information is not impacted. Furthermore, perceptual fluency could reduce this phenomenon for poeple that are accustomed to bullet points (Graf et al., 2017).

We therefore decided to extend our investigation into the impact of the number of bullet points to the size of the company (H1b), in order to understand whether information processing was impacted. This hypothesis allows us to see whether the structural readability of bullet points is independent of the effects that content could have. In addition, the hy-

pothesis helps us to understand whether investors' reactions are motivated by perceptual fluency or whether they are also guided by conceptual fluency.

H1b: *Investor reaction to EPRs' bullet points is influenced by company size.*

Finally, the literature shows that content influences report quality (Thewissen et al., 2021), and that the bullet points contain the essential information (Ho, 2023). In addition, studies on readers' information processing have shown the existence of conceptual fluency, which refers more to the examination of content (Graf et al., 2017).

Our second hypothesis (H2) investigates the impact of bullet point content on investor reaction. Given that most of the content indirectly engages the company's performance, we believe that topics related to performance and finance will make investors react more. The aim of this investigation is to provide answers as to investors' interest in content. In this way, the results contribute to finding out whether content has an impact on information quality. Finally, by comparing the results obtained from our first hypothesis, this second hypothesis allows us to determine the existence of conceptual fluency in readers' information processing.

H2: *Topics related to quantitative data such as finance and performance make investors react more strongly than topics related to qualitative data such as operations, governance and regulations.*

3 Sample selection and data description

This study was carried out with the help of two databases provided by Prof. J. Thewissen based on a sample of 1,725 companies, with data collected from the beginning 2004 to the end of 2021.

The first database contents quantitative data such as the company market value, earning to asset and losses. Additionally, different data that will be important in the investigations are provided, such as the CAR, which is the Cumulative Abnormal Return representing the reactions of investors spread over a measure of time. The second database provides the content of 487,124 bullets present in the EPRs of the selected companies. Each of the bullet point are sorted by document which represent a sample of 54,425 EPRs.

3.1 Topic modeling implementation

Topic modeling is a machine learning method that uses statistical calculations to sort the words in a document into topics. The implementation of topic modeling is essential for this study. In this case, this method sorts the words that make up each bullet point. The result of this process will provide an overview of the topics frequently mentioned in bullet points and allow us to study their weighting in each company document. These datas will form an integral part of our content analysis.

First, the textual data of each bullet point is sorted beforehand. Studies have shown that this initial processing can significantly improve search results. Schofield et al. (2017) demonstrated the impact of pre-processing a text in the application of topic modeling. In this particular case, they show that stopword removal and stemming are crucial steps, as they have impact on the validity of the model. For example, "removing determiners, conjunctions, and prepositions can improve model fit and quality" whereas "stronger stemming treatments produce less consistent topic assignment" (Schofield et al., 2017). The conclusion of this study emphasizes the need to perform these processes carefully, depending on the research framework.

In our case study, the size of the database obliges us to follow the same process. Therefore the textual dataset has been sorted and cleaned to make it usable for the study. The dataset has been sorted as follows:

Steps such as deleting punctuation, numbers, spaces, special characters and stopwords are essential to work with a database containing only the things we're interested in. Afterwards, the words are lemmatized to make them comparable (for example, in the word "ending" the last part in "ing" is removed to get the result "end"). Then words with less than three letters and terms that appear in less than 20 documents are excluded. The most infrequent words in the dataset are also removed. Finally, a sequence that quantifies the importance of words in each document is activated and we keep only the latter. A final clean-up is carried out to remove documents that have become empty, and to remove words that appear less than five times in the entire database.

Cleaning the database enables to retain only words that are relevant to the study to run the LDA (Latent Dirichlet Allocation) on them. Blei et al. (2003) introduced the LDA method in the early 2000s, defining it as "a generative probabilistic model of a corpus"

(Blei et al., 2003, p. 996). They detail this method by stating that "the basic idea is that documents are represented as random mixtures over latent topics, where each topic is characterized by a distribution over words" (Blei et al., 2003, p. 996). The LDA method pays attention to and examines document/data of the set of topics that are likely to have generated the accumulation of words. The set of observations reveals hidden structures by paying attention to the relationship between the grouping of topics as well as the words in a document. LDA consists of two components: the words inside a text (a factor that is already known) and the likelihood of words belonging to a subject, which must be estimated. The method attempts to ascertain, for a given text, how the words pertain to a certain subject. In addition, it tries to identify how many papers relate to a certain subject due to a particular word. LDA makes key assumptions. Some of the assumptions are:

1. Documents cover variety of subjects.
2. Topics include variety of words (or token).

The LDA method can be represented by an equation as follow (Vayansky & Kumar, 2020):

$$p(\theta|\alpha) = \frac{\Gamma(\sum_{i=1}^{k} \alpha_i)}{\prod_{i=1}^{k} \Gamma(\alpha_i)} \theta_1^{z_1-1} ... \theta_k^{z_k-1}$$

Where θ is the Dirichlet variable and α is the number of time a topic is present in a document of the database. This laste parameter is define by k which can be interpreted as the number of words that we want to have in each topic α. Γ(x) is here the gamma function (Blei et al., 2003). Finally, the LDA parameter can be estimated in several ways. In our case, we have chosen the Gibbs method. This method actually refers to a Monte Carlo-chain algorithm, which is a sampling method based on probabilities.

The low number of words remaining in the sort forced us to perform the LDA method for five topics in order to obtain coherent lists. Table 1 shows the five topics identified by the LDA method. Among these, topic 1 shows words related to corporate governance, topic 2 is related to corporate finance, topic 3 to performance, topic 4 to business operations and topic 5 to regulations.

Topic 1	**Topic 2**	**Topic 3**	**Topic 4**	**Topic 5**
Officer	Ratio	Guidance	Order	Store
President	Return	Repurchase	Cent	Pretax
Chief	Deposit	Fullyear	Pipeline	Basic
Strategy	Organic	Sequentially	Thousand	Nonperforming
Key	FFO	Raise	NOI	Regulation
Corporate	Book	Premium	Fuel	Restaurant
Client	Home	Hotel	Rental	Information
Investor	Declare	Rise	Competitive	Domestic
Initiate	Backlog	Pro	Samestore	Litigation
Vice	Noninterest	Expenditure	Subscription	Trade
Enrollment	Liquidity	Exceed	Recur	Claim
Independent	Tangible	Forma	Mobile	Legal
Track	Consecutive	Outlook	Shipment	Franchise
Member	Yeartodate	Dollar	Channel	Environmental
Name	Call	Occupancy	China	Class

Table 1: Topic Modeling Results

Each of these topics represents a sample that will be used in the investigations. The latter are manipulated to determine the weight they have in each bullet point, called gammas. For greater clarity, these last variables have been renamed by the topic they represent: Governance for topic 1, Finance for topic 2, Performance for topic 3, Operation for topic 4 and Regulation for topic 5.

3.2 Data description

Table 2 shows all the data used for the investigations, including the topics obtained using the topic modeling method explained in the previous section. For the sake of clarity, each variable has been described.

Variables	Description
CAR(1,3 or 5)	Cumulative abnormal return observed one, three or five day(s) before and after the announcement. This is the difference between the expected return and the actual return over a given period of time. This variable represents the reaction of investors. This variable is numerical.
Topics(Governance, Finance, Performance, Operations or Regulations)	Numerical variable representing the weight of each topic in each document of the database.
nbrbullets	Number of bullet points in the EPR.
EARNTA	Numerical variable representing the company's earning to asset.
Return	Numerical variable representing the company's return.
Sd	Numerical variable representing the standard deviation of the company's share return. This variable represents the share's volatility on the market.
Loss	Variable that can take the value of 1 or 0. A value of 1 means that the company has made a loss.
Mkvaltq	Numerical variable representing the company's market valuation. It represents the sum of outstanding shares.
Errormedian	Standard error of the median. This variable calculates the deviation from the median of the sample.
NoA	Numerical variable representing the number of analyst that follow the company.
Fyearq	Numeric variable. It represents the fiscal year in which the EPRs were published.
Fqrt	Numeric variable with a value from 1 to 4. It represents the fiscal quarter in which the EPRs were published. The number 1 designates the first quarter.
Risk_monthly	Numerical variable representing the monthly risk linked to the company's share.

Table 2: Data description

3.3 Descriptive statistics

Table 3 provides descriptive statistics for each variable. CAR1, 3 and 5 each have a mean of 0.002, indicating a low investor response in general. A high standard deviation for each of these variables give assumption that good diversity is observed in investors reaction. The average number of bullet points is 5.285, with a high standard deviation as well. This latter observation suggests a non-uniform use of bullet points by companies. The topics each indicate a weighting close to 20% in bullet points, and a low standard deviation too. This last statement suggest that most of the companies all talk about these topics in the same way in their reports. Furthermore, some of the observations are in line with the

literature. For example, the bullet points are not over-used, which suggests good use by companies.

Table 3: Descriptive statistics

Statistic	N	Mean	St. Dev.	Min	Max
car1	32,323	0.002	0.072	−0.703	1.373
car3	32,323	0.002	0.082	−0.872	1.210
car5	32,323	0.002	0.091	−0.930	1.069
Governance	50,829	0.198	0.016	0.147	0.486
Finance	50,829	0.204	0.014	0.101	0.349
Performance	50,829	0.202	0.015	0.101	0.412
Operation	50,829	0.199	0.013	0.138	0.374
Regulation	50,829	0.198	0.012	0.119	0.397
nbrbullets	54,725	5.285	2.561	2	10
EARNTA	54,725	0.002	0.050	−2.521	0.674
sd	54,682	0.012	0.034	0.00000	1.401
loss	54,725	0.209	0.406	0	1
mkvaltq	54,725	15,653.510	48,203.320	4.011	1,678,381.000
errormedian	54,702	−75.666	7,446.733	−820,312.500	38,985.150
NoA	54,725	8.973	6.975	1	48
fyearq	54,725	2,014.014	4.194	2,004	2,021
fqtr	54,705	2.329	1.043	1	4
risk_monthly	54,725	0.080	0.460	−0.979	19.554

4 Method

Hypotheses are investigated using two least square regressions. This model will enable to understand the strength of the relationship between the variables by means of a statistical test. Then, we will be able to observe the effect of the independent variable(s) on the dependent variable. Control variables were also selected on the basis of their potential to bias the results and ensure the validity of the models:

- EARNTA (Company's earning to asset)
- Return (Company's return)
- Sd (Standard deviation of the company's share return)
- Loss (Presence of loss)
- Mkvaltq (Company's market value)
- Errormedian (Standard error of the median)

- NoA (Number of analysts)
- Fyearq (Fiscal year)
- Fqtr (Fiscal quarter)
- Risk_monthly (Monthly risk linked to the share)

4.1 Regressions

The aim of the first investigation is to study the effectiveness and readability of bullet points. To this end, we are studying the impact of the number of bullet points on investor reaction (represented by the Cumulative Abnormal Return). The regression that is performed is the following equation:

$$\begin{aligned} log(1 + carX) = \beta_0 &+ \beta_1 log(1 + nbrbullets) + \beta_2 EARNTA + \beta_3 return + \beta_4 sd \\ &+ \beta_5 loss + \beta_6 log(1 + mkvaltq) + \beta_8 errormedian + \beta_9 log(1 + NoA) \\ &+ \beta_{10} factor(fyearq) + \beta_{11} factor(fqtr) + \beta_{12} risk\_monthly + \epsilon \end{aligned}$$

Where log(1+carX) is the dependent variable and X can take the value of one, three or five (for car1, car3 and car5). β_0 is the intercept and log(1+nbrbullets) is the independent variable. Some of the variables are logarithmic to ensure a linear relationship between the variables. Also, other variables are factorized in view of their ability to categorize our data.

Various tests have been carried out to confirm the accuracy of the model. The results of a correlation table showed poor performance between variables (Appendix 1). The maximum correlations observed were -0.293 and 0.488, irrespective of the dependent variable used. In addition, a Variance Inflation Factor (VIF) test was carried out and showed results not exceeding 2.456, far from the threshold of 5 which considers possible multicollinearity across variables (Appendix 2). Finally, a Breusch-Pagan test was performed and showed the absence of heteroskedasticity.

The second investigation looks at the impact of the topics on investor reaction. A joint investigation with all the topics was envisaged, but a high score on the VIF test showed possible multicollinearity between the variables (Appendix 3). These results demonstrated the risk of having an invalid model and forced us to abandon this approach. Therefore a separated investigation for each topic is needed.

For the second investigation, regression follows the equation:

$$\begin{aligned} log(1+carX) = \beta_0 &+ \beta_1 log(1+Topic) + \beta_2 EARNTA + \beta_3 return + \beta_4 sd + \beta_5 loss \\ &+ \beta_6 log(1+mkvaltq) + \beta_7 errormedian + \beta_8 log(1+NoA) \\ &+ \beta_9 factor(fyearq) + \beta_{10} factor(fqtr) + \beta_{11} risk\_monthly + \epsilon \end{aligned}$$

Where log(1+carX) is the dependent variable and X takes the value of one, three or five (for car1, car3 and car5). β_0 is the intercept and log(1+Topic) is the model's dependent variable and "Topic" can take the value of Governance, Finance, Performance, Operation or Regulation. As for the first investigation, some of the variables are logarithmic to ensure a linear relationship between the variables and some of them are factorized due to their ability to categorize our data.

The control variables were also tested. The correlation table shows a maximum negative correlation of -0.487 and a maximum positive correlation of 0.332 between the variables (Appendix 5). A VIF test was also carried out on each topic, and none showed the existence of multicollinearity (Appendix 4). As for the first investigation, a Breusch-Pagan test also proved the presence of homoskedasticity in our model.

5 Empirical analysis

5.1 Bullet point effectiveness and readability

5.1.1 Bullet points impact

Table 4 shows the results of our first regression investigating the impact of the number of bullet point in EPRs on investors reaction. The results show that investor reaction is significantly linked to a variation in the number of bullet points in the medium term (car3 and car5). Furthermore, the number of bullet points is positively correlated with investors reaction in both cases.

The relationship between the number of bullet points and car3 is significant at 90%. This means that there is less than a 10% chance of being wrong about the fact that a variation in the number of bullet points changes investor reaction. The latter relationship is positive at 0.2%. In other words, an increase in the number of bullet points has a positive impact on investor reaction of 0.2%. This relationship with car3 indicates that investor reaction is observed within a time lapse of three days before and after the EPR. In the

other hand, the relationship between car5 and the number of bullet points is significant at 95%. This relationship is positive by 0.3%, meaning that a 1% increase in the number of bullet points would increase car5 by 0.3%. This stronger relationship than with car3 means that investor reaction is greater over a longer time lapse. In fact, car5 represents the time lapse of five days before and after the release of the report, so only four more days are observed than with car3, and these influence the result by 0.1%.

The results are in line with our hypothesis (H1a: The more the buller points there are in EPRs, the greater the reaction of investors). These results confirm the usefulness of bullet points in EPRs and it also demonstrate that bullet points have an impact in terms of structural readability. Moreover, this phenomenon corroborates Ho's (2023) literature. Our investigation shows that the cognitive effects of bullet points apply in the case of financial reports. What's more, the results allow us to go further in our investigations, since they suggest a gain in the readers' processing fluency. However, the literature has shown the limits of bullet point effectiveness through over-use (Ho, 2023). Future research is needed to clarify whether the relationship between the number of bullet points and investor reaction is still significant for reports with many bullet points.

Table 4: Investors' reaction and presence of bullet points

	Dependent variable:		
	log(1 + car1)	log(1 + car3)	log(1 + car5)
	(1)	(2)	(3)
log(1 + nbrbullets)	0.001	0.002*	0.003**
	(0.001)	(0.001)	(0.001)
EARNTA	0.093***	0.125***	0.138***
	(0.014)	(0.016)	(0.017)
return	−0.013	0.007	0.014
	(0.008)	(0.009)	(0.010)
sd	0.042**	−0.001	0.014
	(0.018)	(0.021)	(0.023)
loss	−0.016***	−0.019***	−0.016***
	(0.001)	(0.002)	(0.002)
log(1 + mkvaltq)	−0.004***	−0.003***	−0.003***
	(0.0003)	(0.0004)	(0.0004)
errormedian	0.00004*	0.00002	−0.00000
	(0.00002)	(0.00002)	(0.00003)
log(1 + NoA)	0.003***	0.003***	0.004***
	(0.001)	(0.001)	(0.001)
risk_monthly	−0.007***	−0.015***	−0.026***
	(0.001)	(0.001)	(0.001)
factor(fqtr)2	0.0003	0.001	−0.002
	(0.001)	(0.001)	(0.001)
factor(fqtr)3	−0.002	−0.001	−0.002
	(0.001)	(0.001)	(0.001)
factor(fqtr)4	−0.002	0.0002	0.001
	(0.001)	(0.001)	(0.002)
factor(fyearq)	Yes	Yes	Yes
Constant	0.020***	0.010*	0.004
	(0.005)	(0.006)	(0.007)
Observations	32,291	32,291	32,291
R^2	0.016	0.018	0.020
Adjusted R^2	0.015	0.017	0.020
Residual Std. Error (df = 32261)	0.072	0.083	0.092
F Statistic (df = 29; 32261)	17.658***	20.018***	23.154***
Note:			*p<0.1; **p<0.05; ***p<0.01

5.1.2 Company size influence

Our first investigation showed the usefulness of bullet points and their potential impact on readability. However, the literature states that company size affects readability (Dodoo & Gu, 2020). In this section, we investigate whether the usefulness of bullet points differs according to company size. To this end, the companies in the sample were separated into two distinct groups. The first group represents companies with a market value below the sample median. The second group comprises companies with a market value above the sample median.

Table 5: Company size effect on bullet points/investor reaction relationship

	Dependent variable:					
	Market value below the median			*Market value above the median*		
	log(1 + car1)	log(1 + car3)	log(1 + car5)	log(1 + car1)	log(1 + car3)	log(1 + car5)
	(1)	(2)	(3)	(4)	(5)	(6)
log(1 + nbrbullets)	0.004**	0.006***	0.007***	0.0001	−0.0002	0.0005
	(0.002)	(0.002)	(0.002)	(0.001)	(0.001)	(0.001)
EARNTA	0.099***	0.143***	0.151***	0.068***	0.072***	0.092***
	(0.019)	(0.023)	(0.026)	(0.022)	(0.025)	(0.027)
return	0.038***	0.052***	0.051***	−0.093***	−0.071***	−0.041***
	(0.013)	(0.015)	(0.017)	(0.011)	(0.012)	(0.013)
sd	0.015	−0.003	0.026	0.070**	−0.014	−0.040
	(0.025)	(0.030)	(0.033)	(0.031)	(0.035)	(0.038)
loss	−0.021***	−0.024***	−0.024***	−0.011***	−0.013***	−0.009***
	(0.002)	(0.003)	(0.003)	(0.002)	(0.002)	(0.002)
log(1 + mkvaltq)	−0.004***	−0.005***	−0.004***	−0.002***	−0.002***	−0.002***
	(0.001)	(0.001)	(0.001)	(0.0004)	(0.0004)	(0.0005)
errormedian	0.00005*	0.00003	0.00001	0.008*	0.031***	0.008
	(0.00003)	(0.00003)	(0.00003)	(0.004)	(0.005)	(0.005)
log(1 + NoA)	0.004**	0.004**	0.006***	0.003***	0.003***	0.004***
	(0.001)	(0.002)	(0.002)	(0.001)	(0.001)	(0.001)
risk_monthly	−0.016***	−0.023***	−0.035***	0.007***	−0.002	−0.013***
	(0.002)	(0.002)	(0.002)	(0.001)	(0.002)	(0.002)
factor(fqtr)2	0.001	0.0002	−0.002	−0.0003	0.001	−0.001
	(0.002)	(0.002)	(0.003)	(0.001)	(0.001)	(0.001)
factor(fqtr)3	−0.001	0.00000	0.0003	−0.002*	−0.002*	−0.003**
	(0.002)	(0.002)	(0.003)	(0.001)	(0.001)	(0.001)
factor(fqtr)4	−0.0003	−0.0002	−0.002	−0.003**	0.001	0.002
	(0.003)	(0.003)	(0.003)	(0.001)	(0.002)	(0.002)
factor(fyearq)	Yes	Yes	Yes	Yes	Yes	Yes
Constant	0.018*	0.006	0.0002	0.016***	0.014***	0.010*
	(0.009)	(0.011)	(0.012)	(0.004)	(0.005)	(0.005)
Observations	13,732	13,732	13,732	18,559	18,559	18,559
R^2	0.023	0.027	0.032	0.018	0.014	0.011
Adjusted R^2	0.021	0.025	0.030	0.017	0.013	0.010
Residual Std. Error	0.086 (df = 13702)	0.101 (df = 13702)	0.114 (df = 13702)	0.060 (df = 18531)	0.067 (df = 18531)	0.072 (df = 18531)
F Statistic	10.990*** (df = 29; 13702)	13.004*** (df = 29; 13702)	15.601*** (df = 29; 13702)	12.739*** (df = 27; 18531)	10.041*** (df = 27; 18531)	7.807*** (df = 27; 18531)
Note:						*p<0.1; **p<0.05; ***p<0.01

Table 5 below shows the results obtained. The results show strong significance between the two variables for companies below the median market value. The number of bullet points is significant at 95% for car1 and shows a positive relationship of 0.4%. This last finding shows that investors react to an increase in the number of bullet points in the very short term. In fact, car1 shows investors' reactions spread over two days (the day of the announcement and the day after the announcement). Furthermore, the relationship between the number of bullet points and investor reaction is highly significant for car3 and car5. Moreover, investors reacted more strongly to an increase in the number of bullet

points than in car1. The positive relationship is 0.6% for car3 and 0.7% for car5.

Results below the median market value are not significant. These results validate the constructed hypothesis (H1b: Investor reaction to EPR bullet points is influenced by company size). The results also corroborate the literature of Dodoo an Gu (2020). This suggests that the readibility of bullet points is affected in this case. This also suggest possible conceptual fluency in information processing which means that bullet point are not seen as an attractive visual in this case (Laura and al., 2017). However, this does not call into question the structural effectiveness of bullet points but calls into question the benefits of data processing described by Ho (2023). To understand this phenomenon further, more detailed research is required.

5.2 Content effectiveness investigations

Table 6: Performance and Finance topic impact on investors reaction

	Dependent variable:					
	Performance topic impact			*Finance topic impact*		
	log(1 + car1)	log(1 + car3)	log(1 + car5)	log(1 + car1)	log(1 + car3)	log(1 + car5)
	(1)	(2)	(3)	(4)	(5)	(6)
log(1 + Performance)	0.153***	0.150***	0.156***			
	(0.036)	(0.041)	(0.046)			
log(1 + Finance)				−0.061*	−0.031	0.008
				(0.034)	(0.039)	(0.043)
EARNTA	0.080***	0.112***	0.126***	0.085***	0.119***	0.134***
	(0.014)	(0.016)	(0.018)	(0.014)	(0.016)	(0.017)
return	−0.006	0.012	0.021**	−0.007	0.012	0.020*
	(0.008)	(0.010)	(0.011)	(0.008)	(0.010)	(0.011)
sd	0.040**	−0.005	0.011	0.039**	−0.004	0.015
	(0.018)	(0.021)	(0.023)	(0.018)	(0.021)	(0.024)
loss	−0.017***	−0.020***	−0.018***	−0.017***	−0.020***	−0.017***
	(0.001)	(0.002)	(0.002)	(0.001)	(0.002)	(0.002)
log(1 + mkvaltq)	−0.004***	−0.003***	−0.003***	−0.004***	−0.003***	−0.003***
	(0.0003)	(0.0004)	(0.0004)	(0.0003)	(0.0004)	(0.0004)
errormedian	0.00004*	0.00002	−0.00000	0.00004*	0.00002	−0.00000
	(0.00002)	(0.00002)	(0.00003)	(0.00002)	(0.00002)	(0.00003)
log(1 + NoA)	0.003***	0.004***	0.005***	0.003***	0.004***	0.005***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
risk_monthly	−0.005***	−0.015***	−0.026***	−0.005***	−0.015***	−0.026***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
factor(fqtr)2	0.0002	0.001	−0.001	0.0002	0.001	−0.001
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
factor(fqtr)3	−0.002	−0.001	−0.001	−0.002	−0.001	−0.001
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
factor(fqtr)4	−0.002	0.0005	0.001	−0.002	0.0004	0.001
	(0.001)	(0.002)	(0.002)	(0.001)	(0.002)	(0.002)
factor(fyearq)	Yes	Yes	Yes	Yes	Yes	Yes
Constant	−0.004	−0.012	−0.018*	0.034***	0.021**	0.008
	(0.008)	(0.010)	(0.011)	(0.008)	(0.009)	(0.010)
Observations	30,097	30,097	30,097	30,097	30,097	30,097
R^2	0.016	0.018	0.020	0.016	0.018	0.020
Adjusted R^2	0.015	0.017	0.019	0.015	0.017	0.019
Residual Std. Error (df = 30067)	0.071	0.082	0.091	0.071	0.082	0.091
F Statistic (df = 29; 30067)	16.857***	19.458***	21.150***	16.339***	19.021***	20.747***
Note:						*p<0.1; **p<0.05; ***p<0.01

Table 6 shows separately the impact of the performance and finance topics on investor reaction.

The impact of the performance topic is highly significant with investor reaction, regardless of the time lapse observed (car1, car3 and car5). Significance is 99%. Furthermore, the presence of the topic is positively related to car1, car3 and car5. This positive relationship is 15.3% for car1, 15% for car3 and 15.6% for car5. This phenomenon means that a positive variation of 1% in the weight of the performance-related topic in bullet points increases investor reaction in the short and medium term by 15% on average. In the other hand, investors react more to the performance topic than to the finance topic. Indeed, only 90% significance is observed for car1. Moreover, the relationship between the two variables is negative, which runs counter to our intuitions which expected that financial topics make investors react stronger. A one-percentage-point increase in the weight of the finance topic reduces investor reaction by 6.1% on the day of the announcement and the day after. No significativity is noticed for car3 and car5.

Table 7 shows the impact of topics related to governance and corporate operations. The latter are studied separately. Our regression results show no significance in either case. Consequently, no conclusion can be drawn as to the impact of our independent variables on investor reaction. However, Table 8 presents the results of the regression studying the impact of the topic on regulations on investor reaction and shows a significant relationship at 95% between the dependent variable and the independent variable. This means that the presence of a regulatory topic does affect investor reaction. The relationship is negative at -9.3% between the two variables in the short term (car1).

These results are not conclusive enough to validate our hypothesis (H2: Topics related to quantitative data such as finance and performance make investors react more strongly than topics related to qualitative data such as operations, governance and regulation). Our expectation were that topics related to performance and finance would make investors react more strongly. To a certain extent, the results indicate that this is the case, but we observed that the topic of regulation also had a significant impact on investors. Our intuition was that qualititative topics impact would be absorbed by quantitative topics impact, given its direct link to performance. However, topics such as performance and regulation suggest to be in line with the conceptual theory (which is linked to quick mental association) (Reber et al., 2004). In other words, this last statement suggests that investors pay more attention to the content of bullet points than to the visual represen-

tation they represent. Finally, our results don't allow us to make a definitive statement on the fact that content adds quality to information, but they do point the way to future research on the topics on which information makes investors react.

Table 7: Governance and Operation topic impact on investors reaction

	Dependent variable:					
	Governance topic impact			*Operation topic impact*		
	log(1 + car1)	log(1 + car3)	log(1 + car5)	log(1 + car1)	log(1 + car3)	log(1 + car5)
	(1)	(2)	(3)	(4)	(5)	(6)
log(1 + Governance)	−0.011	−0.065	−0.076			
	(0.036)	(0.042)	(0.046)			
log(1 + Operation)				−0.013	−0.020	−0.046
				(0.039)	(0.045)	(0.049)
EARNTA	0.086***	0.116***	0.129***	0.087***	0.120***	0.134***
	(0.014)	(0.016)	(0.018)	(0.014)	(0.016)	(0.017)
return	−0.006	0.013	0.021**	−0.007	0.012	0.020*
	(0.008)	(0.010)	(0.011)	(0.008)	(0.010)	(0.011)
sd	0.044**	0.001	0.018	0.044**	−0.001	0.015
	(0.018)	(0.021)	(0.024)	(0.018)	(0.021)	(0.023)
loss	−0.016***	−0.020***	−0.017***	−0.016***	−0.020***	−0.017***
	(0.001)	(0.002)	(0.002)	(0.001)	(0.002)	(0.002)
log(1 + mkvaltq)	−0.004***	−0.003***	−0.003***	−0.004***	−0.003***	−0.003***
	(0.0003)	(0.0004)	(0.0004)	(0.0003)	(0.0004)	(0.0004)
errormedian	0.00004*	0.00002	−0.00000	0.00004*	0.00002	−0.00000
	(0.00002)	(0.00002)	(0.00003)	(0.00002)	(0.00002)	(0.00003)
log(1 + NoA)	0.003***	0.004***	0.005***	0.003***	0.004***	0.005***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
risk_monthly	−0.005***	−0.015***	−0.026***	−0.005***	−0.015***	−0.026***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
factor(fqtr)2	0.0002	0.001	−0.001	0.0002	0.001	−0.001
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
factor(fqtr)3	−0.002	−0.001	−0.001	−0.002	−0.001	−0.001
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
factor(fqtr)4	−0.002	0.0003	0.001	−0.002	0.0003	0.001
	(0.001)	(0.002)	(0.002)	(0.001)	(0.002)	(0.002)
factor(fyearq)	Yes	Yes	Yes	Yes	Yes	Yes
Constant	0.025***	0.026***	0.023**	0.025***	0.019*	0.019
	(0.008)	(0.010)	(0.011)	(0.009)	(0.010)	(0.012)
Observations	30,097	30,097	30,097	30,097	30,097	30,097
R^2	0.015	0.018	0.020	0.015	0.018	0.020
Adjusted R^2	0.014	0.017	0.019	0.014	0.017	0.019
Residual Std. Error (df = 30067)	0.071	0.082	0.091	0.071	0.082	0.091
F Statistic (df = 29; 30067)	16.229***	19.083***	20.840***	16.230***	19.006***	20.777***
Note:						*p<0.1; **p<0.05; ***p<0.01

Table 8: Regulation topic impact on investor reaction

	Dependent variable:		
	log(1 + car1)	log(1 + car3)	log(1 + car5)
	(1)	(2)	(3)
log(1 + Regulation)	−0.093**	−0.048	−0.075
	(0.043)	(0.049)	(0.055)
EARNTA	0.088***	0.120***	0.134***
	(0.014)	(0.016)	(0.017)
return	−0.007	0.012	0.020*
	(0.008)	(0.010)	(0.011)
sd	0.043**	−0.002	0.014
	(0.018)	(0.021)	(0.023)
loss	−0.016***	−0.020***	−0.017***
	(0.001)	(0.002)	(0.002)
log(1 + mkvaltq)	−0.004***	−0.004***	−0.003***
	(0.0003)	(0.0004)	(0.0004)
errormedian	0.00004*	0.00002	−0.00000
	(0.00002)	(0.00002)	(0.00003)
log(1 + NoA)	0.003***	0.004***	0.005***
	(0.001)	(0.001)	(0.001)
risk_monthly	−0.005***	−0.015***	−0.026***
	(0.001)	(0.001)	(0.001)
factor(fqtr)2	0.0002	0.001	−0.001
	(0.001)	(0.001)	(0.001)
factor(fqtr)3	−0.002	−0.001	−0.001
	(0.001)	(0.001)	(0.001)
factor(fqtr)4	−0.002	0.0004	0.001
	(0.001)	(0.002)	(0.002)
factor(fyearq)	Yes	Yes	Yes
Constant	0.040***	0.024**	0.023*
	(0.009)	(0.011)	(0.012)
Observations	30,097	30,097	30,097
R^2	0.016	0.018	0.020
Adjusted R^2	0.015	0.017	0.019
Residual Std. Error (df = 30067)	0.071	0.082	0.091
F Statistic (df = 29; 30067)	16.394***	19.032***	20.812***
Note:			*p<0.1; **p<0.05; ***p<0.01

6 Additional analysis - Diversity impact

Table 9: Diversity impact on investors reaction

	Dependent variable:		
	log(1 + car1)	log(1 + car3)	log(1 + car5)
	(1)	(2)	(3)
log(1 + Diversity)	−0.501**	−0.127	−0.267
	(0.230)	(0.265)	(0.294)
EARNTA	0.089***	0.120***	0.134***
	(0.014)	(0.016)	(0.017)
return	−0.007	0.012	0.020*
	(0.008)	(0.010)	(0.011)
sd	0.042**	−0.002	0.014
	(0.018)	(0.021)	(0.023)
loss	−0.016***	−0.020***	−0.017***
	(0.001)	(0.002)	(0.002)
log(1 + mkvaltq)	−0.004***	−0.003***	−0.003***
	(0.0003)	(0.0004)	(0.0004)
errormedian	0.00004*	0.00002	−0.00000
	(0.00002)	(0.00002)	(0.00003)
log(1 + NoA)	0.003***	0.004***	0.005***
	(0.001)	(0.001)	(0.001)
risk_monthly	−0.005***	−0.015***	−0.026***
	(0.001)	(0.001)	(0.001)
factor(fqtr)2	0.0002	0.001	−0.001
	(0.001)	(0.001)	(0.001)
factor(fqtr)3	−0.002	−0.001	−0.001
	(0.001)	(0.001)	(0.001)
factor(fqtr)4	−0.002	0.0004	0.001
	(0.001)	(0.002)	(0.002)
factor(fyearq)	Yes	Yes	Yes
Constant	0.503**	0.136	0.265
	(0.220)	(0.254)	(0.282)
Observations	30,097	30,097	30,097
R^2	0.016	0.018	0.020
Adjusted R^2	0.015	0.017	0.019
Residual Std. Error (df = 30067)	0.071	0.082	0.091
F Statistic (df = 29; 30067)	16.392***	19.006***	20.775***
Note:			*p<0.1; **p<0.05; ***p<0.01

Based on the methodology of previous research (Deveux, 2021), we realized that measuring the impact of topic diversity on investor reaction can tell us more about the importance of content in bullet points. For this investigation, we calculated Simson's diversity index, which is initially used to measure diversity within a group of individuals.

Table 9 investigate the impact of topics diversity on investors reaction. The results show that topics diversity has a significant impact on investor reaction in the short term (car1, representing the time lapse one day before and after the announcement). Furthermore, the relationship between the diversity index and our dependent variable is strongly negative at -50%. This last finding means that an increase in the diversity of topics in documents drastically reduces investor reaction.

7 Limits

This study is not without limitations. The precise scope of our investigations does not take into account certain variables that should be integrated in future work.

Initially, our research into the effectiveness of bullet points is limited by our sample. Table 3 shows a maximum of ten bullet points in the EPRs of the companies in the sample. This maximum is too low to investigate whether the effectiveness of bullet points is impacted, as Ho (2023) points out. Ho (2023) asserts that the readability of bullet points declines with overuse. Our study therefore does not allow us to take this phenomenon into account.

Secondly, literature indicates the value that investors place on the content before an investment. We have seen that among the topics found in our study, some are dependent on the context which in the company operates. In particular, the topic of regulation is highly dependent on the legislature under which the company operates. In other words, the legal obligations of companies in the United States are not the same as in Europe. For future work, it is imperative to categorize the sample by locality. In our case, we were unable to do this, as almost 90% of the sample are American companies.

Our study is also limited by the analysis of investor information processing. We have seen in the literature that investors' processing fluency is partly subjective (Novemsky et al., 2007). Although our study may suggest that investors follow a conceptual fluency or a perceptual fluency in processing information, it does not allow us to establish any

generality.

Finally, several studies have shown that the LDA topic modeling method is limited. For example, the topic modeling method does not take into account trends over time (Vayansky & Kumar, 2020). This posed complications for the study we carried out, given our sample from 2004 to 2021. Furthermore, this method encounters difficulties when the database includes a large number of words (Vayansky & Kumar, 2020). Several methods have been developed to overcome these problems, and studies using other topic modeling methods would be interesting.

8 Conclusion

In this thesis, we investigated the structural readability of earning press releases through the presence of bullet points. Furthermore, we looked for frequently used topics in bullet points and assess the impact of the content on investors. Our findings enable us to understand the usefulness of textual structure in financial reports and confirm the readability advantages of bullet points mentioned in the literature. Content inspection also sheds light on the quantitative and qualitative data mentioned in bullet points. Finally, analysis of investor reaction to the presence of bullet points and their content give overview on possible existence of information asymmetry in earning press releases.

Using a database containing the content and number of bullet points in the earning press releases of 1,725 companies from 2004 and 2021, we carried out two investigations to analyze the informativeness and usefulness of bullet points. An LDA topic modeling method was used to identify the topics appearing in the bullet points. Using this data, two types of regression were run to observe the impact of bullet points and their content on the cumulative abnormal return representing investor reaction.

Our first study showed that the presence of bullet points caused investors to react within a time lapse ranging from three to five days before and after the publication of an earning press release. We also found that investor reaction was positively correlated with the number of bullet points during this period. Furthermore, our results corroborate the literature and affirm that this relationship is stronger for companies with a lower market value. Finally, the results from investigating the effect of company size suggest that investors rely on conceptual fluency when processing information.

Secondly, the results of topic modeling enabled us to identify five topics frequently used in bullet points. We found that topics related to governance, finance, performance, operations and regulations each account for an average of 20% of the information. By studying the impact of each of these topics we showed that performance-related information was the most significant on investor reaction, and that the relationship was strongly positive. We also demonstrated a negative impact of the presence of financial or regulatory topics on investor reaction. This results were partially in line with the literature. Furthermore, an additional analysis showed that the diversity of topics had a negative short-term impact. These results corroborate the investigation into the effects of company size on readability, and lead us to suspect that investors follow a conceptual fluency when it comes to processing information in bullet points.

However, the study has certain limitations. Firstly, the literature has shown us the limits of the effectiveness of bullet points through their intensive use (Ho, 2023). This phenomenon could not be observed in our sample, which showed a maximum of ten bullet points in earning press releases. In addition, the discovery of certain topics such as regulation made us realize that its involvement was context-dependent, and further research is needed to better understand its real impact. Our investigation is also limited in the interpretation of investors' information processing. The subjective aspect of processing fluency does not allow us to construct generalizations (Novemsky et al., 2007). Finally, studies have shown the limitations of the LDA method (Vayansky & Kumar, 2020). Our study observes the content of earning press releases over a given time lapse, and this method does not allow us to differentiate between textual trends over time.

Finally, our study specifically targets the textual structure of earning press releases. It does not, therefore, enable the phenomenon to be generalized to the entire text or to other financial reports. Research into text structures such as tabulation would complement this study to understand the general impact of structural readability in earning press releases. In the same way, similar studies on other financial reports could provide some answers as to the recurrence of the phenomena we've been looking at.

Bibliography

Alter, A. L., & Oppenheimer, D. M. (2009). Uniting the Tribes of Fluency to Form a Metacognitive Nation. *Personality And Social Psychology Review*, *13*(3), 219-235. https://doi.org/10.1177/1088868309341564

Begley, J., & Fischer, P. E. (1999). Is there information in an earnings announcement delay. *Review Of Accounting Studies*, 3, 347 363.

Blei, D-M., Ng, A-Y,. & Jordan, M-I. (2003). Latent Dirichlet Allocation. The *Journal of Machine Learning Research,* (3), pp. 993-1022. https://www.jmlr.org/papers/volume3/blei03a/blei03a.pdf?ref=http://githubhelp.com

Boudt, K., Thewissen, J., & Torsin, W. (2018). When does the tone of earnings press releases matter ? *International Review Of Financial Analysis, 57*, 231 245. https://doi.org/10.1016/j.irfa.2018.02.002

Brown, S., & Hillegeist, S. A. (2007). How disclosure quality affects the level of information asymmetry. *Review Of Accounting Studies, 12*(2 3), 443 477. https://doi.org/10.1007/s11142-007-9032-5

Cheng, L., Roulstone, D. T., & Van Buskirk, A. (2018). Are Investors Influenced by the Order of Information in Earnings Press Releases. https://www.bauer.uh.edu/departments/accy/research/documents/Roulstone-Are-Investors-Influenced-by-the-Order-of-Information-in-Earnings-Press-Releases.pdf

Christensen, T. E., Fronk, K. E., Lee, J. A., & Nelson, K. K. (2024). Data visualization in 10-K filings. *Journal Of Accounting And Economics, 77*(2 3), 101631. https://doi.org/10.1016/j.jacceco.2023.101631

Davis, A. K., Piger, J. M., & Sedor, L. M. (2012). Beyond the Numbers : Measuring the Information Content of Earnings Press Release Language*. *Contemporary Accounting Research*, *29*(3), 845-868. https://doi.org/10.1111/j.1911-3846.2011.01130.x

Deveux, Benoit. *The Informativeness of ICO Whitepapers, A Topic Modeling Approach.* Louvain School of Management, Université catholique de Louvain, 2021. Prom. : Thewissen, James. http:// hdl.handle.net/2078.1/thesis:29385

Dodoo, R. N. A., & Gu, S. (2020). The Impact of Firm Performance on Annual Report Readability: Evidence from Listed Firms in Ghana. *Journal of Economics, Business, and Accountancy Ventura, 22*(3), 444-454. DOI: 10.14414/jebav.v22i3.2097

Fiske, S. T., & Taylor, S. E. (2013). Social Cognition: From Brains to Culture. https://doi.org/10.4135/9781529681451

Graas, Alexandre. *The Informativeness in Earnings Press Release, A Topic Modeling Approach.* Louvain School of Management, Université catholique de Louvain, 2023. Prom. : Thewissen, James. http:// hdl.handle.net/2078.1/thesis:40109

Graf, L. K. M., Mayer, S., & Landwehr, J. R. (2017). Measuring Processing Fluency : One versus Five Items. *Journal Of Consumer Psychology*, *28*(3), 393-411. https://doi.org/10.1002/jcpy.1021

Henry, E. (2006). Are Investors Influenced by How Earnings Press Releases are Written ? *Social Science Research Network*. https://doi.org/10.2139/ssrn.933100

Ho, A. (2023). Enhancing User Experience: Exploring the Interactions and Visualisation of Data Through Bullet Points and Typographic Design. *Human Interaction and Emerging Technologies, 111*, 985-995. http://doi.org/10.54941/ahfe1004171

Ho, A. (2023b). Types and Role of the Bullet Point System in Reading Experience. *AHFE International*. https://doi.org/10.54941/ahfe1003704

How to Read an 8-K. (s. d.). SEC. https://www.sec.gov/files/readan8k.pdf

Joshi, D., & Pradhan, T. (2018). *A study on the impact of earning news on investors' decision.* Parul University.

Loughran, T., & Mcdonald, B. (2014). Measuring Readability in Financial Disclosures. *The Journal Of Finance, 69*(4), 1643 1671. https://doi.org/10.1111/jofi.12162

Novemsky, N., Dhar, R., Schwarz, N., & Simonson, I. (2007). Preference Fluency in Choice. *Journal Of Marketing Research*, *44*(3), 347-356. https://doi.org/10.1509/jmkr.44.3.347

Nwaobia, A. N., Kwarbai, J., Olajumoke, J., & Ajibade, A. T. (2013). Financial Reporting Quality on Investors' Decisions. *Social Science Research Network.* https://papers.ssrn.com/sol3/Delivery.cfm/SSRN_ID3139529_code1629942.pdf?abstractid=3139529&mirid=1

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

Schofield, A., Magnusson, M., Thompson, L., & Mimno, D. (2017, July). Understanding text *pre-processing for latent Dirichlet allocation. In Proceedings of the 15th conference of the European chapter of the Association for Computational Linguistics* (2), 432-436.

Thewissen, J., Shrestha, P., Torsin, W., & Pastwa, A. M. (2022). Unpacking the black box of ICO white papers : A topic modeling approach. *Journal Of Corporate Finance*, *75*, 102225. https://doi.org/10.1016/j.jcorpfin.2022.102225

Vayansky, I., & Kumar, S. A. (2020). A review of topic modeling methods. *Information Systems, 94*, 101582. https://doi.org/10.1016/j.is.2020.101582

Winkielman, P., Schwarz, N., Fazendeiro, T. A., & Reber, R. (2003). The hedonic marking of processing fluency: Implications for evaluative judgment. In J. Musch & K. C. Klauer (Eds.), The psychology of evaluation. Affective processes in cognition and emotion (pp. 189–217).

Appendix

Appendix 1 - Bullet points investigation - Correlation matrix

	nbrbullets	EARNTA	return	sd	loss	mkvaltq	errormedian	NoA	fyearq	fqtr	risk_monthly
nbrbullets	1	0.005	-0.002	-0.036	-0.018	0.085	-0.010	0.068	0.127	0.048	-0.012
EARNTA	0.005	1	-0.243	-0.414	-0.491	0.066	0.080	0.110	-0.030	0.047	0.077
return	-0.002	-0.243	1	0.250	0.326	-0.100	-0.008	-0.101	-0.0004	-0.032	0.102
sd	-0.036	-0.414	0.250	1	0.256	-0.065	-0.053	-0.077	0.032	-0.057	-0.013
loss	-0.018	-0.491	0.326	0.256	1	-0.092	-0.020	-0.114	0.054	-0.050	-0.095
mkvaltq	0.085	0.066	-0.100	-0.065	-0.092	1	0.003	0.467	0.006	0.052	0.006
errormedian	-0.010	0.080	-0.008	-0.053	-0.020	0.003	1	0.012	-0.0002	0.013	0.012
NoA	0.068	0.110	-0.101	-0.077	-0.114	0.467	0.012	1	-0.035	0.075	-0.008
fyearq	0.127	-0.030	-0.0004	0.032	0.054	0.006	-0.0002	-0.035	1	-0.052	-0.036
fqtr	0.048	0.047	-0.032	-0.057	-0.050	0.052	0.013	0.075	-0.052	1	0.045
risk_monthly	-0.012	0.077	0.102	-0.013	-0.095	0.006	0.012	-0.008	-0.036	0.045	1

Appendix 2 - Bullet points investigation - VIF test

	GVIF	Df	GVIF^(1/(2*Df))
log(1 + nbrbullets)	1.039	1	1.019
EARNTA	1.358	1	1.165
return	1.696	1	1.302
sd	1.142	1	1.069
loss	1.406	1	1.186
log(1 + mkvaltq)	2.455	1	1.567
errormedian	1.004	1	1.002
log(1 + NoA)	2.243	1	1.498
risk_monthly	1.118	1	1.057
factor(fqtr)	1.044	3	1.007
factor(fyearq)	1.694	17	1.016

Appendix 3 - Content investigation (if all topics are in the same regression) - VIF test

Table 10:

	GVIF	Df	GVIF^(1/(2*Df))
log(1 + Governance)	326.377	1	18.066
log(1 + Finance)	362.044	1	19.027
log(1 + Performance)	307.215	1	17.528
log(1 + Operation)	267.585	1	16.358
log(1 + Regulation)	213.827	1	14.623
EARNTA	1.404	1	1.185
return	1.700	1	1.304
sd	1.164	1	1.079
loss	1.414	1	1.189
log(1 + mkvaltq)	2.473	1	1.573
errormedian	1.005	1	1.002
log(1 + NoA)	2.255	1	1.502
risk_monthly	1.122	1	1.059
factor(fqtr)	1.045	3	1.007
factor(fyearq)	1.715	17	1.016

Appendix 4 - Content investigation - VIF test

Table 11: Governance

	GVIF	Df	GVIF^(1/(2*Df))
log(1 + Governance)	1.071	1	1.035
EARNTA	1.380	1	1.175
return	1.699	1	1.303
sd	1.149	1	1.072
loss	1.404	1	1.185
log(1 + mkvaltq)	2.468	1	1.571
errormedian	1.005	1	1.002
log(1 + NoA)	2.248	1	1.499
factor(fyearq)	1.670	17	1.015
factor(fqtr)	1.041	3	1.007
risk_monthly	1.121	1	1.059

Table 12: Finance

	GVIF	Df	GVIF^(1/(2*Df))
log(1 + Finance)	1.049	1	1.024
EARNTA	1.357	1	1.165
return	1.695	1	1.302
sd	1.160	1	1.077
loss	1.413	1	1.188
log(1 + mkvaltq)	2.452	1	1.566
errormedian	1.004	1	1.002
log(1 + NoA)	2.247	1	1.499
factor(fyearq)	1.681	17	1.015
factor(fqtr)	1.042	3	1.007
risk_monthly	1.122	1	1.059

Table 13: Performance

	GVIF	Df	GVIF^(1/(2*Df))
log(1 + Performance)	1.022	1	1.011
EARNTA	1.373	1	1.172
return	1.692	1	1.301
sd	1.144	1	1.070
loss	1.403	1	1.184
log(1 + mkvaltq)	2.452	1	1.566
errormedian	1.004	1	1.002
log(1 + NoA)	2.244	1	1.498
factor(fyearq)	1.666	17	1.015
factor(fqtr)	1.041	3	1.007
risk_monthly	1.120	1	1.058

Table 14: Operation

	GVIF	Df	GVIF^(1/(2*Df))
log(1 + Operation)	1.021	1	1.010
EARNTA	1.357	1	1.165
return	1.692	1	1.301
sd	1.142	1	1.069
loss	1.403	1	1.185
log(1 + mkvaltq)	2.454	1	1.566
errormedian	1.004	1	1.002
log(1 + NoA)	2.247	1	1.499
factor(fyearq)	1.671	17	1.015
factor(fqtr)	1.041	3	1.007
risk_monthly	1.121	1	1.059

Table 15: Regulation

	GVIF	Df	GVIF^(1/(2*Df))
log(1 + Regulation)	1.015	1	1.007
EARNTA	1.354	1	1.164
return	1.692	1	1.301
sd	1.142	1	1.069
loss	1.402	1	1.184
log(1 + mkvaltq)	2.457	1	1.568
errormedian	1.004	1	1.002
log(1 + NoA)	2.248	1	1.499
factor(fyearq)	1.672	17	1.015
factor(fqtr)	1.041	3	1.007
risk_monthly	1.120	1	1.058

Appendix 5 - Content investigation - Correlation matrix

	Governance	Finance	Performance	Operation	Regulation	EARNTA	return	sd	loss	mkvaltq	errormedian	NoA	fyearq	fqtr	risk_monthly
Governance	1	-0.337	-0.327	-0.212	-0.262	-0.287	0.158	0.189	0.211	-0.007	-0.009	-0.045	0.026	-0.025	0.026
Finance	-0.337	1	-0.251	-0.298	-0.125	0.077	-0.083	-0.128	-0.129	0.064	0.004	0.050	0.035	0.039	-0.031
Performance	-0.327	-0.251	1	-0.239	-0.244	0.116	-0.051	-0.019	-0.052	-0.046	0.003	0.056	-0.005	-0.011	0.008
Operation	-0.212	-0.298	-0.239	1	-0.178	0.043	-0.017	-0.013	0.010	-0.032	0.003	-0.074	-0.027	-0.015	0.012
Regulation	-0.262	-0.125	-0.244	-0.178	1	0.091	-0.025	-0.055	-0.067	0.024	-0.0002	0.012	-0.041	0.016	-0.019
EARNTA	-0.287	0.077	0.116	0.043	0.091	1	-0.247	-0.421	-0.487	0.066	0.082	0.110	-0.028	0.047	0.074
return	0.158	-0.083	-0.051	-0.017	-0.025	-0.247	1	0.251	0.332	-0.103	-0.008	-0.098	0.002	-0.030	0.091
sd	0.189	-0.128	-0.019	-0.013	-0.055	-0.421	0.251	1	0.257	-0.070	-0.055	-0.077	0.033	-0.059	-0.014
loss	0.211	-0.129	-0.052	0.010	-0.067	-0.487	0.332	0.257	1	-0.096	-0.020	-0.113	0.050	-0.049	-0.094
mkvaltq	-0.007	0.064	-0.046	-0.032	0.024	0.066	-0.103	-0.070	-0.096	1	0.004	0.487	-0.005	0.055	0.005
errormedian	-0.009	0.004	0.003	0.003	-0.0002	0.082	-0.008	-0.055	-0.020	0.004	1	0.012	-0.0002	0.014	0.013
NoA	-0.045	0.050	0.056	-0.074	0.012	0.110	-0.098	-0.077	-0.113	0.487	0.012	1	-0.035	0.073	-0.006
fyearq	0.026	0.035	-0.005	-0.027	-0.041	-0.028	0.002	0.033	0.050	-0.005	-0.0002	-0.035	1	-0.053	-0.042
fqtr	-0.025	0.039	-0.011	-0.015	0.016	0.047	-0.030	-0.059	-0.049	0.055	0.014	0.073	-0.053	1	0.046
risk_monthly	0.026	-0.031	0.008	0.012	-0.019	0.074	0.091	-0.014	-0.094	0.005	0.013	-0.006	-0.042	0.046	1