Norwegian School of Economics

Bergen, Spring, 2023

Options Speak Louder Than Words: Strategic Negativity in Earnings Calls Prior to Option Grants

Nicholas Palacio, Michael Toivonen

Supervisor: Maximillian Rohrer

Master Thesis in Business Analytics

NORWEGIAN SCHOOL OF ECONOMICS

Acknowledgment

We want to thank our supervisor Maximilian Rohrer for his invaluable guidance, advice, and support throughout the writing of our master thesis. His expertise and feedback played a pivotal role in shaping our research, refining our methodology, and enhancing the quality of our work.

Abstract

We investigate whether executive compensation affects disclosure during earnings conference calls. In particular, we hypothesize that executives who have an upcoming option grant will use overly negative language in earnings calls, intending to temporarily depress their companies stock price and obtain a lower strike price on their options. The sentiment during earnings calls is measured with both dictionary-based approaches as well as with the FinBERT and RoBERTa large language models. Our main finding is that executives use more negative language in conference calls in the quarter preceding their option grant than on average. A causal inference is made by leveraging distinct characteristics in multi-year option schedules and by conducting placebo tests with pseudo-option grant dates to validate our results. We learn that executives that stand to benefit from a temporarily reduced stock price will exhibit opportunistic behavior by adopting an excessively pessimistic tone in earnings calls.

Table of Contents

List of Tables

List of Appendix Figures

1. Introduction

Executive compensation packages are structured with the intention of aligning the interests of executives and shareholders, reducing agency concerns, and promoting long-term growth. In 2020, stock options accounted for an average of 19.1 percent of total compensation, serving as the primary form of equity-based compensation awarded to executives (Tonello & Tay, 2021). These options grant holders the right to purchase a specific number of shares at a predetermined strike price on a future date, usually corresponding to the market price. Despite their intended purpose however, recent literature on option-based compensations' impact on executive opportunism has revealed a different pattern: executives will alter their behaviour or disclosures to maximize their personal compensation. This pattern is caused by executives' foreknowledge of option grant dates or equity vesting periods, incentivising them to manipulate stock prices before these points in time to maximise their personal payoff (Quigley, Ward, Hubbard , & Graffin, 2020; Edmans, Goncalves-Pinto, Groen-Xu, & Wang, 2018).

In this paper, we test the primary hypothesis that executives' sentiment on earnings calls will be overly negative if the call directly precedes an option grant. The motive for using strategic negativity is the potential for a larger payoff on their options driven by the temporary reduction in stock price. After mitigating endogeneity concerns, we find that upcoming option grants influence more negative sentiment from executives in earnings calls. When presented with the opportunity to maximize their compensation, managers will strategically convey a negative tone on earnings calls with the intent to exploit the structure of option grant schedules.

Our analysis is based on a large corpus of conference call transcripts, spanning the years 2006 to 2020[1]. Earnings calls are of particular interest for this analysis because they follow a scheduled release pattern unlike discretionary news disclosures. Existing literature emphasizes that earnings calls are increasingly important in conveying pertinent information to the market (Frankel, Johnson, & Skinner, 1999; Kimbrough, 2005). To develop measures for assessing the sentiment of each call, we employ commonly used dictionary-based approaches, as well as large language models (LLMs) developed by Meta and Google. These methods develop a measure for the tone of earnings calls that we assess in relation to an impending option grant.

[1] The conference call data was generously provided to us by Professor Max Rohrer of the Norwegian School of Economics.

Endogeneity problems come from two potential sources. First, option grants are often awarded based on performance and are not independent of the actions of executives. This may lead to the dependent variable (option grant) and the independent variable (sentiment) being correlated. Second, omitted variables such increased competition, industry shocks and other firm characteristics could both impact the reception of an option grant and the sentiment of an earnings call. To address both sources of endogeneity, our analysis utilizes techniques from Shue & Townsend (2017) for classifying option grant cycles. Option grants can be awarded based on performance or by a predetermined schedule. Using the methods outlined in the literature by Shue & Townsend (2017) and Hall (1999) to classify options grant schedules, we only examine scheduled two year cycle grants, ensuring grant dates are unrelated to the performance or actions of the CEO or CFO.

We use similar economic methods to those discussed in Edmans et al. (2018) which examined the relationship between equity vesting months and the number of discretionary news releases. To validate their results, Edmans et al. (2018) conducted a placebo test with a pseudo equity vesting date and found statistically insignificant coefficients on the explanatory variables, suggesting the true results are not driven by other cyclical events. When examining the influence of an upcoming option grant date on the sentiment of the previous earnings call, we employ multiple linear regressions with the sentiment of an earnings call as the dependent variable and an upcoming option grant as the predictor. We validate our results by conducting a similar placebo test using a pseudo date of 90-days following the scheduled grant to ensure that the firm maintains consistent firm-level characteristics and minimum variation in other control factors. Our expectation is that these pseudo-option grants will show no effect on the tone of the conference calls, suggesting that abnormal negative sentiment is unique to the call preceding an option grant date.

We reduce the threat caused by endogeneity by only considering options that are awarded as part of a schedule, as well as improve the robustness of our results by completing placebo tests and conducting series of comparisons across subsets of the data. These comparisons involve examining the results for the presentation section of the call, where executives have more discretion compared to the question-and-answer period. Additionally, we test whether option grants making up larger proportion of an executive’s total compensation further induces this opportunistic behavior.

As part of our secondary analysis, we look directly at the relative frequency of words that are used in calls preceding option grants. We discover that words that have been previously demonstrated to result in lower stock returns are more commonly used by executives in calls proceeding a grant date (Li, 2006). This extends beyond positive and negative words to terms associated with uncertainty as well. This provides further evidence that executives adopt a more pessimistic and uncertain tone in their communication to elicit a temporary negative market reaction.

Our thesis makes contributions to two bodies of literature: stock options and corporate disclosures. Starting with the former, there have been several papers examining the relationship between stock options and executive behavior or action. Lie (2005), investigates unscheduled option grants, finding negative abnormal stock returns before a grant date and positive afterwards. He suggests that some CEOs engaged in option back-dating where they designate their option grant dates prior to their boards' decision date (Lie, 2005). Daines, McQueen and Schonlau (2018) examine opportunistic behavior surrounding scheduled option grants and finds negative abnormal returns before option grant dates even after regulatory changes to prevent option backdating. The negative abnormal returns are partly caused by the timing and substance of news disclosures (Daines, McQueen, & Schonlau, 2018). Shue & Townsend (2017) examine how scheduled option grants influence the risk-taking behavior of CEOs, finding a 10 percent increase in new options granted leads to a 2.8–4.2 percent increase in equity volatility. These papers show the influence that options have on the decision making and the behavior of CEOs. Our approach varies from Lie (2005) by examining scheduled option grants that are independent of executive performance. We differ from Daines et al. (2018) and from Shue & Townsend (2017) by showing the influence of options extends beyond news releases and risk taking to the disclosures of executives on earnings calls.

Secondly, there is extensive research on corporate disclosures and equity incentives. The volume and tone of corporate disclosures is impacted by compensation plans of executives. Brockman, Khurana and Martin (2008) find an increase in the frequency and magnitude of negative news during the month before repurchasing shares. Edmans et al. (2018) focuses particularly on the CEO, revealing their firms release 20% more discretionary news items in months where their shares are scheduled to vest. Our paper differs by focusing not on disclosures in news releases, but rather on the tone and the language of executives during earnings calls.

In the first section of our analysis, we develop our hypothesis based on previous literature. Then, we provide a summary of the data used for our analysis, followed by sentiment score development. Finally, we present and discuss our findings, as well as acknowledge the limitations of our analysis.

2. Hypothesis Development

In this section, we review the relevant literature within corporate governance and compensation and executive opportunism to develop our hypothesis. This section also draws on seminal research establishing the effects of executive tone on stock market reactions, in particular in earnings calls.

2.1 Option Grant Features & Identificaiton

The most common form of equity-based compensation for executives is stock options, which can be categorized based on their *plan type*. Hall (1999) detailed how option grants cycles can be classified as either a fixed-value or fixed-number. A fixed-value plan stipulates that the options granted to an employee will be the same value each year for the duration of the schedule. For example, a CEO may receive $100,000 in stock options each year for the next two years with minimal variation. Alternatively, a fixed number plan will dictate that an employee receives the same number of stock options, regardless of the price of the option, each year for the duration of the grant schedule.

An option grant schedule refers to the timing and terms for when an executive is granted stock options as part of their compensation package. These options usually are issued in 12-month intervals and can be part of a broader grant *cycle* (Shue & Townsend, 2017). This is consistent with how we classify options discussed in section 3.2.2. A cycle simply refers to the number of scheduled occurrences of options being granted that fall within a plan type, and on average last about two years. Using the CEO example in the previous paragraph, the cycle length would be two years and the grants would be awarded once per year.

The use of option grants by companies seems intuitive: encourage employee retention while ensuring that the option holding executives are incentivized to work towards the long-term success of the company. This shared interest between shareholders and executives is referred to as Agency Theory (Quigley, Ward, Hubbard , & Graffin, 2020). While this seems logical, the literature on executive compensation and stock performance has poked holes in this theory, finding that in the short-term lead-up to an option grant, this may not always be the case.

One reason for this could be executives' prior knowledge of their option grant schedules and plan types. By knowing exactly when they will be awarded stock options, executives have a clear information advantage which they can exploit for their personal gain.

2.2 Company Disclosure & Manipulation

Previous research has demonstrated that market participants responses to mandated reports and voluntary disclosures are influenced by their content and timing (Bagnoli, Kross, & Watts, 2002; Bowen, Davis, & Matsumoto, 2005). Recently, this research has expanded beyond the quantitative content of reports and timing of disclosures to the tone and language used. For instance, Tetlock (2007) found that the market reacted negatively to increased pessimism in a prominent Wall Street Journal column. Similarly, Davis and Tama-Sweet (2012) observed that managers who are more sensitive to stock price fluctuations used less pessimistic language in earnings press releases compared to management discussion and analysis. While the study did not provide conclusive evidence that this led to future abnormal returns, it suggests that managers may avoid pessimistic language to reduce negative market reactions.

This body of literature reveals a notable correlation between the language of company executives in various financial mediums, such as conference calls and news disclosures, and market reactions. This raises questions about the motivations of executives, and if they were to benefit from a negative market reaction would they use more pessimistic language at the expense of shareholders. This leads us to our first hypothesis that this paper aims to test.

(1) Executives of publicly traded companies who have an upcoming option grant in the next fiscal quarter will use overly negative language in earnings calls with the intention to temporarily depress their companies stock price and obtain a lower strike price.

This hypothesis is firmly rooted in the disclosure manipulation and compensation literature, both shortly after equity vesting dates (Edmans, Goncalves-Pinto, Groen-Xu, & Wang, 2018) as well as in the lead up to option grant dates (Shue & Townsend, 2017; Daines, McQueen, & Schonlau, 2018). For example, in his research on equity vesting and discretionary news disclosures, Edmans et al. (2018) found that in vesting months, 57% of CEOs would sell their equity, while the companies they managed would release 20% more discretionary news items, the vast majority of which were positive. This increase in positive news disclosures would temporarily boost the company's stock price, resulting in greater profit when the CEO cashed out. These finding were compounded when considering positive news releases significantly declined in the immediate months following a vesting month, in addition to negative news disclosures being more unlikely during a vesting month. These findings show a clear pattern from executives of selectively choosing the most opportune times to disclose news about their company for personal gain.

Similarly, research on executives engaging in this behaviour in the lead up to an option grant date has only furthered this discussion. For example, just as (Edmans, Goncalves-Pinto, Groen-Xu, & Wang, 2018) found that CEOs release more positive news about their companies in vesting months, there has been a similar and opposite behaviour observed in the month proceeding an option grant (Quigley, Ward, Hubbard , & Graffin, 2020). A similar effect has been observed through significant negative abnormal returns (CARs) in the lead up to a grant date, and positive CARs shortly after (Daines, McQueen, & Schonlau, 2018). Using CARs as an indicator, this further supports the claim that executives manipulate disclosures for personal gain. These findings lead us to expand our analysis, pertinent on if hypothesis 1 is proven to be true, we would then hypothesize:

(1.1) This behaviour will be observed more in the presentation section of earnings calls, where executives have more discretion on what they say.

(1.2) This behaviour will be more prevalent in executives with a higher share of their total compensation that is made up of options.

Hypothesis 1.1 is based on the well-established finding that more informative information is shared during the Q&A section of conference calls, because the presentation section gives CEOs more discretion in what they disclose. Hypothesis 1.2 is inspired by findings that "underpaid" CEOs are more likely to engage in disclosure manipulation (Quigley, Ward, Hubbard , & Graffin, 2020; Matsumoto, Pronk, & Roelofsen, 2011).

3. Data and Variable Selection

In this section, we provide an overview of the data used in our analysis. We describe how we gathered the data, including how the text data was pre-processed. Additionally, we explain the method used to classify option grant schedules and discuss the control variables used in the analysis. A detailed description of every variable used in our analysis can be found in Appendix 1.

3.1 Collection

The data collected for this analysis is categorized into three groups: company conference call transcripts, executive compensation data, and financial performance data. There is an intuitive relation between these datasets: company financial data, such as measures of liquidity and revenue are reported quarterly shortly after an earnings conference call, while stock return data in the lead up to and shortly after a conference call provides insight into how the market values a company at a given point in time.

All data included in this analysis is post-2006 when new accounting standards were put into place in the U.S. by the securities and exchange commission (SEC). These changes were made in part to make disclosures about executive compensation of publicly traded companies more transparent as well as to restrain option backdating (SEC, 2006).

3.1.1 Earnings Call Transcripts

An earnings call is a conference call that is held by publicly traded companies to discuss their financial performance and results for the previous quarter. The participants of the call are the firm's management, as well as analysts that cover the firm. Each call begins with a presentation by the firm's management, usually a reiteration of the press release, disusing key highlights from the previous quarter, followed by a question-and-answer period with analysts.

The conference call data was provided to us by Professor Max Rohrer of the Norwegian School of Economics[2]. This contains quarterly conference call transcripts formatted in a structured

[2] This data was comprised of conference call transcripts collected from Seeking Alpha (2006 to 2016) and Wall Street Horizons (2016 to 2020).

way with the speaker and the section of the conference call tagged. This contains 81,593 calls from 4,980 firms. After filtering for available executive compensation information from Execucomp, and controls from Compustat and the Centre for Research in Security Prices (WRDS, 2023), we were left with 38,935 calls from 1,660 firms.

3.1.2 Stock Returns

The CRSP stock database from Wharton Research Data Services was used for collecting stock price and industry level data (WRDS, 2023). Stock prices were used as a control as they are reflective of investor perception of a company at a given point in time, and industry codes were used to segment model results in the Results and Discussion Section. The variables collected from CRSP were stock closing price, quarterly low and high price, common shares outstanding, total monthly return and GIC sector.

3.1.3 Company Financials

Financial information on the companies with earnings call transcripts was retrieved from the Compustat Database. The retrieved data was merged with the conference call data and used as controls for the econometric modelling. The collected variables include dividend yield, market value, return on assets, capex, and revenue growth.

3.1.4 Executive Compensation

Executive compensation data was collected from Execucomp, a database of executive compensation packages for S&P 1000 companies offered by Wharton Research Data Services (WRDS, 2023). The data includes information on executives’ base salary, bonus and stock options granted, total direct compensation, among other equity-based incentives. This data was used to establish the option plan type (fixed value vs. fixed number), and vesting schedule for CEOs and CFOs. After filtering to only include instances when an option was granted, the sample was reduced to 66,708 observations of options granted, for 12,000 executives between 2006 and 2021.

3.2 Preprocessing

3.2.1 Earnings Calls Transcript Cleaning

To calculate sentiment scores for earnings calls, the transcripts underwent a cleaning process based on established practices in NLP literature. Firstly, the text was tokenized, meaning that it was split into individual words or "tokens," and each token was tagged using the udpipe package in R to eliminate proper nouns. This allowed for the removal of words that did not contribute to the sentiment analysis, such as names of people or companies. Secondly, the snowball stop-word list[3] in the “tm” R package was used to remove unhelpful words. Stopwords are words that are common in language but do not convey much meaning, such as "and," "the," or "a.". Some words from the snowball list could convey tone such as "against," "above," "below," "up," "down," "over," "under," "again," "further," "few," "more," "most," "no," and "not" and were kept in the conference call transcripts. Additionally, all contractions, such as "I’ll" were converted to their uncontracted forms “I will”. Finally, all punctuation and numerical characters were removed, and the remaining text was converted to lower case. These steps ensured that the text was in a standardized format for calculating sentiment scores when using the dictionaries. Lemmatization is commonly used where words are converted to their respective stem (i.e., Winning to win). However, the Loughran McDonald (LM) and Machine Learning dictionaries were constructed using unstemmed words, so the conference calls words were kept in their original form.

3.2.2 Classifying Grant Plan Types and Schedules

Public companies are not required to disclose the plan type or grant schedule for options granted to executives; however, these classifications can be approximated using historical data (Shue & Townsend, 2017). To do this, we first calculated the number of months between option grants. Following practices established in Shue & Townsend (2017) and Hall (1999), only the largest grant was used for classifying plan type for executives that received multiple grants per year. This is because larger grants are more likely to be related to a recurring compensation plan, whereas smaller, more randomized grants are likely related to bonuses or other auxiliary forms of compensation that is not unique to the executive but is common to all

[3] https://stopwords.quanteda.io/reference/data_stopwords_snowball.html

executives in the company. This did not change the sample size significantly as most executives received only one option grant per year.

Of the total options granted, approximately 65% occurred 12 months after a previous option was granted, while 76% fell within a range of 11 to 13 months (Table 1). This increase in percentage for the range one month before and after the 12-month mark was also observed around the 3- and 9-month intervals. This is likely due to two reasons: grants primarily being scheduled in quarterly intervals throughout the year and slight differences in grant dates leading to an under or overestimate of the grant date by one month. For our analysis, we looked at executives with annual option grant cycles.

Table 1: Months between option grant dates

Month	Grants	Percent	Cumulative Percent
1	123	0.28	0.28
2	237	0.53	0.81
3	939	2.11	2.91
4	348	0.78	3.69
5	305	0.68	4.38
6	847	1.90	6.28
7	523	1.17	7.45
8	541	1.21	8.66
9	987	2.21	10.88
10	893	2.00	12.88
11	2,530	5.67	18.56
12	29,006	65.06	83.62
13	2,682	6.02	89.63
14	742	1.66	91.30
15	729	1.64	92.93
16	349	0.78	93.72
17	285	0.64	94.35
18	386	0.87	95.22
19	184	0.41	95.63
20	181	0.41	96.04
21	214	0.48	96.52
22	183	0.41	96.93
23	241	0.54	97.47
24	1,128	2.53	100

This table reports the number of months between option grants. There are a total of 44583 observations in the data. Executive option dates are collected from Execucomp from 2006 to 2020. We consider only the time between grant dates up to 24 months as this is the most common cycle length.

To determine whether an executive's option plan type was fixed value or fixed number, we calculated the percent change in both the face value and the number of options from two consecutive grant dates. The face value of a stock option is simply the number of options granted multiplied by the market price on that day. If the number of shares an executive received on their present grant date was within 1 percent of the previous grant or within the proportions of previous stock splits, they would be classified as being on a fixed-number plan type. Alternatively, an executive would be classified as being on a fixed-value plan type if the face value of the grants they received in the present period was within 3 percent of the previous grant. In scenarios where neither was exactly true, the grant would be classified as "unknown". We allowed these "buffers" to account for methodological differences in how companies calculate the fair value of their options (Shue & Townsend, 2017).

When classifying executive plan types, approximately 27.60 percent of executives were associated with a fixed-number plan, while 15.95 percent were associated with a fixed-value plan (Table 2**)**. These values are likely low due to the conservative thresholds (1 percent and 3 percent) used in classifying plan types, as discussed above. For example, if threshold values for classification were increased to 5 percent for both fixed-number and fixed-value plan types, the proportion of executives under each plan would increase to 34.1 and 19.9 percent respectively. These classification proportions are consistent with those of Shue and Townsend (2017), who used similar threshold for classifying plan type.

Table 2: Plan type summary statistics

Plan Type	Count	Percent
Fixed Number	3,460	27.60
Fixed Value	1,999	15.95
Unknown	7,077	56.45

This table reports the summary statistics for the plan type each executive was classified as. The data was collected from Execucomp and filtered for CEOs and CFOs that spoke on conference calls. Executives that could not be classified as either fixed value or fixed number were assigned an unknown classification.

3.2.3 Variable Selection

This study uses sentiment scores as dependent variables, which are calculated in three different ways: the first employs LMs popular financial annotated dictionary, while the remaining two use Meta's RoBERTa and Google's FinBERT model described in the following section. The study employs option grant information as independent variables, including a binary variable indicating whether executives are expected to receive an option grant in the next quarter (1 if yes, 0 if no) and a variable representing the fair value of the upcoming grant.

To isolate for the impact that upcoming option grants have on sentiment firm-level control variables were used. These include revenue growth, quarterly return, return on assets, dividend yield, and market value. Additionally, earnings surprise or "SUE Score" is retrieved from IBES, which is calculated with the following formula.

$$SUE_t = \frac{EPS_t - mean(Expected(EPS_t^{Analyst}))}{SD(EPS_t)} \quad (1)$$

Table 3 contains a summary of the independent and control variables that are used throughout the analysis. Based on the availability of control variables and filtering executives that we can attribute to an option cycle; we are left with roughly 17,439 calls containing a total of 1530 executives. Each observation is an executive in a particular section of a conference call which explains for the 42,411 observations for 17,439 calls. The table shows that in 6 percent of observations, an executive receives an option grant as part of their grant cycle within three months after the call. Each observation is separated into the respective section which is why half of the observations are the presentation section. This is relevant for testing hypothesis 1.1. The rest of the table summarizes the control variables used in the regression in section 5.

Table 3: Control variable summary statistics

Statistic	N	Mean	St. Dev.	Min	Max
Upcoming Option	42,411	0.06	0.25	0.00	1.00
Presentation Section	42,411	0.50	0.50	0.00	1.00
Options Percent of Compensation	42,411	0.003	0.09	0.00	1.00
Dividend Yield	42,411	1.17	1.44	0.00	37.5
Revenue Growth	42,411	0.04	0.84	-14.74	139.81
Return on Assets	42,411	0.14	0.10	-0.84	0.87
Earnings Surprise	42,411	1.38	5.76	-397.81	253.79
Quarterly Stock Return	42,411	0.12	10.74	-0.91	1,561.50
Fair Value of Option	42,411	12.71	41.06	0.00	1,736.07
CAPEX	42,411	269.00	981.92	0.00	37,985.00
Market Value	42,411	12,047.60	27,354.16	0.00	394,755.70
Shares Owned	42,411	1,702.27	19,047.43	0.00	1,178,771.00

This table reports summary statistics for the control variables used in our analysis. Compensation data was collected from Execucomp, while company financial data was collected from CRSP and Compustat between 2006 and 2020.

A correlation matrix was created to examine the relationship between variables used in our analysis Appendix 2. We used the variance inflation factor (VIF) to assess multicollinearity in our regressions. Based on the VIF scores with the highest being 1.2, we can conclude that our independent variables are not largely influenced by one another. Appendix 3 displays the VIF measure for the independent variables.

4. Sentiment Score Development

This section discusses the methodologies used to calculate sentiment scores for executives in earnings calls. Two approaches were employed: a dictionary-based approach using Loughran and McDonald's (LM) financial sentiment dictionary, and a large language model approach, using two LLM models (RoBERTa and FinBERT) based on Googles seminal BERT model. Finally, a summary of the sentiment scores obtained from each method is provided.

4.1 Dictionary-Based Approach

Annotated dictionaries are commonly used to calculate the sentiment of textual data. The following formula is used to obtain the sentiment of a document of text, where *d* represents the document that is being measured.

$$Sentiment_d = \frac{PositiveWords_d - NegativeWords_d}{TotalWords_d} \tag{2}$$

The LM dictionary is specifically designed for calculating sentiment in financial communications. Unlike previous dictionaries, which often fail to capture the uniqueness of language used in financial reports, Loughran and McDonald offer a more nuanced approach. For example, in previous annotated dictionaries, "tax" is classified as a negative word where in a business context it is a frequent word that does not convey meaningful information. The LM dictionary is comprised of 2355 negative words associated with negative implications to finance, as well as 354 positive words (Loughran & McDonald, 2011). This comprehensive dictionary has gained significant recognition in financial sentiment research, evident from its substantial citation count of 4,440 on Google Scholar. In our calculation of sentiment score, we solely used the LM negative dictionary to calculate the sentiment scores, as we are more interested in the negative tone of the call.

To isolate the impact of option grants on individual sentiment, a distinct sentiment score is calculated for both the CEO and CFO that are participating in each call. The sentiment score of each speaker is calculated using equation 1, where each document represents a speaker of a section in each call. The resulting score reflects the overall sentiment of the speaker during the call, which was then compared to the subsequent option grant to evaluate for a potential causal relationship.

4.2 Large Language Model (LLM) Approach

4.2.1 Model Background: FinBERT & RoBERTa

BERT or Bidirectional Encoder Representations from Transformers is a large language model (LLM) developed by Google AI in 2018 (Devlin, Chang, Lee, & Toutanova, 2018). BERT was trained on two primary corpora; BooksCorpus and English Wikipedia, which together contained a total of 3.3 billion tokens, and was designed to understand the context of words and sentences based on surrounding text. Unlike one-directional LLMs, which read text sequentially (from left to right), BERT utilizes transformers, a type of model architecture, to perform bidirectional context modelling (Vaswani, et al., 2017). Transformers enable the model to incorporate text that comes both before and after a word (left and right side) to better determine its context in a sentence. This helps the model disambiguate complex sentences and capture nuanced relationships between words, regardless of their placement in a sentence. To do this, BERT utilizes two unsupervised training approaches: masked language modelling (MLM) and next sentence prediction (NSP) (Devlin, Chang, Lee, & Toutanova, 2018). MLM involves randomly covering or "masking" some words in the input of the model (15% in the case of BERT), and then having the model predict which of those words are based on the surrounding text. Similarly, next sentence prediction (NSP) involves inputting two sentences into the model and having it try to predict which comes first in a text.

MLM and NSP help BERT learn the relationships between words and sentences, and the context in which they are used (Devlin, Chang, Lee, & Toutanova, 2018). Since its initial release however, NLP researchers fine-tuned or created entirely new models from base BERT, by training their models using domain-specific data, or by altering the models training architecture all together. Two such models that do this are FinBERT and RoBERTa.

FinBERT is a version of the base BERT model, but was trained and fine-tuned on financial texts such as: corporate annual reports (10-Ks and 10-Qs) from the SEC's EDGAR database, analyst reports from the Thomson Investext database and earnings call transcripts from Seeking Alpha (Araci, 2019; Huang, Wang, & Yang, 2021)[4]. In total, FinBERT was trained on 4.9 billion tokens (32% more than BERT) and can classify text as either "positive",

[4] FinBERT was first developed by (Araci, 2019), with later iterations and fine tunning for specific NLP tasks. This analysis will use (Huang, Wang, & Yang, 2021) adaption of FinBERT, which has been fine-tuned for sentiment analysis on conference calls, for all zero-shot sentiment analysis experiments.

"negative" or "neutral". RoBERTa or Robustly Optimized BERT Pretraining Approach is a LLM developed by Facebook AI Research (FAIR) in 2019 (Liu, et al., 2019). The model builds off the architecture of BERT, but with two key changes to improve performance on NLP tasks like sentiment analysis: an increased training dataset and dynamic masking.

RoBERTa is trained on a much larger dataset than BERT, totalling 160 GB or roughly 10 times the size of the base BERT model, enabling more applications to a wider range of texts. This includes the base BERT training data, as well as additional data from the CC-News dataset (63 million English news articles between 2016 and 2019), OpenWebText (23 million URLs and 10 million HTML pages extracted from Reddit)[5] and the STORIES dataset (32 GB subset of CommonCrawl, an archive of internet web pages). In addition to an increased training dataset, RoBERTa also utilizes a dynamic masking technique, where different sets of words are masked out for each example in the training dataset. This differs from BERTs static masking technique, where the same set of words are masked for each example in the training set and helps RoBERTa better understand the bidirectional relationships between words in a sentence (Liao, Zeng, Yin, & Wei , 2020).

When tested on professionally annotated financial texts from the Financial Phrase Bank, both FinBERT and RoBERTa have been shown to outperform dictionary-based models (such as LM's), when classifying sentiment (Pekka , Sinha, Takala, Korhonen, & Wallenius, 2013; Leippold, 2023). However, the application of FinBERT for sentiment analysis on conference call data has largely been untested.

4.2.2 Implementation of the Large Language Models

The RoBERTa[6] and FinBERT[7] models used for this analysis were downloaded from the Hugging Face model hub. Each model was run using a GPU on Googles' Collaboratory. Because FinBERT was trained on a large amount of financial data, including conference call transcripts (the primary data source of interest for this analysis), the base model required minimal fine tuning for this analysis. Experiments that compare the base models of FinBERT and RoBERTa with lexicon-based models were shown to be sufficient (Leippold, 2023). To

[5] https://github.com/jcpeterson/openwebtext
[6] RoBERTA model found at https://huggingface.co/siebert/sentiment-roberta-large-english
[7] FinBERT model found at https://huggingface.co/yiyanghkust/finbert-tone

build off of the base models, several pre-processing steps were applied to the conference call data before applying RoBERTa and FinBERT, which included:

(i) Useing the Transformers package in Python to set up model, tokenize the text, and implement the models from Hugging Face.

(ii) For longer blocks of text above BERTs' 512 token limit, we separated the text into *chunks* of a maximum of 400 words.

Creating chunks of text was used to classifying sentiment primarily for the Presentation Section of earnings calls because these sections involve large components of uninterrupted speech containing an average of 893 tokens per speaker as seen in Table 4[8]. Alternatively, the Q&A section has smaller blocks of text with approximately 2.4 million rows distinct blocks and an average token length of roughly 75. Therefore, most rows in the Q&A section did not require splitting the text into chunks.

Table 4: conference call token count by section

Call Section	Rows	Token Count (avg)	Chunks (avg)
Question and Answer	2,414,083	74.59	0
Presentation	145,610	892.78	2

This table reports the average number of rows (chunks) and tokens per row (chunk) that are in each section of an earnings call. Each chunk is approximately 400 tokens in length. The earnings call data was collected from Seeking Alpha and Wall Street Horizons and encapsulates the period between 2006 and 2020.

After running the RoBERTa and FinBERT models, each chunk of text received a sentiment classification of either positive, negative, or neutral. The chunks were then aggregated using the common keys (*call id, section type, speaker name*) to get the overall sentiment of a single person (CEO or CFO) in a section of a conference call (Presentation or Q&A). The formula for this classification is as follows:

$$LLMSentiment_{ijk} = \frac{PositiveChunks_{ijk} - NegativeChunk_{ijk}}{TotalChunks_{ijk}} \qquad (3)$$

[8] Tokens per speaker = 892 (avg tokens per row/chunk) * 2 (avg chunks per speaker)

Here, *ijk* represents executive *i* during call section *j* (presentation or question and answer period) in earnings call *k*. Chunks were aggregated using a similar approach to Leippold (2023), who aggregated at the sentence level when measuring sentiment using BERT.

It should be noted that while this technique for chunking text adheres to the token processing limits of BERT, there are drawbacks if sentences are cut off as part of a "chunk", which may alter the sentiment of a sentence. To address this, we created a parameter that each chunk must end with a period. This provides more contextual information by allowing both models to reference all of the sentences in each chunk.

4.3 Sentiment Score Summary

Each section of each call is assigned a dictionary-based sentiment score, however when using the LLM models, some of the text data is not interpretable by the model and is unable to be classified. This results in 31,869 LLM sentiment scores and 42,411 LM dictionary sentiment scores. Both sample sizes are sufficient for the analysis in the following section. Table 5 shows a summary of the sentiment score calculations in the analysis. The sole negativity of the LM dictionary is explained using only the negative LM words. RoBERTa tends to score more of the text positively and there is the least deviation between each observation. When FinBERT is used, which is trained on a corpus of financial text, the average tone is still positive but there is higher variation between observations. Comparing the LLM scores to the LM dictionary score requires more analysis beyond the scope of this paper. These are tools used to assess the tone of speakers during an earnings call. By comparing the regression outputs from the different scoring methods and find a similar pattern in the predictor variables, our findings avoid model selection bias.

Table 5: Sentiment score summary statistics by dictionary

Statistic	N	Mean	St. Dev	Min	Max
LM Dictionary	42,411	-1.53	0.98	-5.13	0.00
RoBERTa	31,869	0.78	0.31	-1.00	1.00
FinBERT	31,869	0.53	0.36	-1.00	1.00

This table reports summary sentiment score data for each of the models used. The LM dictionary was created in 2011, while RoBERTa and FinBERT were released for public use in 2019 and 2021, respectively.

5. Analysis and Results

This section outlines the primary and secondary results of the analysis, followed by a discussion on what these results indicate. The primary focus is to isolate for the relationship between conference call sentiment and the presence of an option grant in the subsequent quarter. To mitigate endogeneity concerns and establish a causal relationship, control variables, fixed effects, and placebo testing are employed.

5.1 Primary Results

5.1.1 Upcoming Options Effect on Earnings Call Sentiment

To determine the influence that an upcoming option grant has on the sentiment of an executive in a call, linear regression is used, with control variables included to improve the causal inference. By incorporating *firm*, *fiscal year*, and *quarterly fixed effects*, we control for unobservable firm-level or time specific determinants of sentiment and options grants. The quarterly fixed effect is used to address for the majority of option grants being in the fourth quarter. The model can be summarized by the equation below, where *t-1* represents the sentiment score of the earnings call in the previous quarter to the option grant, and *controls* encompass the variables described in section 3.2.3.

$$\begin{aligned} SentimentScore_{i,t-1} = \alpha &+ \beta_1 * UpcomingOption_{i,t} \\ &+ \beta_2 * OptionFairValue \\ &+ \lambda * Controls_t + FixedEffects + \epsilon \end{aligned} \tag{4}$$

Table 6 displays the results of the regression analysis, conducted on a subset of executives who were assigned to a specific option plan type using the methodology outlined in section 3.2.2. Each observation represents an executive, with sentiment scores calculated for their respective earnings calls. The results of this analysis reveal three key findings regarding the relationship between option grants and sentiment expressed in preceding earnings calls.

Table 6: Regression output of an upcoming option on earnings call sentiment

Dependent Variables:	LM Dictionary Score (1)	RoBERTa Score (2)	FinBERT Score (3)
Upcoming Option	-6.29***	-1.67**	-1.81**
	(2.23)	(0.841)	(0.915)
Fair Value of Option	0.180*	0.031	0.042
	(0.104)	(0.020)	(0.032)
Dividend Yield	-6.99***	-2.11***	-3.11***
	(1.33)	(0.464)	(0.568)
Log(Market Value)	8.05***	1.20	2.10**
	(2.46)	(0.820)	(0.880)
Revenue Growth	1.09***	0.258***	0.172
	(0.380)	(0.059)	(0.139)
Return on Assets	67.9***	10.0**	7.29
	(17.9)	(5.03)	(5.86)
Earnings Surprise	0.780***	0.249***	0.414***
	(0.281)	(0.046)	(0.058)
Quarterly Stock Return	0.012	0.003*	0.004
	(0.009)	(0.002)	(0.003)
Fixed-effects			
Firm Effect	Yes	Yes	Yes
Fiscal Year	Yes	Yes	Yes
Fiscal Quarter	Yes	Yes	Yes
Fit statistics			
Observations	21,970	15,127	15,127
R^2	0.359	0.252	0.350
Within R^2	0.021	0.010	0.018

Clustered (Firm Effect) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

First, whether using the LM dictionary or the two LLM models, an upcoming option grant shows to have a statistically significant influence on the sentiment expressed by executives during a call. The impact of an upcoming option is most pronounced when employing the LM dictionary, as reflected by the largest negative coefficient in the sentiment score calculation. However, it is important to note that we do not ascribe greater importance to the LM calculation of sentiment compared to the LLM models, as they are not scaled in relation to one another. Each model illustrates the relationship between an upcoming option and the overall tone of an earnings call, yielding consistent outcomes across all sentiment measures.

Second, we find that an upcoming option grant has a negative correlation with the sentiment of the preceding call. This finding is consistent with our hypothesis and supported by previous literature on tone and manager incentives. This is also consistent with prior research showing executives tend to release more positive discretionary news items during months when they anticipate selling equity, while those who are more sensitive to stock movements avoid employing pessimistic language (Edmans, Goncalves-Pinto, Groen-Xu, & Wang, 2018). These results demonstrate that executives who stand to benefit from downward stock price movements strategically use more negative language, which is reflected by the inverse relationship between upcoming option grants and sentiment.

Third, we do not find sufficient evidence to support the conclusion that the size of the grant or its fair value has a significant impact on the sentiment of the call. One factor that contributes to this outcome is that larger option grants are typically awarded to higher-paid managers, making it difficult to isolate the specific impact of the grant size on sentiment (Tonello & Tay, 2021). This is pattern is also seen by Shue & Townsend (2017), who found that the marginal effect of new options on risk taking behaviour by CEOs was weaker if the executive's total compensation is higher. This is because when an option grant represents only a small proportion of the executive's total compensation, its influence on their behaviour and sentiment may be limited.

To gain a more meaningful understanding, we examine options as a percentage of total compensation in section 5.2.2, as this approach provides greater insights into the relationship between compensation structure and executive behaviour.

5.1.2 Placebo Test

To validate the robustness of our results, we conduct placebo tests to isolate the causal impact of an upcoming option grant on the tone of the preceding conference call. This approach aligns with the method used by Daines, McQueen, and Schonlau (2018), who created a pseudo-option grant date to determine if the pattern of abnormal returns persists on the fake grant date. Their results proved to be insignificant, suggesting negative abnormal returns are specific to the grant date. By introducing a pseudo grant date occurring three months after the scheduled grant date, we maintain consistent firm-level characteristics and minimize variations in other control factors. This approach validates our results and supports a causal relationship between option-based compensation and sentiment. The results show that the pseudo grant date is not statistically significant in influencing the sentiment score of the preceding conference call, lending support to Hypothesis 1 (Table 7).

Table 7: placebo regression of an upcoming option on sentiment of earnings calls

Dependent Variables:	LM Dictionary Score	RoBERTa Score	FinBERT Score
	(1)	(2)	(3)
Upcoming Pseudo Option	-1.22	-0.116	-0.732
	(2.36)	(0.835)	(0.908)
Fair Value	0.160	0.017	0.032
	(0.105)	(0.019)	(0.030)
Dividend Yield	-7.02***	-2.12***	-3.11***
	(1.34)	(0.466)	(0.569)
Log(Market Value)	8.14***	1.21	2.15**
	(2.47)	(0.822)	(0.882)
Revenue Growth	1.10***	0.261***	0.175
	(0.383)	(0.059)	(0.140)
Return on Assets	69.3***	10.3**	7.61
	(17.8)	(5.03)	(5.87)
Earnings Surprise	0.771***	0.247***	0.410***
	(0.281)	(0.046)	(0.058)
Quarterly Stock Return	0.014	0.005***	0.005*
	(0.010)	(0.002)	(0.003)
Fixed-effects			
Firm Effect	Yes	Yes	Yes
Fiscal Year	Yes	Yes	Yes
Fiscal Quarter	Yes	Yes	Yes
Fit statistics			
Observations	21,970	15,127	15,127
R^2	0.359	0.251	0.350
Within R^2	0.020	0.009	0.018

Clustered (Firm Effect) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

5.2 Secondary Results

Recognizing the influence of options on executive sentiment during calls, we conduct a comparative analysis to better understand the magnitude of this behavior. This section focuses on testing the two secondary hypotheses: first, whether executives exhibit increased strategic negativity in the presentation section where they possess greater discretion, and second, whether executives with a total compensation comprised of higher proportion of options are more inclined to engage in strategic negativity prior to option grant dates.

5.2.1 Presentation Versus Question & Answer Period

To examine the validity of hypothesis 1.1, we conduct a regression analysis similar to section 5.1.1, but with the inclusion of two additional variables. The first of these variables, $Presentation^{Binary}$, is a binary indicator that assumes a value of one when referring to the presentation section of the call, while the second binary variable, $Presentation^{Binary}_{Option}$ takes the value of one for the presentation section preceding an option grant date. This measures the direct impact that an upcoming has on the sentiment of the presentation section of an earnings call. The equation that models this relationship is:

$$\begin{aligned} SentimentScore_{i,t-1} = \alpha &+ \beta_1 * UpcomingOption_{i,t} \\ &+ \beta_2 * Presentation^{binary} \\ &+ \beta_3 * Presentation^{binary}_{Option} \\ &+ \lambda * Controls_t + FixedEffects + \epsilon \end{aligned} \tag{5}$$

The results for this regression are presented in Table 8. We can infer that an upcoming option has no significant influence on the sentiment of the presentation section compared to the rest of the call. This does not provide sufficient evidence to support hypothesis 1.1, that strategic negativity in earnings calls is more prevalent in the presentation section. We find that on average, the presentation section has a more positive tone than the question-and-answer period. This is consistent with the findings of Huang, Wang, & Yang (2021), who found that the Q&A section of conference calls conveyed more important information, as the presentation section is usually just a reiteration of the quarterly press releases (Price, Doran, Peterson, & Bliss, 2012). Like the previous regression analysis, a placebo test was conducted, finding no significant relation between pseudo dates and the sentiment of the presentation section of a call (Appendix 4).

Table 8: Regression of upcoming option grants on presentation sentiment

Dependent Variables:	LM Dictionary (1)	RoBeRTa (2)	FinBERT (3)
Upcoming Option	-4.64*	-2.42**	-1.22
	(2.70)	(1.02)	(0.876)
$\text{Presentation}^{binary}$	5.19***	31.6***	36.9***
	(1.67)	(0.684)	(0.736)
$\text{Presentation}^{binary}_{Option}$	0.279	1.24	0.389
	(3.29)	(1.20)	(1.22)
Dividend Yield	-7.72***	-2.11***	-3.14***
	(1.23)	(0.447)	(0.533)
Log(Market Value)	8.11***	1.02	1.70**
	(2.44)	(0.718)	(0.840)
Revenue Growth	1.21**	0.283***	0.276
	(0.592)	(0.093)	(0.194)
Return on Assets	63.4***	11.5***	9.71
	(17.0)	(4.43)	(6.07)
Earnings Surprise	0.727***	0.153***	0.243***
	(0.262)	(0.052)	(0.085)
Quarterly Stock Return	0.026***	0.005***	0.007***
	(0.005)	(0.0008)	(0.001)
Fixed-effects			
Firm Effect	Yes	Yes	Yes
Fiscal Year	Yes	Yes	Yes
Fiscal Quarter	Yes	Yes	Yes
Fit statistics			
Observations	41,231	31,160	31,160
R^2	0.263	0.375	0.446
Within R^2	0.012	0.272	0.302

Clustered (Firm Effect) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

5.2.2 Options as a Percentage of Total Compensaiton

To assess the validity of hypothesis 1.2, we introduce a variable representing options as a percentage of total compensation to determine whether this amplifies the effects observed in hypothesis 1. This hypothesis drew inspiration from prior research, specifically the work of Shue and Townsend (2017), which provided evidence that CEOs with a higher proportion of their compensation in the form of options tend to exhibit greater risk-taking behaviour. Hypothesis 1.2 posits that executives who receive a greater proportion of their total compensation through option grants are more inclined to engage in sentiment manipulation. To quantify this effect, we utilize the following equation:

$$
\begin{aligned}
SentimentScore_{i,t-1} = \alpha &+ \beta_1 * UpcomingOption_{i,t} \\
&+ \beta_2 * PercentofTotalCompensation \\
&+ \beta_3 * PercentofTotalCompensation_{Option} \\
&+ \lambda * Controls_t + FixedEffects + \epsilon
\end{aligned} \tag{6}
$$

We use relevant variables collected from Execucomp, which include the *face value of options* (market price * number of options) and *total direct compensation* which encompasses components such as salary, bonus, restricted stock grants, long-term incentive plans payouts and value of option grants.

Table 9 presents the results of the regression and provides insight into the sensitivity of the models used. Specifically, when employing the LM dictionary to compute the sentiment score of earnings calls, no significant relationship is observed between options as a percentage of total compensation and the tone. However, the use of LLMs yields more insightful outcomes. By utilizing either RoBERTa or FinBERT to calculate the sentiment score, the results align with hypothesis 1.2. Notably, the trend of declining sentiment scores in earnings calls preceding an option grant becomes more pronounced when executives have a higher proportion of their salary determined by option grants. This is consistent with the findings of Shue & Townsend (2017), who found that CEOs with a higher proportion of their total compensation made up of options were three to five times more likely to engage in risk-taking behaviour, which could increase the value of their options through the increased volatility of the stock. Finally, placebo results for this test also demonstrate insignificance, lending support to a causal association between upcoming options and sentiment scores in earnings calls (Appendix 5).

Table 9: Options as a percentage of of total compensation results

Dependent Variables:	LM Dictionary Score	RoBERTa Score	FinBERT Score
	(1)	(2)	(3)
Upcoming Option	-5.78***	-1.56*	-1.77*
	(2.20)	(0.840)	(0.914)
Percent of Total Compensation	-4.78	1.49*	6.69***
	(10.1)	(0.882)	(2.35)
Percent of Total Compensation$_{Option}$	2.70	-6.52***	-4.47**
	(8.61)	(0.745)	(1.98)
Dividend Yield	-7.07***	-2.13***	-3.12***
	(1.34)	(0.466)	(0.570)
Log(Market Value)	8.15***	1.20	2.13**
	(2.46)	(0.824)	(0.885)
Revenue Growth	1.09***	0.259***	0.173
	(0.381)	(0.059)	(0.140)
Return on Assets	68.3***	10.2**	7.51
	(18.3)	(5.04)	(5.88)
Earnings Surprise	0.778***	0.248***	0.412***
	(0.280)	(0.045)	(0.058)
Quarterly Stock Return	0.026***	0.006***	0.007***
	(0.005)	(0.0008)	(0.001)
Fixed-effects			
Firm Effect	Yes	Yes	Yes
Fiscal Year	Yes	Yes	Yes
Fiscal Quarter	Yes	Yes	Yes
Fit statistics			
Observations	21,970	15,127	15,127
R^2	0.357	0.251	0.349
Within R^2	0.017	0.009	0.016

Clustered (Firm Effect) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

5.2.3 Subject Content Analysis

Analyzing each of the previous regression results, there is a consistent pattern of executives exhibiting a more negative tone during earnings calls prior to an option grant. We can contextualize this by comparing the relative frequency of terms in call with and without an upcoming option grant. To do this, we use the LM dictionary of positive and negative words, as well as LM's other dictionary catered to certain and uncertain terms[9]. We do this to get a better understanding of the subject matter of the earnings calls, as well as how the language and tone exhibited by executives changes in anticipation of an upcoming option grant compared to periods without one.

[9] The dictionaries were retrieved from: https://sraf.nd.edu/loughranmcdonald-master-dictionary/

The inclusion of test around certain and uncertain words was due to their predictive power of future returns (Li, 2006). In particular, Li (2006) found that companies displaying a greater rise in the usage of uncertain and risky language in their annual reports tend to experience relatively negative returns, as opposed to companies with minimal increases in such linguistic patterns. This aligns with our hypothesis that executives, motivated to temporarily depress their stock price, employ language that elicits a negative market reaction.

Table 10 presents the words from the LM dictionaries that occur in at least 10% of all the calls, with an absolute change in relative frequency exceeding 5%. This shows that the percentage change in the relative frequency of these words between the calls preceding an option grant and calls without. This shows that the uses of negative and uncertain words are more prevalent in the earnings calls that precede an option grant. The table illustrates that the previously observed trend extends beyond negative terms to uncertain terms as well.

Table 10: Term frequency increase (decrease) in the quarter before option grant

LM Negative		LM Positive		LM Positive		LM Uncertain		LM Certain	
Word	Percent	Word	Percent	Word	Percent	Word	Percent	Word	Percent
impairment	47.50%	efficiencies	17.30%	improved	5.60%	assumptions	40.60%	highest	8.20%
restructuring	34.70%	enhance	13.90%	profitability	5.40%	assuming	19.50%	always	-7.10%
declined	14.60%	leadership	12.10%	achieve	-5.10%	roughly	16.10%	never	-9.30%
declines	12.50%	improving	10.50%	optimistic	-5.10%	approximately	15.50%	definitely	-12.40%
decline	12.30%	achieved	10.10%	confident	-6.20%	anticipate	12.40%	-	-
loss	11.70%	effective	10.00%	pleased	-6.90%	volatility	9.30%	-	-
negative	9.40%	successfully	9.10%	excited	-7.70%	nearly	8.50%	-	-
volatility	9.30%	efficiency	8.60%	good	-8.50%	probably	-7.70%	-	-
unfavorable	8.50%	benefit	8.50%	efficient	-8.50%	may	-8.90%	-	-
closed	5.90%	improve	8.50%	advantage	-8.80%	might	-11.50%	-	-
challenges	5.50%	highest	8.20%	great	-9.30%	-	-	-	-
critical	-6.30%	improvements	7.50%	exciting	-10.80%	-	-	-	-
force	-7.10%	enable	6.90%	happy	-13.20%	-	-	-	-
difficult	-10.80%	benefited	6.90%	excellent	-16.10%	-	-	-	-

5.3 Discussion and Limitations

5.3.1 Discussion

Our analysis shows that company executives strategically deploy more negative language in conference calls in the quarter proceeding an option grant date. This is consistent with the literature on disclosure manipulation and compensation, supporting the theory that executives will act opportunistically when given advanced knowledge of compensation grant schedules, creating an information symmetry imbalance (Shue & Townsend, 2017; Daines, McQueen, & Schonlau, 2018; Quigley, Ward, Hubbard , & Graffin, 2020). While previous research has primarily focused on the utilization of news disclosures by executives to engage in this behaviour, our findings demonstrate that executives also strategically employ negative tone as a tool to depress their companies' stock price when they stand to benefit from it. Both dictionary-based and large language model approaches for sentiment classification in this analysis demonstrate the same inverse relationship between upcoming option grants and tone.

In our first regression analysis, LM’s dictionary appears to show the most significant relationship between an upcoming option and sentiment. One explanation for this, discussed in Cao et al. (2023), is that executives, aware of the significance of Loughran and McDonalds influential paper, consciously avoid using negative words from the LM dictionary in their disclosures. Appendix 6 shows the frequency of positive and negative LM words used by executives in conference calls from 2006 to 2020. While it is evident that the use of negative LM words has declined since the release of the LM dictionary, we contend that there are other exogenous factors, both macro and firm level, that can be attributed to this effect rather than deliberate avoidance of LM words by executives[10]. Placebo tests help to validate the robustness of our results and support a causal relationship between compensation and sentiment.

We expand the breadth of our analysis to consider the presentation versus the question & answer period of conference calls, the proportion of options to total compensation, and the

[10] The average frequency of LM negative words in our call corpus decreased by 20.3 percent from 2011 to 2020. Future research may investigate market conditions during this time period and the role they play in the decreased word frequency of LM negative words.

words used in the calls. This is to better understand the extent and context to which executives strategically deploy negative sentiment, as well as where and when it was most pronounced.

First, we find that there is not significant evidence to suggest that this behaviour is more prevalent in the presentation section of earnings calls. This suggests that there is not an outsized difference in the use of abnormally negative language by executives across different sections of a call, contrary to our expectations for hypothesis 1.1. This could be due to presentation sections often being reiterations of quarterly press releases, which primarily discuss facts of the quarterly results (Price, Doran, Peterson, & Bliss, 2012).

While the size of option grants does not have a statistically significant influence on the sentiment in earnings calls, a significant relationship emerges when comparing options as a percentage of total compensation. Notably, both RoBERTa and FinBERT effectively capture this relationship. As discussed in Shue & Townsend (2017), executives earning a greater share of their total compensation in the form of options will be more prone to engage in risk taking behaviour. The volatility of a stock and a lower strike price both increase the value of an option. Shue & Townsend (2017) show that managers who rely more heavily on options for their compensation try to increase the value of options through increasing the volatility of their stock. We show a similar pattern of executives attempting to increase the value of their options through the use of overly pessimistic language in earnings calls intending to reduce the strike price.

Lastly, when analyzing the content of executives' discussions during calls, we find a clear increase in the use of negative and uncertain words in the quarter prior to an option grant, compared to when there is not. Terms such as “declined”, “unfavourable” and “volatility” appear with frequencies that are 14.6, 8.3 and 9.3 percent higher, respectively, in calls proceeding an option grant. This pattern extends to uncertain terms as well, with words such as “assumptions” and “roughly” appearing 40.6 and 16.1 percent more frequently, respectively. These findings provide further support for hypothesis 1.

To conclude, our results support the hypothesis that executives of publicly traded companies who have an upcoming option grant in the next fiscal quarter will use overly negative language in earnings calls with the intention to temporarily depress their companies stock price and obtain a lower strike price.

5.3.2 Limitations

The impact of the structure of compensation packages on the timing and content of corporate disclosures has garnered extensive attention from researchers. One of the primary concerns in this area of research is endogeneity, which poses significant risks affecting both sides of the economic problem: managerial compensation and disclosures. Our analysis focuses on the sentiment expressed in earnings calls and is less affected by the timing of predetermined, scheduled of calls and options grants. However, it is important to acknowledge the limitations of our analysis.

Sentiment quantifies the tone of a body of text but the nuances of the LLMs lead to sentiment classifications that are difficult to interpret. For example, while we can discern that a BERT score of -0.6 indicates greater negativity compared to a score of -0.1, the specific implications concerning the content discussed during the call remain unclear. This introduces a degree of imprecision in our results. Although we establish a causal relationship between option grants and increased negativity in conference calls, we are unable to fully understand level of negativity that options grants cause. Another limitation is the differing sentiment scores from each sentiment measure. The majority of the conference calls have similar sentiment classifications from RoBERTa and FinBERT but few calls have diverging sentiment scores depending on the method used. Future research on comparing the results of these models on annotated financial text, would allow us to use the model that is proven to be the most accurate. This topic could benefit from leveraging emerging language models better suited for sentiment scoring, especially those with higher token limits to reduce the need to break down long-form documents into smaller chunks[11].

Another limitation stems from the concentration of scheduled option grant dates within the fourth quarter. Although we address this concern using quarterly fixed effects in our economic analysis, it gives rise to endogeneity problem in section 5.3. While negative and uncertain words are more frequent in earnings calls preceding option grant dates, this higher frequency could be influenced by the fiscal quarter. Using additional robustness techniques would help

[11] https://help.openai.com/en/articles/7127966-what-is-the-difference-between-the-gpt-4-models

isolate for the precise impact that option grants have on frequency of words used in earnings calls, independent of the quarter in which they fall.

To better understand the relationship between compensation and the tone used by executives in conference calls, further analysis in equity vesting months would be valuable. This period represents the time where executives are allowed to sell their shares. Expanding the analysis to include this period would explore whether executives adopt an overly positive tone in calls during scheduled equity vesting months. This research would build upon the findings of Edmans et al. (2018), by extending the influence that equity vesting month have on the tone and content of earnings calls beyond strategic new releases of CEOs. Analysing tone in both the lead up to grant dates and in equity vesting months would provide further evidence that executives strategically manipulate their tone in earnings calls when they stand to benefit from temporary stock movements.

6. Conclusion

In this paper we explore the impact that upcoming option grants have on the tone of prior earnings calls. We find that in calls before option grant dates, executives exhibit a higher degree of pessimism compared to calls without an impending grant. Additionally, we find that executives who have a receive higher proportion of their total compensation through option grants engage in this behaviour more frequently. We reduced endogenous risks by ensuring that option grants are independent of the actions of executives by only considering grants that can be classified to a schedule. To ensure the robustness of our analysis and provide evidence for causal relationship, placebo testing, and fixed effects were used to account for firm- and time-dependent characteristics. These methods mitigated the threat of endogeneity and allowed us to extract the effect of an option grant on the tone of an earnings call. From analysing the content of the calls, we found that the frequency of words associated with negative market reactions increases (LM negative and uncertain words) with an upcoming option grant.

The main implication of our results is that executives on earnings call strategically modify the tone of the call to be more negative when they stand to benefit from a temporally lower stock price. The tone of conference calls impacts stock price, evidenced by Price et. al. (2012), and speaking in an exaggeratedly negative way, has consequences for stakeholders who make decisions based on the information disclosed in these calls. Our results suggest that the structure of an executive's compensation package has a causal effect on their tone during conference calls, which leads to agency concerns. This may necessitate the need for boards of public companies to re-evaluate how compensation packages are structured to reduce information asymmetry for the benefit of shareholders.

7. Appendix

Appendix 1:Variable definitions

Variable	Definition
Upcoming Option	An indicator that is 1 if there is an option in the next quarter from Execucomp
Upcoming Pseudo Option	An indicator that is 1 if there was an option in the previous quarter used in the placebo tests
Fair Value of Option	The fair value of an option on the grant date from Execucomp
Options Percent of Percent of Total Compensation	The face value of the option grant divided by total direct compensation from Execucomp
Presentation	An indicator that is 1 if the section of the call is the presentation section from the conference call data
Dividend Yield	The dividend yield measured at the end of the previous fiscal year from Compustat
Log(Market Value)	The logarithm of the market value of the firm from Compustat
Revenue Growth	The growth in revenue as a percentage over the previous quarter from Compustat
Return on Assets	The net income of the firm divided by the average total assets for the quarter from Compustat
Earnings Surprise	An average of the difference between reported earnings and analysts expected earnings of a firm from I/B/E/S
Quarterly Stock Return	The quarterly growth in stock price in the quarter from CRSP
Firm Effect	The unique Global Company Key for the firm from Compustat
Fiscal Year	The fiscal year of the earnings call from Compustat
Fiscal Quarter	The fiscal quarter of the earnings call from Compustat

Appendix 2: Correlation matrix for variables used in analysis

Variable	Firm Effect	Fiscal Year	Fiscal Quarter	Upcoming Option	Fair Value of Option	Dividend Yield	Log (Market Value)	Revenue Growth	Return on Assets	Earnings Surprise	Quarterly Returns
Firm Effect	-	-	-	-	-	-	-	-	-	-	-
Fiscal Year	0.028***	-	-	-	-	-	-	-	-	-	-
Fiscal Quarter	-0.011	-0.101***	-	-	-	-	-	-	-	-	-
Upcoming Option	-0.008	0.002	0.202***	-	-	-	-	-	-	-	-
Fair Value of Option	0.027***	-0.013*	-0.002	0.005	-	-	-	-	-	-	-
DividendYield	-0.204***	0.027***	0.014*	0.016*	-0.030***	-	-	-	-	-	-
Log(Market Value)	-0.172***	0.152***	-0.004	0.018**	0.139***	0.245***	-	-	-	-	-
Revenue Growth	0.007	0.009	0.016*	-0.002	-0.001	-0.017*	-0.022**	-	-	-	-
Return on Assets	-0.069***	-0.017*	-0.012	-0.002	0.033***	0.068***	0.258***	-0.065***	-	-	-
Eearnings Surprise	0.004	0.027***	0	-0.015*	0.002	-0.041***	0.041***	0.008	0.057***	-	-
Quarterly Returns	0.008	-0.007	0.003	-0.002	0.022***	-0.008	0	0	0.001	-0.007	-

Appendix 3: Variance inlation factor of independent variables

Variables	VIF
Upcoming Option	1.00
Fair Value of Option	1.03
Dividend Yield	1.07
Log(Market Values)	1.16
Revenue Growth	1.00
Return On Assets	1.08
Earnings Surprise	1.01
Quarterly Returns	1.00
Presentation Binary Variable	1.00
Percent of Total Compensation	1.00

Appendix 4: Placebo regression for presentation vs. Q&A

Dependent Variables:	LM Dictionary (1)	RoBeRTa (2)	FinBERT (3)
Upcoming Pseudo Option	-0.892	-1.29	-1.24
	(2.66)	(1.07)	(0.927)
$\text{Presentation}^{binary}$	5.29***	31.6***	36.8***
	(1.69)	(0.690)	(0.749)
$\text{Presentation}^{binary}_{Option}$	-1.23	1.06	1.42
	(3.28)	(1.24)	(1.31)
Dividend Yield	-7.73***	-2.12***	-3.15***
	(1.23)	(0.447)	(0.533)
Log(Market Value)	8.07***	1.01	1.70**
	(2.45)	(0.718)	(0.841)
Revenue Growth	1.21**	0.283***	0.277
	(0.593)	(0.093)	(0.194)
Return on Assets	63.7***	11.6***	9.78
	(17.0)	(4.42)	(6.07)
Earnings Surprise	0.728***	0.154***	0.243***
	(0.262)	(0.052)	(0.085)
Quarterly Stock Return	0.026***	0.006***	0.007***
	(0.005)	(0.0008)	(0.001)
Fixed-effects			
Firm Effect	Yes	Yes	Yes
Fiscal Year	Yes	Yes	Yes
Fiscal Quarter	Yes	Yes	Yes
Fit statistics			
Observations	41,231	31,160	31,160
R^2	0.263	0.375	0.446
Within R^2	0.012	0.272	0.302

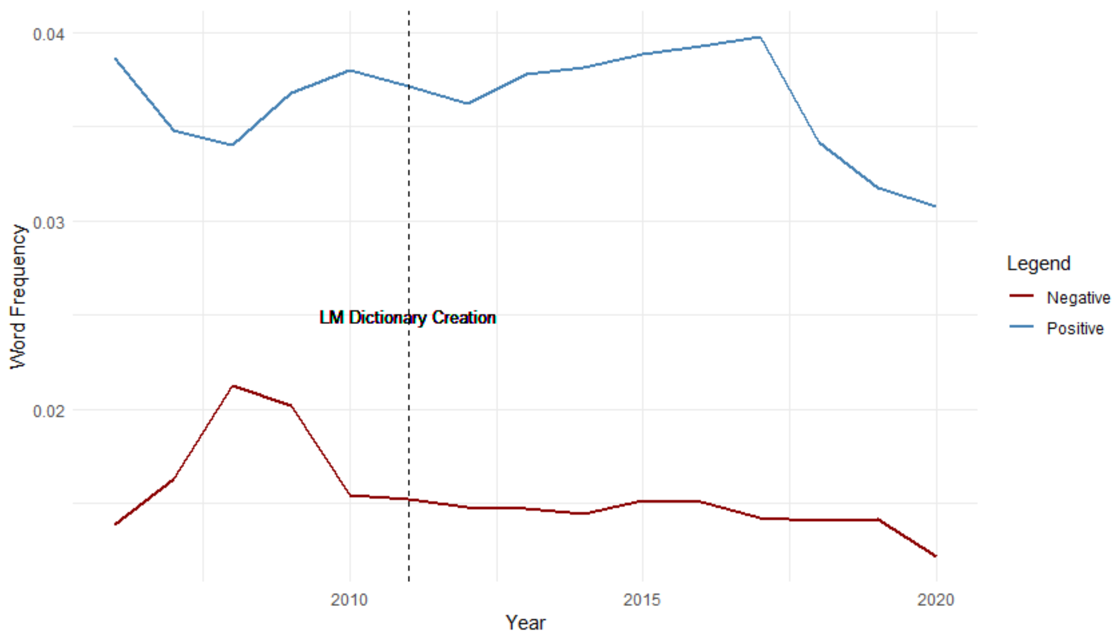
Clustered (Firm Effect) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Appendix 5: Placebo regression for options as a proportion of total compensation

Dependent Variables:	LM Dictionary Score	RoBERTa Score	FinBERT Score
	(1)	(2)	(3)
Upcoming Pseudo Option	1.40	-0.392	-0.408
	(2.57)	(0.919)	(0.926)
Percent of Total Compensation	-0.531	1.40***	3.44***
	(4.12)	(0.267)	(0.668)
Percent of Total Compensation$_{Option}$	-923.8	122.1	-105.2
	(488.7)	(114.5)	(179.5)
Dividend Yield	-7.08***	-2.14***	-3.13***
	(1.34)	(0.467)	(0.571)
Log(Market Value)	8.11***	1.20	2.12**
	(2.46)	(0.825)	(0.885)
Revenue Growth	1.09***	0.260***	0.174
	(0.382)	(0.059)	(0.140)
Return on Assets	68.6***	10.3**	7.60
	(18.3)	(5.03)	(5.89)
Earnings Surprise	0.780***	0.249***	0.413***
	(0.280)	(0.046)	(0.058)
Quarterly Stock Return	0.027***	0.006***	0.007***
	(0.005)	(0.0009)	(0.001)
Fixed-effects			
Firm Effect	Yes	Yes	Yes
Fiscal Year	Yes	Yes	Yes
Fiscal Quarter	Yes	Yes	Yes
Fit statistics			
Observations	21,970	15,127	15,127
R^2	0.357	0.251	0.349
Within R^2	0.017	0.009	0.0161

Clustered (Firm Effect) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Appendix 6: Frequency of LM positive and negative words from 2006 to 2020.



8. References

Araci, D. T. (2019). FinBERT: Financial Sentiment Analysis with Pre-trained. *Association for Computational Linguistics (ACL).* Florence .

Bagnoli, M., Kross, W., & Watts, S. (2002). The Information in Mangement's Expected Earnings Report Date: A Day Late, a Penny Short. *Journal of Accounting Research*, Vol. 40, No. 5, 1275-1376.

Bowen, R., Davis, A., & Matsumoto, D. (2005). Emphasis on Pro Forma versus GAAP Earnings in Quarterly Press Releases: Determinants, SEC Intervention, and Market Reactions. *The Accounting Review*, Vol. 80 No. 4 1011-1038.

Brockman, P., Khurana, I., & Martin, X. (2008). Voluntary disclosure around share repurchases. *Journal of Financial Economics*, Vol. 89 pp 175-191.

Cao, S., Jiang, W., Yang, B., & Zhang, A. (2023). How to Talk When a Machine Is Listening: Corporate Disclosure in the Age of AI. *The Review of Financial Studies*, https://doi.org/10.1093/rfs/hhad021.

Daines, R., McQueen, G., & Schonlau, R. J. (2018). Right on Schedule: CEO Option Grants and Opportunism. *Journal of Financial and Quantitative Analysis*, pp. 53: 1025-1058.

Davis, A., & Tama-Sweet, I. (2012). Managers' Use of Language Across Alternative Disclosure Outlets: Earnings Press Releases versus MD&A. *Contemporary Accounting Research*, Vol. 29, No. 3, 804-837.

Devlin, J., Chang, M.-W., Lee, K., & Toutanova, K. (2018). BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding.

Edmans, A., Goncalves-Pinto, L., Groen-Xu, M., & Wang, Y. (2018). Strategic News Releases in Equity Vesting Months. *The Review of Financial Studies*, Volume 31, Issues 11, Pages 4099-4141.

Frankel, R. M., Jennings, J. N., & Lee, J. A. (2021). Disclosure Sentiment: Machine Learning vs Dictionary Methods. *Management Science*, 1-47.

Frankel, R., Johnson, M., & Skinner, D. (1999). An Empirical Examination of Conference Calls as a Voluntary Disclosure Medium. *Journal of Accounting Research, 37(1)*, 133-150.

Garcia, D., Hu, X., & Rohrer, M. (2022, August). *The Colour of Finance Words*. Retrieved from SSRN: https://dx.doi.org/10.2139/ssrn.3630898

Hall, B. J. (1999). The Design of Multi-Year Stock Option Plans. *Journal of Applied Corporate Finance*, 97-106.

Huang, A. H., Wang, H., & Yang, Y. (2021). FinBERT: A Large Language Model for Extracting Information from Financial Text. *Contemporary Accounting Research*, 1 - 36.

Kimbrough, M. D. (2005). The Effect of Conference Calls on ANalyst and Market Underreaction to Earnings Accouncements. *The Accouning Review*, Vol. 80 No. 1, pp. 189-219.

Larcker, D. F., & Zakolyukina, A. A. (2011). Detecting Deceptive Discussions in Conference Calls. *Journal of Accounting Research, 50(2)*, 495-540.

Lee, J. (2022, June 1). *Vesting: What it is and how it works*. Retrieved from Carta.com: https://carta.com/blog/what-is-stock-vesting/

Leippold, M. (2023). *Sentiment Spin: Attacking Financial Sentiment with GPT-3.* Zurich: Swiss Finance Institute .

Li, F. (2006). Do Stock Market Investors Understand the Risk Sentiment of Corporate? *Unpublish Paper, SSRN eLibrary*, Available At: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=898181.

Liao, W., Zeng, B., Yin, X., & Wei , P. (2020). An improved aspect-category sentiment analysis model for text sentiment analysis based on RoBERTa. *Applied Intelligence*, 3522–3533.

Lie, E. (2005). On the Timing of CEO Stock Option Awards. *Management Science*, Vol. 51, Is. 5, pp: 679-849.

Liu, Y., Ott, M., Goyal, N., Du, J., Joshi, M., Chen, D., . . . Stoyanov, V. (2019). RoBERTa: A Robustly Optimized BERT Pretraining Approach. *Empirical Methods in Natural Language Processing*, (pp. 1-13). Hong Kong.

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.

Matsumoto, D., Pronk, M., & Roelofsen, E. (2011). What Makes Conference Calls Useful? The Information Content of Managers' Presentations and Analysts' Discussion Sessions. *The Accounting Review*, 1383-1414.

Medhat, W., Hassan, A., & Korashy, H. (2014). Sentiment Analysis Algorithms and Applications: A Survey. *Ain Shams Engineering Journal, 5(4)*, 1093-1113.

Mishev, K., Gjorgjevikj, A., Vodenska, I., Chitkushev, L. T., & Trajanov, D. (2020). Evaluation of Sentiment Analysis in Finance: From Lexicons to Transformers. *IEEE Access*, 1-23.

Mohey El-Din, D. (2016). Enhancement Bag-of-Words Model for Solving the Challanges of Sentiment Analysis. *International Journal of Advanced Computer Science and Applications, 7(1)*.

Pekka , M., Sinha, A., Takala, P., Korhonen, P., & Wallenius, J. (2013). Good Debt or Bad Debt: Detecting Semantic Orientations in Economic Texts. *American Society for Information Science and Technology*, 1-15.

Price, S., Doran, J., Peterson, D., & Bliss, B. (2012). Earnings conference calls and stock returns: The incremental informativeness of textual tone. *Journal of Banking and Finance*, Vol. 36, No. 4, pp. 992-1011 .

Quigley, T., Ward, A., Hubbard , T., & Graffin, S. D. (2020). Unintended Consequences: Information Releases and CEO Stock Option Grants. *Academy of Management*, 155–180.

SEC. (2006, December 22). *SEC Amends Executive Compensation Disclosure To More Closely Align With FAS 123R.* Retrieved from https://www.sec.gov/news/press/2006/2006-219.htm

Shue, K., & Townsend, R. (2017). How Do Quasi-Random Option Grants Affect CEO Risk Taking. *The Journal of Finance*, 2551-2588.

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

Tonello, M., & Tay, O. (2021, October 7). *CEO and Executive Compensation Practices in the Russell 3000 and S&P 500*. Retrieved from Harvard Law School Forum on Corporate Governance: https://corpgov.law.harvard.edu/2021/10/07/ceo-and-executive-compensation-practices-in-the-russell-3000-and-sp-500/

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., . . . Polosukhin, I. (2017). Attention Is All You Need. *Advances in Neural Information Processing Systems*, (pp. 6000–6010). Long Beach.

WRDS. (2023). *CRSP Stock Database Structure*. Retrieved from Wharton University of Pennsylvania: https://wrds-www.wharton.upenn.edu/pages/grid-items/crsp-stock-database-structure/