

Louvain School of Management

The Informativeness in Earnings Press Release

A Topic Modeling Approach

Author(s): Graas Alexandre
Supervisor(s): Thewissen James
Academic year 2022-2023
Dissertation for the master of GEST2 MS/GE Master [120] en sciences
de gestion, à finalité spécialisée
Corporate Finance
Daytime schedule

The Informativeness in Earnings Press Release

A Topic Modeling Approach

Alexandre Graas

Louvain-School of Management (LSM)

alexandre.graas@student.uclouvain.be

Promotor: James Thewissen

UCLouvain

james.thewissen@uclouvain.be

ABSTRACT

The objective of this paper is to show the impact of the thematic content of earnings press releases on the probability of companies meeting or beating analysts' expectations. Using an algorithm that extracts a set of words representing a main topic from a collection of documents, we empirically determined and quantified 6 themes in a sample of 56,071 earnings press releases between 2004 and 2021. We find that the topics produced by this algorithm are semantically meaningful and that this gives us clear indications about the topics discussed in the earnings press of companies. We show in this paper which topics positively or negatively impact the probability of firms to meet or beat expectations. In addition to the meet or beat expectations variable, we focus on earnings manipulation to examine potential similarities in the impact of topic content on these two variables. In addition, we analyze the diversity of topics present in earnings press. We find that firms that meet or beat expectations tend to cover more topics in their earnings press. On the contrary, firms that practice earnings manipulation tend to cover less. Overall, this paper sheds light on the topics covered in earnings press and how investors are sensitive to them.

Keywords: Earnings press; Topic modeling; Thematic content; Information value

I would like to warmly thank James Thewissen, Professor of Finance at UCLouvain, for his supervision both in the writing of this paper and in the master's degree. The numerous projects and discussions over the past two years have allowed me to develop my understanding of a wide range of financial topics and especially my critical thinking skills. His experience and valuable advice were indispensable to the achievement of my Master Thesis.

Contents

1	Introduction	3
2	Literature review	8
2.1	Earnings Press Release	8
2.2	Meet or Beat Expectations	10
2.2.1	Topic Diversity in Meet or Beat Expectations	11
2.3	Earnings Management	12
2.3.1	Topic Diversity in Earnings Management	14
3	Research Question	15
3.1	Hypothesis	17
3.1.1	<i>Meet or Beat Expectations Hypothesis</i>	17
3.1.2	<i>Earnings Management Hypothesis</i>	17
4	Data & Sample selection	18
4.1	Topic Modeling Implementation	18
4.2	Statistic Descriptive	21
5	Empirical Analysis	22
5.1	Meet or Beat Expectations Interpretation	23
5.1.1	<i>Topics in Meet or Beat Expectations</i>	23
5.1.2	<i>Topic Diversity in Meet or Beat Expectations</i>	26
5.2	Earnings Management Interpretation	27
5.2.1	<i>Topics in Earnings Management</i>	27
5.2.2	<i>Topic Diversity in Earnings Management</i>	28
5.3	Empirical Evaluation of Hypothesis	29
5.3.1	<i>Hypothesis 1 : The topics of the companies that MBE are more financial while the topics of the companies that don't MBE are less financial?</i> . . .	29

5.3.2	<i>Hypothesis 2 : Are the topics that are significant when it is MBE always the same as when it is not MBE?</i>	30
5.3.3	<i>Hypothesis 3: Companies that meet or beat expectations tend to cover more topics in their earnings press releases?</i>	31
5.3.4	<i>Hypothesis 4: The topics of companies that use EM are less financial while the topics of companies that do not use EM are more financial? . .</i>	32
5.3.5	<i>Hypothesis 5: Companies that use earnings management tend to cover less topics in their earnings press releases?</i>	33
6	Conclusion	34

1 Introduction

The quarterly earnings season is a busy time for public companies and their investors. Indeed, firms issue press releases several weeks after the end of each quarter to announce their results for the period. These earnings press releases, however voluntary, are an important way for companies to communicate their financial performance to the public before filing their official financial statements with the U.S. Securities and Exchange Commission (SEC). Earnings press releases are an important form of spontaneous communication for companies to reduce the information asymmetry between management and stakeholders.

Recent research provides strong evidence that companies are managing both earnings and expectations to report earnings that meet or beat analysts' earnings forecasts. The SEC has also expressed its concern that the pressure to meet expectations influences the quality of financial reporting (Levitt Jr, 1998). The main feature of earnings releases is that they are voluntary, unregulated disclosures, giving management almost total discretion over what information to disclose to investors. These disclosures are therefore more difficult to regulate or challenge in court and give management the opportunity to more subtly manipulate market participants' perception of the company's future performance.

Meeting expectations gives companies more credibility in the eyes of investors. This will generate a lot of interest from investors and analysts. Indeed, for them, it is possible that a company that meets expectations is perceived as less risky. Some companies will then manipulate earnings in order to meet analysts' expectations. These manipulations will not necessarily allow companies to have higher future earnings than companies that do not meet expectations, but it will bring more credibility to the company.

In addition to manipulating earnings, other elements such as the tone of the press release or the way it is written can also influence the behavior of investors and analysts and the credibility of the company. The tone of earnings press releases (i.e. the difference between the amount of positive and negative words) is a reliable signal of future earnings and that the market reacts positively to the tone in the short and long term (Miller & Piotroski, 2000).

However, it is also interesting to look at companies that do not meet analysts' expectations. Indeed, some companies do not manage their earnings, which forces them to publish press releases that do not meet the expectations of investors and analysts. This can sometimes have negative consequences for companies. In general, negative results cause larger fluctuations than positive results, and large surprises tend to have the greatest impact. There is no formula for predicting market reactions, so it is impossible to know how much a large deviation or a positive result will affect a company's share price. Reporting earnings below consensus estimates, even by a small margin, can cause investors to penalize these companies by lowering the share price, for example.

Furthermore, it is also interesting to know why some companies do not meet these expectations. Indeed, certain elements such as the health crisis for example, or certain topics can influence the results of companies and lead them to publish earnings press releases below expectations. Therefore, we think it would be interesting to focus on the reasons for this non-meeting by looking at the topics that come up the most in the earnings press releases being below expectations. This leads us to construct the following research question which will guide our work, "what are the topics discussed in earnings press when firm performance is below expectations?".

We believe that this research could provide information on the topics covered in the press releases and that these could illustrate whether a company tends to meet or beat the analysts' expectations. To do this, we think it is useful to perform topic modeling by linking thematic content with miss or beat expectations.

First, we provide valuable insights into the relationship between meeting or beating expectations and the presence of financial topics in earnings press releases. We reveal a negative correlation between the binary variable MBE and the dependent variable Financial, indicating that companies with a higher proportion of financial topics in their press releases are less likely to meet or beat analyst expectations. This finding challenges the common perception that companies that emphasize financial topics in their communication are more likely to achieve positive performance outcomes. The results

contribute to the existing literature by highlighting the importance of considering the thematic content of press releases when assessing performance expectations. The study suggests that companies that focus on financial issues may have difficulties meeting or beating expectations. This implies that financial indicators alone may not be sufficient indicators of future performance and that a broader understanding of the topics covered in earnings releases is crucial. By highlighting the relationship between financial themes and meeting or exceeding expectations, this study provides empirical evidence that can inform investors, analysts, and other stakeholders in their decision-making processes. It underscores the need for a comprehensive analysis that considers both the financial and non-financial aspects of a company's communication when evaluating it.

Second we show that the important themes in earnings press releases differ by performance outcome. For firms that fail to meet or beat expectations, most of the themes in their earnings releases are significant at a high confidence level, indicating the importance of various non-financial factors in explaining underperformance. In contrast, for companies that meet or beat expectations, the significant themes are distinct. The social theme is highly significant, suggesting that companies that emphasize social aspects in their communication are more likely to beat expectations. In addition, the research&development theme shows some significance, although at a lower level of confidence. These results contribute to the literature by highlighting the divergent themes associated with meeting or beating expectations versus underperforming. They suggest that different factors may be driving the outcomes in these two scenarios. Understanding the specific themes that are significant to each outcome can help investors and analysts adjust their analysis and expectations accordingly. The study also highlights the importance of considering non-financial aspects, such as social and research&development factors, in addition to financial indicators when assessing performance expectations. By recognizing the distinct themes associated with different outcomes, stakeholders can better understand the underlying dynamics and drivers of business performance.

Third, we contribute to the existing literature by highlighting the relationship be-

tween meeting or beating expectations and the topic diversity of earnings press releases. The positive correlation observed suggests that companies that strive to meet or beat analysts' expectations are more likely to provide detailed information on various aspects of their business. In doing so, they seek to highlight their strong financial performance and build investor confidence. This finding is consistent with the idea that positive financial results can attract media and investor attention, and companies can strategically disclose more specific information to capitalize on this positive momentum. By broadening the range of topics covered, companies may be able to provide a complete picture of their strategy and performance, ultimately enhancing their credibility and transparency. Overall, these results shed light on the motivations behind the choice of topics covered in earnings releases and offer valuable insights into the communication strategies of companies that meet or beat expectations.

The results of this study also provide an interesting perspective on the relationship between earnings management and the thematic content of earnings releases. While the hypothesis predicted a significant association between earnings management and financial themes, the results indicate the opposite. There was no significant correlation between earnings management and the financial dependent variable, suggesting that the presence of financial themes in earnings press releases is not influenced by the practice of earnings management. However, the study identifies a significant positive correlation between earnings management and environmental theme. This suggests that companies may engage in earnings management practices when addressing environmental issues in their earnings releases. This finding challenges the assumption that the use of earnings management is solely financially motivated. The study highlights the need for further research to better understand the complex relationship between the presence of environmental topics and the motivations behind earnings management practices.

Finally, this study shows a negative correlation between the use of earnings management and the diversity of topics covered in earnings press releases. This suggests that companies that use earnings management techniques are more likely to disclose less in-

formation and cover a narrower range of topics. These findings provide insight into the disclosure strategies of companies that practice earnings management. The study suggests that companies may engage in earnings management to avoid disclosing information that could cast doubt on their financial health, raise suspicion, or compromise their ability to monitor market expectations. By reducing their disclosure, companies can mitigate the risk of revealing adverse information and maintain a certain level of opacity. These results underscore the importance of considering the selective nature of information disclosed in earnings releases and provide insight into the communication choices made by firms involved in earnings management practices.

Overall, this study sheds light on the relationship between meeting or beating expectations, earnings management, and the thematic content of earnings releases. The results provide insight into the factors that influence the choice of themes and disclosure strategies in corporate communications. By exploring these relationships, this research contributes to the existing literature on earnings press releases and provides valuable information to investors, analysts, and various stakeholders.

The rest of the paper is organized as follows. The first section introduced our research question. The second section reviews the literature regarding the earnings press, the concept of meeting or beating expectations and the earnings management as well as a short description of the research question. The third section will be a presentation of our hypothesis. The fourth section highlights the data used and the methodological approach. The fifth section will be an empirical analysis. This paper will be closed with a short conclusion with a critical view.

2 Literature review

2.1 Earnings Press Release

Earnings press releases present a company's performance in the form of an official public statement. Concretely, their purpose is to communicate information to investors, analysts or other stakeholders in a narrative and numerical way. Earnings press releases contain the information that corresponds to a given period (every quarter in our case). This information typically includes an introductory paragraph of some information followed by a statement from managers. In addition, it contains the key elements that investors need such as net income, sales, earnings per share, dividends and cash flow. They are also part of a company's overall disclosure strategy. In a certain way, it means that third party perceptions of these documents can be biased by managers since the content of press releases is discretionary. Indeed, this gives them the opportunity to present the information in the way they want.

We will examine the different aspects of earnings press releases, as well as their usefulness for investors, their impact on financial markets and the different practices used by companies in their communication of these reports. Earnings press releases are also regularly subject to controversy and criticism by financial regulators.

Financial analysts and investors consider earnings press releases to be an essential source for evaluating a company's financial performance. In addition, according to a study conducted by [Graham & Harvey \(2001\)](#), investors place a high value on earnings press releases when making investment decisions, and management's comments on the company's results are considered particularly useful. In addition, earnings press releases are often used to assess how well a company is being managed. There is a real interest for companies to provide complete and transparent financial information. Indeed, according to a research done by [Christensen et al. \(2010\)](#), Companies that have more transparent financial disclosure practices have better financial performance than those that disclose by omitting information.

Earnings press releases also play a major role in a company's stock price. A study carried out by [Francis & Soffer \(1997\)](#) showed that earnings press releases have a greater impact on the financial markets than annual reports or other types of financial communications. Companies publishing earnings press releases with positive financial results observe an increase in their share price. In contrast, companies releasing negative or below expected results see a decline in their stock price. For example, a study carried out by [Johnson & Schwartz Jr \(2005\)](#) showed that companies that do not meet investors' expectations in their earnings press release can suffer a drop in their share price of up to 10%.

Despite this, studies have shown that the quantitative information that was provided to stakeholders was an incomplete picture of the real business' performance ([Boudt, Thewissen, & Torsin, 2018](#); [Davis & Tama-Sweet, 2012](#); [Henry, 2008](#)). For this reason, it is important to provide additional qualitative information in order to provide other fundamental business-related information that cannot be understood through numbers. However, the tone of the narrative is tailored to the message the executive wants to deliver. Actually, the market responds positively to the tone set in the press releases in the three days around the publication date of these documents ([Davis & Tama-Sweet, 2012](#); [Henry, 2008](#)) but this trend is reversed over the long term ([Huang, Teoh, & Zhang, 2014](#)). The prospect theory can justify this approach since it explains that investors focus on gains rather than losses when the financial performance of a company is positive ([Henry, 2008](#)). Also, the tone of these press releases helps to understand the reaction of investors, even more so when information asymmetry is high. Clearly, the tone used in press releases attached to quantitative results is a relevant market signal to be taken into account. Furthermore, the tone and topics used in press releases are related, as we note that positive words are used before important events in order to boost the share price ([Huang, Teoh, & Zhang, 2014](#)). These elements therefore become strategic tools to modulate the information provided.

2.2 Meet or Beat Expectations

Meeting or beating expectations is a key concept in our research. We will examine the different definitions and perspectives of this term, as well as its importance for companies. Meet or beat analysts' forecasts of earnings is a notion well entrenched in today's corporate culture (Bartov, Givoly, & Hayn, 2002). From corporate boards' deliberations to financial press reports and Internet chats, emphasis is placed on whether a company meets its earnings forecasts (Bartov, Givoly, & Hayn, 2002). According to a research by Bouri et al. (2017), financial analysts often set forecasts for companies using historical company data and market trends. However, these forecasts can sometimes be inaccurate, which can result in the company outperforming or underperforming expectations.

Anecdotal evidence indicates that there are strong incentives to achieve analysts' earnings forecasts to protect a company's stock price (Payne & Robb, 2000). Studies have found that firms experience significant gains in stock price when they MBE and significant declines in stock price when they do not (Bartov, Givoly, & Hayn, 2002). In addition, a research by Call et al. (2013) shows that firms that consistently meet or exceed financial analysts' expectations have lower financing costs and are more likely to receive investment from institutional investors. This can also lead to a better credit rating for the company. Numerous examples in the stock markets have also shown the impact of disappointing results on the price of a stock. There are examples every day of stocks that go down 25% to 50% because they missed analysts' earnings estimates by a few cents (Lee, Yu, & Zhang, 2007). Other reasons, such as managers' compensation contracting, reputation with stakeholders, and career concerns, also provide incentives for managers to report higher earnings that meet or exceed analyst forecasts (Graham et al., 2005; Healy & Wahlen, 1999).

Therefore, both the apparent benefits to meet or beat analyst earnings forecasts and potential significant costs of missing analysts' expectations provide a rationale for why managers are well motivated to manage earnings upward to exceed the threshold: analysts' earnings expectations (Bhojraj et al., 2009). In this case, managers use judgment

in financial reporting and in structuring transactions to alter financial reports, either to mislead some stakeholders about the underlying economic performance of the firm or to influence contractual outcomes that depend on reported accounting figures. The way managers manage the company's earnings is complex and is done in a variety of different ways (playing with accruals, managing actual earnings, etc.). Nonetheless, a research by [Hribar & Yang \(2016\)](#) shows that companies seeking to meet analysts' expectations may be tempted to manipulate their financial results to meet expectations, which can ultimately be detrimental to the company in the long term.

In some particular cases, some companies have an interest in not meeting or beating expectations. For example, when union organizations are present in companies, they may tend to underperform. Indeed, if results turn out to be positive, unions may claim better working conditions, better wages or even higher bonuses. Some studies even show that unionized firms are more likely to narrowly miss earnings forecasts than non-unionized control groups ([Bova, 2013](#)).

There is therefore significant evidence that meet or beat expectations play a key role in the perception of the company by its stakeholders. In our paper, we will focus on companies that do not meet their expectations and find out which topics are discussed in the earnings press releases of underperforming companies.

2.2.1 Topic Diversity in Meet or Beat Expectations

In addition to analyzing the thematic content present in earnings press releases, we then measured the relationship between companies' meet or beat expectations and the diversity of thematic content. It is crucial for a company to alert investors when it meets or beats the expectations of the financial markets. Earnings press releases are frequently used for this purpose. These reports can cover a wide range of topics, from the company's financial performance to recent changes in its products or markets. A positive ratio of topic diversity indicates that the earnings press releases of companies that meet or beat expectations contain a wide variety of information. It also means that these companies

tend to cover more topics and therefore provide more information to stakeholders. This allows them to get a better idea of the company and its activities. In contrast, when the diversity of topics is negative, the company tends to provide less information.

Prior studies have shown that the diversity of topics in earnings reports can be influenced by several factors and that the company's willingness to highlight its accomplishments is a major motivation. [DeFond & Hung \(2003\)](#); [Dichev et al. \(2013\)](#) have shown that companies that beat analysts' expectations tend to highlight specific achievements, such as increases in sales or market share gains. As well, [Han & Tan \(2010\)](#) also made this finding, suggesting that these companies seek to highlight their positive performance and enhance their reputation with investors. These results indicate that companies try to influence investors' perceptions through reports by trying to convince them through achievements. Other similar studies have shown that investors are more likely to buy shares in these companies as a result of positive information.

Other factors such as the tone of press releases, governance, gender of the CEO, institutional ownership, quality of the auditor, and conservatism in financial reporting can influence the diversity of topics in earnings press. One factor of particular interest to us is conservatism in financial reporting. In their study, [Rogers et al. \(2011\)](#) found that the degree of conservatism in financial reporting is lower for firms that beat expectations than for those that meet or miss them. [Nollet et al. \(2016\)](#) study confirms this result by finding that companies that beat expectations more often communicated on non-financial topics such as corporate social responsibility and environmental sustainability. While companies that missed expectations tend to communicate on more traditional financial topics such as earnings. We will see later if our study confirms these results.

2.3 Earnings Management

Earnings management is also a key concept in our research. Indeed, this concept is closely linked to the notion of meet or beat expectations. As mentioned earlier, given the importance for companies to meet or beat market expectations, companies can use earnings

management to positively influence their financial results. There are other reasons for companies to use this concept. Many different methods are also used to implement it. However, this practice is not favorably perceived by financial institutions. Companies practicing earnings management are subject to sanctions and significant negative consequences.

Earnings management is the choice of a manager of accounting policies or other actions, including voluntary earnings forecasting, voluntary disclosure, and estimation of accruals, to affect earnings intentionally (Man, Wong, et al., 2013). When a company practices earnings management, it discredits the financial information it provides to its stakeholders and market players. Moreover, the information included in financial statements is a crucial resource for investors. Indeed, in order for them to make the best investment decision, it is necessary to get the right and relevant information. Most prior studies have identified two main methods in earnings management, accrual management and manipulation of actual economic activities. Other studies have found that there are other more complex methods. A study by Roychowdhury (2006) shows that firms can also use more sophisticated earnings management methods, such as the use of stock options and forward contracts to alter financial results. Firms can also manipulate cash flows to influence financial results.

There are several explanations for the motivation of companies to undertake this practice. The first is the need to satisfy the expectations of the stock market players. A study by Graham et al. (2005) shows that companies that publish financial results in line with market expectations benefit from a positive valuation of their shares. Another important motivation is when executive compensation is linked to the company's financial results. In cases where managers are paid on the basis of performance, the incentive to practice management becomes even more likely to be used. Performance-based compensation schemes can provide incentives for managers to make accounting decisions that increase profits in a favorable way (Healy, 1985). Another important aspect of executive compensation is when executives are compensated through company stock. A study by

McVay et al. (2006) finds that in this case, executives are more likely to manage earnings to meet analysts' forecasts before selling their shares. This earnings management can also be explained for tax purposes. For example, when corporate tax rates are expected to be higher, the future tax deduction for deferred compensation may be favorable compared to the immediate tax deduction for cash compensation (Guay, Core, & Larcker, 2002). Therefore, the use of stock-based compensation should be less costly for firms with low marginal tax rates (Guay, Core, & Larcker, 2002). Indeed, studies by P. M. Dechow et al. (1996) and Bryan et al. (2000) find that stock option use is higher for firms with low marginal tax rates.

Earnings management can have very harmful consequences for the reputation of firms and managers who practice it. The use of this method is also inconsistent with the desire of financial institutions to make information transparent and relevant. Companies are therefore also exposed to legal proceedings. A study by Healy & Wahlen (1999) shows that companies that practice earnings management can suffer negative long-term consequences, such as higher borrowing costs and reduced investor confidence.

Although earnings management can have significant damage to the company, the practice remains common because the stakes and the need to meet forecasts are high. The motivations and methods used are multiple and some may include financial manipulation or unethical behavior.

2.3.1 Topic Diversity in Earnings Management

In addition to analyzing the thematic content of earnings presses of companies that practice earnings management, we will analyze the topic diversity. Earnings management companies may choose to focus on different topics in their earnings press releases in order to present their financial results in a positive way. As mentioned earlier, positive topic diversity indicates that earnings management companies' earnings press releases contain a wide variety of information. If the relationship is negative, the topic diversity is limited.

Previous studies have looked at the variety of subjects covered in the earnings press

releases of companies using earnings management. These studies have examined whether companies that practice earnings management cover a wider range of topics than companies that do not. Several studies have found that companies practicing earnings management use earnings press releases to communicate on a wider range of topics (Maglio, Rey, Agliata, & Lombardi, 2020; Kim, Li, & Li, 2014)

Earnings management firms have different motivations for communicating a wide range of topics in their earnings press releases. First, there is a motivation to hide variations in their financial results by using the topic diversity to mask earnings management practices. The aim is to avoid suspicion from stakeholders and to give an image of transparency in their published information. Another motivation is to drown compromising information in a flood of information. When reports contain a lot of information, it is more difficult for stakeholders to detect evidence of earnings management practices.

Therefore, topic diversity of earnings press releases of firms practicing earnings management has been studied by a number of researchers. The literature suggests that companies use earnings press releases to manipulate information and distract investors from potential negative news.

3 Research Question

In the context of our research question, we found it important to develop a descriptive approach around the key concepts related to our research question: Earnings press release, earnings management and meet or beat expectations. Earnings press releases and earnings management are the triggers for MBE. Depending on their content, stakeholders will form their opinion on the company's situation and the market will take this information into account. A study by Zhang et al. (2019) found that corporate financial results press releases that contain positive information tend to elicit stronger reactions from investors, while press releases that contain negative information tend to be ignored or downplayed. In other words, investors react more strongly to positive news than negative news, which may reflect their overall sentiment toward the company. Nowadays, there is an abun-

dance of information circulating non-stop, so the market may have anticipated sentiment at times.

Studies already conducted focus on tone or readability to assess whether companies will meet or beat expectations. [Huang et al. \(2014\)](#) found that companies can manipulate the tone of press releases to influence investor and financial analyst reaction. The results suggest that companies can influence investors' perceptions of their financial performance by using tone management techniques, such as using emotional language and structuring press releases to highlight positive aspects. Another study by [Arslan-Ayaydin et al. \(2016\)](#) revealed that executives with equity incentives tend to issue press releases with a more positive tone, even if the financial results are mitigated. These studies underestimate the importance of the context and type of information available in publications.

Our paper will therefore take into account not only the amount of information offered but also the type of information present in earnings press. In other words, what topics are present when companies are below expectations and therefore what content their earnings press contains. Conversely, we will observe the difference in content with companies that meet or beat expectations. Concerning earnings management, we will analyze the representation of our topics in companies practicing earnings management when they meet or beat expectations. It will also be interesting to look at the significance of these topics in this scenario. This will allow, in addition, a comparison with the results obtained for companies that meet or do not meet or beat expectations. It is therefore very interesting to see which topics present in the earnings press can allow us to understand its effects in a meaningful way. Depending on the topics in the earnings press of the companies, stakeholders could get a better idea of what might happen to stock prices in the future and then make better decisions.

To achieve this, we will use a holistic and replicable method, namely topic modeling, which identifies the topics covered without imposing restrictions based on what we expect to find ex ante. This theme modeling allows us to highlight the non-technical topics and to illustrate the great heterogeneity of the topics covered in the earnings press. We

will therefore model topics to summarize the content of the earnings press into individual topics (Thewissen, Shrestha, Torsin, & Pastwa, 2021). Then, we will test the impact of the topics on the influence of the earnings press on whether or not expectations are met. Finally, we will test the impact of the topics on the influence of earnings management.

3.1 Hypothesis

3.1.1 *Meet or Beat Expectations Hypothesis*

The first 3 hypothesis focus on the meet or beat expectations variable. The objective is to investigate the differences in the thematic content of earnings press releases for companies that meet or beat expectations and those that do not meet or beat expectations.

As discussed in the literature review, prior research has suggested that earnings press releases are often used to influence investor and analyst perceptions, and that companies often use specific terms to present their results favorably. Thus, there may be a significant difference in the topics used in companies' earnings press releases depending on their financial performance. Knowing this, we decided to study 3 hypothesis concerning this aspect.

- ***Hypothesis 1:*** The topics of the companies that MBE are more financial while the topics of the companies that don't MBE are less financial? (Difference in the topics discussed, technical topics on one side, general on the other)
- ***Hypothesis 2:*** Are the topics that are significant when it is MBE always the same as when it is not MBE?
- ***Hypothesis 3:*** Companies that meet or beat expectations tend to cover more topics in their earnings press releases?

3.1.2 *Earnings Management Hypothesis*

The other two hypotheses focus on the earnings management variable. The objective is to discover the differences in the thematic content of earnings press releases for companies

that practice earnings management.

Hypothesis 5 suggests that firms that use earnings management tend to cover more topics in their earnings press releases. Earnings management firms are likely to want to minimize exposure to analysts and investors to avoid raising suspicion. This could result in less transparent and therefore less informative press releases. The literature has shown that companies that use earnings management tend to provide more information in their earnings press releases than companies that do not use it in order to make the earnings manipulation technique less detectable.

- **Hypothesis 4:** The topics of companies that use EM are less financial while the topics of companies that do not use EM are more financial? (Difference in the topics covered, general topics on one side, technical on the other)
- **Hypothesis 5:** Companies that use earnings management tend to cover less topics in their earnings press releases?

4 Data & Sample selection

This section describes the methodology and variables used in our empirical test. We explain the method by which we were able to obtain and map the topics contained in earnings press. To do so, we collected several thousand earnings press releases from different companies available on the SEC website. Following this, we created a database of these earnings press releases to perform our analysis. Then, we detail the set of variables used to perform our model using a table.

4.1 Topic Modeling Implementation

We used the topic modeling method to discover the main topics present in the earnings press releases. These topics included a group of words that often appear together. In fact, this method helps to discover the hidden structure of a document collection (Bardes & Bainwad, 2017). The method of topic modeling is used to group together words that

appear in a similar context and conversely to separate words that are not related to each other.

In the context of our research, we used this method to identify the topics that regularly appeared in earnings press releases when companies' results were below expectations. These expectations are determined by analysts.

We performed an empirical analysis using a Latent Dirichlet Allocation (LDA) tool in the R studio program. This technique was applied to textual documents such as earnings press releases to deduce the distribution of topics. These topics were represented by a distribution of words. Thanks to the words that emerged with LDA, we were able to highlight semantic clusters that we used to determine the topics that came up the most in the earnings press releases. The analysis of financial earnings press releases is an important field of research for financial analysts and investors, as these releases contain key information about the performance of companies. Several studies have applied LDA to the analysis of financial earnings press releases to identify key topics discussed in these documents. Two studies conducted by [Schumaker & Chen \(2009b\)](#) and [Schumaker & Chen \(2009a\)](#) applied the LDA method to financial earnings press releases to identify the dominant topics discussed in the press releases. The results showed that the most prevalent topics were future earnings forecasts, economic trends, and product and market information. These results showed that the dominant topics identified by LDA were significant predictors of future stock returns. A more recent study by [Aziz et al. \(2022\)](#) applied the topic modeling approach to financial data from US companies and identified topics such as financial performance, investor relations, competition and corporate governance. The authors found that topic modeling can help investors and financial analysts better understand financial data and make more accurate decisions. The topic modeling equation is as follows:

$$\begin{aligned}
p(\theta, z, w|\alpha, \beta) &= \prod_{d=1}^D p(\theta_d|\alpha) \prod_{n=1}^{N_d} p(z_{d,n}|\theta_d) p(w_{d,n}|z_{d,n}, \beta) \\
&= \prod_{d=1}^D \frac{\Gamma(\sum_{k=1}^K \alpha_k)}{\prod_{k=1}^K \Gamma(n_{d,k} + \alpha_k)} \prod_{k=1}^K \theta_{d,k}^{n_{d,k} + \alpha_k - 1} \prod_{n=1}^{N_d} p(z_{d,n}|\theta_d) \prod_{n=1}^{N_d} p(w_{d,n}|z_{d,n}, \beta) \quad (1) \\
&= \prod_{d=1}^D \frac{\Gamma(\sum_{k=1}^K \alpha_k)}{\prod_{k=1}^K \Gamma(n_{d,k} + \alpha_k)} \prod_{k=1}^K \theta_{d,k}^{n_{d,k} + \alpha_k - 1} \prod_{n=1}^{N_d} \prod_{k=1}^K (\phi_{k,w_{d,n}})^{z_{d,n,k}}
\end{aligned}$$

where θ_d is the probability distribution of topics for document d , $z_{d,n}$ is the topic assigned to each word $w_{d,n}$ in document d , $\phi_{k,w}$ is the probability distribution of words for topic k , α is the Dirichlet concentration parameter for the topic probability distribution, and β is the Dirichlet concentration parameter for the word probability distribution (Blei, Ng, & Jordan, 2003). The notation $\Gamma(\cdot)$ is the gamma function.

On the practical side, it is not an easy task to arrive at a set of words that make sense all together. It is essential to run the code many times to clean up the data with stopwords. Knowing that the more stopwords there are, the more precise the code will be and the more topics will be proposed with a coherent set of words. Choosing which words to remove is also a complicated and decisive task. Indeed, at first sight, some words or abbreviations may have no meaning while on the contrary they may be very relevant. When we obtain a set of coherent words, we have to label the topics and group some of them under the same cluster. This is a crucial step because we must be sure that these words are used in the same sense. If this is not the case and we group them together, the results could be biased.

To study the informational value of the thematic composition of the earnings press releases, we use the K topics obtained from the LDA model as input variables in a multifactor model. Specifically, we examine the relationship between the thematic content of the earnings press releases of several companies and whether they meet or beat the expectations.

4.2 Statistic Descriptive

The [Table 1](#) represents and explains the six clusters composed of specific topics. This table also presents summary statistics for the various explicative and control variables that will be used in the regressions. For ease of interpretation, we have manually grouped the 20 subjects into 6 categories, representing distinct themes.

Table 1: **Variable Description**

Variable	Description	Min	1st Qu.	Median	Mean	3rd Qu.	Max
<i>Dependent Variable</i>							
Environmental	Proportion of words that include topics about environmental issues in earnings press releases	0.00004	0.00159	0.00430	0.15987	0.13338	0.99873
Research&Development	Proportion of words that include topics about research&development in earnings press releases	0.00001	0.00077	0.00250	0.14362	0.11575	0.99816
Social	Proportion of words that include topics about social in earnings press releases	0.00001	0.00066	0.00132	0.009205	0.00530	0.99837
Financial	Proportion of words that include topics about monetary matters in earnings press releases	0.00001	0.00203	0.01645	0.26567	0.57427	0.99796
Entertainment	Proportion of words that include topics about entertainment in earnings press releases	0.00000	0.00035	0.00069	0.05385	0.00290	0.99108
Property	Proportion of words that include topics about property in earnings press releases	0.00000	0.00033	0.00062	0.04771	0.00142	0.99696
<i>Explicative Variable</i>							
MBE	Binary variable that takes the value of 1 if the company MBE and 0 if the company doesn't MBE	0.0000	0.0000	1.0000	0.6786	1.0000	1.0000
EM	Binary variable that takes the value of 1 if the company applies EM and 0 if the company doesn't apply EM	0.0000	0.0000	0.0000	0.48	1.0000	1.0000
Loss	Binary variable that takes the value of 1 of the ROA is negative and 0 if the ROA is positive	0.0000	0.0000	0.0000	0.1414	0.0000	1.0000
<i>Control Variable</i>							
Return	Return of the company	0.008294	0.054965	0.077163	0.092977	0.111470	1.180068
EARNTA	Pre-tax earnings	-1.371666	0.002023	0.005851	0.007754	0.018042	0.674329
Lev	Leverage	0.0000	0.0360	0.1291	0.1884	0.2794	3.9048
Sd	Standard Deviation that measures how dispersed the data is in relation to the mean	0.000002	0.000834	0.003184	0.009045	0.008409	1.400647
Errormedian	Allows to interpret the binary variable, if it is greater than 0 then the company MBE	-8928.571	0.000	0.001	-0.157	0.002	2000.000
LNoA	Log of Net operating Asset related to its activities, less any liabilities directly related to its activities	0.000	1.386	2.079	1.945	2.708	3.871
Lmarketsize	Log of Market capitalisation of the company	1.389	6.577	7.927	8.086	9.495	14.619
Sic	Standard Industrial Classification code of the companies	100	3674	6020	5211	6282	9997
Fyearq	Fiscal Year	2004	2009	2013	2013	2017	2021
Fqtr	Fiscal Quarter	1.000	1.000	2.000	2.331	3.000	4.000

5 Empirical Analysis

This section provides the results addressing our research questions. We proceed in three steps. Before analyzing the results, we first discuss the evaluation of the topic modeling approach. We then provide details on the topics contained in the earnings press releases. Finally, we test which thematic content of press releases is most present in earnings press releases using a linear regression model.

$$\begin{aligned}
 Lm(Topic) = & \beta_0 + \beta_1 MBE + \beta_2 Loss + \beta_3 Return + \beta_4 Lev + \beta_5 Sd \\
 & + \beta_6 Return + \beta_7 Errormedian + \beta_8 \log(NoA) \\
 & + \beta_9 \log(Marketsize) + \beta_{10} Sic + \beta_{11} Fyearq + \beta_{12} Fqtr + \epsilon
 \end{aligned} \tag{2}$$

$$\begin{aligned}
 Lm(Topic) = & \beta_0 + \beta_1 EM + \beta_2 Loss + \beta_3 Return + \beta_4 Lev + \beta_5 Sd \\
 & + \beta_6 Return + \beta_7 Errormedian + \beta_8 \log(NoA) \\
 & + \beta_9 \log(Marketsize) + \beta_{10} Sic + \beta_{11} Fyearq + \beta_{12} Fqtr + \epsilon
 \end{aligned} \tag{3}$$

In both equations, $Lm(Topic)$ represents the thematic content of earnings press releases, and the other variables are financial and non-financial control variables that could influence the thematic content. The coefficients β_i measure the effect of each variable on the thematic content, and ϵ is the error term.

In addition to analyzing the topics contained in the earnings press releases, we also tested the diversity of the topics present in our sample using the Shannon index. The Shannon index is a widely used measure of diversity that takes into account both the number of different topics present in a document and their relative frequencies. The ratio of each species in the sample's logarithm is used to calculate the Shannon index (Shannon, 1948).

$$TopicDiversity = - \sum_{k=1}^{20} p_k \ln(p_k) \tag{4}$$

Where $TopicDiversity$ represents the Shannon index, 20 is the total number of topics in the sample, p_k represents the proportion of the k topics in the sample, and \ln represents the natural logarithm function.

5.1 Meet or Beat Expectations Interpretation

5.1.1 Topics in Meet or Beat Expectations

Table 2: MBE according to the topics

<i>Dependent variable:</i>							
	<i>Panel A: Topics</i>					<i>Panel B: ShannonIndex</i>	
	Environmental	R&D	Social	Financial	Entertainment	Property	TopicDiversity
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
MBE	-0.011*** (0.002)	0.004* (0.002)	0.008*** (0.002)	-0.012*** (0.003)	-0.007*** (0.001)	-0.003* (0.002)	0.009** (0.004)
Loss	-0.029*** (0.004)	0.013*** (0.004)	0.038*** (0.003)	-0.077*** (0.004)	0.003 (0.002)	0.027*** (0.002)	0.038*** (0.006)
Return	-0.006 (0.022)	-0.116*** (0.021)	-0.092** (0.019)	0.416*** (0.025)	0.124*** (0.012)	-0.044*** (0.014)	-0.329*** (0.035)
EARNTA	0.220*** (0.038)	0.702*** (0.037)	-0.760*** (0.032)	-0.765*** (0.042)	0.102*** (0.021)	0.225*** (0.024)	0.799*** (0.060)
Lev	-0.018*** (0.006)	-0.127*** (0.005)	-0.019*** (0.005)	-0.270*** (0.006)	0.154*** (0.003)	0.240*** (0.004)	0.195*** (0.009)
Sd	-0.035 (0.050)	0.263*** (0.048)	0.518*** (0.042)	-0.931*** (0.056)	-0.043 (0.027)	0.173*** (0.032)	0.419*** (0.079)
Errormedian	0.00001 (0.00003)	0.00000 (0.00003)	-0.00003 (0.00002)	0.0001** (0.00003)	-0.00000 (0.00001)	-0.00003* (0.00002)	-0.00000 (0.00004)
log(NoA)	0.014*** (0.002)	0.001 (0.002)	0.003* (0.002)	0.061*** (0.002)	-0.001 (0.001)	-0.086*** (0.001)	-0.008** (0.003)
log(marketsize)	-0.004*** (0.001)	-0.004*** (0.001)	0.006*** (0.001)	-0.022*** (0.001)	-0.008*** (0.0005)	0.026*** (0.001)	-0.015*** (0.001)
Factor(sic, 1, 1)	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Factor(fyearq)	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Factor(fqtr)	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	54,522	54,543	54,526	54,522	54,977	54,790	54,515
R ²	0.249	0.195	0.216	0.449	0.156	0.220	0.095
Adjusted R ²	0.248	0.195	0.215	0.448	0.156	0.219	0.095
Residual Std. Error	0.255 (df = 54484)	0.248 (df = 54505)	0.216 (df = 54488)	0.283 (df = 54484)	0.139 (df = 54939)	0.162 (df = 54752)	0.406 (df = 54477)
F Statistic	487.089*** (df = 37; 54484)	357.165*** (df = 37; 54505)	405.292*** (df = 37; 54488)	1,199.147*** (df = 37; 54484)	275.060*** (df = 37; 54939)	417.146*** (df = 37; 54752)	154.858*** (df = 37; 54477)

Note:

*p<0.1; **p<0.05; ***p<0.01

Concerning the regression of the **Environmental** topic on the explanatory variables, we can see that the explanatory variable MBE is negatively correlated with the dependent variable Environmental. This means that the presence of the topic Environmental in the earnings press releases tends to be below the analysts' forecasts. Indeed, it is -0.011 which means that the topics reflecting the terms Environmental in the earnings press releases will be less likely to meet the analysts' forecasts if the binary explanatory variable MBE increases. The explanatory variable MBE is highly significant at the 99% level. On

the other hand, when we focus on the explanatory variable Loss, we see that it is also negatively correlated with the dependent variable Environmental. This illustrates that environmental-related topics tend to be present in the earnings press releases of companies that do not make losses. Its value of -0.029 means that the presence of an additional proportion of the explanatory variable Loss will decrease the dependent variable Environmental by 0.029. The explanatory variable Loss is highly significant at the 99% level.

Concerning the regression of the **Research&Development** topic, we can see that the explanatory variable MBE of the regression is weakly significant at a threshold of 90%. This MBE variable is positively correlated with the dependent variable R&D. This means that the increase of the explanatory variable MBE will have a positive impact on the terms reflecting the R&D topic in the earnings press releases and therefore this topic will tend not to be below the analysts' expectations. Indeed, the presence of an additional proportion of the explanatory variable MBE will increase the 0.004 probability of this topic meeting or beating expectations. The second explanatory variable, Loss, which aims to announce the presence of a gain or loss, is highly significant at the 99% threshold. Indeed, the dependent variable R&D and the variable Loss are positively correlated, which means that the presence of topics related to R&D in earnings press releases reflects a loss. We note that the presence of these topics could significantly increase the probability of loss for the company by 0.013.

Regarding the **Social** topic, we can see in the table above that the explanatory variable MBE is positively correlated with the dependent variable Social, which means that the presence of health-related terms makes it easier to meet or exceed analysts' expectations. Indeed, the explanatory variable MBE is 0.008 and is highly significant at a 99% threshold. Secondly, the explanatory variable Loss shows a positive value of 0.038 which means that it is positively correlated with the dependent variable Social. In other words, the topics related to Social tend to be present in those earnings press releases of companies that make losses. Indeed, its value is 0.038 so the presence of an additional proportion of the explanatory variable Loss will increase the dependent variable Social by

0.038. This explanatory variable is highly significant at the 99% level.

In the column concerning the topic **Financial**, we can see that the explanatory variable MBE of this LM regression is negatively correlated with the dependent variable Financial. Indeed, the explanatory variable MBE is -0.0012, which means that increasing the proportion of the variable MBE will decrease the probability that the forecasts are met or beaten in the Financial topics present in the earnings press releases. This explanatory variable MBE in this model is highly significant at the 99% level. This means that when companies meet or beat expectations, they focus less on financial topics. On the contrary, they focus more on other information. It also means that investors are looking for information about more than just financial topics. On the other hand, we see that the explanatory variable Loss is also negatively correlated with the dependent variable Financial. This means that topics related to Financial tend to be present in the earnings press releases of companies that do not make losses. Its value is -0.077 which means that the presence of an additional proportion of the explanatory variable Loss will decrease the dependent variable Financial by 0.077. The explanatory variable Loss is highly significant at the 99% level.

When looking at the column dealing with the **Entertainment** topic, we can see that the explanatory variable MBE is highly significant at the 99% level. We can also see that it is negatively correlated with the dependent variable Entertainment, which means that increasing the explanatory variable MBE will decrease the probability that the Entertainment topic in the earnings press releases will meet the analysts' expectations. Indeed, the explanatory variable Meet or Beat the Expectations in this model is -0.007. The second explanatory variable Loss is positively correlated with the dependent variable Entertainment, which reflects losses reported in earnings press releases when topics related to this variable are present. However, given the insignificance of this variable, it may be preferable not to take this interpretation into account.

Concerning the regression using the LM model of the **Property** topic, we can see that the explanatory variable is negatively correlated with the dependent variable Prop-

erty. This means that increasing the explanatory variable MBE will increase the probability of being below analysts' forecasts for the terms present in the earnings press releases. Indeed, the explanatory variable MBE in this model is -0.003 and is weakly significant at the 90% level. Second, the explanatory variable Loss and the topic Property are positively correlated. In other words, topics related to Property tend to be present in the earnings press releases of loss-making firms. In fact, it is 0.027, which means that the presence of an additional proportion of the explanatory variable Loss will increase the dependent variable Property by 0.027. The explanatory variable Loss is highly significant at the 99% level.

Following the linear regression tests carried out, we can now test the veracity of our 2 hypotheses. In order to test our first hypothesis, we performed two linear regressions. We have thus defined 2 dependent variables in order to verify our hypothesis: **Financial and Property**. These different regressions allowed us to regress the independent variable MBE on the two dependent variables.

5.1.2 *Topic Diversity in Meet or Beat Expectations*

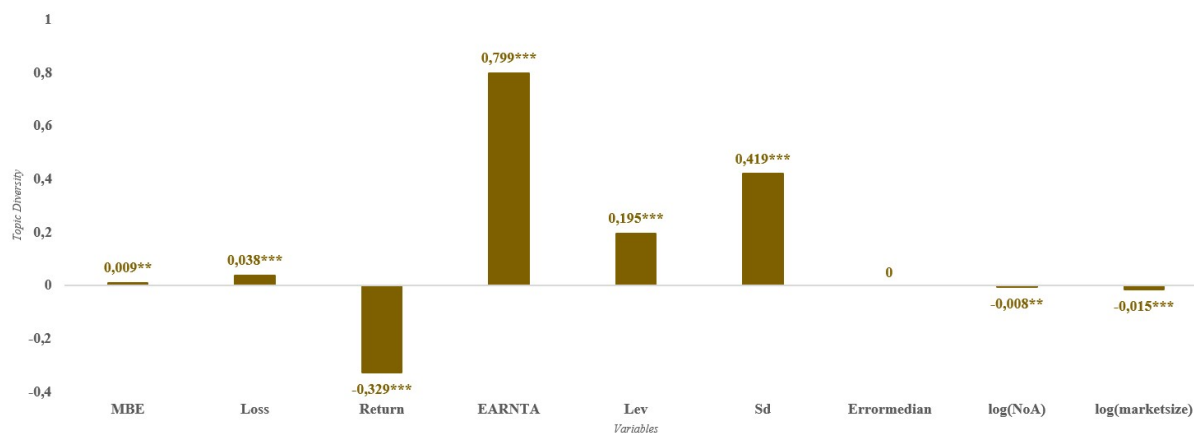


Figure 1: Topic Diversity in MBE

About the Topic Diversity on the explanatory variables, **Figure 1** shows that the MBE variable is positively correlated with the dependent variable TopicDiversity. This means that when companies meet or beat expectations, it is more likely to find a larger diversity

in the thematic content addressed. This variable is highly significant at the 99% level.

Looking at the explanatory variable Loss, we find that it is also positively correlated at 99% with the dependent variable TopicDiversity. This indicates to us that topic diversity in thematic content tends to be observed more often in companies incurring losses. Moreover, we can also notice that the variable Return is negatively correlated at 99% with the dependent variable TopicDiversity. This means that TopicDiversity is less likely to be observed for companies with a positive return.

5.2 Earnings Management Interpretation

5.2.1 Topics in Earnings Management

Table 3: EM according to the topics

	Dependent variable:						
	Environmental	R&D	Social	Financial	Entertainment	Property	TopicDiversity
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
EM	0.019*** (0.005)	-0.005 (0.005)	-0.005 (0.004)	-0.002 (0.002)	-0.003 (0.002)	-0.001 (0.001)	-0.020*** (0.006)
loss	-0.057*** (0.008)	-0.007 (0.008)	0.037*** (0.007)	0.017*** (0.004)	-0.001 (0.004)	0.001 (0.002)	-0.007 (0.010)
return	-0.066 (0.047)	-0.044 (0.047)	-0.055 (0.039)	0.106*** (0.021)	0.102*** (0.023)	0.021* (0.012)	-0.211*** (0.059)
EARNTA	0.065 (0.073)	0.535*** (0.073)	-1.039*** (0.060)	0.245*** (0.032)	0.145*** (0.036)	0.060*** (0.019)	0.375*** (0.092)
lev	0.021* (0.011)	-0.188*** (0.012)	-0.049*** (0.009)	-0.036*** (0.005)	0.187*** (0.006)	-0.013*** (0.003)	0.086*** (0.014)
sd	-0.176* (0.099)	0.186* (0.100)	0.085 (0.082)	0.202*** (0.044)	0.018 (0.049)	0.122*** (0.026)	0.178 (0.124)
errormedian	-0.00000 (0.00003)	0.00001 (0.00003)	0.00002 (0.00003)	-0.00001 (0.00002)	-0.00001 (0.00002)	-0.00000 (0.00001)	-0.00000 (0.00004)
log(NoA)	-0.010** (0.004)	0.017*** (0.004)	-0.005 (0.003)	0.018*** (0.002)	0.001 (0.002)	0.004*** (0.001)	-0.045*** (0.005)
log(marketsize)	-0.004* (0.002)	-0.008*** (0.002)	0.019*** (0.002)	-0.004*** (0.001)	-0.016*** (0.001)	-0.002*** (0.001)	-0.016*** (0.003)
Factor(sic, 1, 1)	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Factor(fyearq)	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Factor(lqtr)	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	17,858	17,858	17,855	17,852	17,946	17,882	17,851
R ²	0.220	0.098	0.212	0.338	0.147	0.056	0.076
Adjusted R ²	0.219	0.097	0.211	0.337	0.145	0.054	0.074
Residual Std. Error	0.311 (df = 17826)	0.314 (df = 17826)	0.258 (df = 17823)	0.139 (df = 17820)	0.156 (df = 17914)	0.083 (df = 17850)	0.392 (df = 17819)
F Statistic	162.482*** (df = 31; 17826)	62.686*** (df = 31; 17826)	154.899*** (df = 31; 17823)	293.497*** (df = 31; 17820)	99.359*** (df = 31; 17914)	33.857*** (df = 31; 17850)	47.298*** (df = 31; 17819)

Note:

*p<0.1; **p<0.05; ***p<0.01

Concerning the regression of our topics on the explanatory variable EM (Earnings management), we can observe that the **Environmental** topic is the only relevant topic.

We find that the explanatory variable EM is positively correlated with the dependent variable Environmental. This means that the presence of the Environmental topic in earnings releases tends to be associated with earnings management. Indeed, it is 0.019, which means that themes reflecting the terms Environmental in earnings press releases will be more likely to be in the presence of earnings management if the binary explanatory variable EM increases. The explanatory variable EM is highly significant at the 99% level. On the other hand, when we focus on the explanatory variable Loss, we find that it is negatively correlated with the dependent variable Environmental. This shows that Environmental topics tend to be present in the earnings releases of companies that do not incur losses. Its value of -0.057 means that the presence of an additional proportion of the Loss explanatory variable will decrease the dependent variable Environmental by 0.057. The explanatory variable Loss is highly significant at the 99% level.

5.2.2 *Topic Diversity in Earnings Management*

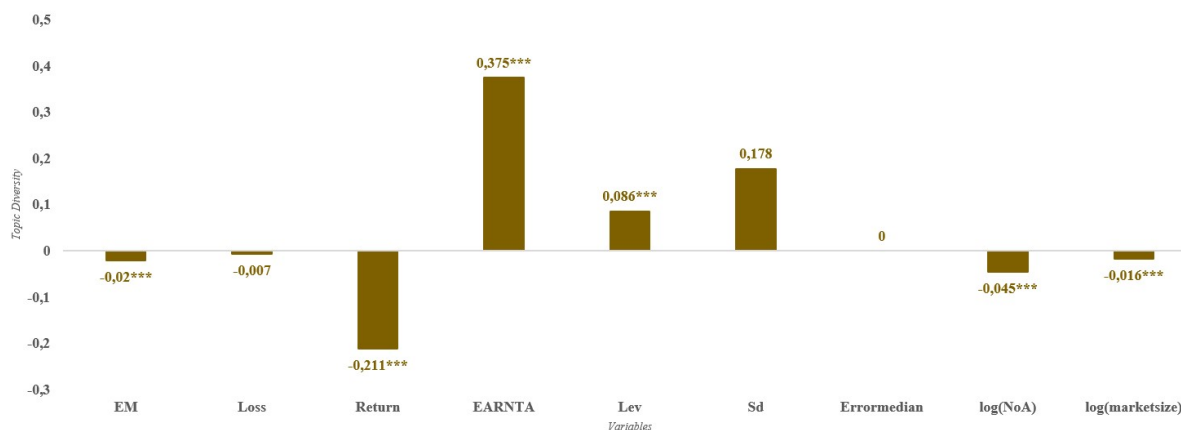


Figure 2: Topic Diversity in EM

About the Topic Diversity on the explanatory variables, **Figure 2** shows that the EM variable is negatively correlated with the dependent variable TopicDiversity. This means that when companies practice earnings management, it is less likely to find a larger diversity

in the thematic content addressed. This variable is highly significant at the 95% level.

When we look at the Return variable, we observe that it is 99% negatively correlated with the dependent variable TopicDiversity. This allows us to say that TopicDiversity is more likely to be observed for companies with a positive return.

5.3 Empirical Evaluation of Hypothesis

5.3.1 *Hypothesis 1 : The topics of the companies that MBE are more financial while the topics of the companies that don't MBE are less financial?*

This first model allows us to study the presence of Financial issues when a company meets or beats analysts' expectations. To do this, we regressed our binary variable MBE and the control variables on the dependent variable Financial. This represents the total proportion of words that deal with financial topics in earnings press releases.

$$\begin{aligned} Lm(\text{Financial}) = & \beta_0 + \beta_1 \text{ MBE} + \beta_2 \text{ Loss} + \beta_3 \text{ Return} + \beta_4 \text{ Lev} + \beta_5 \text{ Sd} \\ & + \beta_6 \text{ Return} + \beta_7 \text{ Errormedian} + \beta_8 \log(\text{NoA}) \\ & + \beta_9 \log(\text{Marketsize}) + \beta_{10} \text{ Sic} + \beta_{11} \text{ Fyearq} + \beta_{12} \text{ Fqtr} + \epsilon \end{aligned}$$

As we saw earlier, the MBE variable in this LM regression is negatively correlated with the dependent variable of Financial at a significant confidence level of 99%. Therefore, this infers our first hypothesis. Indeed, the regression showed that increasing the proportion of the variable MBE will decrease the probability that the forecasts are met or beaten in the Financial topics present in the earnings press releases.

The second regression allowed us to observe the relationship between the fact that earnings press releases are below investors' expectations and the presence of topics related to Property discussed in these earnings press releases. To do this, we kept the same control

variables as before and regressed them on the dependent variable Property.

$$\begin{aligned} Lm(Property) = & \beta_0 + \beta_1 MBE + \beta_2 Loss + \beta_3 Return + \beta_4 Lev + \beta_5 Sd \\ & + \beta_6 Return + \beta_7 Errormedian + \beta_8 \log(NoA) \\ & + \beta_9 \log(Marketsize) + \beta_{10} Sic + \beta_{11} Fyearq + \beta_{12} Fqtr + \epsilon \end{aligned}$$

As shown in the [Table 2](#), the explanatory variable MBE in this model is -0.003 and is weakly significant at a 90% confidence level. We can see that the explanatory variable is negatively correlated with the dependent variable Property. This means that increasing the explanatory variable MBE will increase the probability of being below analysts' forecasts. Again, this infers the first hypothesis. We can therefore conclude that subjects dealing with finance generally are not particularly more likely to meet or beat analysts' expectations. On the contrary, our analyses have shown that earnings press releases containing financial topics tend to be below expectations.

5.3.2 Hypothesis 2 : Are the topics that are significant when it is MBE always the same as when it is not MBE?

Thanks to our analysis, we are able to answer the second hypothesis we posed. Indeed, we notice that most of the topics present in the earnings press releases when companies do not meet expectations are significant at the 99% threshold. However, there is an exception for the Property topic which is significant at a threshold of only 90%. As for the topics found in the earnings press releases of companies that meet or beat the expectations, we find the topic Social which is significant at a 99% threshold. The R&D topic is also significant, but less so at the 90% level. Therefore, we can say that the topics found in earnings press releases that are significant when companies meet or beat the expectations are not the same as when companies do not meet or beat the expectations.

5.3.3 Hypothesis 3: Companies that meet or beat expectations tend to cover more topics in their earnings press releases?

Using our model with the Shannon index, we are able to address this third hypothesis. To do this, we first computed our shannon index, renamed TopicDiversity, and then regressed our binary variable MBE and the control variables on the dependent variable TopicDiversity.

$$TopicDiversity = - \sum_{k=1}^6 p_k \ln(p_k)$$

$$\begin{aligned} Lm(TopicDiversity) = & \beta_0 + \beta_1 MBE + \beta_2 Loss + \beta_3 Return + \beta_4 Lev + \beta_5 Sd \\ & + \beta_6 Return + \beta_7 Errormedian + \beta_8 \log(NoA) \\ & + \beta_9 \log(Marketsize) + \beta_{10} Sic + \beta_{11} Fyearq + \beta_{12} Fqtr + \epsilon \end{aligned}$$

As we observed in [Figure 1](#), the MBE variable is 95% positively correlated with the independent variable TopicDiversity. We can therefore say that when companies meet or beat expectations, it is more likely to find a larger diversity in the thematic content covered.

We can also reflect on this result and try to understand why companies would have this tendency. First, companies may want to emphasize their strong financial performance by explaining various aspects of their business in detail. This is especially true for companies that meet or beat investor and analyst expectations. Positive financial results can also attract media and investor attention, which can persuade companies to disclose more specific information about their performance. In addition, companies that meet or beat expectations may be seen as more trustworthy and transparent, which could strengthen investor confidence in the company. Therefore, broadening the range of topics covered in their earnings press releases may be a strategy to build investor confidence and support their long-term interest. With this in mind, the company can use the positive momentum of having met or beaten expectations to talk about its future targets and plans. In this case, press releases can cover a wider range of topics.

In conclusion, companies that meet or beat analyst expectations tend to cover more topics in their earnings press releases than companies that do not. They may do so for a number of reasons, including to improve their performance and credibility in the eyes of stakeholders, to provide detailed information about their strategy and performance, or to ride the positive wave. Companies generally want to increase the impact of their earnings press releases, which may explain the observed trend.

5.3.4 *Hypothesis 4: The topics of companies that use EM are less financial while the topics of companies that do not use EM are more financial?*

To solve this fourth hypothesis, we regressed our binary variable EM and the control variables on the different independent variables that are our topics.

$$\begin{aligned} Lm(Topic) = & \beta_0 + \beta_1 EM + \beta_2 Loss + \beta_3 Return + \beta_4 Lev + \beta_5 Sd \\ & + \beta_6 Return + \beta_7 Errormedian + \beta_8 \log(NoA) \\ & + \beta_9 \log(Marketsize) + \beta_{10} Sic + \beta_{11} Fyearq + \beta_{12} Fqtr + \epsilon \end{aligned}$$

As we saw in [Table 3](#), the explanatory variable EM is not significantly correlated with the dependent variable Financial. This means that the dependent variable does not have a significant effect on the explanatory variable we are trying to explain. In other words, there is not enough evidence to say that the dependent variable is significantly associated with the explanatory variable.

The explanatory variable EM is only significantly correlated with the dependent variable Environmental. The significativity is positive at the 99% level. This means that companies are more likely to practice earnings management when the thematic content of their earnings press releases is based on the environment. Although this is our only significant topic, it is relevant.

It is likely that some companies with adverse environmental impacts will try to improve their reputation by highlighting their environmental commitment, which may involve mentioning environmental issues in their earnings releases. These companies may

try to avoid the negative effects that poor financial results could have on their reputation and image by using earnings management.

However, the practice of earnings management is not unique to companies that include environmental topics in their earnings press releases. Regardless of the topics covered in their earnings releases, companies may have various motivations for manipulating their financial statements.

Moreover, even in the literature there is no clear link between the environmental topic in earnings press releases and the practice of earnings management. Some studies have found a positive correlation between the two, while others have found a negative correlation or no correlation at all. Nevertheless, according to a study conducted by [Brown et al. \(2015\)](#), companies that use earnings management tend to emphasize the non-financial aspects of their business in their earnings press releases.

5.3.5 *Hypothesis 5: Companies that use earnings management tend to cover less topics in their earnings press releases?*

To answer our final hypothesis, as in hypothesis 3, we first computed our Shannon index, renamed TopicDiversity, and then regressed our binary variable EM and the control variables on the dependent variable TopicDiversity.

$$TopicDiversity = - \sum_{k=1}^6 p_k \ln(p_k)$$

$$\begin{aligned} Lm(TopicDiversity) = & \beta_0 + \beta_1 EM + \beta_2 Loss + \beta_3 Return + \beta_4 Lev + \beta_5 Sd \\ & + \beta_6 Return + \beta_7 Errormedian + \beta_8 \log(NoA) \\ & + \beta_9 \log(Marketsize) + \beta_{10} Sic + \beta_{11} Fyearq + \beta_{12} Fqtr + \epsilon \end{aligned}$$

As we observed in [Figure 2](#), the EM variable is 99% positively correlated with the independent variable TopicDiversity. We can therefore say that when companies use earnings management, it is more likely to find less diversity in the thematic content covered.

There are several reasons for this result. First, by reducing their disclosures, firms

can avoid disclosing anything that might cast doubt on their financial health or long-term plan. Indeed, companies that use earnings management have a vested interest in reducing the amount of management information available. The goal is to reduce the risk of disclosing information that could raise suspicions (P. Dechow, Ge, & Schrand, 2010). In addition, companies that practice earnings management may tend to be less transparent and disclose less information about their past and future performance. This may help them avoid disclosing information that could be used as evidence against them or raise suspicion. Finally, companies that use earnings management techniques may try to avoid disclosing too much information that could compromise their ability to monitor market expectations. According to a study (P. M. Dechow et al., 2010), companies that use earnings management generally give less information about forecasts and management, which can help them avoid raising expectations among investors and analysts.

In conclusion, companies that use earnings management tend to cover fewer topics in their earnings releases to avoid disclosing information that might cast doubt on their performance, to avoid disclosing too much information that could be interpreted negatively or used against them, and to avoid jeopardizing their ability to control market expectations.

6 Conclusion

In the literature review, we have highlighted several elements that influence the earnings press releases of companies that do not meet analysts' expectations. Several studies have shown that the tone used can have an impact on the market's response to earnings press releases. However, no study has yet been conducted on the influence of earnings press releases' topics in beating or meeting analysts' expectations. This is the reason for our research question. This question contributes to highlighting the different topics discussed in the earnings press releases when they are below analysts' forecasts and when firms practice earnings management.

We then considered two hypotheses that seemed relevant. The first three hypothe-

ses focus on the MBE variable while the last two hypotheses focus on the EM variable. For the first hypothesis, we considered what topics were present in earnings press releases when companies met or beat analysts' expectations. Then, we determined if it was more financial or non-financial topics that we found in earnings press releases of companies meeting or beating expectations. Afterwards, we tested whether the topics that were significant in earnings press releases when companies did MBE were all the same. For the third hypothesis, we determined whether companies that meet or beat expectations tend to address more topics in their press reports. For the fourth hypothesis, we considered what topics were present in earnings press releases when companies used earnings management. Then, we determined if it was more financial or non-financial topics that we found in earnings press releases of companies using earnings management. For the fifth hypothesis, we determined whether companies that use earnings management tend to cover fewer topics in their earnings press.

In order to do our analysis, we used topic modeling to highlight topics from different earnings press releases. We then manually formed 6 groups of similar topics to use them as dependent variables to perform the regressions. Finally, we selected several control variables to conduct linear regression tests.

In general, we find that the topics we studied are mostly present in earnings press releases in companies that do not meet expectations. In other words, the six topics we looked at are not regularly found in earnings press releases in companies that meet expectations. However, this is not the case for the Health and Tech Provider topics which are more easily found in the earnings press releases of companies that meet or beat the analysts' expectations. For companies practicing earnings management, we can see that these companies tend to discuss the Environmental topic in their earnings press. In terms of diversity of topics, we observe that companies that meet or beat expectations tend to address more topics in their earnings press. In contrast, companies that practice earnings management are less talkative in their earnings press.

To conclude this article, we will make a general review and critique by not limiting

ourselves to the boundaries of this research. If we put the results of our study into practice in the real world, we can put some limits to our findings. Indeed, just because the topic of the Environment is addressed in a company's earnings press does not necessarily mean that it will be below expectations. The company could address the subject of the Environment by saying that it is involved in sustainable projects and environmental protection. In this case, the topic of pollution would probably come up via the topic of modeling without being addressed negatively. A second limitation is that the topics emerging from the topic modeling are labeled by ourselves. Therefore, there could be divergences in the interpretation of some words, which would result in different clusters. Although we can say that the topics can give investors an indication of how well the company is doing against analysts' expectations, it is not the only way to form an opinion.

Moreover, the company writes its own earnings press. The management team could therefore be tempted to manipulate the topics discussed or to play on words to influence the readers' feelings. Therefore, it is essential to take into account other variables and data (quantitative and qualitative data) to get the best possible picture of the company and its future plans.

References

- Arslan-Ayaydin, Ö., Bishara, N., Thewissen, J., & Torsin, W. (2020). Managerial career concerns and the content of corporate disclosures: An analysis of the tone of earnings press releases. *International Review of Financial Analysis*, *72*, 101598.
- Arslan-Ayaydin, Ö., Boudt, K., & Thewissen, J. (2016). Managers set the tone: Equity incentives and the tone of earnings press releases. *Journal of Banking & Finance*, *72*, S132–S147.
- Arslan-Ayaydin, Ö., Thewissen, J., & Torsin, W. (2021). Disclosure tone management and labor unions. *Journal of Business Finance & Accounting*, *48*(1-2), 102–147.
- Aziz, S., Dowling, M., Hammami, H., & Piepenbrink, A. (2022). Machine learning in finance: A topic modeling approach. *European Financial Management*, *28*(3), 744–770.
- Ball, R., & Brown, P. (2019). Ball and brown (1968) after fifty years. *Pacific-Basin Finance Journal*, *53*, 410–431.
- Barde, B. V., & Bainwad, A. M. (2017). An overview of topic modeling methods and tools. In *2017 international conference on intelligent computing and control systems (iciccs)* (pp. 745–750).
- Bartov, E., Givoly, D., & Hayn, C. (2002). The rewards to meeting or beating earnings expectations. *Journal of accounting and economics*, *33*(2), 173–204.
- Bernard, V. L., & Thomas, J. K. (1989). Post-earnings-announcement drift: delayed price response or risk premium? *Journal of Accounting research*, *27*, 1–36.
- Bhojraj, S., Hribar, P., Picconi, M., & McInnis, J. (2009). Making sense of cents: An examination of firms that marginally miss or beat analyst forecasts. *The Journal of Finance*, *64*(5), 2361–2388.
- Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003). Latent dirichlet allocation. *Journal of machine Learning research*, *3*(Jan), 993–1022.

- Boudt, K., Thewissen, J., & Torsin, W. (2018). When does the tone of earnings press releases matter? *International Review of Financial Analysis*, 57, 231–245.
- Bouri, E., de Boyrie, M. E., & Pavlova, I. (2017). Volatility transmission from commodity markets to sovereign cds spreads in emerging and frontier countries. *International Review of Financial Analysis*, 49, 155–165.
- Bova, F. (2013). Labor unions and managementâs incentive to signal a negative outlook. *Contemporary Accounting Research*, 30(1), 14–41.
- Brown, L. D., Call, A. C., Clement, M. B., & Sharp, N. Y. (2015). Inside the âblack boxâ of sell-side financial analysts. *Journal of Accounting Research*, 53(1), 1–47.
- Bryan, S., Hwang, L., & Lilien, S. (2000). Ceo stock-based compensation: An empirical analysis of incentive-intensity, relative mix, and economic determinants. *The Journal of Business*, 73(4), 661–693.
- Call, A. C., Chen, S., & Tong, Y. H. (2013). Are analysts' cash flow forecasts naïve extensions of their own earnings forecasts? *Contemporary Accounting Research*, 30(2), 438–465.
- Christensen, J., Kent, P., & Stewart, J. (2010). Corporate governance and company performance in australia. *Australian accounting review*, 20(4), 372–386.
- Davis, A. K., & Tama-Sweet, I. (2012). Managersâ use of language across alternative disclosure outlets: Earnings press releases versus md&a. *Contemporary Accounting Research*, 29(3), 804–837.
- Dechow, P., Ge, W., & Schrand, C. (2010). Understanding earnings quality: A review of the proxies, their determinants and their consequences. *Journal of accounting and economics*, 50(2-3), 344–401.
- Dechow, P. M., Hutton, A. P., & Sloan, R. G. (1996). Economic consequences of accounting for stock-based compensation. *Journal of accounting research*, 34, 1–20.

-
- Dechow, P. M., Myers, L. A., & Shakespeare, C. (2010). Fair value accounting and gains from asset securitizations: A convenient earnings management tool with compensation side-benefits. *Journal of accounting and economics*, 49(1-2), 2–25.
- Dechow, P. M., & Schrand, C. M. (2004). *Earnings quality*. Research Foundation of CFA Institute Charlottesville, VA.
- DeFond, M. L., & Hung, M. (2003). An empirical analysis of analysts' cash flow forecasts. *Journal of accounting and economics*, 35(1), 73–100.
- Dichev, I. D., Graham, J. R., Harvey, C. R., & Rajgopal, S. (2013). Earnings quality: Evidence from the field. *Journal of accounting and economics*, 56(2-3), 1–33.
- Foster, G., Olsen, C., & Shevlin, T. (1984). Earnings releases, anomalies, and the behavior of security returns. *Accounting Review*, 574–603.
- Francis, J., & Soffer, L. (1997). The relative informativeness of analysts' stock recommendations and earnings forecast revisions. *Journal of Accounting Research*, 35(2), 193–211.
- Graham, J. R., & Harvey, C. R. (2001). The theory and practice of corporate finance: Evidence from the field. *Journal of financial economics*, 60(2-3), 187–243.
- Graham, J. R., Harvey, C. R., & Rajgopal, S. (2005). The economic implications of corporate financial reporting. *Journal of accounting and economics*, 40(1-3), 3–73.
- Guay, W. R., Core, J. E., & Larcker, D. F. (2002). Executive equity compensation and incentives: A survey. *Available at SSRN 276425*.
- Han, J., & Tan, H.-T. (2010). Investors' reactions to management earnings guidance: The joint effect of investment position, news valence, and guidance form. *Journal of Accounting Research*, 48(1), 81–104.
- Healy, P. M. (1985). The effect of bonus schemes on accounting decisions. *Journal of accounting and economics*, 7(1-3), 85–107.

-
- Healy, P. M., & Wahlen, J. M. (1999). A review of the earnings management literature and its implications for standard setting. *Accounting horizons*, 13(4), 365–383.
- Henry, E. (2008). Are investors influenced by how earnings press releases are written? *The Journal of Business Communication (1973)*, 45(4), 363–407.
- Hesarzadeh, R., Bazrafshan, A., & Rajabalizadeh, J. (2020). Financial reporting readability: Managerial choices versus firm fundamentals. *Spanish Journal of Finance and Accounting/Revista Española de Financiación y Contabilidad*, 49(4), 452–482.
- Hribar, P., & Yang, H. (2016). Ceo overconfidence and management forecasting. *Contemporary accounting research*, 33(1), 204–227.
- Huang, X., Teoh, S. H., & Zhang, Y. (2014). Tone management. *The Accounting Review*, 89(3), 1083–1113.
- Johnson, W. B., & Schwartz Jr, W. C. (2005). Are investors misled by *apro forma* earnings? *Contemporary Accounting Research*, 22(4), 915–963.
- Kasznik, R., & McNichols, M. F. (2002). Does meeting earnings expectations matter? evidence from analyst forecast revisions and share prices. *Journal of Accounting research*, 40(3), 727–759.
- Kim, Y., Li, H., & Li, S. (2014). Corporate social responsibility and stock price crash risk. *Journal of Banking & Finance*, 43, 1–13.
- Lee, Y. W., Yu, T., & Zhang, T. (2007). *Do corporations manipulate earnings to meet or beat analysts' forecasts? evidence from pension plan assumption changes* (Tech. Rep.). Citeseer.
- Levitt Jr, A. (1998). The numbers game. *The CPA Journal*, 68(12), 14.
- Livnat, J., & Mendenhall, R. R. (2006). Comparing the post-earnings announcement drift for surprises calculated from analyst and time series forecasts. *Journal of accounting research*, 44(1), 177–205.

- Lopez, T. J., & Rees, L. (2002). The effect of beating and missing analysts' forecasts on the information content of unexpected earnings. *Journal of Accounting, Auditing & Finance*, 17(2), 155–184.
- Maglio, R., Rey, A., Agliata, F., & Lombardi, R. (2020). Connecting earnings management and corporate social responsibility: A renewed perspective. *Corporate Social Responsibility and Environmental Management*, 27(2), 1108–1116.
- Man, C.-k., Wong, B., et al. (2013). Corporate governance and earnings management: A survey of literature. *Journal of Applied Business Research (JABR)*, 29(2), 391–418.
- McNichols, M. F., & Stubben, S. R. (2008). Does earnings management affect firms' investment decisions? *The accounting review*, 83(6), 1571–1603.
- McVay, S., Nagar, V., & Tang, V. W. (2006). Trading incentives to meet the analyst forecast. *Review of Accounting Studies*, 11, 575–598.
- Mendenhall, R. R. (2004). Arbitrage risk and post-earnings-announcement drift. *The Journal of Business*, 77(4), 875–894.
- Miller, G. S., & Piotroski, J. D. (2000). Forward-looking earnings statements: Determinants and market response. *Available at SSRN 238593*.
- Nollet, J., Filis, G., & Mitrokostas, E. (2016). Corporate social responsibility and financial performance: A non-linear and disaggregated approach. *Economic Modelling*, 52, 400–407.
- Payne, J. L., & Robb, S. W. (2000). Earnings management: The effect of ex ante earnings expectations. *Journal of Accounting, Auditing & Finance*, 15(4), 371–392.
- Puffer, S. M., & Weintrop, J. B. (1991). Corporate performance and ceo turnover: The role of performance expectations. *Administrative Science Quarterly*, 1–19.
- Rendleman Jr, R. J., Jones, C. P., & Latane, H. A. (1982). Empirical anomalies based on unexpected earnings and the importance of risk adjustments. *Journal of Financial*

-
- economics*, 10(3), 269–287.
- Rogers, J. L., Van Buskirk, A., & Zechman, S. L. (2011). Disclosure tone and shareholder litigation. *The Accounting Review*, 86(6), 2155–2183.
- Ronen, J., Fogel, H., & Yaari, V. (2011). *The reporting strategy to meet or beat thresholds* (Tech. Rep.). Working Paper, New York University, University of Toronto, and Morgan State â|.
- Roychowdhury, S. (2006). Earnings management through real activities manipulation. *Journal of accounting and economics*, 42(3), 335–370.
- Schumaker, R. P., & Chen, H. (2009a). A quantitative stock prediction system based on financial news. *Information Processing & Management*, 45(5), 571–583.
- Schumaker, R. P., & Chen, H. (2009b). Textual analysis of stock market prediction using breaking financial news: The azfin text system. *ACM Transactions on Information Systems (TOIS)*, 27(2), 1–19.
- Sengupta, P. (1998). Corporate disclosure quality and the cost of debt. *Accounting review*, 459–474.
- Shannon, C. E. (1948). A mathematical theory of communication. *The Bell system technical journal*, 27(3), 379–423.
- Skinner, D. J., & Sloan, R. G. (2002). Earnings surprises, growth expectations, and stock returns or don't let an earnings torpedo sink your portfolio. *Review of accounting studies*, 7(2-3), 289–312.
- Thewissen, J., Shrestha, P., Torsin, W., & Pastwa, A. (2021). *Unpacking the black box of ico white papers: a topic modeling approach* (Tech. Rep.). Université catholique de Louvain, Louvain Finance (LFIN).
- Yan, B., Arslan-Ayaydin, Ö., Thewissen, J., & Torsin, W. (2021). Does managerial ability affect disclosure? evidence from earnings press releases. *Asian Review of Accounting*.

Zhang, S., Aerts, W., & Pan, H. (2019). Causal language intensity in performance commentary and financial analyst behaviour. *Journal of Business Finance & Accounting*, 46(1-2), 3-31.