



UNIVERSIDADE CATÓLICA PORTUGUESA

Is it profitable to trade solely upon financial narratives?

Portfolio creation based on textual analysis

Final Assignment in the modality of Dissertation presented to Universidade
Católica Portuguesa to obtain the degree of Master in Finance

by

Diogo Gouveia

Under the supervision of
Paulo Alves, Ph.D.

Católica Porto Business School, Universidade Católica do Porto
March 2024

Acknowledgments

I would like to express my gratitude to my guiding professor, Paulo Alves, for his invaluable guidance, unwavering support and insightful mentorship throughout the completion of my master's thesis in finance.

Resumo

O estudo tem como objetivo compreender se a rentabilidade pode ser alcançada exclusivamente através da narrativa financeira. Para o efeito, foram estudadas *word count*, *Flesch score*, *Fog index*, *causal reasoning*, *forward-looking indicators*, *uncertainty*, e *net tone*. Estas características das narrativas por sua vez foram examinadas em três diferentes secções: o documento completo, as mensagens do Chairman e do CEO.

Foram construídos dez portefólios para cada característica, segurando o mesmo número de posições *short* e *long*, predefinidas pelos melhores e piores desempenhos identificados na narrativa, com base na literatura disponível. Contudo, os resultados indicam que, com base nos critérios de construção dos portefólios, um suporte exclusivo nas medidas narrativas não é uma abordagem viável para alcançar a rentabilidade, conduzindo inclusive à destruição de valor.

Palavras-chave: Narrativa financeira, análise textual, construção de portefólio, *word count*, *Flesch Score*, *causal reasoning*, *forward looking*, *uncertainty*, *net tone*

Abstract

The research aims to understand if profitability can be achieved relying solely on textual analysis metrics, such as word count, Flesch score, Fog index, causal reasoning, forward-looking indicators, uncertainty, and net tone. The characteristics of financial narratives were examined across three different sections of the annual report, the entire document, the Chairman's and the CEO's statements.

Ten portfolios for each characteristic were constructed with equal weighting of short and long positions, determined by the top and bottom performers identified within the narrative, based on the literature available. However, findings indicate that, based on the portfolio construction criteria, exclusive dependence on narrative measures is not a viable approach for achieving profitability and mostly leads to value erosion.

Keywords: Financial narrative, textual analysis, portfolio creation, word count, Flesch Score, causal reasoning, forward looking, uncertainty, net tone

Table of Contents

Acknowledgments.....	iii
Resumo.....	v
Abstract.....	vii
List of Tables.....	xii
List of Graphs.....	xiii
1. Introduction.....	1
2. Literature Review	3
3. Methodology - portfolio creation criteria.....	11
4. Data Collection.....	15
5. Results	16
5.1 Entire document.....	16
5.1.1 Word count.....	16
5.1.2 Flesch Score	17
5.1.3 Fog Index	18
5.1.4 Causal Reasoning	20
5.1.5 Forward looking	21
5.1.6 Uncertainty	23
5.1.7 Net tone.....	24
5.1.8 Summary.....	26
5.2 Chairman statement	28
5.2.1 Word count (Chair)	28
5.2.2 Flesch Score (Chair).....	29
5.2.3 Fog Index (Chair).....	30
5.2.4 Causal Reasoning (Chair).....	32
5.2.5 Forward looking (Chair).....	33
5.2.6 Uncertainty (Chair)	34
5.2.7 Net tone (Chair)	35
5.2.8 Summary (Chair)	37
5.3 CEO statment.....	39
5.3.1 Word count (CEO).....	39
5.3.2 Flesch Score (CEO)	40
5.3.3 Fog Index (CEO)	41
5.3.4 Causal Reasoning (CEO)	42

5.3.5 Forward looking (CEO)	43
5.3.6 Uncertainty (CEO)	44
5.3.7 Net tone (CEO).....	45
5.2.8 Summary (CEO).....	46
6. Top and Bottom performers.....	48
7. Conclusions	50
References	54
Appendix	60

List of Tables

Table 1: Portfolio positions.....	13
Table 2: Portfolio averages (entire document).....	26
Table 3: Portfolio averages (Chairman statement).....	37
Table 4: Portfolio averages (CEO statement).....	46
Table 5: Descriptive statistics	56

List of Graphs

Figure 1: Word count (entire document).....	16
Figure 2: Flesch Score (entire document)	17
Figure 3: Fog Index (entire document)	18
Figure 4: Causal reasoning (entire document)	20
Figure 5: Forward looking (entire document)	21
Figure 6: Uncertainty (entire document).....	23
Figure 7: Net tone (entire document).....	24
Figure 8: Portfolio averages (entire document).....	26
Figure 9: Word count (Chairman Statement)	28
Figure 10: Flesch Score (Chairman Statement	29
Figure 11: Fog Index (Chairman Statement).....	30
Figure 12: Causal reasoning (Chairman Statement).....	32
Figure 13: Forward looking (Chairman Statement).....	33
Figure 14: Uncertainty (Chairman Statement)	34
Figure 15: Net tone (Chairman Statement)	35
Figure 16: Portfolio averages (Chairman Statement)	37
Figure 17: Word count (CEO Statement).....	39
Figure 18: Flesch Score (CEO Statement)	40
Figure 19: Fog Index (CEO Statement)	41
Figure 20: Causal reasoning (CEO Statement)	42
Figure 21: Forward looking (CEO Statement).....	43
Figure 22: Uncertainty (CEO Statement).....	44
Figure 23: Net tone (CEO Statement).....	45
Figure 24: Portfolio averages (CEO Statement).....	46
Figure 25: Top performer (average of portfolios)	48
Figure 26: Second best performer (average of portfolios)	48
Figure 27: Bottom performer (average of portfolios)	48
Figure 28: Second worst performer (average of portfolios).....	48

1.Introduction

Financial narratives, also known as qualitative financial data, are non-numerical material that is included in financial reports and statements. This data provides context and allows for a more in-depth view of a company's financial performance, operations, and strategy. Management's discussion and analysis (MD&A), industry analysis, risk factors, and qualitative assessments of a company's financial health can all be included in financial narratives. In contrast to quantitative financial data, which is numerical and objective, financial narratives are subjective and can reveal insights that numbers alone cannot capture. As investors seek a more thorough understanding of a company's financial condition and future prospects, financial narratives have become increasingly significant in financial analysis and investment decision-making.

While the financial narrative can be subjective in nature, it can provide useful insights that quantitative financial data alone cannot capture. Financial narratives are relevant, according to research, and can have a major impact on the market's opinion of a company's success and future performance (Johnson and Tuckett, 2017).

According to research, firms with more informative and well-written MD&A sections in their annual reports are more likely to have positive stock returns (Brown and Tucker, 2011; Hsieh et al., 2015). Additionally, the quality of disclosure in the MD&A section can influence analysts' profit projections and recommendations (Ali, Klasa and Ye, 2014). Furthermore, research has shown that financial narratives, such as risk factor disclosures, can impact credit ratings (Kim et al., 2012).

Being said, the study of the financial narrative is significant for various reasons. One of the most important reasons is that numbers do not provide the "full picture" of a company's financial performance and future prospects. Quantitative data, such as financial statements, can only reveal so much about a company's operations, strategy, and risks. In contrast, qualitative data provides a more complete and detailed insight of a company's financial status.

Furthermore, qualitative data, such as textual information, can supplement quantitative data by providing additional insights that numerical statistics may not capture. Financial narratives, such as those found in annual reports' management discussion and analysis (MD&A) sections, can give context for the data and explain the variables that have influenced a company's financial success.

Another motivation for studying the subject is that textual data can be revealed more quickly than numerical data. Companies, for example, can distribute earnings press releases and MD&A sections ahead of their financial statements, giving investors and analysts a glimpse of the company's financial performance. Moreover, a huge amount of unstructured data is available from a variety of sources, including company filings, business websites, social media, analyst reports, and news. This data can be analysed and interpreted to provide useful insights into a company's financial status and market trends.

These findings imply that financial narratives do matter and can have a considerable impact on financial markets when taken collectively. Being said, when making investment decisions, investors, analysts, and other financial stakeholders must pay special attention to qualitative financial data.

2.Literature Review

Research states that improved disclosure quality tends to draw attention from analysts and enhance information intermediation (Lang and Lundholm, 1996; Lehavy et al., 2011). This suggests that as the quality of disclosure increases, analysts are more likely to cover the firm, more accurate information. The anticipated outcome enhances stock liquidity and reduces capital costs for the firm, aligning with the expectation that more analyst following leads to these favourable outcomes (Beyer et al., 2010).

Existing research continuously highlights the importance of narrative disclosures for market participants' decision-making processes, and subsequent market reactions have been investigated. Henry (2006) adds considerably to this discussion by illustrating how the market reacts to the verbal components of earnings press releases. She emphasizes that the incorporation of linguistic information improves the prediction accuracy of market returns, outperforming the contribution of financial information alone.

The manipulation of disclosures adds another degree of complexity, raising worries that such acts may entail the suppression of critical information, thereby affecting the company's valuation and resulting in economic implications (Kothari et al., 2009). Another study provides a deeper understanding of management tactics by understanding how managers strategically employ financial narratives to impact readers' perceptions (Merkl-Davies and Brennan, 2007).

Li (2010) found that forward-looking information has positive relation between forward-looking disclosures, especially in the MD&A segment, and future business performance. This shows that including forward-looking information in key portions of financial reports might predict a company's future

success. Furthermore, data from other report parts, such as profits, supports the idea that forward-looking information helps the market forecast future earnings (Hussainey et al. 2003).

Davis et al. (2015) investigated optimism, on the management level, during earnings conference calls. The researchers applied the word lists from Henry (2008) and Loughran and McDonald (2011) to assess sentiment in these reports. Specifically, the word lists from Henry (2008) and Loughran and McDonald (2011) were useful in determining the tone of the narrative. The study's findings demonstrated a positive relation between managers' tone, as represented by the aforementioned word lists, and future operating performance. This implies that managers' verbal expressions of optimism during earnings conference calls, identified by the provided word lists, can serve as indicators of favourable future operational results.

Furthermore, the positive relationship between the tone of textual content and future stock returns is evident within different contexts, including news media articles (Tetlock, 2007) and the MD&A sections (Feldman et al., 2009).

According to Aerts and Zhang (2014), causal reasoning takes up a significant amount of the content in management commentary in the annual reports. This is a critical component that explains the cause-and-effect connections between business successes and the performance results that go along with them. Aerts and Zhang (2014) point out that explanations can go beyond unintentional causes and take needs and motivations into account.

Reasoning is especially important for publicly traded companies, especially for those that must adhere to strict rationality standards. In these situations, it becomes imperative to demonstrate the application of sound reasoning since it is supposed to demonstrate organizational competence and promote credibility (Sonenshein et al., 2011). The study by Sonenshein et al. (2011) highlights how transparent and well-reasoned communication contributes to the perception of

reliability and credibility within the financial community, emphasizing the significance of rational discourse in enhancing trustworthiness.

Although, on the other side of the spectrum, managers frequently utilize causal language as a strategic means to distance themselves from poor performance, effectively mitigating concerns about their managerial competence among investors (Asay et al., 2018).

The employment of high causal reasoning in annual reports presents a nuanced scenario with both positive and negative implications, contingent on the context. On one hand, it serves as a strategic tool for managers to justify organizational performance and allay investors' concerns (Tessarolo, 2010; Aerts and Zhang, 2014). Conversely, it can be leveraged to obscure adverse information, thereby increasing the risk of a substantial decline in stock prices (Kong, 2021).

This manipulation of causal language is notably prevalent when explaining unfavourable outcomes, often emphasizing external, unstable, and uncontrollable causes (Bettman, 1983). Consequently, while high causal reasoning may enhance the perceived plausibility of reported performance, it can also increase information asymmetry and the potential for negative outcomes.

Research in the field of finance has consistently shown that high uncertainty, particularly in the form of information uncertainty, can lead to lower stock returns (Jiang, 2004). This is further supported by studies that have found a negative impact of uncertainty on stock returns (Pereira, 2019). In fact, stocks with high uncertainty about risk, as measured by the volatility, have been found to underperform those with low uncertainty about risk (Baltussen, 2014). These findings suggest that high uncertainty, whether in the form of narrative or economic uncertainty, can indeed lead to lower stock returns.

When it comes to complexity, Ertugrul et al. (2017) found that banks typically charge higher interest rates on loans given to companies whose 10-K report file

sizes are larger, attributing this to complexity on the disclosure, suggesting that it occurs because the firms may use complex annual reports to obfuscate negative information, which in turn affects borrowing costs. Ertugrul et al. (2017) further support their claim by showing a positive correlation between the size of a 10-K file and the risk of a crash later, highlighting the possible effects of information complexity in financial reporting on the risk profile of borrowing firms.

Among the various measures employed to assess the readability and complexity of financial narratives, the Fog Index stands out as both well-known and subject to considerable critique.

The Fog Index, also referred to as the Gunning Fog Index, serves as an important tool for gauging the readability of written text. Developed by Robert Gunning in 1952, this index quantifies the level of complexity inherent in written language.

The calculation of the Fog Index revolves around two key components: the average sentence length and the proportion of complex words. Its formula is expressed as follows:

$$\text{Fog Index} = 0.4 * (\text{Average sentence length} + \text{Percentage of complex words})$$

To determine the average sentence length, simply divide the total word count by the number of sentences within the text. The percentage of complex words involves the identification of words considered complex or challenging, typically being those comprising three or more syllables.

The resulting Fog Index value offers an approximation of the grade level required to comprehend the text upon the initial reading. Therefore, a lower Fog Index suggests a text that is more accessible and easier to understand, whereas higher scores indicate greater complexity and content that may prove more challenging to comprehend.

In the realm of finance literature, Li (2008) fundamental research in uncovered crucial findings. In his paper, the author finds that firms reporting lower earnings

tend to present annual reports characterized by high Fog Index values or high word counts, thus being harder to read (Li, 2008).

This discovery, as underscored by Bloomfield (2008), may be attributed to underperforming companies needing to resort to verbosity and longer sentences to comprehensively convey their financial situation to investors. Furthermore, Li's work established a strong correlation between annual report readability and earnings persistence, highlighting the link between linguistic attributes in these reports and actual firm performance. Li (2008) concludes that "annual reports of firms with poor performance are more difficult to read" and companies whose annual reports exhibit greater readability experience higher earnings persistence in the next one to four years.

Subsequent studies, building upon Li (2008)'s pioneering work, have adopted the Fog Index as a metric for assessing annual report readability. Biddle, Hilary, and Verdi (2009) expanded the investigation using Fog measures, demonstrating that firms characterized by high reporting quality, as measured by the Fog Index and other variables, tend to exhibit greater efficiency in capital investment, concluding in their words that the findings "are consistent with the idea that financial reporting quality serves a role in mitigating information frictions that ultimately hamper investment efficiency".

A notable observation comes from Guay et al. (2015), who explored the challenges posed by less readable annual reports, employing various readability measures. They found that companies dealing with readability concerns often offset this issue by issuing additional managerial forecasts, such as earnings per share, sales, and cash flows. Being so, citing the paper, the conclusion suggested that "a more nuanced view of how financial statement complexity affects the mosaic of public information about the firm."

Turning to investor behaviour, Miller (2010) provided a valuable insight on the trading habits of small investors, where he discovered that investors trade

significantly fewer shares of companies with high Fog Index values and word counts, particularly around the 10-K filing date, implying that less readable annual reports pose a significant challenge, especially for less experienced investors.

Also, in line with the previous topic, Lawrence (2013) delved into the impact of readability on retail investors' stock holdings. Lawrence found that individual investors tend to allocate more capital to firms with concise and readable annual reports, as measured by the Fog Index, being this effect was less pronounced among high-frequency trading individuals.

In the context of cross-listed firms, Lundholm et al. (2014) conducted a comparative analysis between foreign firms listed on U.S. exchanges and domestic firms, revealing that foreign firms tend to produce more readable business documents, as evidenced by lower Fog Index values. The authors argued that companies outside of U.S have a strong incentive to enhance document readability to attract U.S. investors.

Lehavy et al. (2011) explored the link between annual report readability, based upon the Fog Index, and analyst dispersion. Their findings indicated that more readable annual reports were associated with lower analyst dispersion and greater earnings forecast accuracy. Additionally, it was observed correlation between 10-K readability and analyst coverage, with firms characterized by higher Fog Index values attracting more analysts to cover their stock, suggesting that “analyst behaviour is related to the readability of firms’ communication.

The influence of readability further extended to analyst reports and the impact on investor behaviour. De Franco et al. (2015) delved into this getting a compelling association out of it: more readable analyst reports were linked to significantly higher trading volume during a three-day window surrounding the analyst report date.

Despite research proving correlation of the Fog index and stock returns, the measure is highly criticized due to the limitations it depicts.

One of the most praised critiques revolves around the oversimplification when accessing the narrative complexity. Being argued that the Fog Index relies primarily on word and sentence length, which may not be adequate to capture the subtle aspects of narrative quality, and so the oversimplification could lead to an incomplete understanding and not capture the entire complexity of the text.

Another common critic is that dictionary-based approaches, such as Fog Index, do not capture the contextual meaning of words within financial narratives resulting in misinterpretations, as the measures may not distinguish the intended meaning of certain terms in their specific context.

While the Fog Index has traditionally been employed in grading K-12 textbooks, its application to business documents is still doubtful. Jones and Shoemaker (1994) noted that "there is no consensus about how closely the readability measures reflect the actual comprehension process". They further highlight that even if the words within each sentence were randomly reordered, the text would yield an identical Fog Index value, rendering it completely unintelligible.

Loughran and McDonald (2014) exposed the weakness in the Fog Index, revealing that the percentage of complex words (those with more than two syllables) accounted for a significant portion of the Fog Index value. Moreover, an increase in the proportion of complex words was associated with reduced readability according to the Fog Index. However, they emphasized that many of the most frequently occurring "complex" words in business documents, such as financial, company, operations, management, employees, and customers, are easily comprehensible to investors. Consequently, they argued that syllable counts serve as an inadequate measure of readability for business documents.

The insights drawn from Loughran and McDonald (2016) challenge understanding of readability, particularly in the context of business documents, raising questions about its relevance and effectiveness as a measure. The authors argue that the traditional concept of readability, which relies on word and sentence length as indicators of comprehension at different grade levels, does not adequately align with the assessment of business documents.

One of the primary criticisms directed at the commonly used measure of readability, the Fog Index, undergoes several key issues. Firstly, the measurement of sentence length in business documents is considered unreliable, making it an unreliable metric for understanding readability. Additionally, it becomes evident that many multisyllabic words commonly found in business documents are not inherently challenging and are easily understood by the intended audience (ex: liability).

Loughran and McDonald (2016) delve deeper into this matter and propose an alternative perspective. They provide evidence that the use of financial jargon is positively associated with their metrics of information integration. This discovery contradicts the traditional interpretation of readability, suggesting that the inclusion of terminology tailored for financially advanced readers may enhance the effectiveness of documents.

Given these findings, it is proposed that the fundamental concept of readability, which was originally borrowed from other disciplines, be better understood, and evaluated within the context of information complexity.

Guay et al. (2016) conducted a comprehensive analysis, combining six distinct indexes to investigate their relationships. Notably, their research demonstrated a high correlation (0.96) between the Fog Index and the Flesch Index.

Rudolf Flesch established the Flesch Readability Score (Flesch, 1949) to assess the readability of a piece of writing. It uses the average sentence length and number of syllables per word to determine how easy or difficult a piece of text is

to read. A lower Flesch score suggests that the text is simpler to read, whilst a higher number implies that it is more difficult to understand. This narrative, like the FOG index, is a measure of complexity and readability, so results for Flesch and Fog can be seen as comparable.

3. Methodology - portfolio creation criteria

According to the available literature, there is enough evidence and empirical studies to support the significance and importance of financial narratives in the financial sector. Nonetheless, the question of whether trading simply on financial narrative measures is profitable remains vague. The goal of this research is to investigate the potential profitability of trading using financial narrative measures, being the goal to evaluate the viability of a portfolio trading solely using financial narrative indicators.

The findings of this study will contribute to the literature on financial narratives by investigating the practical implications of their relevance and importance in finance. Also, the research will give investors and practitioners insights into the potential profitability of trading using textual analysis methods.

To address this question, it was constructed a portfolio creation model that picks the stocks according to the chosen metric.

As articulated by Alves et al. (2016), the existing scene of extensive automated analysis centred on annual report narratives predominantly concentrates on Form 10-K filings submitted by U.S. registrants via the Securities and Exchange Commission's EDGAR system. The intrinsic qualities of these 10-Ks, such as batch retrieval, plain text formatting, and a uniform reporting structure, make

them favourable to systematic research. However, it is important to note that the ambit of 10-Ks represents just a fragment of the of firms' annual report disclosures. A substantial proportion of registrants complement their 10-K submissions with non-standardized reports marked by visual elements, photographs, and supplementary narratives, such as shareholder letters. This multi-faceted documentation, frequently disseminated as PDF files, takes on heightened significance beyond the confines of the United States. On the other hand, despite its frequency, the automated analysis of these integral documents is hindered by barriers, resulting in a scarcity of comprehensive insights.

To summarize the review literature and understand the implications of the narrative measures, and the impact on the portfolio positions:

The word count of financial measures the number of words within a text and can be perceived as a proxy for complexity. According to the literature review, (Li, 2008) higher word count measures were found to have correlation with lower stock returns.

The Flesch score is a readability formula that combines sentence length and syllable count to determine a text's ease of comprehension. Given the high correlation between Fog Index and Flesch Score Guay, Samuels, and Taylor (2016), and they both address complexity, a higher Flesch Score can be perceived as a lead for lower stock returns.

The Fog Index is a readability metric that calculates the difficulty level of a document based on sentence length and word complexity and it is usually measured using a dictionary-based approach. Based on literature (Li, 2008) a higher Fog Index was associated with reading complexity, thus leading to lower stock returns.

Causal Reasoning can be understood as the frequency of causal language markers such as "because", "as a result," and "due to" in any textual data. Although mixed results were found when considering causality, for this study it

will be considered that firms with high causality demonstrate organizational competence and promote credibility (Sonenshein et al., 2011), thus improving the quality of the disclosures.

Forward-looking can be measured in financial narratives by counting the proportion of forward-looking statements. High forward-looking was understood to have a correlation with higher stock returns (Li, 2010).

The uncertainty can be measured by counting the frequency of uncertain words thus high uncertainty was found to be linked with lower stock returns (Jiang, 2004).

As a textual analysis method, tone can be measured by pairing the frequency of positive against negative terms. The higher the tone measures, as they can serve as indicators of favorable future operational results Davis et al. (2015), the higher the expected stock returns.

Measures	Higher values	Lower values
wordcount	Short	Long
flesch	Short	Long
fog	Short	Long
causal	Long	Short
forwardlooking	Long	Short
uncertainty	Short	Long
net_tone	Long	Short

Table 1: Portfolio positions

The dataset employed in this research contains diverse narrative measures, such as word count, flesch score, fog index, causal reasoning, forward-looking measures, uncertainty measures, and net tone.

The research focus is on companies listed on the London Stock Exchange (LSE), aiming to provide insights into their financial reporting practices.

Considering the relatively limited importance of narrative measures outside the United States, particularly in locations where text retrieval poses challenges,

this study focuses on UK annual reports. Consequently, the analysis is confined to UK stocks.

It is imperative to emphasize that the selected narrative measures were retrieved across three specific sections of annual reports: the entire reports, the CEO statement, and the Chairman statement. This segmentation ensures a comprehensive examination of narrative measures within distinct contexts.

Drawing inspiration from Sloan's (1996) foundational exploration of trading strategies grounded in reported accruals, the methodology of this study involves the creation of ten distinct portfolios for each examined narrative, aiming to mitigate the influence of outliers associated with individual portfolios, thereby enabling the evaluation of averages across the ensemble of ten portfolios. Each portfolio adopts a balanced composition, with half of its positions taking long positions and the remaining half assuming short positions. The selection criteria for these positions are determined based on the narrative analysis, identifying the top and bottom performers within the dataset and then taking the position according to the metric within focus.

Consequently, each portfolio consists of a total of 42 active stock positions, evenly distributed with 21 long positions and 21 short positions. This approach ensures a comprehensive exploration of the trading strategy's performance in response to the narrative measures under examination, doubling down on the hypothesis of profitability, tackling both sides of the narrative.

To mitigate the challenge posed by disparate disclosure dates, a weighting mechanism based on reporting periods was applied to determine the distribution of the ten portfolios. The starting period of the positions within each portfolio was then scheduled for the subsequent quarter, with the portfolios strategically divided across quarters. This approach led to the beginning of one portfolio in Q2, followed by four portfolios in Q3, three portfolios in Q4, and concluding with two portfolios in Q1 of the subsequent year.

In alignment with the reported quarters, the construction of positions started in the 2nd Quarter of 2007, and all positions were systematically closed at the conclusion of 2018.

In response to the dynamic landscape of financial reporting, the model integrates an annual recalibration process. During the recalibration of portfolios, the previous balance transferred to the next set of chosen stocks. The realignment of portfolios is executed on an annual basis, synchronized with the starting date of the initial positions. This accounts for the early release of financial reports and ensures the continued relevance and accuracy of the portfolio composition throughout the study period.

4. Data Collection

This research focus is on companies listed on the London Stock Exchange (LSE), aiming to provide insights into their financial reporting practices. The study encompasses annual reports with fiscal year-ends spanning from 2006 to 2018, ensuring a comprehensive analysis over a significant timeframe. The textual data for annual reports is sourced from the CFIE dataset, as curated by El-Haj et al. (2020).

The information pertaining to daily stock prices and the benchmark employed in the analysis was sourced from Refinitiv Eikon Datastream, where extreme outliers were cleaned to ensure a robust analysis. Reporting dates were systematically extracted from Refinitiv Eikon (Reuters).

The sample comprises 14,504 annual reports across 2,148 firms and 13 years.

5. Results

5.1 Entire document

5.1.1 Word Count

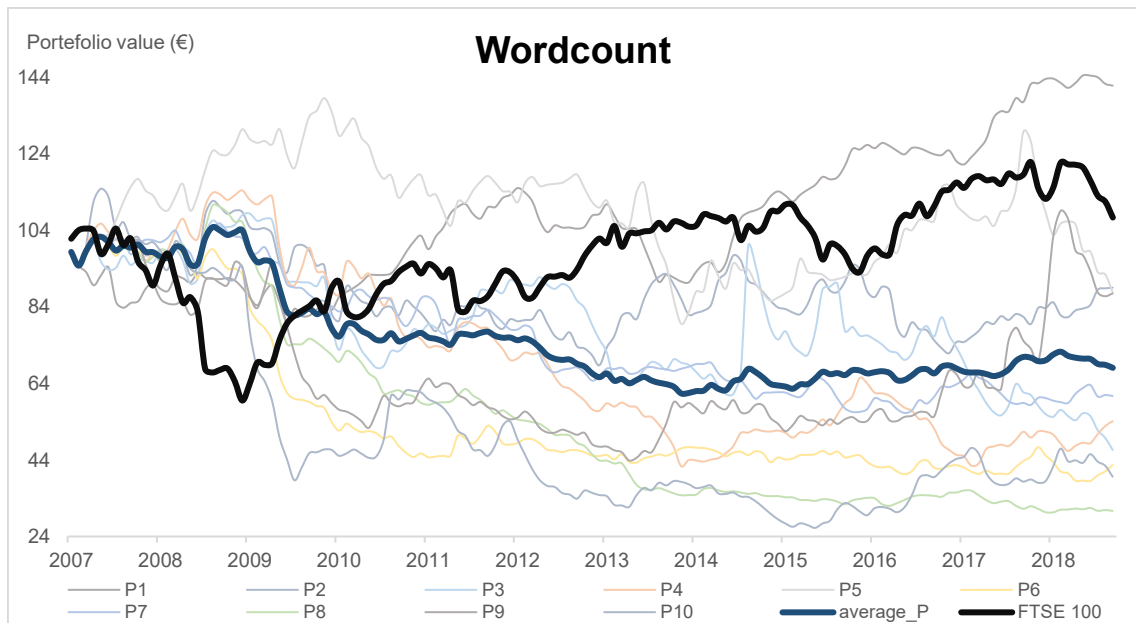


Figure 1: Word count (entire document)

When trading solely upon word count measures, among the picked portfolios, Portfolio (P) 1 was the only to deliver positive returns of around 42,4%, managing to outperform the benchmark by 32.1% at the closing date. The remaining portfolios suffered significant losses, and thus none of them surpassed the FTSE 100, averaging a loss of 36.6%.

P1 was able to thrive with a combination of wins both in the short and long positions in almost identical magnitudes.

As predicted, all the portfolios outperformed the index amid the subprime crises, given that the portfolio structure acts in a way as an edge with the short positions. Despite the positive returns in the year 2008 and the start of 2009, a series of bad chosen stocks, combining losses on short and long positions,

dragged 8 of the portfolios to below the FTSE 100 levels, where they remained until the end of the holding period.

Notably, Portfolios P3, P6, P8, and P10 incurred losses exceeding 55%, indicating a severe underperformance, hinting that volatility was a constant present in the portfolio, with both long and short positions experiencing losses. On average, long positions recorded a loss of 28.3%, while short positions incurred an average loss of 36.3%.

5.1.2 Flesch Score

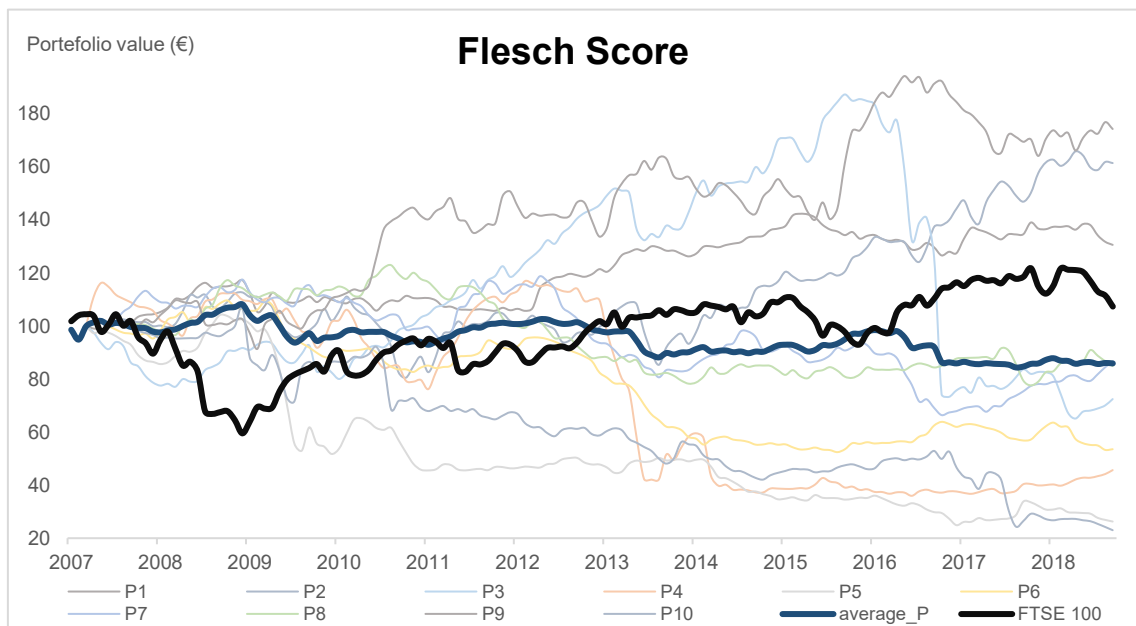


Figure 2: Flesch Score (entire document)

It is observed that three portfolios demonstrated profitable performance when paired against the chosen benchmark: P1 generated a return of 21.6%, P2 achieved a return of 50.3%, and P9 yielded the highest return at 62.3%. On the other hand, four portfolios experienced significant losses, with two of them exceeding 50% losses and two others incurring losses exceeding 75%.

The winning performance of P2 was driven by a 162.9% return on short positions, despite losing value on long positions. On the other hand, P9's impressive performance was primarily attributed to an exceptional 142.2% return on long positions with marginal gain on the shorts taken. Portfolio 1 achieved moderate success, with positive returns in both long and short positions.

Despite the considerable performance thought the first year, the shortfall of P3 around 2016 was mainly driven off the short stock picks, which increased in value, leading to a loss of 62.2% on short positions in that year of which the portfolio was not able to recover from.

When considering the average performance of all portfolios, there was an average loss of 14.3%, falling short of the FTSE 100 by an average of 20.3%.

Furthermore, it is essential to acknowledge that volatility was a prevalent factor across the portfolios under evaluation.

5.1.3 Fog Index

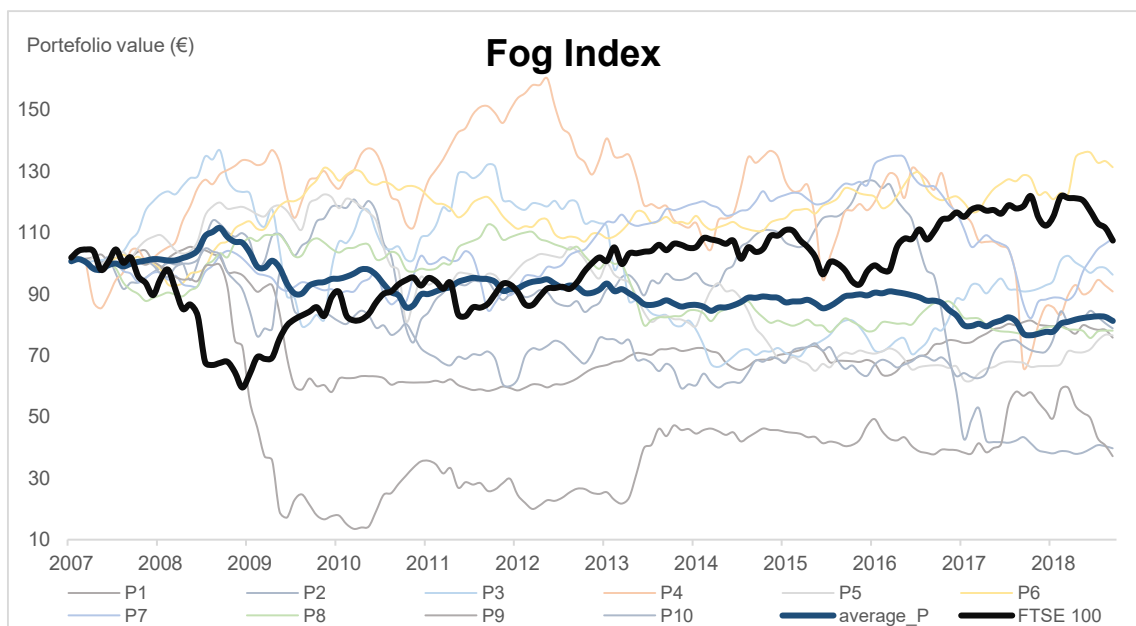


Figure 3: Fog Index (entire document)

When considering portfolio trading based on the fog index, only two portfolios, namely P6 and P7, demonstrated superior performance compared to the benchmark, achieving returns of 22.2% and 0.6%, respectively. In contrast, the majority of the remaining portfolios experienced value depreciation, with two of those portfolios incurring losses exceeding 60%. Also, only the two portfolios that outperformed the FTSE 100 managed to get positive returns.

Notably, although the portfolios exhibited volatility, it was not as pronounced as the previous 2 studied narrative metrics. On average, the portfolios experienced a collective loss of 18.6%, and when paired against the index they fell short on average 24.3%.

These losses were primarily attributed to the long positions within the portfolios. Only P6 managed to generate positive returns from its long positions, while all other portfolios suffered losses, averaging 27.4%. In contrast, the short positions in five portfolios yielded positive returns. However, the gains from short positions were overshadowed by the substantial losses incurred in the long positions, except in the cases of P6 and P7.

In summary, the results of this portfolio analysis underscore the challenges and complexities inherent to the Fog measure, as is highly criticized. While two portfolios outperformed the benchmark, the majority faced significant losses primarily in their long positions, demonstrating what the literature review predicted on this particular metric.

5.1.4 Causal Reasoning

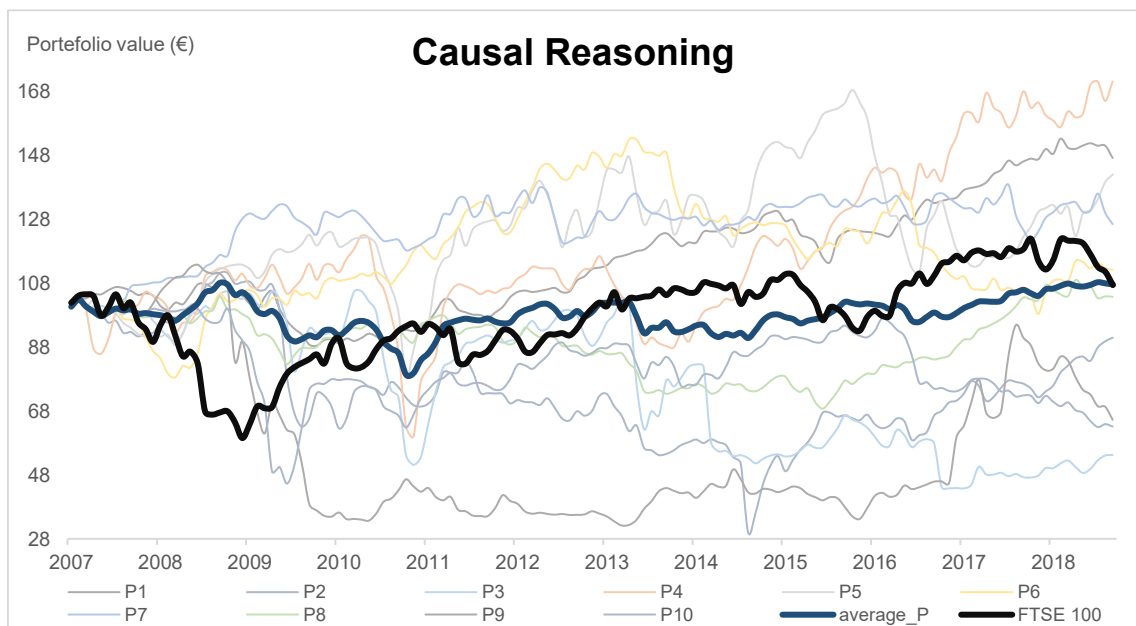


Figure 4: Causal Reasoning (entire document)

Among the ten portfolios under scrutiny, when trading exclusively based upon casual reasoning metrics, six portfolios demonstrated positive returns, with two of them exhibiting returns exceeding 40%, and a remarkable outlier (P4) achieving a remarkable increase of 72.3%. Despite six of them producing positive returns only 5 managed to outperform the benchmark, with the highest outperformance achieved by P4 being 59.3%.

Conversely, within the subset of portfolios yielding losses, the most substantial decline amounted to 45.5%, thus being outperformed by the index by 49.4%.

Short positions were the main driver for the returns, contributing positively, exhibiting an average return of 31.2%. In contrast, long positions yielded less favourable results, averaging a loss of 15.9%, losing value on seven of the portfolios. Nevertheless, the winners regain this value via short positions.

When comparing these findings with alternative narrative measures, it becomes evident that this particular aspect of portfolio performance exhibits comparatively lower volatility. Moreover, when we consider the overall average return across all portfolios, it emerges as a modest but notable profit, standing at 7.6%, actually outperforming the benchmark by 0.2%. While this figure may appear diminutive, it attests to the potential for positive returns within the portfolio selection.

5.1.5 Forward Looking

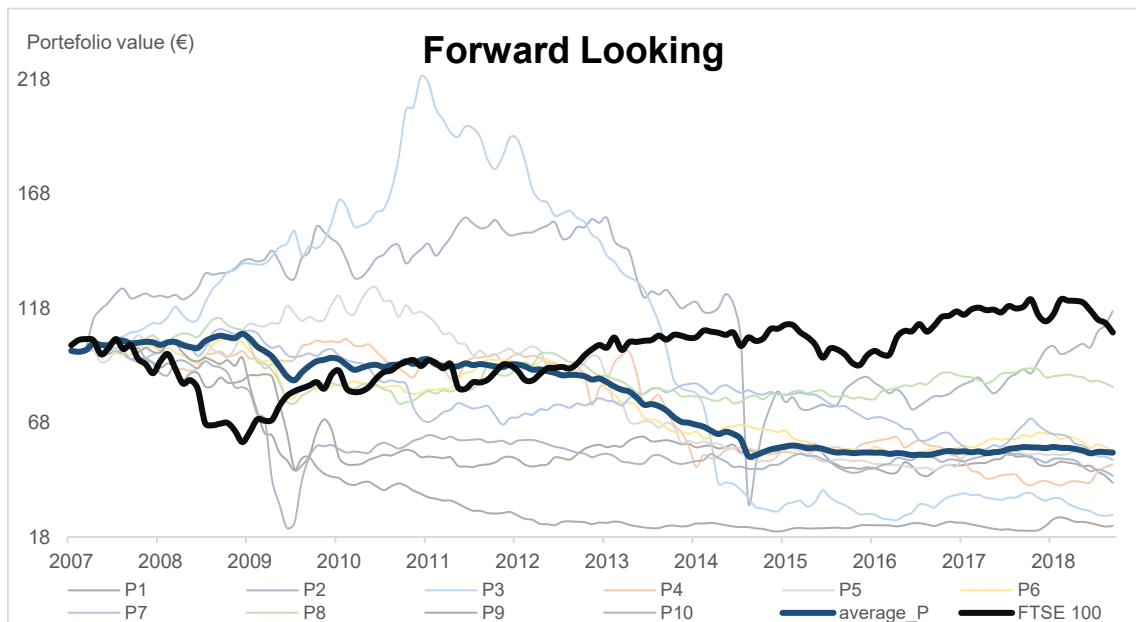


Figure 5: Forward looking (entire document)

When trading solely on forward looking measures, among the ten portfolios examined, only a single portfolio, P2, yielded positive returns, exhibiting a return of 19,4%, narrowly outperforming the benchmark by 8.7%. This achievement, however, stood in contrast to the clear trend of substantial losses witnessed throughout the portfolio set.

The severity of the losses become evident when looking deeper into the results. Out of the 9 losing portfolios, the lowest loss was 17.0%. The remaining ones incurred losses of significant magnitude, superior to 45%, and 2 of them with losses above 72%. The P9 endured the most severe setback when paired against the index, with a staggering loss of 78.7%. When we aggregate the performance of all portfolios, their average loss amounted to a substantial 48.9% against the FTSE 100, hinting to the significant erosion of value.

While portfolios P2 and P3 exhibited promising returns initially, capitalizing on short positions during the aftermath of the subprime crisis, they experienced a shift in 2013 and 2014, declining in value, attributed to a set of highly negative short positions, combined with unfavourable long positions.

Notably, none of the long positions yielded positive returns, resulting in an average loss of 54.8%. Even the most modest loss within this category was still significant, standing at 27.5%. A similar pattern emerged for short positions, with only two out of the ten portfolios achieving positive returns. On average, the short positions incurred a loss of 35.5%. Noteworthy, however, is the exceptional outlier, portfolio P2, the only one securing a positive return, returning a remarkable 98.4% gain in its short positions, thereby offsetting the losses incurred in its long positions.

In summary, with an average loss of 44.3%, the portfolios exhibited an underperformance of 48.1% relative to the benchmark.

5.1.6 Uncertainty

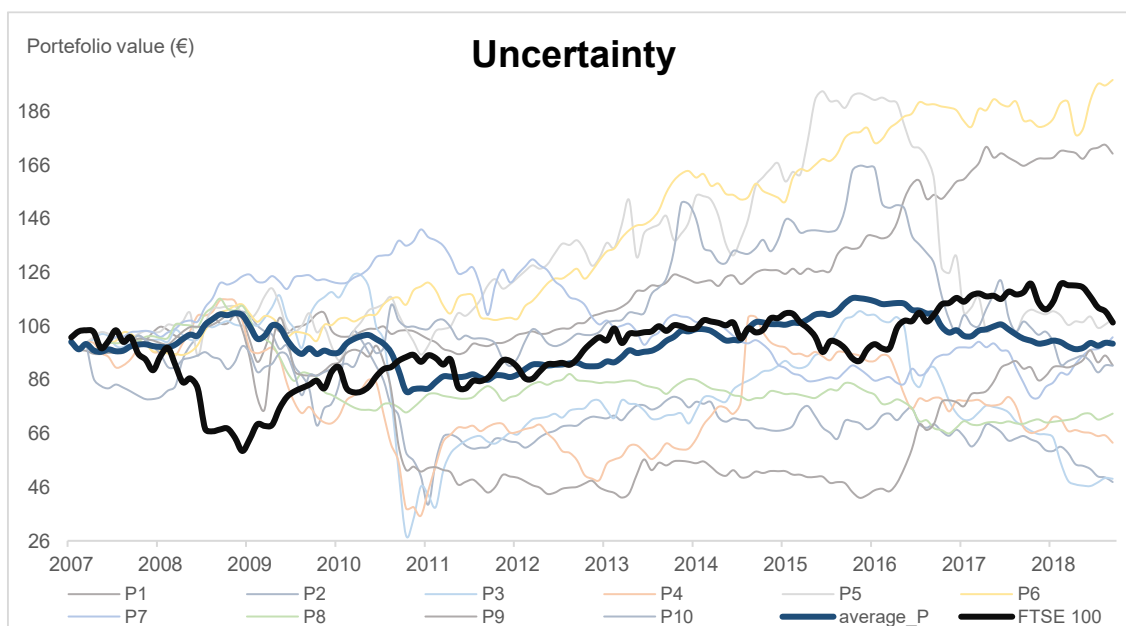


Figure 6: Uncertainty (entire document)

Upon close examination, it becomes evident that a subset of the portfolios exhibited promising returns when set against the benchmark, P1 and P6. Only four portfolios managed to secure positive returns. Despite the positive returns only three of them managed to outperform FTSE 100, namely, P1 achieving a return of 58.6%, and P6 which achieved a remarkable return of 84.2% above the index and P5, narrowly surpassing the benchmark by 1.4%. In contrast, the remaining portfolios incurred losses, with two portfolios experiencing losses exceeding 54% when against the FTSE 100.

Delving into the dynamics of portfolios P1 and P6, it is noteworthy that what led to their relevance were the positive returns achieved across both long and short positions. P1 was particularly influenced by a robust return of 93.3% on short positions, while Portfolio P6 thrived on a substantial 117.0% return from long positions.

P5 exhibited a noteworthy performance trajectory, consistently outperforming the benchmark until approximately 2016. However, during that year, a set of short positions lead to a notable loss of 55%, though it should be noted that prior successes had also been derived from short positions.

Further analysis reveals that, on average, short positions across all portfolios returned a modest gain of 2.5%, while long positions sustained an average loss of 5.8%. When aggregated, the performance of all portfolios relative to the FTSE 10, averages a negative return of 7.4%.

5.1.7 Net tone

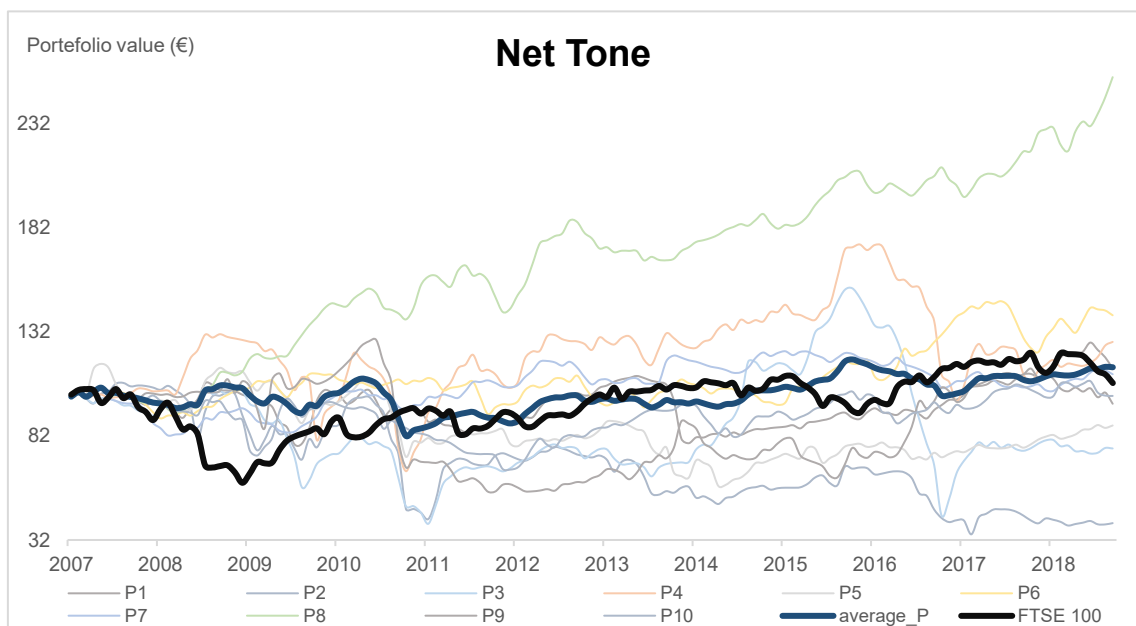


Figure 7: Net tone (entire document)

Considering only net tone measures as trading ground, an examination of the ten portfolios revealed that six of them exhibited positive returns. When paired against the benchmark five of them outperformed it. Notably, P8 emerges as a noteworthy outlier, yielding a remarkable return of 161.5%, representing an outperformance of the index by a substantial 136.6%. It is important to emphasize

that Portfolio 8 incurred losses on its long positions but was able to capitalize on declining companies through short positions, benefiting by the short part of the portfolio by 347.5%.

Analysing the long positions individually, we observe an average loss of 9.7%, with only three portfolios generating positive returns. On the other hand, the short positions delivered an average return of 39.6%, with just three portfolios reporting losses.

When considering the collective performance of all portfolios, the average return was 15.0%. This average return surpasses the benchmark by an impressive margin of 7.0%. It is worth highlighting that Portfolio 8 exerts a substantial influence on this average. Nevertheless, it is noteworthy that a majority of six portfolios achieved positive returns, representing the highest number among all the narrative metrics considered.

5.1.8 Summary

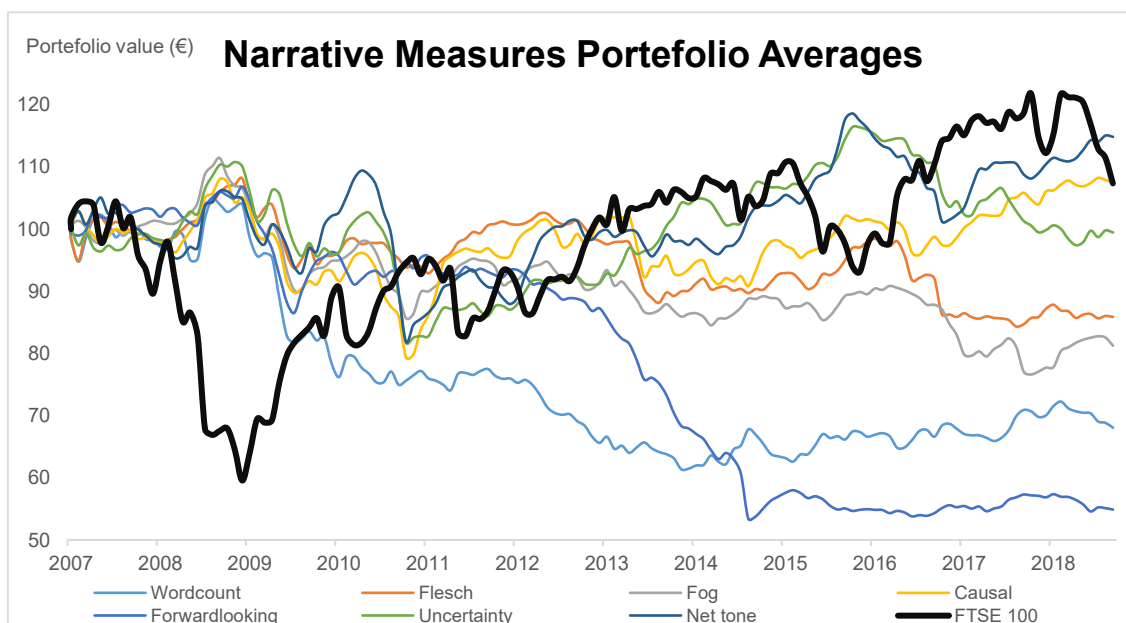


Figure 8: Portfolio averages (entire document)

average_P	abr/07	dez/07	dez/08	dez/09	dez/10	dez/11	dez/12	dez/13	dez/14	dez/15	dez/16	dez/17	dez/18
Wordcount	100	99	104	84	75	78	69	64	67	66	67	71	68
Cum_returns	n.a.	(0,5%)	3,9%	(16,5%)	(25,0%)	(22,5%)	(30,9%)	(36,2%)	(33,1%)	(33,8%)	(33,4%)	(29,1%)	(31,9%)
alpha	n.a.	(2,4%)	53,9%	(0,8%)	(19,5%)	(10,7%)	(26,3%)	(38,8%)	(35,5%)	(32,3%)	(39,6%)	(40,2%)	(36,6%)
Flesch	100	100	106	97	95	101	101	90	90	95	92	85	86
Cum_returns	n.a.	(0,4%)	5,9%	(2,7%)	(4,8%)	0,6%	1,4%	(10,2%)	(9,6%)	(5,0%)	(8,0%)	(15,3%)	(14,1%)
alpha	n.a.	(2,2%)	56,8%	15,5%	2,3%	15,9%	8,1%	(13,9%)	(12,9%)	(2,8%)	(16,6%)	(28,6%)	(20,0%)
Fog	100	100	111	93	90	95	92	88	89	89	88	77	81
Cum_returns	n.a.	(0,2%)	11,4%	(7,3%)	(10,5%)	(5,2%)	(7,8%)	(12,1%)	(11,2%)	(10,7%)	(12,2%)	(22,9%)	(18,7%)
alpha	n.a.	(2,0%)	65,0%	10,1%	(3,9%)	9,2%	(1,7%)	(15,6%)	(14,4%)	(8,6%)	(20,4%)	(35,0%)	(24,3%)
Causal	100	100	108	92	86	97	99	96	93	102	99	106	108
Cum_returns	n.a.	(0,3%)	8,1%	(8,4%)	(14,0%)	(3,1%)	(0,7%)	(4,3%)	(6,7%)	2,1%	(1,2%)	5,9%	7,6%
alpha	n.a.	(2,1%)	60,0%	8,8%	(7,6%)	11,6%	5,8%	(8,2%)	(10,0%)	4,5%	(10,5%)	(10,8%)	0,2%
Forwardlooking	100	103	106	93	93	93	89	73	54	55	54	57	55
Cum_returns	n.a.	2,7%	6,0%	(7,2%)	(6,6%)	(6,9%)	(11,2%)	(26,7%)	(46,2%)	(45,0%)	(45,7%)	(42,7%)	(45,1%)
alpha	n.a.	0,8%	56,9%	10,2%	0,4%	7,3%	(5,4%)	(29,7%)	(48,1%)	(43,7%)	(50,8%)	(51,7%)	(48,9%)
Uncertainty	100	99	110	98	90	86	92	101	107	115	110	101	99
Cum_returns	n.a.	(1,5%)	10,4%	(2,2%)	(10,5%)	(14,0%)	(7,6%)	0,6%	7,5%	14,5%	10,4%	1,3%	(0,5%)
alpha	n.a.	(3,3%)	63,4%	16,1%	(3,9%)	(0,9%)	(1,5%)	(3,5%)	3,7%	17,2%	0,0%	(14,6%)	(7,4%)
Net tone	100	101	106	97	90	91	101	99	104	118	106	109	115
Cum_returns	n.a.	0,7%	6,1%	(3,1%)	(9,8%)	(9,3%)	1,3%	(0,9%)	3,5%	17,6%	6,1%	8,8%	14,9%
alpha	n.a.	(1,2%)	57,1%	15,1%	(3,1%)	4,5%	8,0%	(4,9%)	(0,1%)	20,4%	(3,9%)	(8,3%)	7,0%
FTSE 100	100	102	68	84	93	87	94	104	104	98	110	119	107
Cum_returns	n.a.	1,9%	(32,5%)	(15,8%)	(6,9%)	(13,2%)	(6,2%)	4,3%	3,7%	(2,3%)	10,4%	18,6%	7,4%

Table 2: Portfolio averages (entire document)

Among the narrative measures calculated taking into consideration the whole annual report, the bottom performer was Forward-looking, which exhibited an average loss of 45.1% when compared to the FTSE 100 index. Subsequently, word

count measures also demonstrated a substantially negative performance, averaging a loss of 31.9%.

Following closely was the fog index, which registered an average loss of 24.3%, although possible to argue that the outcome could be somehow predicted, in line with all the critics surrounding this narrative and the application in finance, as noted in the literature review.

Furthermore, Flesch score also exhibited a negative performance, resulting in an average loss of 20.0%.

In the other spectrum, the portfolio constructed based upon "Uncertainty" measures yielded returns that approached neutrality, but when paired against the index, still experienced a loss of 7.4%.

Positive returns are observed when trading based on "Causal reasoning" measures, however, only achieving a marginal gain of 0.2% relative to the index on the closing date of our study.

Finally, the portfolio constructed upon Net tone measures was the better performer among the seven narrative measures studied. Notably, the analysis revealed that, on average, portfolios aligned with this narrative exhibited results closely in line with the index. This measure at the end of the holding period had outperformed the FTSE on average by 7.0%.

Although, it seems critical to highlight the performance of one outlier, within net tone measures. One particular portfolio managed to outperform the index by 136.6%, significantly boosting the overall average performance of all the ten portfolios. If excluded from the calculation, the average return would be negative of 2.1%.

5.2 Chairman statement

5.2.1 Word Count (Chair)

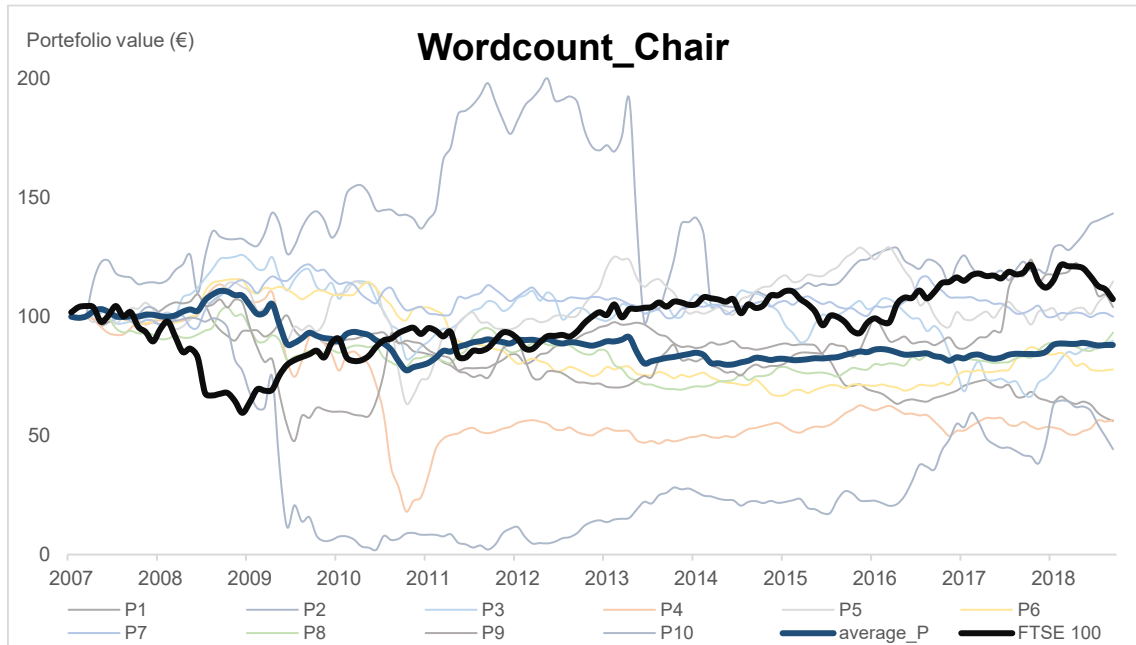


Figure 9: Word count (Chairman Statement)

When constructed upon the word count in the chairman's statement, four out of the ten portfolios demonstrated positive returns, with three portfolios surpassing the benchmark. Conversely, three portfolios experienced losses exceeding 47%, with the most adversely affected being P10, which incurred a substantial loss of 56.5%. Notably, in 2008, P10 witnessed a precipitous decline of 92.4%, primarily attributed to a series of backfiring short positions. Despite positive returns from long positions during this period, the significant losses from short positions overshadowed the overall performance.

Upon closer scrutiny of the top performer, P2, it is evident that the portfolio exhibited a robust start, particularly in the initial three years. A significant upswing was observed in 2011, followed by a notable downturn in 2013,

resulting in a 42.9% loss primarily influenced by adverse short positions, which had previously contributed substantially to its value.

Remarkably, when assessing the ten stocks, long positions incurred an average loss of 22.6%, while shorts yielded a positive return of 2.4%. Considering both positions, portfolios, on average, experienced a 10.1% loss of value, underperforming the index by 16.2%.

5.2.2 Flesch Score (Chair)

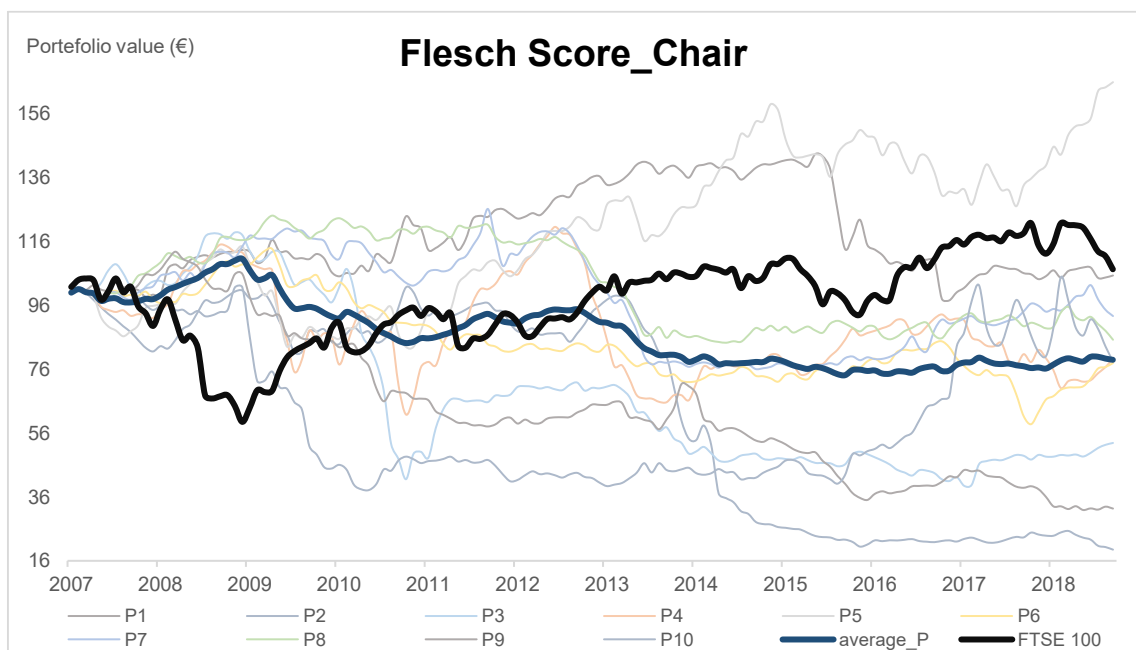


Figure 10: Flesch Score (Chairman Statement)

When it comes to portfolio construction based Flesch score calculation of the chairman statement, within the examined cohort of ten portfolios, only one portfolio, namely P5, manifested positive returns, achieving a noteworthy return of 63.9% and surpassing the benchmark by 54.6%. However, this singular success sharply contrasts with the prevailing pattern of substantial losses evident across the entire portfolio spectrum.

Among the nine portfolios that incurred losses, six experienced significant declines exceeding 25%, with three of them suffering losses surpassing 50%. Particularly notable were the severe losses incurred by P2 and P9, amounting to 81.9% and 67.0% respectively, dismantling value across both long and short positions, though with slightly greater magnitude observed in the short positions.

While P1 initially demonstrated promising returns by capitalizing on short positions, a reversal in 2015-2016 resulted in significant losses in short positions, thereby neutralizing the overall portfolio performance.

Aggregate analysis across all portfolios reveals an average loss of 27.9% on long positions and 14.4% on short positions. Consequently, the collective performance of the portfolios culminated in a substantial average loss of 26.1% relative to the FTSE 100, underscoring a pronounced erosion of value.

5.2.3 Fog Index (Chair)

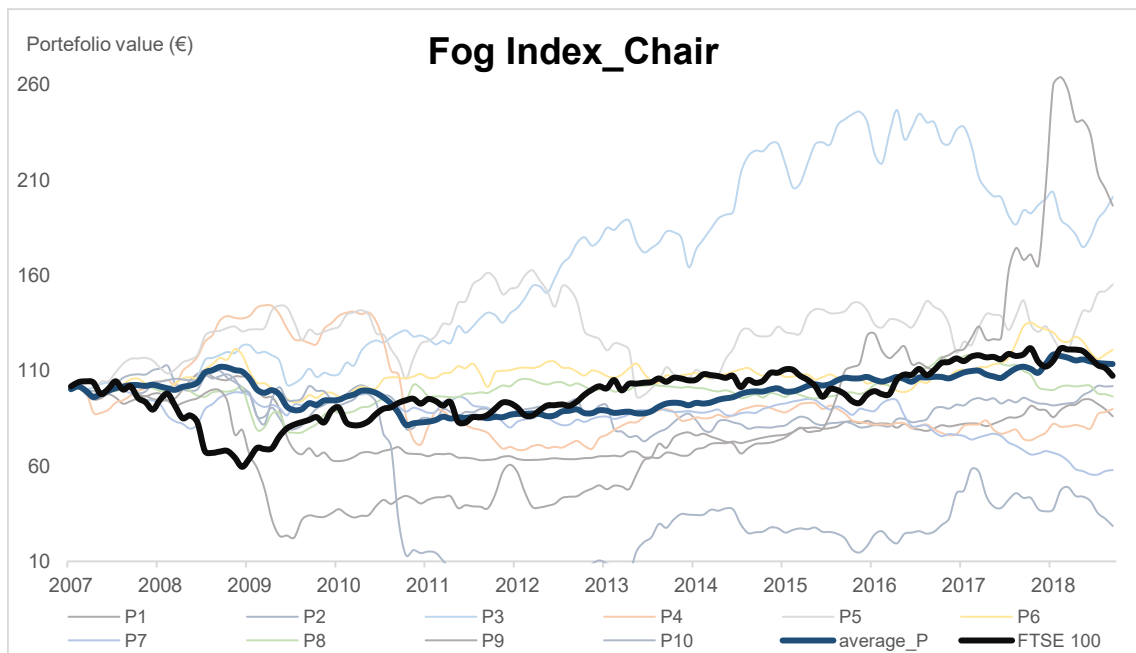


Figure 11: Fog Index (Chairman Statement)

Regarding a trading strategy based on fog index measures solely upon the chairman statement, five out of the ten portfolios exhibited positive returns, with only four surpassing the benchmark. Notably, the standout performers were P3 and P9, demonstrating returns of 104.1% and 92.5% respectively. P3 accrued value exclusively through short positions, while P9 derived its value exclusively from long positions portion of the portfolio.

Furthermore, P9, upon further inspection experienced, as previously stated significant benefits from its short positions, however, a notable setback occurred in 2017, resulting in losses attributed mostly to declining value in the short positions undertaken.

Conversely, the underperformers were identified as P7 and P10, incurring losses of 40.7% and 71.8% respectively. P7 witnessed negative returns in both long and short positions, whereas P10 sustained its losses exclusively from the short positions undertaken.

Despite the inherent volatility observed, the portfolios, on average, yielded a return of 13.8% over the holding period, surpassing the FTSE 100 index by 5.7%.

5.2.4 Causal Reasoning (Chair)

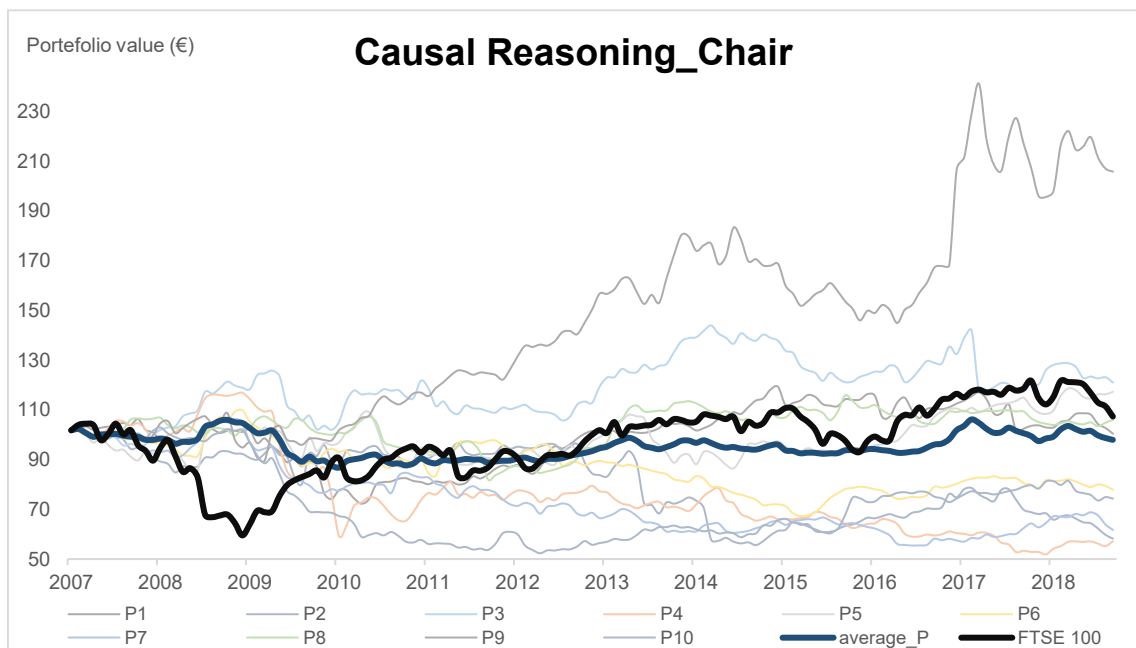


Figure 12: Causal Reasoning (Chairman Statement)

When considering a trading strategy built upon causal reasoning measures solely based on the chairman statement, four out of the ten portfolios demonstrated positive returns, with only three surpassing the benchmark. Notably, P9 emerged as the singular standout portfolio with positive returns, amassing a remarkable 91.6% return above the index. It is noteworthy that P9 achieved these gains primarily through long positions, garnering additional value from short positions. The year 2016 proved particularly favourable for P9, resulting in a substantial 31.5% return solely from long positions.

Conversely, on the opposite end of the performance spectrum, three portfolios—P2, P4, and P7—experienced losses exceeding 40% of their value. These losses were incurred across both short and long positions. Specifically, P4 suffered a notable decline of 34.5% in 2009, attributable mostly to short positions. Similarly, P2 encountered a loss of 36.9% in 2013, primarily due to the short positions undertaken.

In terms of averages, long positions yielded a return of 13.4%, while short positions incurred a loss of 18.6%. Consequently, when assessing overall performance, the portfolios, on average, incurred a 2.6% loss of value, demonstrating an underperformance relative to the benchmark of 8.8%.

5.2.5 Forward Looking (Chair)

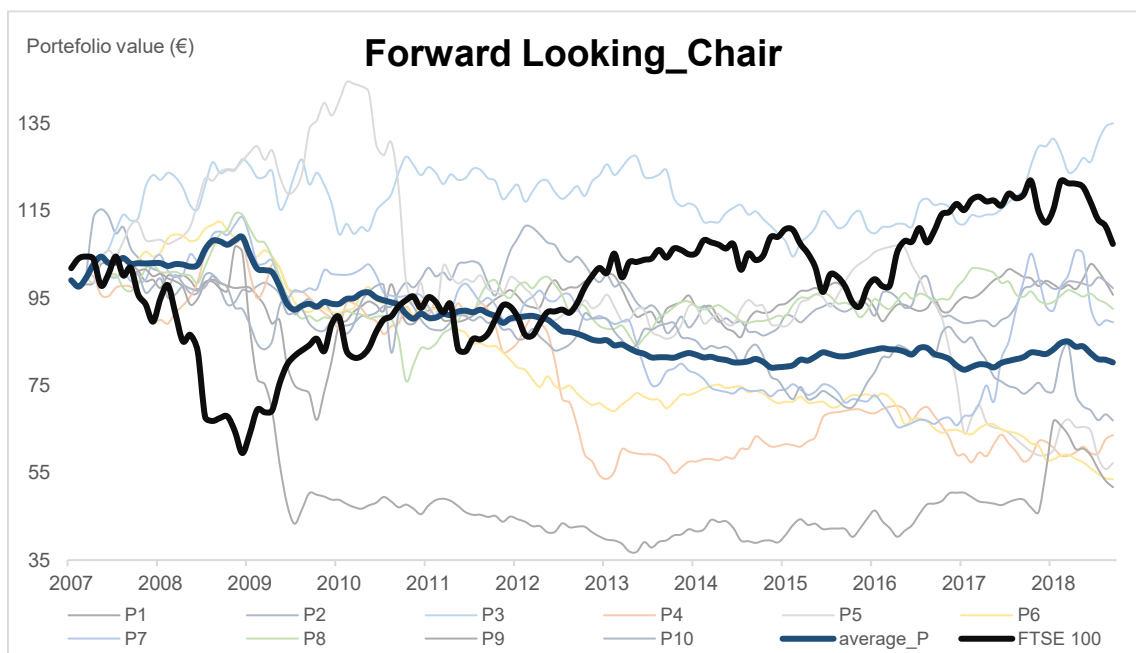


Figure 13: Forward looking (Chairman Statement)

When it comes to forward-looking measures, focusing exclusively on the chairman's statement, a solitary portfolio, P3, exhibited positive returns, surpassing the FTSE 100 by 25.7%. Noteworthy is the fact that P3's performance was primarily attributed to gains from short positions, set against losses incurred on long positions.

Contrarily, the remaining nine portfolios collectively recorded an average loss of 25.6%, with this decline observed symmetrically in both long and short

positions. It is noteworthy that four portfolios sustained losses exceeding 40% of their value.

The most substantial loss, evident in P9, stemmed from particularly unfavourable short positions in 2008, resulting in a halving of the portfolio's value, a setback from which it has been unable to recover since.

In a comprehensive evaluation of overall performance, long positions yielded a negative return of 17.1%, while short positions incurred a more significant loss of 21.8%. The average performance across all portfolios culminated in a value destruction of 19.5%, indicating an underperformance against the benchmark by 25.2%.

5.2.6 Uncertainty (Chair)

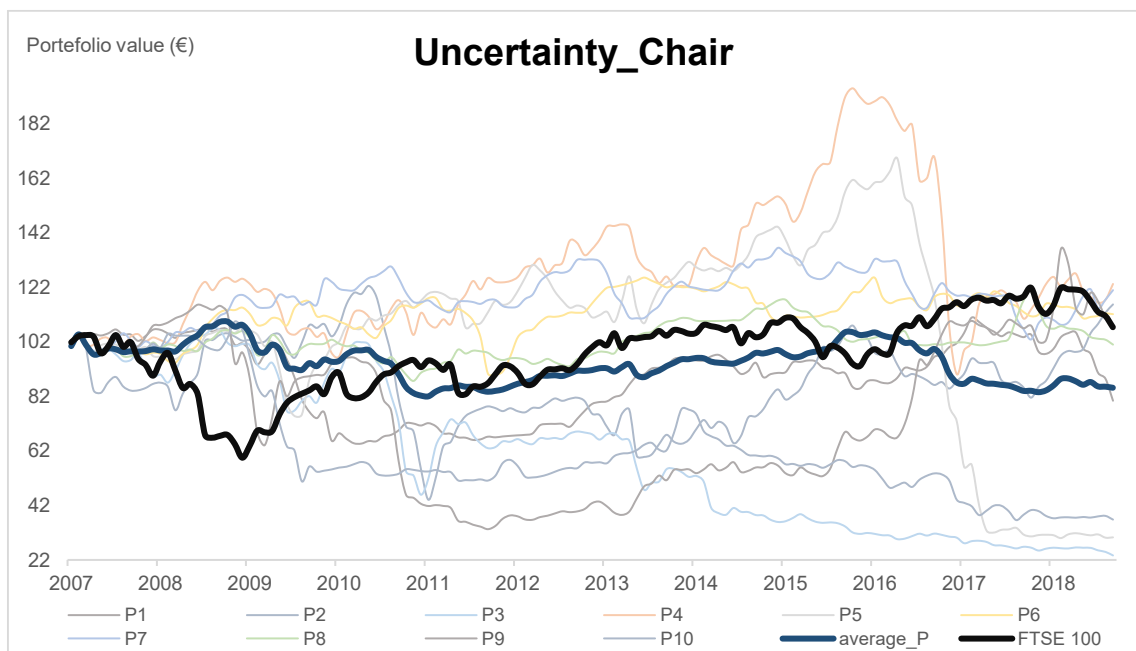


Figure 14: Uncertainty (Chairman Statement)

When examining uncertainty measures derived from chairman statements, notable losses were incurred by three portfolios: P3, P5, and P10, which experienced declines of 76.6%, 69.7%, and 63.4% of their respective values.

Intriguingly, only four out of the ten portfolios registered losses, yet the magnitude of these losses significantly influenced the overall average. Consequently, when considering the collective average, portfolios incurred a loss of 14.8%.

The portfolios that performed the worst experienced value erosion in both long and short positions, with the magnitudes of losses being almost equivalent. On average, long positions incurred a 15.5% loss, while short positions experienced a loss of 14.1%.

In comparison to the FTSE 100, portfolios, on average, underperformed by 20.7%, although it is crucial to note that this average is notably impacted by the performance of the bottom-performing portfolios.

5.2.7 Net tone (Chair)

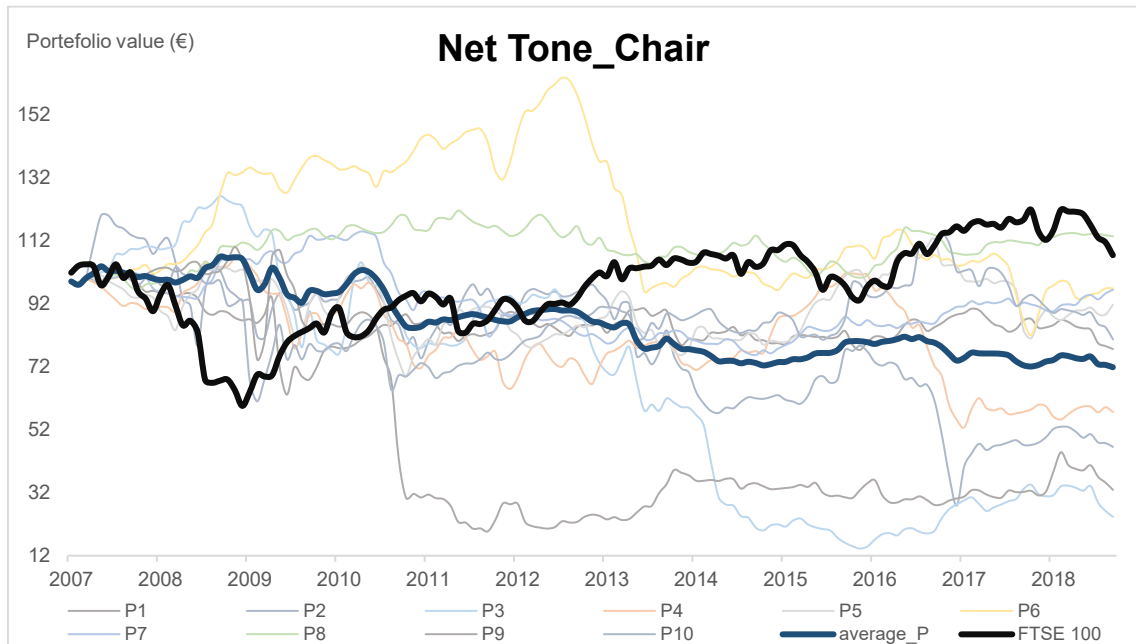


Figure 15: Net tone (Chairman Statement)

In the analysis of net sentiment portfolios derived exclusively from chairman statements, it becomes evident that only one portfolio managed to outperform

the index, although by a marginal margin of 5.6%. Notably, this outperforming portfolio, P8, stood as the sole contributor to positive returns, amounting to 13.4%.

On the other hand, one of the portfolios, P6, displayed promising initial performance but encountered a setback in 2012, incurring a 41.4% loss primarily attributable to unfavourable short positions.

Two portfolios, P3 and P9, emerged as significant underperformers, experiencing losses in both long and short positions, with the latter positions exhibiting a more pronounced magnitude.

Upon an assessment of averages, long positions incurred a 17.5% loss of value, while short positions suffered a more substantial decline of 38.4%. In aggregate, the portfolios, on average, sustained a 27.9% loss of value, thereby underperforming the benchmark by 33.1%.

5.2.8 Summary (Chair)

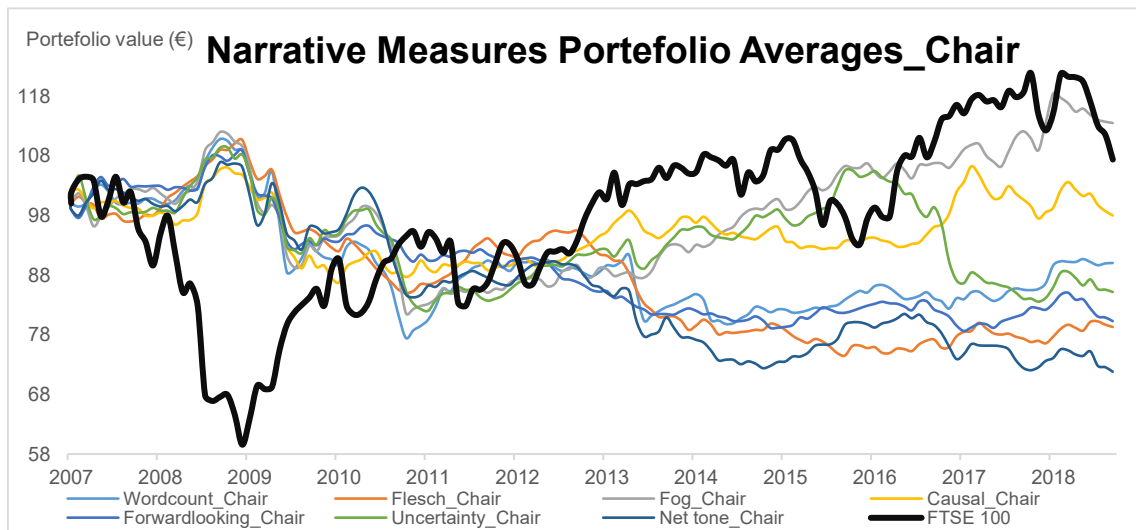


Figure 16: Portfolio averages (Chairman Statement)

average_P	abr/07	dez/07	dez/08	dez/09	dez/10	dez/11	dez/12	dez/13	dez/14	dez/15	dez/16	dez/17	dez/18
Wordcount_Chair	100	100	111	93	81	90	89	83	83	84	84	86	90
Cum_returns	n.a.	0,1%	10,8%	(6,8%)	(18,7%)	(9,6%)	(11,0%)	(17,4%)	(17,3%)	(16,1%)	(16,0%)	(14,5%)	(10,0%)
alpha	n.a.	(1,8%)	64,0%	10,7%	(12,7%)	4,2%	(5,2%)	(20,7%)	(20,3%)	(14,2%)	(23,9%)	(27,9%)	(16,2%)
Flesch_Chair	100	97	109	96	86	94	95	81	79	74	77	77	79
Cum_returns	n.a.	(3,0%)	9,0%	(4,3%)	(14,4%)	(5,8%)	(4,5%)	(19,0%)	(21,2%)	(25,5%)	(22,8%)	(22,7%)	(20,7%)
alpha	n.a.	(4,8%)	61,3%	13,6%	(8,1%)	8,5%	1,8%	(22,3%)	(24,0%)	(23,8%)	(30,1%)	(34,8%)	(26,1%)
Fog_Chair	100	102	112	93	88	85	90	93	99	106	107	112	114
Cum_returns	n.a.	2,5%	12,0%	(7,1%)	(12,1%)	(14,9%)	(10,3%)	(7,3%)	(0,8%)	6,3%	6,8%	12,1%	13,5%
alpha	n.a.	0,6%	65,8%	10,4%	(5,6%)	(2,0%)	(4,4%)	(11,1%)	(4,4%)	8,8%	(3,2%)	(5,5%)	5,7%
Causal_Chair	100	99	106	91	88	89	92	95	94	94	96	101	98
Cum_returns	n.a.	(0,8%)	5,7%	(8,8%)	(11,6%)	(11,2%)	(7,9%)	(4,9%)	(6,0%)	(6,3%)	(3,9%)	0,6%	(2,0%)
alpha	n.a.	(2,6%)	56,5%	8,4%	(5,0%)	2,3%	(1,8%)	(8,8%)	(9,4%)	(4,1%)	(13,0%)	(15,2%)	(8,8%)
Forwardlooking_C	100	103	108	94	93	92	87	81	81	82	82	81	80
Cum_returns	n.a.	2,9%	7,9%	(6,3%)	(6,6%)	(8,5%)	(13,1%)	(18,5%)	(18,9%)	(18,3%)	(17,8%)	(18,5%)	(19,7%)
alpha	n.a.	0,9%	59,8%	11,3%	0,3%	5,4%	(7,4%)	(21,8%)	(21,8%)	(16,4%)	(25,5%)	(31,3%)	(25,2%)
Uncertainty_Chair	100	99	109	94	90	84	91	93	98	106	99	84	85
Cum_returns	n.a.	(1,5%)	9,3%	(5,8%)	(9,9%)	(16,3%)	(8,6%)	(6,9%)	(2,1%)	5,5%	(0,6%)	(16,0%)	(14,8%)
alpha	n.a.	(3,3%)	61,9%	11,9%	(3,2%)	(3,6%)	(2,6%)	(10,7%)	(5,5%)	8,0%	(10,0%)	(29,2%)	(20,7%)
Net tone_Chair	100	101	107	96	88	88	89	81	73	80	80	72	72
Cum_returns	n.a.	1,1%	6,9%	(4,0%)	(12,1%)	(12,5%)	(10,6%)	(19,1%)	(26,9%)	(20,3%)	(20,4%)	(27,6%)	(28,2%)
alpha	n.a.	(0,8%)	58,2%	14,0%	(5,6%)	0,8%	(4,7%)	(22,4%)	(29,5%)	(18,5%)	(27,9%)	(39,0%)	(33,1%)
FTSE 100	100	102	68	84	93	87	94	104	104	98	110	119	107
Cum_returns	n.a.	1,9%	(32,5%)	(15,8%)	(6,9%)	(13,2%)	(6,2%)	4,3%	3,7%	(2,3%)	10,4%	18,6%	7,4%

Table 3: Portfolio averages (Chairman Statement)

When considering the narrative measures exclusively targeting the Chairman's statement within the annual reports, the Net Tone emerged as the least favourable performer, demonstrating an average loss of 33.1% in comparison to the FTSE 100 index. Subsequently, the Flesch Score also exhibited a considerable negative performance, averaging a loss of 20.7%, resulting in an underperformance against the benchmark of 26.1%.

Closely following were the forward-looking measures, recording an average loss of 25.2% when set against with the index. The portfolio constructed based on uncertainty measures, initially promising, experienced a downturn in 2016 due to unfavourable short positions, resulting in underperformance against the benchmark by 20.7%.

The portfolio constructed upon word count measures also witnessed value erosion, falling short of the FTSE 100 on average by 16.2%. On the contrary, the portfolio based on causal reasoning measures approached a neutral return, yet when compared to the index, still incurred a loss of 8.8%.

Notably, the portfolio constructed upon the Fog Index emerged as the more favourable performer among the seven narrative measures, despite acknowledged and consistent criticisms of the metric and its application in finance. The analysis revealed that, on average, portfolios aligned with this narrative closely mirrored the index. At the conclusion of the holding period, this measure outperformed the FTSE on average by 5.7%, reflecting a return of 13.8%.

However, it is crucial to underscore the inherent volatility within the portfolios, characterized by two significant winners that substantially influenced the average. Simultaneously, certain portfolios experienced substantial value erosion. Importantly, when considering the median, the return turns negative by 7.6%.

5.3 CEO statement

5.3.1 Word Count (CEO)

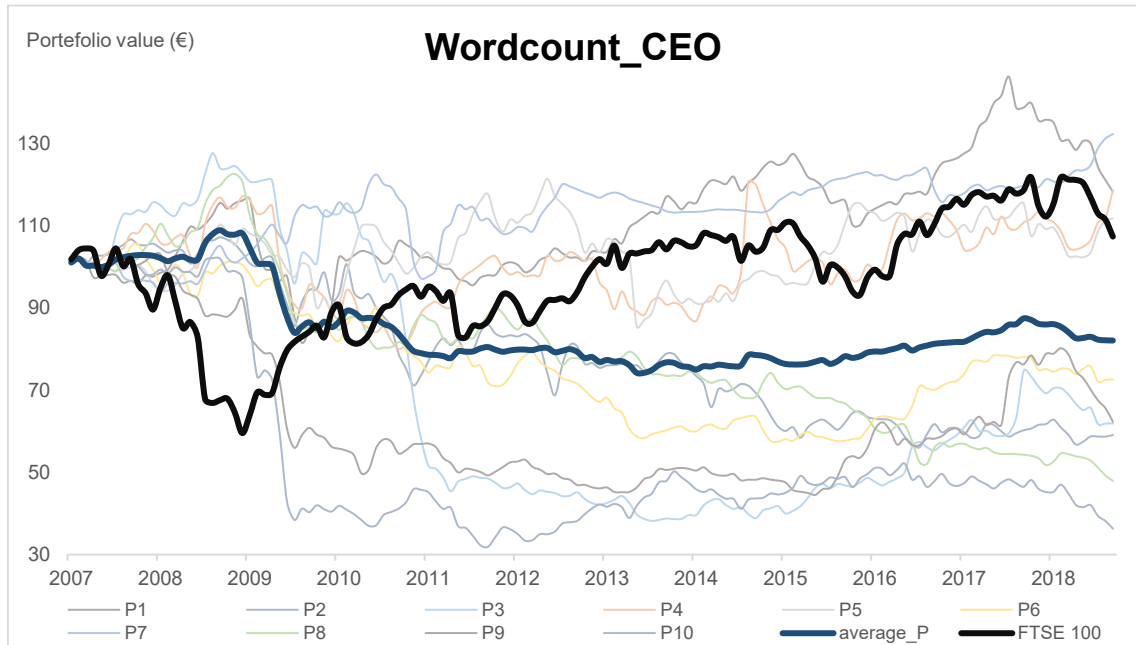


Figure 17: Word count (CEO Statement)

When it comes to portfolio construction upon of word count considering exclusively the CEO's statement within the MD&A segment, four out of the ten portfolios exhibited positive returns, concurrently surpassing the benchmark. Conversely, five portfolios incurred losses exceeding 40% of their value. Notably, the lowest performer, P10, experienced a staggering 99.0% loss in the value of the short segment of the portfolio, nearly eradicating the entire short portion of the portfolio.

Analysing the average performance across all portfolios, long positions demonstrated a notable downturn with an average loss of 29.7% of their value. In contrast, short positions incurred a comparatively modest loss of 7.0%. The consolidation of both positions resulted in an average loss of 18.3% of the portfolio's value, indicating an underperformance against the FTSE 100 by 23.5%.

5.3.2 Flesch Score (CEO)

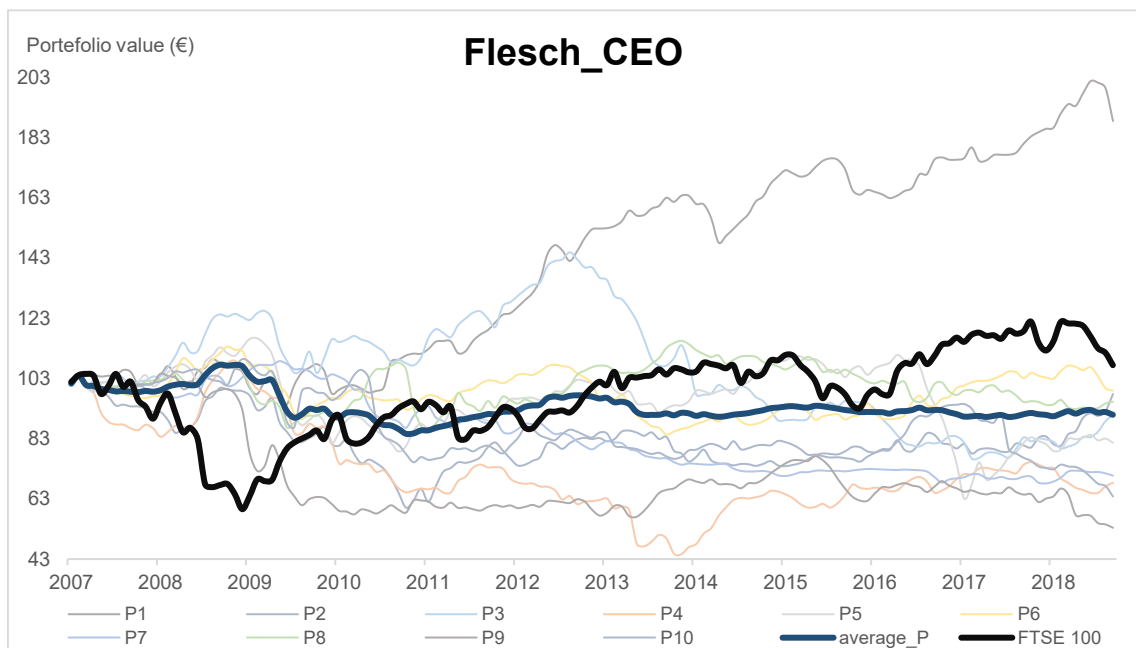


Figure 18: Flesch Score (CEO Statement)

In the examination of the Flesch score based on CEO statements, a notable standout portfolio emerges, namely P1, which recorded a remarkable return of 87.4%, surpassing the benchmark by 75.5%. While a portion of the value in this particular portfolio was derived from long positions, the predominant source of value derived from short positions, contributing a substantial 131.3% return on the short side of the portfolio.

Only one additional portfolio generated positive returns, however, it failed to outperform the FTSE 100. When considering the overall performance of the portfolios their value erosion averaged 7.5%, primarily attributable to negative returns on long positions, which averaged (16.8)%, while short positions yielded positive, albeit marginal, returns of 1.7%. Against the FTSE 100 Flesch score portfolio underperformed it by 13.8%.

Also, it is important to note if excluded P1, the average portfolio return would be of negative 18,1%, where short positions lost on average 12.7% and longs 23.5%.

5.3.3 Fog Index (CEO)

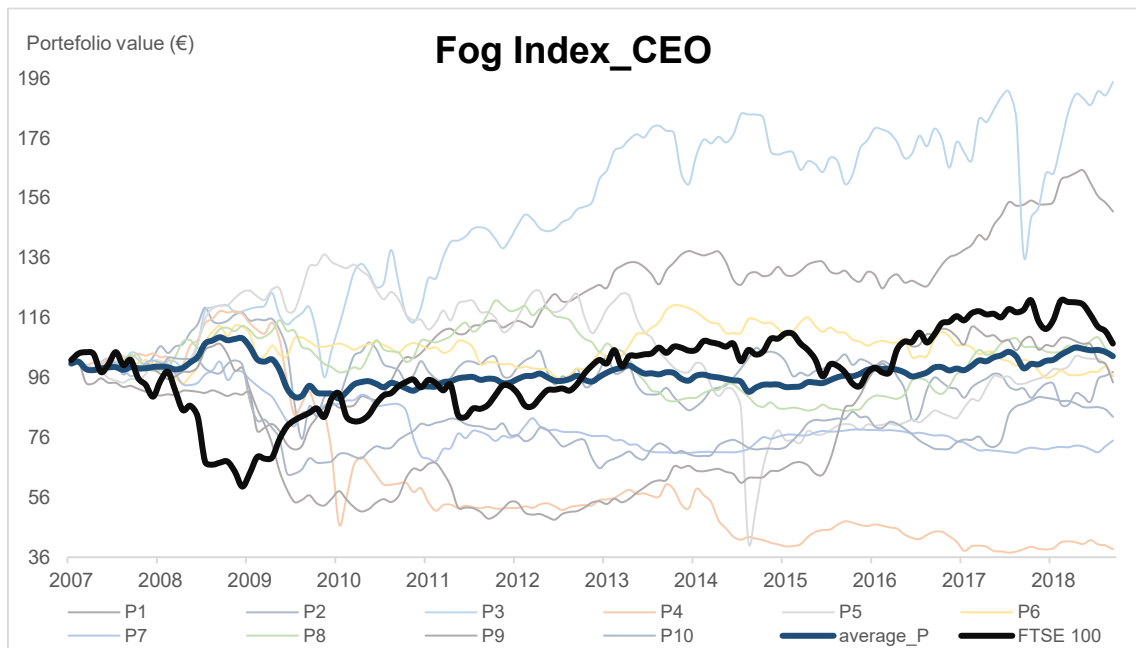


Figure 19: Fog Index (CEO Statement)

When it comes to portfolio construction upon Fog Index as a measure for portfolio creation based solely upon the CEO statement, a limited number of portfolios, specifically three, recorded positive returns. Notably, two portfolios, P1 and P3, demonstrated considerable high returns of 51.6% and 89.6% respectively. The robust performance of P1 is primarily attributed to the success of long positions, while P3 derived its value from short positions. Despite three portfolios yielding positive returns, only the aforementioned two were able to outperform the FTSE 100.

When considering the average returns, the portfolios collectively returned 2.3%, although this figure is significantly skewed by the exceptional performance

of the top portfolios. Excluding these top performers, the average returns turn negative to 14.7%. In comparison to the index, portfolios, on average, underperformed by 3.9%. These findings shed light on the diverse outcomes and challenges associated with employing the Fog Index as a measure for portfolio creation.

5.3.4 Causal Reasoning (CEO)

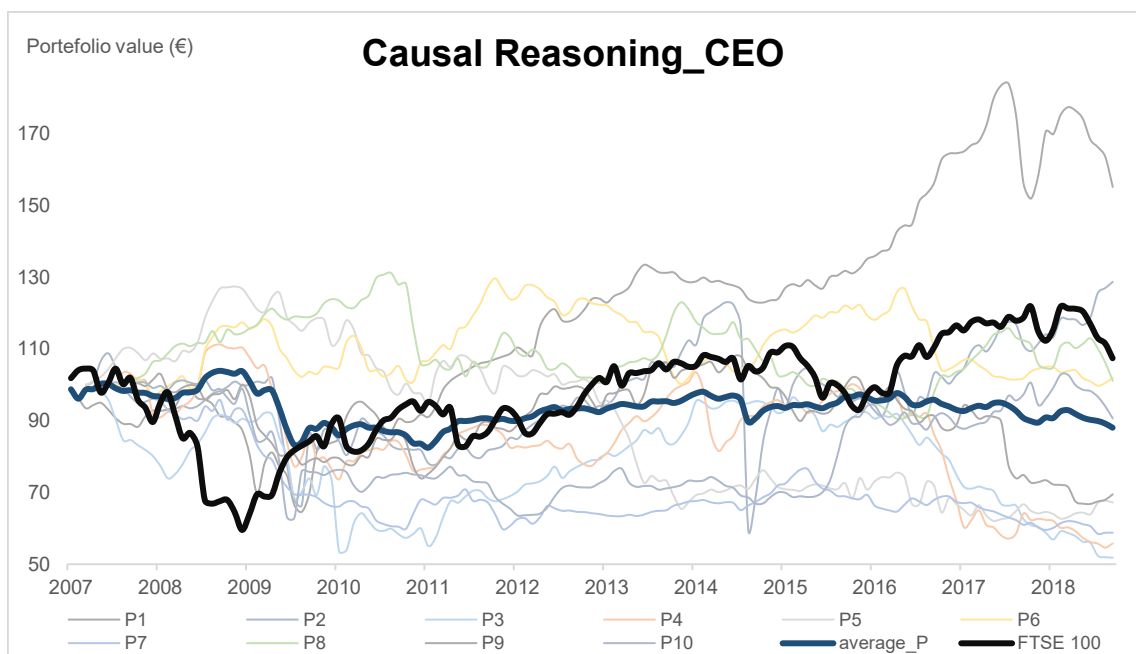


Figure 20: Causal reasoning (CEO Statement)

When it comes to causal reasoning measures, specifically analysing the CEO statements, only three portfolios yielded positive returns, with merely two surpassing the FTSE 100. Notably, one standout performer was P1, which achieved substantial value, particularly in 2016, primarily attributed to the success of long positions.

Conversely, two portfolios, namely P3 and P4, emerged as noticeable underperformers, reporting value erosion of 48.9% and 45.3%, respectively, experiencing losses in both long and short positions of comparable magnitude.

When examining the average performance across all portfolios, long positions incurred a loss of 14.8%, and short positions suffered a 10.1% decline in value. Consequently, the portfolios, on average, lost 12.5%, translating to an 18.0% underperformance against the benchmark.

5.3.5 Forward Looking (CEO)

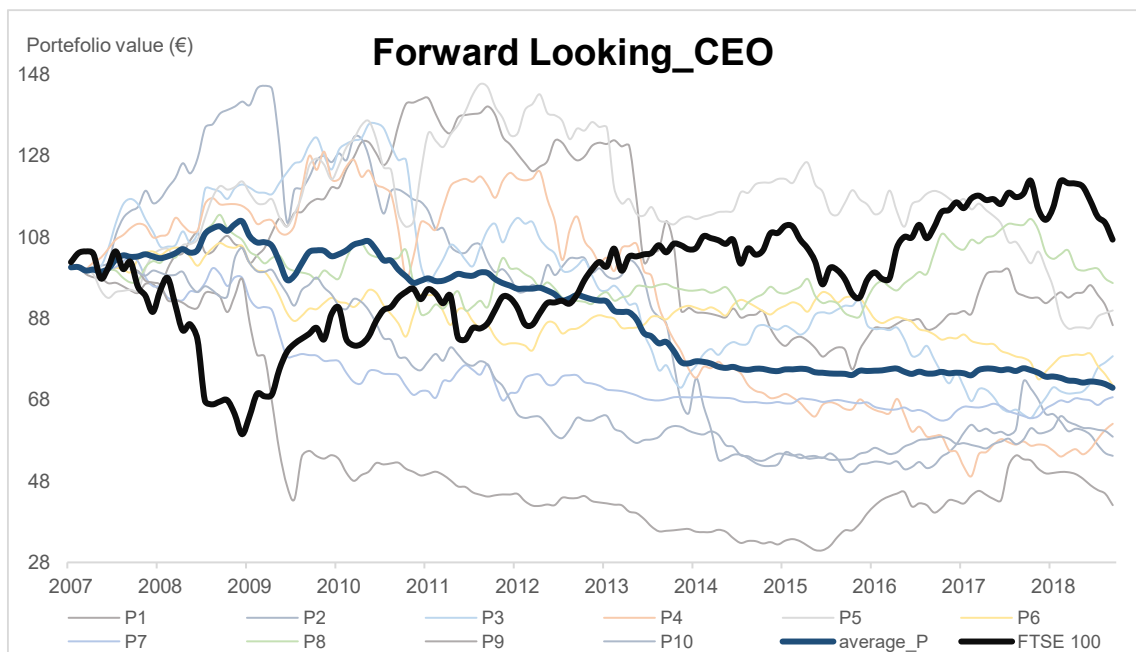


Figure 21: Forward looking (CEO Statement)

In terms of forward-looking measures exclusively derived from CEO statements for portfolio construction, all 10 portfolios exhibited negative returns, culminating in an average loss of 29.4%. Specifically, long positions witnessed an average decline of 40.0%, while short positions incurred a loss of 18.8%.

Examining the poorest performing portfolio, P1, which incurred a 58.6% loss of its value, revealed that the majority of losses were attributed to short positions. In 2008, the short positions suffered a 67.3% loss, and in 2012, a 51.8% loss. Coupled with losses in long positions, this led to an overall underperformance of 60.8% when measured against the benchmark.

Relative to the index, the portfolios, on average, underperformed by 34.0%.

5.3.6 Uncertainty (CEO)

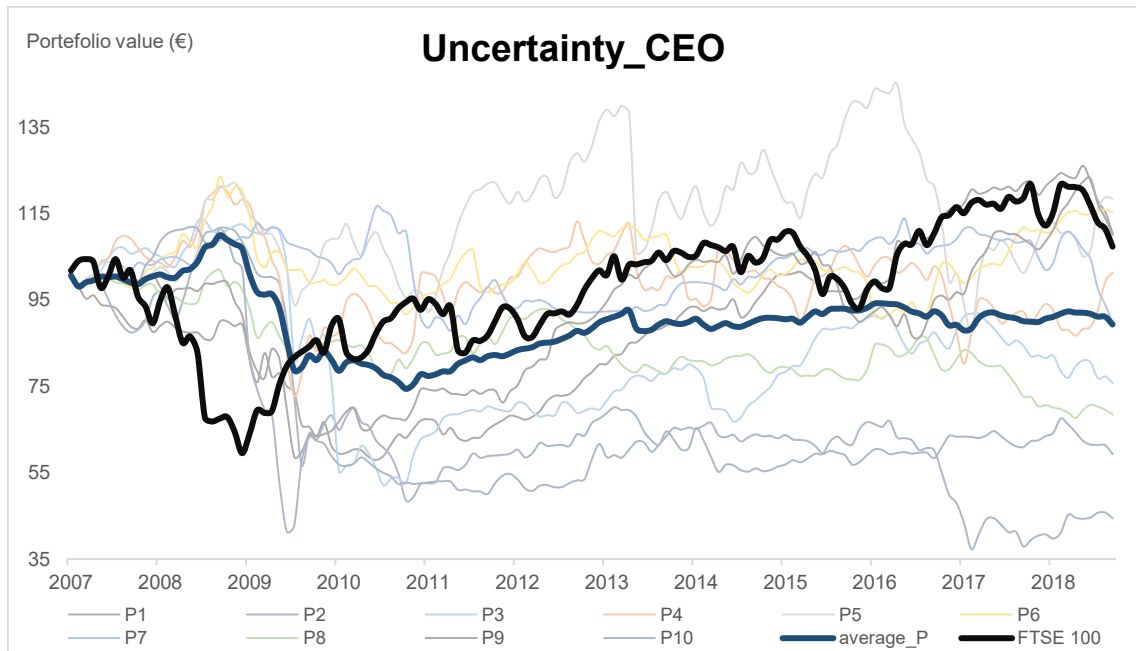


Figure 22: Forward looking (CEO Statement)

In the realm of uncertainty measures derived from CEO statements, five out of the ten portfolios demonstrated positive returns, although only four managed to marginally surpass the FTSE100 index. Notably, two portfolios, P2 and P10, incurred substantial losses of 55.9% and 40.9% respectively. P2 witnessed a decline in value across both short and long positions, while in the case of P10, although losses occurred in the long positions, the predominant impact was from the short positions.

In terms of averages, the portfolios experienced an 11.1% loss of value, with declines evident in both long and short positions. Consequently, the portfolios, on average, underperformed the benchmark by 16.8%.

5.3.7 Net tone (CEO)

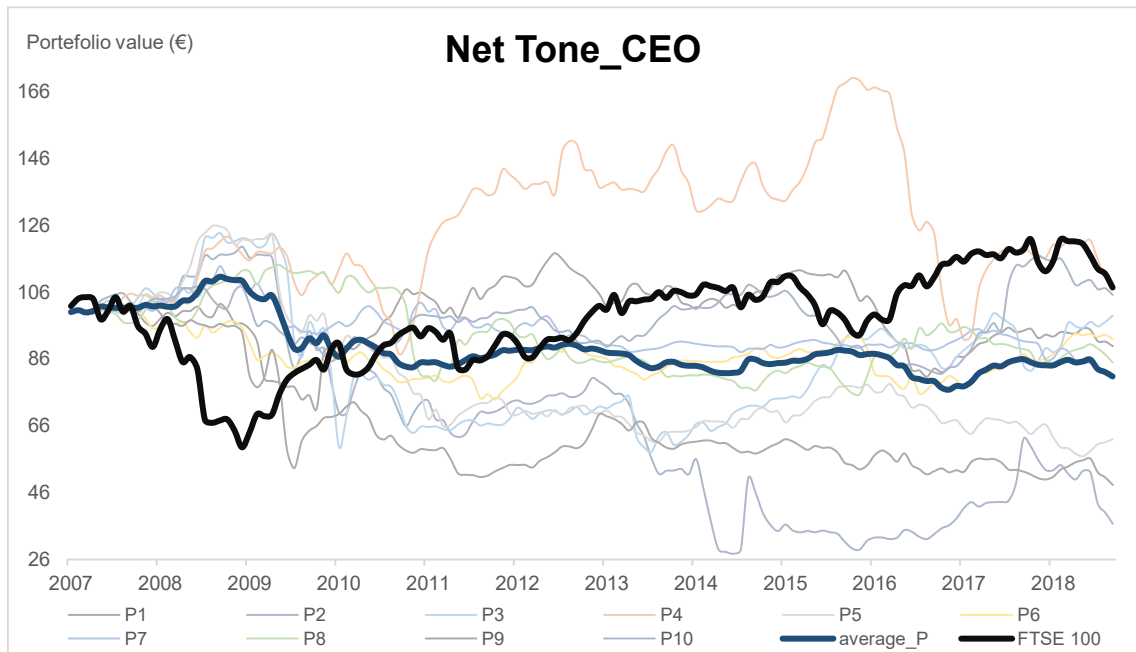


Figure 23: Net tone (CEO Statement)

When looking into net tone measures derived exclusively from CEO statements, only two portfolios exhibited positive returns, with only one, P4, managing to surpass the benchmark, albeit by a narrow margin of 1.1%. P4 demonstrated a robust performance in 2010 and 2011 through long and short positions, respectively. However, a setback occurred in 2011, resulting in a 30.6% loss attributed to a series of short positions, bringing its performance in close proximity to the FTSE100.

Upon analysing the averages, the portfolios collectively incurred a 19.4% loss of value, with losses evident in both long and short positions to an identical extent. Consequently, the portfolios, on average, underperformed the index by 24.8%.

5.3.8 Summary (CEO)

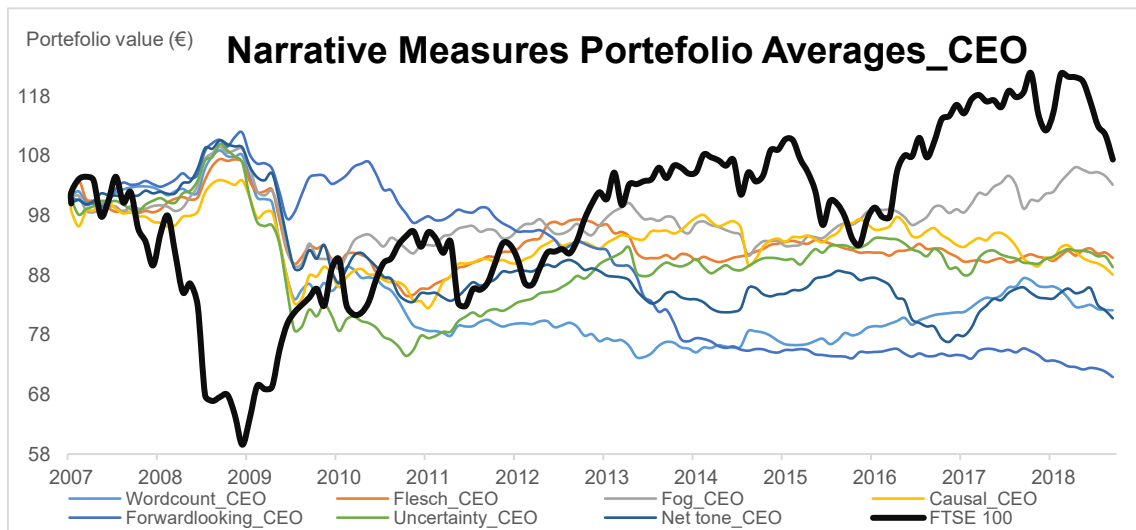


Figure 24: Portfolio averages (CEO Statement)

average_P	abr/07	dez/07	dez/08	dez/09	dez/10	dez/11	dez/12	dez/13	dez/14	dez/15	dez/16	dez/17	dez/18
Wordcount_CEO	100	103	109	87	84	80	79	77	79	78	81	87	82
Cum_returns	n.a.	2,7%	8,9%	(13,5%)	(16,3%)	(19,5%)	(20,6%)	(23,2%)	(21,4%)	(21,7%)	(18,7%)	(12,5%)	(17,9%)
alpha	n.a.	0,8%	61,2%	2,8%	(10,1%)	(7,3%)	(15,4%)	(26,4%)	(24,2%)	(19,9%)	(26,4%)	(26,3%)	(23,5%)
Flesch_CEO	100	99	107	93	86	91	97	91	92	92	92	91	91
Cum_returns	n.a.	(1,5%)	7,5%	(7,1%)	(13,8%)	(8,9%)	(2,7%)	(8,6%)	(8,0%)	(7,5%)	(7,5%)	(8,9%)	(9,1%)
alpha	n.a.	(3,3%)	59,1%	10,3%	(7,5%)	5,0%	3,8%	(12,4%)	(11,3%)	(5,4%)	(16,2%)	(23,2%)	(15,3%)
Fog_CEO	100	100	110	93	94	96	97	98	93	97	99	99	103
Cum_returns	n.a.	(0,4%)	9,5%	(6,7%)	(6,4%)	(4,3%)	(3,4%)	(2,2%)	(7,3%)	(3,3%)	(0,6%)	(0,8%)	3,2%
alpha	n.a.	(2,3%)	62,1%	10,8%	0,5%	10,2%	3,0%	(6,2%)	(10,6%)	(1,0%)	(9,9%)	(16,4%)	(3,9%)
Causal_CEO	100	99	104	88	87	91	93	95	91	97	96	91	88
Cum_returns	n.a.	(1,5%)	3,9%	(12,1%)	(13,2%)	(9,3%)	(6,6%)	(4,5%)	(9,3%)	(3,3%)	(4,1%)	(9,2%)	(11,9%)
alpha	n.a.	(3,3%)	53,9%	4,4%	(6,8%)	4,4%	(0,4%)	(8,4%)	(12,5%)	(1,1%)	(13,1%)	(23,5%)	(18,0%)
Forwardlooking_C	100	103	111	104	100	99	94	82	76	74	74	76	71
Cum_returns	n.a.	3,2%	10,7%	4,3%	(0,0%)	(0,8%)	(6,3%)	(17,8%)	(24,4%)	(25,6%)	(25,6%)	(24,3%)	(29,1%)
alpha	n.a.	1,2%	63,8%	23,9%	7,4%	14,2%	(0,1%)	(21,2%)	(27,1%)	(23,9%)	(32,6%)	(36,2%)	(34,0%)
Uncertainty_CEO	100	99	110	82	76	82	88	90	90	93	92	90	89
Cum_returns	n.a.	(1,0%)	10,0%	(17,8%)	(24,0%)	(18,0%)	(12,2%)	(9,9%)	(9,8%)	(7,0%)	(7,6%)	(9,9%)	(10,7%)
alpha	n.a.	(2,8%)	62,8%	(2,4%)	(18,4%)	(5,6%)	(6,4%)	(13,6%)	(13,0%)	(4,8%)	(16,3%)	(24,1%)	(16,8%)
Net tone_CEO	100	101	111	92	85	87	90	85	86	88	79	86	81
Cum_returns	n.a.	1,3%	10,6%	(7,8%)	(15,2%)	(13,3%)	(10,0%)	(15,0%)	(14,1%)	(11,5%)	(20,6%)	(14,1%)	(19,2%)
alpha	n.a.	(0,5%)	63,8%	9,5%	(8,9%)	(0,2%)	(4,1%)	(18,5%)	(17,2%)	(9,4%)	(28,1%)	(27,6%)	(24,8%)
FTSE 100	100	102	68	84	93	87	94	104	104	98	110	119	107
Cum_returns	n.a.	1,9%	(32,5%)	(15,8%)	(6,9%)	(13,2%)	(6,2%)	4,3%	3,7%	(2,3%)	10,4%	18,6%	7,4%

Table 4: Portfolio averages (CEO Statement)

When it comes to the narrative measures exclusively derived from the CEO statements within the annual reports, the forward-looking measures emerged with the least favourable performance, revealing an average loss of 34.0% in comparison to the FTSE 100 index.

Subsequently, the Net Tone measures also demonstrated considerable negative performance, with an average loss, underperforming the benchmark by 24.8%.

Following closely in the evaluation were the word count measurements, which exhibited an average loss of 23.5% when juxtaposed with the index.

Moreover, the portfolio constructed upon causal reasoning measures faced value erosion, falling short of the FTSE 100 on average by 18.0%.

Similarly, the Flesch Score metrics underperformed the FTSE 100, resulting in significant value erosion and an underperformance against the benchmark by 15.3%.

Notably, the portfolio constructed upon the Fog Index managed to achieve marginal positive returns of 3.2%, but it still fell short of the index by 3.9%, despite consistent criticisms of the metric and its application in finance.

Once again, the influence of a top-performing portfolio that gained 81.4% of value contributes to the overall average. However, volatility remains prevalent, and when considering the median, the return turns negative at 10.1%.

6. Top and Bottom performers

This section will provide a more in-depth analysis of both the top two (Net tone and Fog_Chair) and bottom two (Forward looking and Wordcount) performers, aiming to understand the yearly catalysts driving their performance.

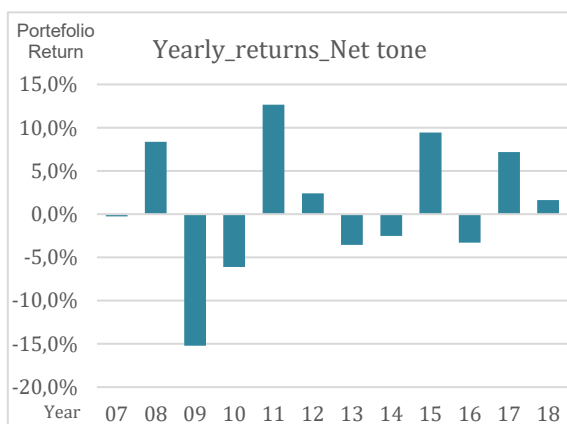


Figure 25: Top performer (average of portfolios)



Figure 26: Second best performer (average of portfolios)

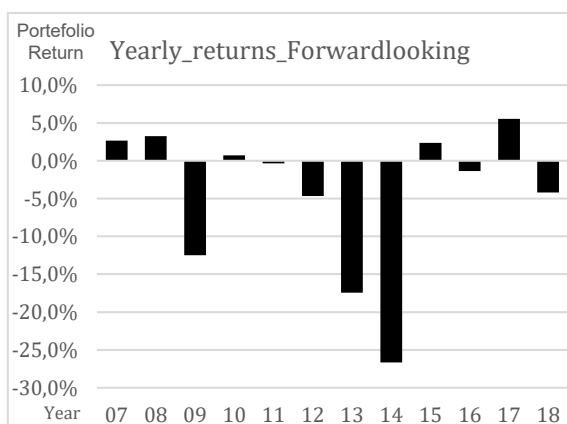


Figure 27: Bottom performer (average of portfolios)

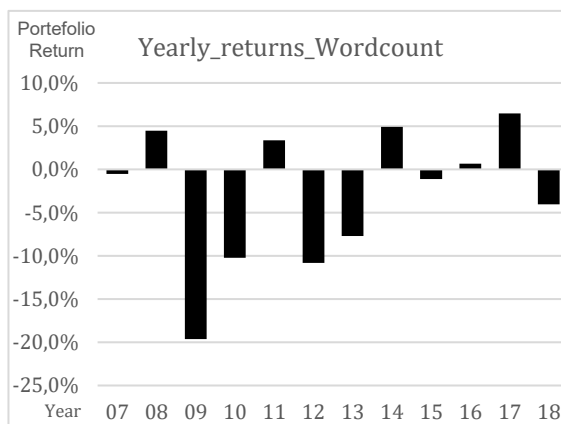


Figure 28: Second worst performer (average of portfolios)

Among the various narratives analysed, Net Tone emerged as the top performer, yielding an average return of 15.0% (7.0% above the index). The average of the portfolios started strong in 2008 capitalizing on short positions (12.7%). Although 2009 and 2010 the averages of the portfolios incurred in losses,

it recouped in 2011 through mostly successful short positions. After this, notably it didn't incur in losses higher than 3.5%, thus consolidating its value with another 9.7% gain in 2015 and 7.2% in 2017.

Six portfolios reported end out returning positive, with five outperforming the benchmark. Although it's important to stand-out one portfolio that achieved a return of 161.5% through successful long and short positions through the holding period, thus boosting the average significantly.

The second-best performer, Fog Index when considering the Chairman's statements, achieved a return of 13.8% (5.7% above the index). Following a trajectory like Net Tone, it started in 2008 with a 5.5% return, experienced losses in 2009 and 2010, mainly due to two portfolios with substantial losses in short positions. In 2012, it returned 11.7% through successful long positions and achieved a 13.6% return in 2015 by capitalizing on both short and long positions. Despite a 9.9% loss in 2016, a single portfolio boosted average returns by 5.6% in 2018 through short positions.

The worst performer was forward looking measures considering the entire report, with an average loss of 44.3%. In the depicted graph (Figure 27), is possible to notice that the decline was primarily concentrated within a consecutive three-year span—2012, 2013, and 2014. Notably, the most recent year, 2014, stands out with a staggering loss exceeding 25%. This loss in value is attributed to a series of significantly negative short positions, compounded by unfavourable long positions.

Wordcount, the second-worst performer when examining the complete annual report, experienced positive returns in 2008. However, in 2009, it incurred a substantial loss of approximately 20%, attributed to poorly chosen stocks that resulted in losses on both short and long positions. This negative trend persisted, with another 10% loss in 2010, maintaining these diminished levels until the end of the holding period.

7. Conclusions

There were investigated seven distinct narrative factors: Word count, Flesch score, Fog index, Causal reasoning, Forward-looking, Uncertainty, and Net tone. Each of the narratives were studied within different annual reports segments being considered: the overall report, the Chairman statement, and the CEO statement. The objective of the master thesis was to assess the performance of such measures in a real-world financial scenario, building upon existing literature that had already demonstrated their correlation with financial outcomes.

It is observable across all the narratives lies in the unique nature of portfolio composition, which involves both short and long positions. This composition inherently results in short positions serving as an effective hedge against market fluctuations. This phenomenon is particularly evident when we examine the performance of all portfolios during critical financial events, namely the sub-prime crisis and the 2015-2016 market selloff.

During the sub-prime crisis, is possible to see that all portfolios consistently outperformed the FTSE 100 index. A similar trend emerged during the 2015-2016 market selloff, where the portfolios, once again, demonstrated an overall positive performance during this year, primarily driven by the strength of short positions.

Although the research aimed to understand how narrative measured portfolios would perform against the benchmark, several factors stand in opposition to the reliability of taking such approach.

Firstly, even when focusing on the net tone measure on the whole annual report, it owes its performance primarily to a single portfolio. It's essential to consider that this could be attributed to sheer luck rather than a consistently effective strategy. If this outlier was excluded, the average portfolio performance would lead to negative returns and thus underperforming the benchmark index.

A similar observation can be made regarding the fog index applied to chairman statements, where it's important to emphasizing the intrinsic volatility within the portfolios. This volatility is characterized by two prominent winners that exerted a substantial influence on the average. Concurrently, certain portfolios witnessed notable value erosion. Of significant note is the fact that, when accounting for the median, the return turns negative.

Furthermore, a critical limitation of the study lies in the disregard of trading fees from the calculations. The returns by itself on the hypothetical world of no fees overall underperformed the index, adding the fees on transactions, being a considerable number of transactions per year per portfolio (with 42 stocks each), would also add downside pressure, eroding even further the already weak returns. Additionally, fees associated with maintaining a margin account, requested to maintain the short positions was not factored. These fees would accrue over the entire holding period of the portfolio and add an additional layer of cost.

Digging deeper into the matter of trading fees, it's important to recognize that they would encompass several extra costs, including brokerage commissions, bid-ask spreads, and market impact costs. Moreover, in the context of portfolios that involve short selling, the potential for incurring borrowing costs should not be overlooked.

Additionally, the issue of possible margin calls was not factored, meaning that in certain cases would be needed an additional injection capital to maintain the required margin levels, which adds another layer of pressure, even though considering the volatility without proper stop losses in place.

For future research, it could be interesting to understand the impact of implementing stop-loss mechanisms within the portfolio strategies. Upon a more in-depth analysis of the results, there are recorded instances where certain positions experienced extreme negative returns due to sudden surges in specific

stocks, resulting, in some cases, in returns plummeting to three, four, or even five times the initial investment on the particular stock. It is arguable that implementing stop-loss orders could help mitigate some of the downside volatility.

The addition of stop-losses could be a potential safeguard against steep losses by triggering the sale of positions when predefined price thresholds are breached. This could serve as a vital risk management tool, especially considering the extensive dataset examined, which included 2149 stocks, including micro and small caps with considerable betas.

Although, it is possible to argue that in some scenarios, there were observed surges followed by corrections, resulting in a near-zero returns scenario, so it can tackle both spectrums of the returns. This highlights the complexity of devising an optimal stop-loss strategy, as excessively tight levels may result in premature exits and missed opportunities, adding then another interesting topic for further research by modelling different thresholds for the stops and observe how the portfolios would behave.

Furthermore, an intriguing prospect for future research lies in the restriction of the sample by market capitalization. This strategic modification holds the potential to address lower betas, thereby mitigating the pronounced volatility observed throughout the study. This approach may contribute to the development of more cohesive portfolios, potentially avoiding instances where sudden shorts result in substantial value erosion.

To conclude, this study understands the informative potential of narrative measures in the realm of obtaining valuable insights of stocks. The measures can indeed provide a nuanced perspective that contributes to a more comprehensive understanding of the market.

However, the takeaway from the research is that relying exclusively upon narrative measures as the foundation for portfolio construction, under the

outlined strategy, is not viable. The analysis indicates that such an approach is prone to significant downturns in most cases. Therefore, while narratives offer valuable supplementary information, managing a portfolio requires a more diversified and robust approach that considers various factors, such as risk management techniques and broader market dynamics.

References

- Aerts, W., & Zhang, S. (2014). Management's causal reasoning on performance and earnings management. *European Management Journal*, 32(5), 770–783.
- Ali, A., Klasa, S., & Ye, Q. (2014). Industry concentration and corporate disclosure policy. *Journal of Accounting and Economics*, 58(2-3), 240-264.
- Alves, P., El-Haj, M., Rayson, P., Walker, M., & Young, S. (2016). Heterogeneous narrative content in annual reports published as PDF files: Extraction, classification, and incremental predictive ability. *SSRN Electronic Journal*.
- Asay, H. S., Libby, R., & Rennekamp, K. (2018). Firm performance, reporting goals, and language choices in narrative disclosures. *Journal of Accounting and Economics*, 65(2-3), 380–398.
- Baltussen, G., van Bekkum, S., & van der Grient, B. (2018). Unknown Unknowns: Uncertainty About Risk and Stock Returns. *Journal of Financial and Quantitative Analysis*, 53(4), 1615–1651.
- Bettman, J. R., & Weitz, B. A. (1983). Attributions in the Board Room: Causal Reasoning in Corporate Annual Reports. *Administrative Science Quarterly*, 28(2), 165.
- Beyer, A., Cohen, D. A., Lys, T. Z., & Walther, B. R. (2010). The financial reporting environment: Review of the recent literature. *Journal of Accounting and Economics*, 50(2-3), 296–343.

- Biddle, G. C., Hilary, G., & Verdi, R. S. (2009). How does financial reporting quality relate to investment efficiency? *Journal of Accounting and Economics*, 48(2-3), 112–131.
- Bloomfield, R. (2008). Discussion of “Annual report readability, current earnings, and earnings persistence.” *Journal of Accounting and Economics*, 45(2-3), 248–252.
- Brown, S., & Tucker, J. W. (2011). Large-Sample Evidence on Firms’ Year-over-Year MD&A Modifications. *Journal of Accounting Research*, 49(3), 309–346.
- 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.
- De Franco, G., Kothari, S. P., & Verdi, R. S. (2011). The Benefits of Financial Statement Comparability. *Journal of Accounting Research*, 49(4), 895–931.
- Ertugrul, M., Lei, J., Qiu, J., & Wan, C. (2017). Annual Report Readability, Tone Ambiguity, and the Cost of Borrowing. *Journal of Financial and Quantitative Analysis*, 52(2), 811–836.
- El-Haj, M., Alves, P., Rayson, P., Walker, M., & Young, S. (2020). Retrieving, classifying and analysing narrative commentary in unstructured (glossy) annual reports published as PDF files. *Accounting and Business Research*, 50(1).

- Feldman, R., Govindaraj, S., Livnat, J., & Segal, B. (2009). Management's tone change, post earnings announcement drift and accruals. *Review of Accounting Studies*, 15(4), 915–953.
- Flesch, R. (1949). The art of readable writing.
- Guay, W., Samuels, D., & Taylor, D. (2016). Guiding through the Fog: Financial statement complexity and voluntary disclosure. *Journal of Accounting and Economics*, 62(2-3), 234–269.
- Henry, E. (2008). Are Investors Influenced By How Earnings Press Releases Are Written? *Journal of Business Communication*, 45(4), 363–407.
- Hsieh, C.-C., Hui, K. W., & Zhang, Y. (2015). Analyst Report Readability and Stock Returns. *Journal of Business Finance & Accounting*, 43(1-2), 98–130.
- Hussainey, K., Schleicher, T., & Walker, M. (2003). Undertaking large-scale disclosure studies when AIMR-FAF ratings are not available: the case of prices leading earnings. *Accounting and Business Research*, 33(4), 275–294.
- Jiang, G., Lee, C. M. C., & Zhang, G. Y. (2004). Information Uncertainty and Expected Returns. *SSRN Electronic Journal*.
- Johnson, S., & Tuckett, D. (2017). Narrative Decision-Making in Investment Choices: How Investors Use News About Company Performance. *SSRN Electronic Journal*.

- Jones, M. J., & Shoemaker, P. A. (1994). Accounting Narratives: A Review of Empirical Studies of Content and Readability. *Journal of Accounting Literature*, 13(1), 142–184.
- Kim, Y., Park, M. S., & Wier, B. (2012). Is Earnings Quality Associated with Corporate Social Responsibility? *The Accounting Review*, 87(3), 761–796.
- Kothari, S. P., Li, X., & Short, J. E. (2009). The Effect of Disclosures by Management, Analysts, and Business Press on Cost of Capital, Return Volatility, and Analyst Forecasts: A Study Using Content Analysis. *The Accounting Review*, 84(5), 1639–1670.
- Kong, D., Shi, L., & Zhang, F. (2021). Explain or conceal? Causal language intensity in annual report and stock price crash risk. *Economic Modelling*, 94, 715–725.
- Lang, M. H., & Lundholm, R. J. (1996). Corporate Disclosure Policy and Analyst Behavior. *The Accounting Review*, 71(4), 467–492.
- Lawrence, A. (2013). Individual investors and financial disclosure. *Journal of Accounting and Economics*, 56(1), 130–147.
- Lehavy, R., Li, F., & Merkley, K. (2011). The Effect of Annual Report Readability on Analyst Following and the Properties of Their Earnings Forecasts. *The Accounting Review*, 86(3), 1087–1115.
- Li, F. (2008). Annual report readability, current earnings, and earnings persistence. *Journal of Accounting and Economics*, 45(2-3), 221–247.

- Loughran, T., & McDonald, B. (2011). When Is a Liability Not a Liability? Textual Analysis, Dictionaries, and 10-Ks. *The Journal of Finance*, 66(1), 35–65.
- Loughran, T., & McDonald, B. (2014). Measuring Readability in Financial Disclosures. *The Journal of Finance*, 69(4), 1643–1671.
- Loughran, T., & McDonald, B. (2016). Textual Analysis in Accounting and Finance: A Survey. *Journal of Accounting Research*, 54(4), 1187–1230.
- Lundholm, R. J., Rogo, R., & Zhang, J. L. (2014). Restoring the Tower of Babel: How Foreign Firms Communicate with U.S. Investors. *The Accounting Review*, 89(4), 1453–1485.
- Merkel-Davies, D. M., & Brennan, N. M. (2007). Discretionary disclosure strategies in corporate narratives: Incremental information or impression management? *Journal of Accounting Literature*, 26, 116-194.
- Miller, G. S. (2006). The Press as a Watchdog for Accounting Fraud. *Journal of Accounting Research*, 44(5), 1001–1033.
- Pereira, C.C. (2019). Textual analysis in finance.
- Sloan, R. G. (1996). Do stock prices fully reflect information in accruals and cash flows about future earnings? *The Accounting Review*, 71(3), 289-315.
- Sonenshein, S., Herzenstein, M., & Dholakia, U. M. (2011). How accounts shape lending decisions through fostering perceived trustworthiness. *Organizational Behavior and Human Decision Processes*, 115(1), 69–84.

Tessarolo, I. F., Pagliarussi, M. S., & Luz, A. T. M. da. (2010). The justification of organizational performance in annual report narratives. *BAR - Brazilian Administration Review*, 7(2), 198–212.

Tetlock, P. (2007). Giving Content to Investor Sentiment: The Role of Media in the Stock Market. *The Journal of Finance*, 62(3), 1139–1168.

Appendix

Variable name	N	Mean	St. Dev	Min	P1	P5	Median	P95	P99	Max
wordcount_Front_2	14.504	17.107	16.255	502	1.314	2.531	11.065	50.413	72.281	198.900
flesch_Front_2	14.504	34,785	19,720	(262,111)	(20,691)	8,889	37,876	51,885	60,552	249,180
fog_Front_2	14.504	24,790	22,817	-	17,240	18,937	22,054	31,827	49,933	904,436
causal_Front_2	14.504	0,014	0,025	-	0,004	0,008	0,014	0,018	0,021	3,000
forwardlooking_Front_2	14.504	0,012	0,004	-	0,004	0,007	0,012	0,018	0,022	0,143
uncertainty_Front_2	14.504	0,008	0,003	-	0,002	0,004	0,007	0,013	0,018	0,127
net_tone_H6_Front_2	14.504	0,352	0,225	(1,000)	(0,337)	(0,048)	0,380	0,665	0,783	1,000
wordcount_Chair	12.329	1.046	846	3	137	325	882	2.217	3.816	28.331
flesch_Chair	12.210	46,881	15,429	(187,105)	6,884	31,003	47,746	62,695	71,453	260,733
fog_Chair	12.334	21,375	22,737	-	14,318	16,211	19,740	24,564	38,617	931,028
causal_Chair	12.050	0,012	0,006	0,001	0,002	0,004	0,012	0,022	0,028	0,169
forwardlooking_Chair	12.149	0,018	0,008	0,000	0,004	0,008	0,017	0,030	0,038	0,222
uncertainty_Chair	11.328	0,005	0,003	0,000	0,001	0,001	0,005	0,011	0,016	0,104
net_tone_H6_Chair	12.136	0,683	0,248	(0,833)	(0,111)	0,200	0,733	1,000	1,001	1,001
wordcount_CEO	7.248	1.924	1.453	13	183	435	1.595	4.300	7.266	26.002
flesch_CEO	7.174	43,290	16,830	(199,781)	(1,182)	26,909	44,228	59,973	69,714	258,537
fog_CEO	7.248	22,465	37,761	6,079	14,695	16,332	19,884	24,547	46,831	1.638,556
causal_CEO	7.152	0,013	0,007	0,001	0,003	0,005	0,012	0,022	0,028	0,324
forwardlooking_CEO	7.146	0,016	0,008	0,000	0,003	0,006	0,015	0,027	0,035	0,353
uncertainty_CEO	6.860	0,005	0,005	0,000	0,001	0,001	0,004	0,011	0,020	0,221
net_tone_H6_CEO	7.162	0,699	0,236	(0,867)	(0,200)	0,263	0,750	0,971	1,008	1,010
Stock Returns	5.295.755	0,0	0,1	(1,0)	(0,1)	(0,0)	0,0	0,0	0,1	99,0

Table 5: Descriptive statistics