The Power of Words: Analysing the Impact of Managerial Sentiment in Earnings Conference Calls on Stock Price Reactions

Author: Finnegan Rachfahl
University of Twente
P.O. Box 217, 7500AE Enschede
The Netherlands

ABSTRACT

Earnings conference calls (ECCs) are a primary channel through which companies share their financial results and future expectations with investors and the public. While the numerical results matter, the tone and sentiment used by managers may influence investors' perceptions and market behaviour. Although previous research suggests that the sentiment expressed during these calls can influence investor behaviour, there is limited large-scale research using state-of-the-art NLP (natural language processing) methods. By applying the FinBERT model to a large dataset of over 15,000 ECC transcripts from S&P 500 companies spanning 2010 to 2018, this thesis explores the relationship between managerial sentiment during ECCs and short-term stock price reactions. Sentiment scores are extracted for the Full Call, the Presentation, and the Q&A. These are matched with the closing stock price of the corresponding firm for each of the three days after the call date. A market-adjusted event study methodology is used to calculate cumulative abnormal returns (CAR) for each firm, isolating firm-specific reactions from market-wide movements to assess the effect of sentiment. Regression analysis is conducted to assess whether sentiment scores significantly predict abnormal returns while controlling for firm sector. This analysis finds that positive sentiment in the Full Call and Presentation is positively and significantly associated with CAR, while the Q&A section shows stronger individual effects but less consistent relationships across sectors. The relationship between sentiment and CAR is stronger in Technology, Communication Services, and Consumer Discretionary. Industries such as Utilities and Energy show little to no sensitivity to sentiment. In summary, this thesis offers new insights into the signalling power of managerial sentiment in financial communication and highlights the potential of combining advanced NLP techniques with market data to better understand investors' behaviour.

Graduation Committee members: Lingbo Shen; Xiaohong Huang

Keywords
Earnings conference calls, managerial sentiment, stock price reaction, NLP, behavioural finance, signalling theory

During the preparation of this work, the author used Grammarly Free Version to proofread the thesis and ChatGPT for image generation and assistance in maintaining methodological coherence. After using these tools, the author reviewed and edited the content as needed and takes full responsibility for the content of the work.

1. INTRODUCTION

1.1 Situation and complication

Earnings conference calls (ECCs) are a primary channel through which companies share their financial results and future expectations with investors and the public. While earnings reports focus on the hard numbers, ECCs give investors insight into how executives feel about those numbers. As Tucker et al. (2024, p. 3) stated, "Sentiment provides guideposts to investors beyond the words of the disclosure alone". Typically, ECCs are quarterly events held by top-level-management to communicate the company's financial results as well as outlook to investors, media, and the public. They usually consist of prepared statements as well as a Q&A with senior executives and analysts.

Since information is limited, these calls provide investors and media representatives with valuable insights and crucial information paired with the earnings statements, usually released around the same time. The significance of these calls is exemplified by a high trade volume and return variance before, during, and after these calls, according to Frankel et al. (1996). This indicates the influence these calls can have on traders.

Therefore, understanding the impact of tone in ECCs is increasingly relevant as investors look for faster, data-driven ways to interpret any managerial communication.

1.2 Research objectives and questions

Previous research has shown that while the numerical results in earnings reports are crucial, the tone and language used by management during ECCs can significantly influence investor perception and market reactions. Mayew and Venkatachalam (2012) demonstrated that sentiment in earnings conference calls can predict future stock returns and profitability.

While several studies have used NLP (Natural Language Processing) methods to analyse stock price reaction after ECCs, these studies often used small samples or minor indexes, and few focused on automated sentiment analysis with actual market response, especially over many years and companies. For instance, a study by Loughran and McDonald (2011) focused on 10-K filings and creating finance-specific dictionaries. While Tucker, Xia and Smelcer (2019) analysed fund disclosure documents. There is no quantitative research for major indices like the S&P 500, FTSE 100, DAX or similar, as well as little to no research using modern NLP models like FinBERT or AI-assisted tools. Most studies still rely on creating their dictionaries or using the Loughran and McDonald dictionary (Loughran & McDonald, 2011), both of which are traditional and outdated. These include Fleiss and Cui (2021), who analysed S&P 500 filings, and Gao et al. (2022), who used it to model sentiment for predicting major index movements. Jennings et al. (2021) concluded, more advanced machine learning models outperform LM-based approaches.

This research aims to fill this gap by examining whether sentiment expressed by executives during ECCs has a measurable impact on short-term stock price movements, even in major markets like the S&P 500, using the state-of-the-art NLP techniques.

The main research question is to what extent managerial sentiment in earnings conference calls influences short-term stock price reactions, and if so, how do these effects differ across the 12 Global Industry Classification Standard sectors (GICS)? To answer these questions, the research first uses a finance-specific NLP model called FinBERT to analyse the sentiment of earnings conference calls transcripts of the S&P 500 companies between 2010 and 2018. Then, this research examines the relationship between the sentiment scores and the cumulative abnormal return (CAR) over a three-day event window after the earnings conference call.

1.3 Contribution

ECCs are among the most timely and direct communication tools companies have, yet their linguistic tone is often overlooked in traditional stock price valuation models. Prior research has primarily relied on traditional dictionary-based methods like the Loughran and McDonald (LM) dictionary, which fails to capture context, negation and tone. In times where communication from companies to investors becomes more direct and open, it becomes increasingly important to take the expressed sentiment into account and analyse it with modern NLP tools like FinBERT. In this context, ECCs play a crucial role as they portray a prepared Presentation, as well as a reactive Q&A where investors and analysts can raise questions. The lack of analysis not only of ECCs in general but also the dissection of them into the Full Call, Presentation, and Q&A, as well as sector-specific effects, led to the motivation for this research. This thesis is grounded in and contributes to behavioural finance, natural language processing, signalling theory and market microstructure by analysing a comprehensive dataset of earnings conference calls transcripts over 8 years of the most prestigious index worldwide, the S&P 500.

This research applied FinBERT at scale to over 15,000 ECCs, one of the largest applications of this model in a financial event study setting. It introduces a section level sentiment breakdown by Full Call, Presentation, Q&A, GICS Sector, and year allowing a differentiation between scripted and unscripted speech over 8 years and 12 different industries and matches this with a cumulative abnormal return (CAR) event study method, demonstrating that positive managerial sentiment is a statistically significant predictor of short-term CAR. This model shows that the Q&A section shows stronger individual effects but varies significantly by sector, while the presentation shows more consistent but smaller effects across industries, suggesting that investors respond more to structured messaging than to real-time Q&A exchanges. The second model demonstrates that the effect of sentiment on returns varies by industry. A stronger relationship between sentiment and return is linked to sectors like Technology, Communication Services and Consumer Discretionary. While there are weak or no effects in sectors like Utilities and Energy. Showing that only certain sectors meaningfully influence returns, revealing that investor reaction is context and industry sensitive. It highlights how modern NLP methods can improve the signal-to-noise ratio in sentiment-based return predictions.

Through combining FinBERT with a formal event study and a regression framework, providing large-scale findings on how managerial sentiment in earnings conference calls relates to short-term stock price reactions, this thesis offers a replicable and scalable method for future ECC studies.

By addressing this, the study deepens the understanding of how investors respond not just to financial results, but to the way those results are communicated.

Beyond academic contributions, this research is also relevant for investors, analysts and algorithmic traders who increasingly rely on sentiment data to inform short-term decisions (Bagate et al., 2022). In an environment where speed and automated interpretation are critical, this thesis offers up-to-date and practical insights, helping market participants better anticipate and respond to short-term price movements driven by sentiment-rich company communication.

2. THEORETICAL FRAMEWORK

2.1 Efficient Market Hypothesis EMH

The Efficient Market Hypothesis established by Eugene Fama (1970), states that financial markets are informationally efficient, meaning that asset prices, mostly share prices, at any given time, reflect all available information. Stating that investors are not capable of beating the overall market movement with clever stock picks or timing, but only with higher risk advocacy. Based on this theory, new information disclosed during the earnings conference calls should be immediately priced in by the market, eliminating the possibility of abnormal returns through further interpretation of such content. This assumption treats financial information as purely objective and numerical, which in most cases is true.

However, earnings conference calls are an exception to the mostly numerical and objective data. Here, the way information is communicated, particularly through verbal tone, emphasis and emotional cues, can influence perception beyond the factual content itself. If this proves true and investors react differently to the same numerical results based on tone and sentiment, this may suggest an inconsistency within the EMH theory, challenging its assumptions. This creates an opportunity for behavioural interpretations of market response so called Behavioural Finance.

2.2 Behavioural Finance

Behavioural finance questions the rationality of investors, suggesting that market participants may react to emotional and psychological cues such as tone and language, rather than solely to financial data. Showing that emotional cues can lead to biased decision making (Thaler, 1993).

A recent study by Alta'any et al. (2024) found that managerial tone and sentiment in ECCs can predict future financial performance, indicating that tone can influence investor expectations and behaviour.

These findings underline the importance of psychological influences like managerial sentiment and its analysis when evaluating events like earnings conference calls, where tone can influence investor behaviour and subsequently market movement.

2.3 Signalling Theory

Signalling Theory goes back to Spence, who introduced it in 1973. The core idea is that two parties are communicating with each other, where one of the two parties has more or better information than the other. This happens in business, like in negotiations, sales, insider trading or investor communication, but also in day-to-day situations at work or school. The main point is that this information is not explicitly stated but rather communicated through observable cues. These cues can be vocal, lie in sentiment or even communicated physically through mimic and gesture. The theory focuses on how the party with the better information conveys these cues so that certain people understand without it being obvious.

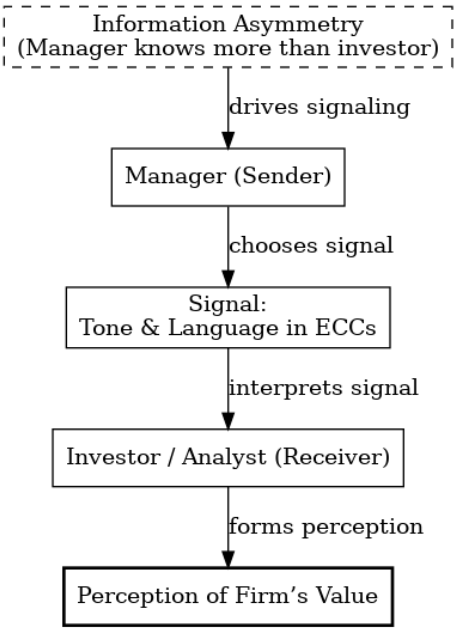
Ahmed et al. (2025) used signalling theory in their research about executive compensation and share buyback, analysing U.S. firms from 2000 to 2020, finding that share buybacks are interpreted as signals of undervaluation, but only credible signals when management is financially aligned with shareholders. This means share repurchase announcements only signal undervaluation credibly when executives have high wealth sensitivity, for example, their compensation being tied to stock performance. This reflects the core of signalling theory that the signal, in this case, the repurchase, must be costly or risky to fake.

AlGhazali et al. (2024) applied signalling theory to dividend signalling in global markets. By analysing over 11,000 firm-years from over 25 countries between 2001 and 2018 they concluded that dividend changes are used as signals of management's confidence in future earnings, with dividend increases predicting higher future earnings but only in countries with strong investor protections, while in weak governance environments dividend changes are less informative or even misleading. This again draws back to the core of signalling theory that the signal must be costly or risky to fake, which the latter only applies if strong investor protection is guaranteed.

In the context of ECCs, one actor, the sender, usually being corporate management, possesses more information than external stakeholders such as investors, analysts, and media. In this situation, during the earnings conference calls, the manager has the power to decide how he conveys the expectations and company performance, besides the hard financial data, through observable cues such as language and tone. It is then up to the receiver to interpret these signals (Spence, 1973). In earnings conference calls, sentiment may serve as an intentional or subconscious signal about management confidence, uncertainty or strategic outlook needing to be interpreted by investors, media, and analysts. And thus, influencing their decision-making, resulting in market movement. Regarding ECCs, the costliness of sentiment as a signal arises from reputational damage, legal exposure, the risk of investor backlash, and negative stock price reactions if the tone is later found to contradict the actual performance. These factors create a natural incentive to minimise exaggerating, making managerial tone a potentially credible signal of internal expectations or confidence.

ECCs in the context of signalling theory are further explained in Figure 1. Illustrating the information asymmetry starting with the manager knowing more than the investor, usually insider information about future projects, company workflows or acquisitions. Things besides hard facts, like news or financial statements. He then proceeds consciously or unconsciously to decide how to convey this Signal. In the context of ECCs, this is restricted to tone and language. The receiver now must catch these cues and decide how to interpret them. This interpretation likely differs between but also within groups of actors, like investors or analysts. These then form their perception of the firm's value, not only influenced by the hard data but by these subtle cues initially conveyed by the manager.

Figure 1 Signalling Theory Conceptual Model



2.4 Prior Research on Managerial Tone and Market Reaction

Several studies have investigated the relationship between managerial tone and market responses. Mayew and Venkatachalam (2012), found that Vocal arousal measured from managers during earnings conference calls predicted stock returns and future profitability even after controlling for earnings and other financial metrics. They analysed around 1,600 earnings conference calls from S&P 1500 firms between 2003 and 2007 by applying vocal recognition software to CEO speech during these calls to extract “Vocal arousal” as their key variable.

Price, S. M., Doran, J. S., Peterson, D. R., & Bliss, B. A. (2012) went even further textually analysing over 2,300 ECCs from 2001 to 2006 using a domain specific sentiment dictionary and regression analysis to conclude that the tone of ECCs heavily supplements earnings figures information and is significantly associated with short term abnormal returns and trading volume. They also found the Q&A section had more predictive power than prepared statements.

Additionally, Loughran, T., & McDonald, B. (2011) introduced financial sentiment dictionaries like the LM Dictionary as well as demonstrated that tone in annual reports (10-k filings) affects both stock returns and volatility, as well as that traditional general-purpose dictionaries like Harvard IV misclassify financial text. They analysed over 50,000 10-K filings from public U.S. companies between 1994 and 2008 using their own Financial Sentiment Dictionary.

While these studies established the relevance of sentiment in financial communication, they often focus on limited sample sizes, non-finance-specific NLP or written disclosures. This thesis builds on their foundation by extending the analysis to a large-scale dataset of spoken earnings conference calls using modern and finance-specific NLP tools.

2.5 Hypotheses

Based on the theoretical background and prior research, this thesis proposes the following two hypotheses.

H1: A higher positive sentiment score of earnings conference calls is associated with higher cumulative abnormal returns (CAR) in the three trading days following the call.

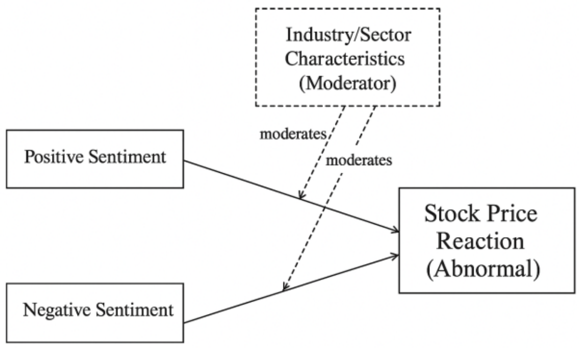
This hypothesis is based on behavioural finance and signalling theory, suggesting that investor behaviour is influenced by emotional cues conveyed by managers during ECCs, which investors and analysts can pick up and thus drive short-term price movements.

H2: The relationship between sentiment scores and short-term CAR differs across sectors.

The second hypothesis explores whether a moderator, like Sector or subsectors, influences H1, considering that sectors vary in their sensitivity and reaction to company communication. Previous research has demonstrated that the informativeness of tone in ECCs may depend on sector-specific characteristics. Price et al. (2012) found that tone conveyed during conference calls had a significantly greater impact on stock returns in sectors where qualitative forward-looking statements are more critical to valuation. Industries like technology or consumer cyclicals, compared to regulated or resource-driven sectors like utilities or energy. Such findings suggest that the informativeness of sentiment may vary depending on how forward-looking or regulated a sector is.

Figure 2 illustrates the conceptual framework assisting these hypotheses. It shows how positive and negative sentiment influence stock price reaction and how this relationship is potentially moderated by industry or sector characteristics. Positive and negative sentiment expressed by managers during ECCs are the independent variables, with stock price reaction in the form of cumulative abnormal return (CAR) as an outcome, and industry characteristics potentially moderating the strength or direction of the relationship.

Figure 2 Hypotheses 1 and 2 Conceptual Model



3. METHODOLOGY

3.1 Research Design

This study employs a quantitative, explanatory research design to examine the relationship between managerial sentiment during earnings conference calls (ECCs) and short-term stock price reactions. By analysing a large data sample with over 15,000 earnings conference call transcripts, the research aims to determine whether the tone and sentiment of these transcripts have predictive value for cumulative abnormal returns CAR in the days following the earnings disclosure. The dataset consists of all available ECC transcripts of S&P 500 companies between 2010 and 2018 to ensure relevance, consistency, as well as avoid major events influencing the stock market and sentiment like the 2008 financial crisis and COVID-19. The decision to use earnings conference calls is based on their often-overlooked importance in conveying information otherwise left out in typical financial statements. This makes them an ideal setting for studying how language-based signals affect market perceptions.

A natural language processing model trained specifically on financial language (FinBERT) is used to compute sentiment scores for each earnings conference call. These scores are then put in relation to cumulative abnormal returns (CAR) over a short event window following the call, using an event study approach. This allows for isolating the incremental impact of sentiment from broader market movements and fundamentals. Missing stock price data resulted in the exclusion of 679 observations from the CAR analysis, reducing the sample from 16,608 to 15,929 observations. The three-day event window Day (0 to 3) was chosen to capture the immediate market reactions, trying to minimise noise from unrelated events that might occur over longer periods. Outliers beyond three standard deviations were identified but retained to preserve the natural distribution of market reactions.

3.2 Sentiment Analysis

Sentiment analysis identifies and quantifies emotions, opinions, or tone within a text. In finance, it is used to interpret the emotional subtext of company communication, like the managerial sentiment in ECCs. There are different ways to analyse sentiment, like bag-of-words or dictionary-based techniques, for example, the Loughran-McDonald dictionary. These rely on counting predefined positive and negative terms without context. They do not consider word order, syntax, or semantic meaning, which can lead to misclassifying financial terms like liability and risk, which are neutral or technical. The simple logic behind these approaches brings limitations such as the inability to detect sarcasm, negotiation, or complex phrasing. As well as poor handling of domain-specific financial language.

Therefore, this thesis employs FinBERT, a transformer-based NLP model fine-tuned on financial text. It classifies text into three sentiment categories, positive, neutral, and negative, as well as returns a confidence score P_{pos}, P_{neu}, P_{neg}, for each of the three classes, always summing up to 1. The model is context sensitive and better at handling finance-specific phrasing like missed guidance vs. loss compared to general NLP models.

The Weighted Sentiment score was constructed to transform FinBERT's categorical output into a single continuous numerical score usable in the regression analyses.

(1) $\text{Weighted Sentiment} = (+1) \cdot P_{pos} + (0) \cdot P_{neu} + (-1) \cdot P_{neg}$

This formula compresses the three-class output into a single sentiment value ranging from –1 to +1, where:

–1 = strongly negative (100% negative)

0 = neutral (100% neutral or equal mix of positive and negative)

+1 = strongly positive (100% positive)

Each Presentation, Q&A, or Full Call is split into sentences. Then the weighted sentiment scores for all sentences are averaged to get one comprehensive sentiment score, which is used in the regression analyses.

The advantages of FinBERT are that it is a financial domain-specialised version transformer-based language model, built on the BERT (Bidirectional Encoder Representations from Transformers) architecture and is pretrained on a large set of financial texts. Specifically trained to classify sentiment in financial documents such as earnings conference call transcripts, annual reports, and analyst commentary. Compared to general Natural Language Processing models (NLPs), FinBERT provides more accurate sentiment analysis by taking industry-specific language and terminology into account. (Araci, 2019)

3.3 Event Study Method

An event study is a statistical method used to assess the impact of a specific event on the value of a firm by comparing actual returns with expected (normal) returns over a defined time window. In this study, the event is the ECC with the event window spanning the day of the call (Day 0) till 3 days later (Day +3) to capture the immediate market response. The expected return, also called the normal return, refers to the return a stock is predicted to earn based on its historical correlation with the market, assuming no event occurs. Expected returns are estimated using a 60-day estimation window, ending 3 trading days before the event. It is estimated using the market model, which assumes a linear relationship between the stock's return and the return of a market index, in this case, the S&P 500, calculated as:

(2) $R_{it} = \alpha_i + \beta_i \cdot R_{mt} + \varepsilon_t$

Where:

R_{it} is the return of stock i at time t

R_{mt} is the return of the market index at time t

α_i and β_i are estimated via ordinary least squares (OLS) over the estimation window

The abnormal return (AR) is the difference between the actual and expected return for each day in the event window.

Abnormal returns (AR) for each event date are computed using the market model:

(3) $AR_{it} = R_{it} - (\alpha_i + \beta_i \times R_{mt})$

where R_{it} is the return of firm i at time t, R_{mt} is the market return at time t, and α_i and β_i are estimated over the defined estimation window.

These abnormal returns are then aggregated to form the CAR, capturing firm-specific price reactions potentially driven by sentiment.

The cumulative abnormal return (CAR) over the event window is then calculated as:

(4) $CAR = \sum (R_{it} - AR_{it})$ from Day 0 to Day +3

This method isolates the portion of the return potentially attributable to new information (e.g., sentiment signals) rather than market-wide movements. The CAR serves as the dependent variable in regression models assessing the influence of managerial sentiment, as quantified by FinBERT.

3.4 Regression Analysis

The primary regression was performed using an ordinary least squares (OLS) regression analysis with the dependent variable being the cumulative abnormal return (CAR) and the independent variable being the weighted sentiment scores derived from the Full Call, the Presentation and the Q&A section. This first regression model tests Hypothesis 1, whether more positive sentiment is associated with higher short-term abnormal returns.

The sector-specific regression was conducted separately for each GICS Sector to test Hypothesis 2, assessing whether sentiment impacts differ across industries. This approach isolates the within-sector effects, allowing for a thorough comparison of managerial sentiment and its effect on CAR. The sector regression used the same three variables (Full Call, Presentation, Q&A) as the main regression.

Sentiment variables were created by applying FinBERT to each of the three sections (see 3.2). Both regressions were conducted using OLS under the assumption of linearity and normally distributed residuals. No additional control variables like firm size or volatility were included in the baseline models to isolate the pure effect of sentiment. The output of the regression models was reported with coefficients, standard errors, t-values, and significance levels. Additional statistics include R-squared, F-statistics and residual standard errors.

4. DATA

The dataset includes over 15,000 earnings conference call transcripts (ECC) from 487 companies listed in the S&P 500 index, covering the period from January 2010 to December 2018. The ECC includes a prepared presentation segment as well as a Q&A for each call. The sentiment classifications (positive, neutral, negative) as well as their correlating confidence score were extracted and calculated for the Q&A, the presentation and the full ECC and logged with the call date and company ticker and put in a new document. With the yfinance API, each date of ECC was matched with their respective GICS sector and sub-industry classification, as well as the end-of-day stock price for day+0 till day+3 after the call date of the respective ticker and the S&P 500 end-of-day price.

Using the yfinance API, historical daily closing prices were retrieved for both the company and the S&P 500 index from 60 trading days before to three days after each ECC. These were used to calculate cumulative abnormal returns (CAR) over the event window (Day 0 to Day +3), using a standard market model with the 60-day estimation window.

The sentiment classifications paired with the model's confidence score were combined to compute a continuous weighted sentiment score ranging from –1 (strongly negative) to +1 (strongly positive). This allows for a nuanced, numerical representation of sentiment intensity and direction.

Table 1 shows the summary statistics. From the near-zero mean of CAR (0.0024) can be inferred that there is no general trend in abnormal returns following ECCs. The sentiment score of the Q&A is centred near 0, indicating that the tone is often neutral, while the presentation section tends to have the most positive sentiment (avg. ≈ 0.49). The overall tone of calls is generally positive but varies by sector and year. All sentiments show small standard deviations, suggesting low variability across calls. The difference between 16,608 sentiment observations and 15,929 CAR observations is due to missing stock price data, mergers & acquisitions, delisting or similar issues during the event window, preventing the calculation of abnormal returns for those cases.

Table 1 Descriptive Statistics - CAR and Weighted Sentiment Scores

Statistic	N	Mean	St. Dev.	Min	Max
CAR (Day 0 to 3)	15,929	0.0024	0.0577	-0.4002	0.7418
Weighted Full	16,608	0.3409	0.4652	-0.9986	1.0000
Weighted Presentation	16,608	0.4875	0.5476	-1.0000	1.0000
Weighted Q&A	16,608	0.1713	0.3717	-0.9978	1.0000

In Figures 3, 4 and 5 (see Appendix), the sentiment is broken down by sector and years across three sections of the earnings conference call (Presentation, Q&A and Full Call)

Figure 3 shows the positive sentiment by sector and year of the Presentation section. It shows consistently higher positive sentiment across all sectors compared to Q&A and Full Call. Sectors like Basic Materials and Consumer Cyclical frequently have positive proportions above 60%. Technology shows increasing positivity from 2013 onward (0.44 to 0.67 in 2017), possibly reflecting sector optimism and innovations. Energy shows notably lower positive sentiment, especially between 2010 and 2012, with an increase in 2014 that may be related to the annexation of Crimea by Russia and resulting sanctions, increasing the demand for oil and gas from American companies. Utilities show the lowest consistent sentiment, with lows around 0.28 in 2010 and remaining below 0.44 through 2010 till 2018. This might reflect the cautious tone typical of regulated industries. Healthcare maintains a stable, slightly increasing sentiment throughout the years. This increase may go back to the implementation of the Affordable Care Act (ACA) throughout the 2010s, as well as an ageing population driving demand for healthcare and eldercare.

Presentation sections are scripted and designed to reassure investors, resulting in the uniformly high positive scores. The sector differences likely reflect industry-specific performance cycles and geopolitical events.

Figure 4 illustrates the positive sentiment by sector and year for the Q&A. Across all sectors, the sentiment in the Q&A part is substantially lower than in the Presentation. Most values stay in the 0.10 to 0.30 range, indicating far less positivity. It reflects the more spontaneous, less rehearsed nature of this segment, where executives are cautious about being too optimistic to investor questions.

Communication Services is an outlier with consistently higher positive sentiment peaking at 0.41 in 2014 and reaching 0.40 again in 2018. Possible explanations include that firms in this sector may use forward-looking language as well as tease new technologies and products. Consumer Defensive and Healthcare show a steady increase in positive sentiment over time, ending in 2018 at 0.33 and 0.26, respectively. This mirrors the effect we have seen in the Presentation. Energy ranks among the lowest in Q&A positivity, dropping to as low as 0.02 in 2015, which could be tied to global oil price crashes from mid-2014 to early 2016.

Utilities keep showing extremely low positivity across the board with values between 0.00 to 0.07, consistent with their conservative communication style.

The notable uptick in 2014-2016 in Technology, Healthcare and Consumer Cyclical suggests an improved economic upward trend starting in the middle of the decade.

In Figure 5, the positive sentiment by sector and year is shown for the Full Call. The overall sentiment is moderate, falling between the extreme highs of the Presentation and the lows of the Q&A. Most sectors stay within the range of 0.20 to 0.50, with a few exceptions that break higher, like Communication Services, Consumer Cyclical and Consumer Defensive. Energy remains persistently low, with a peak in 2014 directly tied to the oil crisis and capital uncertainty. In general, the Full Call sentiment reflects the effects seen in the Presentation and Q&A parts.

5. RESULTS

5.1 Regression Analysis Results

This section presents the results of the ordinary least squares (OLS) regression examining the relationship between the managerial sentiment, classified as Weighted Full (Presentation + Q&A), Weighted Presentation and Weighted Q&A, and the cumulative abnormal return (CAR) over a three-day event window following the ECCs.

Table 2 shows that the CAR positively but weakly correlates with all sentiment scores. The strongest correlation being (0.06) reinforces that sentiment alone does not explain most of the CAR variance. Although the correlations are weak, the strongest correlation exists between CAR and the whole ECC (Weighted Full). This aligns with our results from the regression analysis.

Table 2 Correlation Matrix

	Weighted Full	Weighted Presentation	Weighted Q&A	CAR (Day 0 to 3)
Weighted Full	1	0.190	0.242	0.060
Weighted Presentation	0.190	1	-0.041	0.035
Weighted Q&A	0.242	-0.041	1	0.036
CAR (Day 0 to 3)	0.060	0.035	0.036	1

Table 3 shows the results of the OLS regression with the dependent variable being CAR over Day 0-3, and the independent variable being the weighted sentiment from the full call, presentation, and Q&A.

While all three independent variables are positively significant, the full call has the largest effect (0.006), showing that the whole earnings conference call sentiment positively impacts CAR and ECCs, generally influencing the immediate stock price reaction. Comparing the Presentation and the Q&A, the latter seems to have the larger effect given its relatively high coefficient (0.0040 vs. 0.0029), indicating that investors put a higher importance on Q&A sentiment compared to the prepared Presentation. However, the sector-specific analysis reveals that while the Q&A effects are stronger on average, they vary considerably across industries, whereas Presentation effects are smaller but more consistent across sectors.

Table 3 OLS Regression - Effect of Sentiment on CAR

	Dependent variable:
	CAR (Day 0 to 3)
Weighted Full	0.0060^{***}
	(0.0010)
Weighted Presentation	0.0029^{***}
	(0.0009)
Weighted Q&A	0.0040^{***}
	(0.0013)
Constant	-0.0017^{**}
	(0.0007)
Observations	15,929
R^2	0.0048
Adjusted R^2	0.0046
Residual Std. Error	0.0576 (df = 15925)
F Statistic	25.6607^{***} (df = 3; 15925)
Note:	*p<0.1; **p<0.05; ***p<0.01

Table 4 presents sector-specific OLS regressions assessing how sentiment influences CAR. Consumer Defensive has a strong positive effect (0.0105) and is highly significant, suggesting that investors react strongly to the overall tone. This may be attributed to the critical role these firms play in the overall economy, particularly in times of uncertainty.

For Financial Services, only the Full sentiment is significant (0.0056) while presentation and Q&A alone are not that important.

Industrials and Technology show similar significant positive effects for Full sentiment (0.0085), indicating a strong investor reaction. However, while all components (Full, Presentation, Q&A) are significant in Industrials (0.0044, 0.0056), only the Full sentiment is significant in Technology.

All sentiment variables are significant in the Consumer Cyclical sector, indicating this sector is very tone-sensitive, especially to the Q&A parts of the ECC (0.0081). Of all the sectors, the Consumer Cyclical regression shows the strongest statistical fit (F-statistic), insinuating that sentiment is especially predictive of CAR for firms in this sector.

In Real Estate ECCs, the Presentation is highly significant, having the largest overall coefficient (0.0059), even surpassing Full and Q&A together, highlighting the importance of prepared remarks.

Energy, Communication Services, Basic Materials, Healthcare and Utilities show weaker or no statistically significant results, possibly due to industry characteristics resulting in lower tone sensitivity.

The Unknown sector has a large positive coefficient for the Full sentiment (0.0179) but is not statistically significant, likely due to a high standard error and a small sample size.

These findings suggest that the influence of sentiment on CAR varies notably across different sectors.

Table 4 OLS Regression by Sector – Effect of Sentiment on CAR

	Dependent variable:											
	CAR (Day 0 to 3)											
	Basic Materials	Communication Services	Consumer Cyclical	Consumer Defensive	Energy	Financial Services	Healthcare	Industrials	Real Estate	Technology	Unknown	Utilities
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Weighted Full	0.0022	0.0090	0.0051	0.0105^{***}	-0.0020	0.0056^{***}	0.0041	0.0085^{***}	-0.0044^{*}	0.0085^{**}	0.0179	0.0044
	(0.0044)	(0.0082)	(0.0031)	(0.0032)	(0.0047)	(0.0021)	(0.0030)	(0.0027)	(0.0026)	(0.0034)	(0.0261)	(0.0031)
Weighted Presentation	0.0002	-0.0087	0.0064^{**}	0.0032	0.0038	0.0014	0.0010	0.0044^{**}	0.0059^{***}	0.0031	-0.0149	0.0011
	(0.0038)	(0.0077)	(0.0029)	(0.0028)	(0.0031)	(0.0016)	(0.0025)	(0.0021)	(0.0018)	(0.0029)	(0.0169)	(0.0018)
Weighted Q&A	0.0039	-0.0028	0.0081^{**}	0.0035	-0.0065	0.0015	0.0028	0.0056^{*}	-0.0014	0.0020	0.0559	0.0060
	(0.0060)	(0.0090)	(0.0036)	(0.0037)	(0.0070)	(0.0027)	(0.0035)	(0.0033)	(0.0038)	(0.0040)	(0.0471)	(0.0054)
Constant	0.0034	0.0085	-0.0049^{*}	-0.0067^{***}	0.0006	-0.0011	-0.0016	-0.0040^{**}	-0.0016	0.0014	-0.0104	-0.0019^{*}
	(0.0033)	(0.0069)	(0.0026)	(0.0025)	(0.0023)	(0.0012)	(0.0019)	(0.0019)	(0.0013)	(0.0024)	(0.0137)	(0.0011)
Observations	633	471	2,432	1,177	753	2,075	1,990	2,054	1,153	2,213	103	875
R^2	0.0013	0.0047	0.0065	0.0142	0.0032	0.0048	0.0019	0.0108	0.0104	0.0046	0.0286	0.0056
Adjusted R^2	-0.0034	-0.0017	0.0053	0.0117	-0.0007	0.0034	0.0004	0.0094	0.0078	0.0032	-0.0008	0.0022
Residual Std. Error	0.0505 (df = 629)	0.0814 (df = 467)	0.0725 (df = 2428)	0.0508 (df = 1173)	0.0483 (df = 749)	0.0398 (df = 2071)	0.0582 (df = 1986)	0.0554 (df = 2050)	0.0315 (df = 1149)	0.0718 (df = 2209)	0.0982 (df = 99)	0.0254 (df = 871)
F Statistic	0.2798 (df = 3; 629)	0.7375 (df = 3; 467)	5.2905^{***} (df = 3; 2428)	5.6366^{***} (df = 3; 1173)	0.8122 (df = 3; 749)	3.3242^{**} (df = 3; 2071)	1.2446 (df = 3; 1986)	7.4674^{***} (df = 3; 2050)	4.0137^{***} (df = 3; 1149)	3.3715^{**} (df = 3; 2209)	0.9717 (df = 3; 99)	1.6318 (df = 3; 871)
Note:												$^{*}p^{**}p^{***}p<0.01$

6. DISCUSSION

6.1 Summary of Key Findings

As seen in the main regression table (Table 1), all three sentiment scores (Full, Presentation, Q&A) have a positive, statistically significant relationship with the cumulative abnormal return over 3 days (CAR). This confirms the first hypothesis, “H1: A higher positive sentiment score of earnings conference call is associated with higher cumulative abnormal returns (CAR) in the three trading days following the call.”

The Weighted Full score has the strongest overall effect, suggesting that investors and media react most to the general tone of the ECC. While the Q&A sentiment has a slightly stronger effect than the presentation, indicating that investors respond more to unprepared and open communication than to the rigid presentation part.

The descriptive statistics show that the CAR values are centred around zero, confirming no systematic bias in abnormal returns. The sentiment scores are also clustered around zero, reflecting a neutral tone, especially in the Q&A. The small standard deviations across sentiment scores indicate low variability in tone across calls.

The correlation matrix table reveals a positive but weak correlation between CAR and sentiment variables, and the strongest correlation between CAR and Weighted Full, further supporting the regression findings.

The sector-specific regression shows bigger differences in the regression effects of the three sentiment variables and CAR. Consumer Cyclical has the strongest relationship between Q&A sentiment and CAR (0.0081). This may go back to firms in this sector being highly sensitive to demand expectations, so investors want real-time predictions for upcoming quarters often only addressed in the Q&A.

In the Real Estate sector, the strongest relationship exists between Presentation and CAR. Real Estate is a rather slow-changing market, so short-term changes discussed usually in the Q&A are less relevant, while the prepared and structured remarks in the Presentation hold the highest value to investors.

For the Industrials sector, all three components are significant, marking a broad tone sensitivity, meaning investors value both scripted outlooks and unscripted remarks.

Companies in the technology sector show only in the Full call a significant effect. Same goes for the Consumer Defensive sector, but here the Full call tone is very significant, showing their stock price adjustments to market conditions, caused by the nature of the goods being sold by these firms.

Healthcare, Utilities, Energy and Basic Materials are stable markets and show weak or no statistical significance. Stock performance in these sectors is often driven by external variables like regulations, laws, commodity prices and taxes. Since sentiment is less impactful, investors may prefer hard data and forecasts.

Overall, these findings show that the effect of sentiment on short-term stock price changes is not uniform, and many sectors have specific characteristics making some parts more important while others not. These characteristics suggest investor preferences and market reactions based on industries.

6.2 Discussion of Unexpected Results

Sentiment only explains a small fraction of CAR variance (R-squared 0.005). This is common for financial regression analysis and suggests that other factors like macroeconomics, earnings figures and industry news play a bigger role. Still, it does not invalidate the findings but shows that sentiment is just one of many variables in stock price performance.

The coefficients were statistically significant but economically small. This may reflect market efficiency of investors already pricing in expectations, drawing back to the efficient market hypothesis EMH. Nonetheless, with big enough volume, even small economic movements can be relevant, especially for bigger or institutional investors or option trading. While the R^2 of 0.005, indicating sentiment explains only 0.5% of return variance, appears low, this is typical for financial models, where stock prices are influenced by numerous unpredictable factors. Nonetheless, with big enough volume, even small economic movements can be relevant, especially for bigger or institutional investors and algorithmic trading at scale. For instance, a 0.4 percentage point increase in CAR is a substantial effect when applied to large portfolios or leveraged positions. The consistent statistical significance across multiple sectors suggests the relationship is robust, even if sentiment is just one component of the multitude of factors surrounding earnings announcements.

6.3 Theoretical Implications

The results mostly match prior research. Like Loughran & McDonald (2012) and Price et al. (2012), it shows that the overall tone in ECCs positively predicts abnormal returns, supporting the broader view that sentiment adds incremental information beyond hard earnings numbers.

Price et al. (2012) found the impact of the Q&A to carry unique market value. The same pattern can be found in this analysis, especially for Consumer Cyclical and Industrials sector companies, where Q&A tone shows a stronger CAR effect than Presentation. However, this pattern does not hold consistently across all sectors.

New contributions include sector-level variations. Most prior studies focus on the overall relationship between sentiment and stock price changes, while the applied sector-level regression reveals distinct sensitivity patterns. This highlights tone effects as context-dependent and expands on the research about general tone literature.

By comparing Full Call, Presentation and Q&A tone across multiple industries, this thesis demonstrates that unscripted tone is most influential in demand-sensitive sectors like Consumer Cyclical and Industrials, while scripted tone matters more in stable industries like Healthcare, Utilities, Energy and Basic Materials. An approach not commonly employed in prior studies, and if not to this extent.

6.4 Practical Implications

Tone can be a market signal, as confirmed by these results. Especially, the Q&A part is significantly associated with abnormal returns in the days following the call. Institutional investors such as quantitative or sentiment-based funds could integrate automated ECC tracking and sentiment analysis into short-term trading models, specifically focusing on sectors that show stronger tone sensitivity, like Consumer Cyclical and Industrials. The ease of information available, as well as executing NLP models and regression analysis, also makes it suitable for individual investors trading options around earnings conference calls. While tone matters, the low R^2 suggest it's one of many factors, so sentiment alone should not drive investment decisions.

Companies can use this latest information to better prepare and adjust their communication strategies, especially during Q&A sessions, to more accurately reflect their financial performance. This may allow them to influence investor decision-making and so change their company's stock price. Firms should tailor their focus on Presentation or Q&A during ECCs based on their sector. In stable sectors like Real Estate and Healthcare, clarity and tone in prepared remarks seem more impactful, while in demand-sensitive sectors, confidence and transparency are key in the Q&A.

The low explanatory power (R^2 of 0.005) should not discourage practical application, as even small but consistent effects are valuable in competitive financial markets where marginal advantages compound over time.

6.5 Limitations

This thesis focuses exclusively on short-term cumulative abnormal returns (CARs) over a three-day event window.

This means the analysis captures immediate market reaction, but long-term investor behaviour or delayed effects are not considered.

The sentiment scores were extracted using FinBERT, a pretrained financial language model. Results depend on FinBERT's characteristics, accuracy, and limitations, such as sensitivity to tone context misinterpretation, failure to detect sarcasm or complex sentiment.

The regression does not control for other important variables like market capitalisation, industry volatility, past performance longer than 60 days, earnings surprise or financial statements. This limitation was primarily due to data access constraints, as the university's subscription to financial databases did not include historical firm-level financial data API access required for these control variables. While this could result in omitted variable bias, this choice was made intentionally to focus on sentiment return relationships, though future research with access to more comprehensive financial databases would benefit from including these control variables to test robustness.

Further, OLS regression assumes a linear relationship between sentiment and CAR, but the true relationship might be nonlinear or interact with other variables, potentially oversimplifying a more complex nonlinear relation. While data limitations included the focus on S&P 500 firms only, consisting of large American firms, thereby excluding smaller firms or non-US markets, which may limit generalisability.

7. ACKNOWLEDGEMENTS

I would like to thank my supervisor, Dr Lingbo Shen, for the guidance, feedback, and valuable insights provided during the completion of this thesis. I also thank my parents for their consistent support throughout my studies.

8. REFERENCES

Ahmed, W., Taffler, R., & Tosun, O. K. (2025). Executive compensation and the credibility of share buyback announcements. *Review of Quantitative Finance and Accounting*. https://doi.org/10.1007/s11156-025-01398-1

AlGhazali, A., Al-Yahyaee, K. H., Fairchild, R., & Guney, Y. (2024). What do dividend changes reveal? Theory and evidence from a unique environment. *Review of Quantitative Finance and Accounting*, *62*(2), 499–552. https://doi.org/10.1007/s11156-023-01211-x

Alta'any, M., Kayed, S., Meqbel, R., & Albitar, K. (2024). Speaking success: Managerial tone in earnings conference calls and financial performance. *Corporate Governance: The International Journal of Business in Society*, *25*(2), 311–336. https://doi.org/10.1108/CG-09-2023-0381

Araci, D. (2019, August 27). *FinBERT: Financial Sentiment Analysis with Pre-trained Language Models*. arXiv.Org. https://arxiv.org/abs/1908.10063v1

Bagate, R., Joshi, A., Trivedi, A., Pandey, A., & Tripathi, D. (2023). Survey on Algorithmic Trading Using Sentiment Analysis. In B. Pati, C. R. Panigrahi, P. Mohapatra, & K.-C. Li (Eds.), *Proceedings of the 6th International Conference on Advance Computing and Intelligent Engineering* (pp. 241–252). Springer Nature. https://doi.org/10.1007/978-981-19-2225-1_22

Bhattacharya, S. (1979). Imperfect Information, Dividend Policy, and 'The Bird in the Hand' Fallacy. *The Bell Journal of Economics*, *10*(1), 259–270. https://doi.org/10.2307/3003330

Cui, H., & Fleiss, A. (2021). Forecasting Stock Price Changes Using Natural Language Processing. *ERN: Asset Price Forecasts (Topic)*. https://doi.org/10.2139/ssrn.3936026

Fan, W., Gao, R., Wang, Y., Xiao, H., Zhang, H., & Cui, S. (2022). Integrating the sentiments of multiple news providers for stock market index movement prediction: A deep learning approach based on evidential reasoning rule. *Inf. Sci.*, *615*, 529–556. https://doi.org/10.1016/j.ins.2022.10.029

Frankel, R. M., Johnson, M. F., & Skinner, D. J. (1997). *An Empirical Examination of Conference Calls as a Voluntary Disclosure Medium* (SSRN Scholarly Paper No. 2748). Social Science Research Network. https://papers.ssrn.com/abstract=2748

Jennings, J. N., Lee, J. A., & Frankel, R. (2021). Disclosure Sentiment: Machine Learning vs. Dictionary Methods. *Manag. Sci.*, *68*, 5514–5532. https://doi.org/10.1287/mnsc.2021.4156

Kantos, C., Joldzic, D., Mitra, G., & Thi Hoang, K. (2022). *Comparative Analysis of NLP Approaches for Earnings Calls* (SSRN Scholarly Paper No. 4210529). Social Science Research Network. https://doi.org/10.2139/ssrn.4210529

Kirtac, K., & Germano, G. (2024). Sentiment trading with large language models. *Finance Research Letters*, *62*, 105227. https://doi.org/10.1016/j.frl.2024.105227

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. https://doi.org/10.1111/j.1540-6261.2010.01625.x

Mayew, W. J., & Venkatachalam, M. (2012). The Power of Voice: Managerial Affective States and Future Firm Performance. *The Journal of Finance*, *67*(1), 1–43. https://doi.org/10.1111/j.1540-6261.2011.01705.x

Price, Doran, Peterson, & Bliss. (2012). Earnings conference calls and stock returns: The incremental informativeness of textual tone. *Journal of Banking & Finance*, *36*(4), 992–1011. https://doi.org/10.1016/j.jbankfin.2011.10.013

Spence, M. (1973). Job Market Signaling*. *The Quarterly Journal of Economics*, *87*(3), 355–374. https://doi.org/10.2307/1882010

Thaler, R. H. (1993). *Advances in behavioral finance*. Russel Sage Foundation.

Tucker, A. M., Xia, Y., & Smelcer, S. N. (2024). Text, tone, and legal language: Analyzing mutual fund disclosure sentiment. *American Business Law Journal*, *61*(1), 57–86. https://doi.org/10.1111/ablj.12239

9. APPENDIX

Figure 3 Heatmap Positive Sentiment

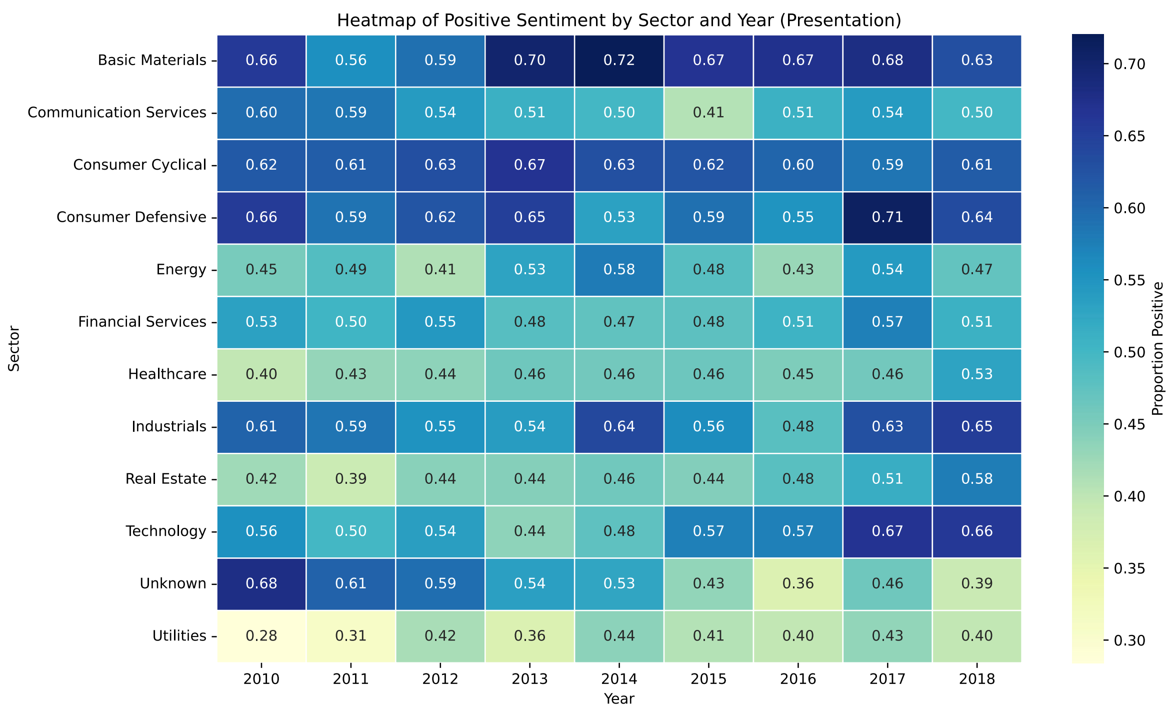


Figure 4 Heatmap Q&A

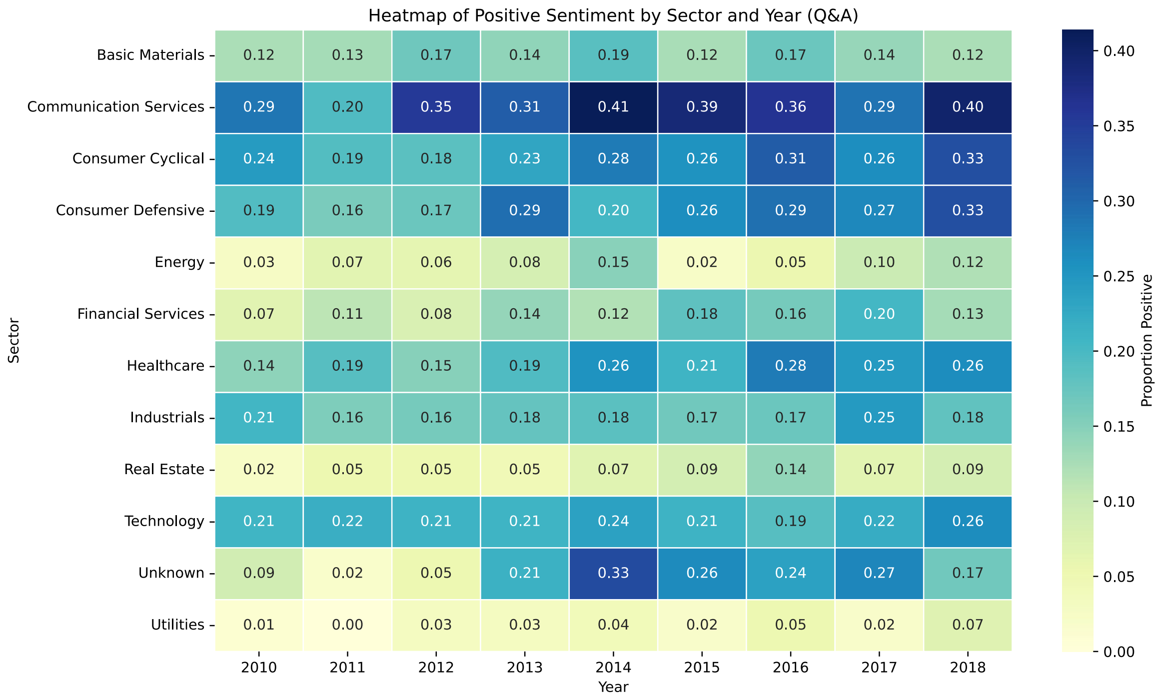


Figure 5 Heatmap Full Call