THE PENNSYLVANIA STATE UNIVERSITY
SCHREYER HONORS COLLEGE

DEPARTMENT OF FINANCE

Natural Language Processing Sentiment Analysis of S&P500 Earnings Calls and Abnormal Stock Returns

WARREN SIPE
SPRING 2022

A thesis
submitted in partial fulfillment
of the requirements
for baccalaureate degrees
in Finance and Philosophy
with honors in Finance

Reviewed and approved* by the following:

Randall Woolridge
Professor of Finance
Thesis Supervisor

Brian Davis
Professor of Finance
Honors Adviser

* Electronic approvals are on file.

ABSTRACT

Corporate earnings calls offer investors in the public markets the opportunity to hear from executives directly about performance, risks, and forward-looking strategy. This direct interaction with management reveals information about corporate operations which would not otherwise be publicly available. A potential untapped source of this information is the tone those executives employ in the call and the sentiment about past and future performance which it represents. This thesis will consist of a natural language processing (NLP) sentiment analysis of randomly selected S&P500 companies since 2016. Sentiment scores for these earnings calls will be analyzed to check for correlation with equity performance, both forward and backward looking, over a variety of timeframes. While this analysis does not indicate any utility of such analysis in predicting excess returns, it does reveal factors influencing call sentiment and provides avenues for further research.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

Dr. Woolridge,

For your assistance and guidance throughout this project and my entire time at Penn State. It has been a true privilege to work with you for the last four years.

Dr. Davis,

For your mentorship throughout the thesis process and your valuable feedback on my ideas.

Drs. Rajtmajer and Grady,

For your help in developing the code utilized in this project.

My family,

For your love and support always.

Chapter 1

Introduction & Literature Review

Corporate earnings calls offer investors in the public markets the opportunity to hear from executives directly about performance, risks, and forward-looking strategy. This direct interaction with management reveals information about corporate operations which would not otherwise be publicly available. A potential untapped source of this information is the tone those executives employ in the call and the sentiment about past and future performance which it represents. This information can be captured quantitatively through a technique known as sentiment analysis, which is enabled by natural language processing (NLP) technology.

A recent body of financial literature has focused on applying sentiment analysis of various data sets to predict stock returns. Tabari et al. 2018 illustrated that there is a causal relationship between stock price moves and the sentiment of a body of Tweets related to company events. Mehtab and Sen 2019 uses deep learning and natural language processing sentiment analysis to develop a predictive model for price movements in the Indian NIFTY 50 index. Xing, Cambria, and Welsch 2017 provides a general survey of the application of NLP in financial forecasting. These papers, along with a collection of others, indicate the potential viability of using natural language processing sentiment analysis of a variety of data sets to predict stock returns.

There has, however, been very little application of sentiment analysis to earnings calls. De Amicis, Falconieri and Tastan 2021 applied sentiment analysis to earnings calls to discover differences in sentiment based on the gender of executives on the calls but did not apply this analysis to stock returns. There are currently no analyses which attempt to predict returns on the basis of sentiment analysis of earnings calls[1]; this thesis will explore the viability of that method.

Chapter 2 will provide a general outline of natural language processing and sentiment analysis as well as a more specific discussion of the capabilities of the sentiment analysis tool utilized in this research.

Chapter 3 will outline the data collection and analysis methodology employed in the study, starting with the earnings call data and moving on to a discussion of returns data and risk adjustment.

Chapter 4 will discuss the variety of statistical analyses performed to attempt to derive insights from the data, both forward- and backward-looking.

Chapter 5 will interpret those statistical results and attempt to make applications towards practical market understandings.

[1] An upcoming paper in The Web Conference (WWW) contains a similar analysis; see Medya et al. 2022

Chapter 2

An Overview of Natural Language Processing & Sentiment Analysis

Natural language processing is an area of computer and information sciences dedicated to mechanical and algorithmic interpretation of language in its “standard” or natural form, aimed at allowing computers to understand text much as humans do. The discipline combines linguistic principles with statistical and machine learning methodologies to achieve a variety of models which allow computer programs to extract insight from unstructured text. Natural language processing has been directed towards a large number of use cases, including but certainly not limited to speech recognition, part of speech tagging, word sense disambiguation, named entity recognition, co-reference resolution, sentiment analysis, and natural language generation (IBM).

Development of Natural Language Processing

The development of natural language processing began far before the first computers, with the development of a science of linguistics. One of the pioneers of such science, Swiss linguist Ferdinand de Saussure, advocated a systemized conception of language in a series of lectures between 1906 and 1911. Saussure identified that the meaning of language was not defined by a simple relationship between a sound or a set of written characters and a concept; rather, it is the interrelation of groups of these signifiers whose interactions produce the meaning which we interpret as language (Foote). German linguist Ludwig Wittgenstein advanced this concept by introducing language games and by refining mathematical representations of the function of language, most notably in his *Philosophical Investigations* published posthumously in 1953 (Wittgenstein). Alan Turing's famous "Turing Test" also proposed the possibility that a computer might be able to understand and imitate human language and thought (Foote).

However, the first true seminal work in natural language processing did not come until Noam Chomsky's 1957 *Syntactic Structures*, in which it was argued that grammatical structure would need to be translated into a new form in order for computers to understand it. This paper introduced a style of grammar known as Phase-Structure Grammar which is still used by computers to understand natural language. Since that genesis, natural language processing technology has been gradually refined to more accurately fulfill a greater variety of use cases all while using fewer resources. Major innovations include the introduction of machine learning and of neural networks, both of which radically expanded the capabilities of natural language processing. Apple's Siri assistant was introduced in 2011, in one of the first commercially successful deployments of a public-facing natural language processing technology (Foote).

Tokenization

The first step of most natural language processing applications, including the one utilized in this research, is to tokenize the inputted text. To tokenize is "Given a character sequence and a defined document unit, tokenization is the task of chopping it up into pieces, called tokens, perhaps at the same time throwing away certain characters, such as punctuation" (Manning, Raghavan and Schütze). In many systems, text is tokenized into individual words but this is not always the case. A token can be any grouping of characters which represent a "useful semantic unit for processing" (Manning, Raghavan and Schütze). While this may seem straightforward, the interpretation of things like symbols, white spaces, hyphenations, and other normal constituents of language are less straightforward. Therefore, it is essential to tokenize properly before having the computer interpret any given text.

Sentiment Analysis

Sentiment analysis is an application of natural language processing which aims to quantify the degree to which a given text sample is positive or negative. After tokenizing a document, sentiment analysis software evaluates both the sentiment of individual tokens and the sentiment implications of the relations between those tokens. This allows sentiment analysis software to differentiate between degrees of intensity of similar sentiments and to detect more complex ideas (Mejova 5). For example, consider the following sentences:

My local Italian restaurant is terrible and nothing they serve is delicious.

My local Italian restaurant is not good and most meals are not tasty.

Although theses sentences express very similar ideas structured similarly, the sentiment which they contain is clearly different. A sentiment library codes in sentiment scores for the body of words in a language to allow for different assessments of similar words with different intensities. Another important capability of sentiment analysis is to understand the implications of context:

My local Italian restaurant is terrible and nothing they serve is delicious.

My local Italian restaurant is delicious and nothing that they serve is terrible.

My local Italian restaurant is terribly delicious.

Although the second sentence uses exactly the same words as the first, the meaning and sentiment of the two phrases are opposite. The third sentence represents a more complicated case which effective sentiment analysis would need to capture: figurative speech can invert the meaning of words as well.

As discussed above, the first step in the development is sentiment analysis software is the composition of sentiment libraries. These libraries are typically constituted manually, with human coders assigning a level of positive or negative sentiment to each individual word. From there, linguists and computer scientists teach the program basic grammar and rules to help it understand the text. More sophisticated sentiment analysis programs use techniques like part-of-speech tagging or machine learning to increase the ability to recognize complex lingual elements (Mejova 8-9).

Natural Language Toolkit & TextBlob

The analysis of this thesis utilizes the Natural Language Toolkit (NLTK) in Python for tokenization. The NLTK is a set of program modules and data sets which aim to enable both symbolic and statistical natural language processing applications (Bird and Loper 1). The particular nltk.tokenize method called in this analysis, nltk.wordpunct_tokenize, splits text based on whitespace and punctuation (nltk.org). The tokenization code also uses the NLTK corpus to remove garbage words from the tokenized text.

```
>>> from nltk.tokenize import wordpunct_tokenize
>>> wordpunct_tokenize(s)
['Good', 'muffins', 'cost', '$', '3', '.', '88', 'in', 'New', 'York', '.',
'Please', 'buy', 'me', 'two', 'of', 'them', '.', 'Thanks', '.']
```

Figure 1 NLTK Tokenization Example

The sentiment analysis program utilized is TextBlob, a Python library for processing textual data. The TextBlob.sentiment method returns a namedtuple which contains both text polarity and subjectivity, of which only the polarity is utilized in this analysis (textblob.io).

Chapter 3

Data & Methodology

Earnings Call Collection, Cleaning & Sentiment Analysis

The first set of data analyzed in this thesis consisted of a set of transcripts of recent earnings calls gathered from SeekingAlpha. A random sample of 10 S&P500 companies was selected and the last 15 transcripts of the earnings calls of each company with at least one year of trading data since their release as of February 15, 2022, were collected (i.e. the most recent possible call included would have been held February 15, 2021). 2 of these samples caused errors in the sentiment analysis code and were thus removed from the data set, leaving a total sample of 148 calls.

The calls were then subjected, first manually during the collection process and then mechanically in the sentiment analysis program, to cleaning to remove noise. A Python code was used for both this data cleaning and for the ultimate sentiment analysis. The cleaning included the manual removal of standardized language which was present in every iteration of a company's calls and the mechanical removal of URLs, email addresses, special characters, garbage words[2], and empty rows from the data.

The cleaned data was then tokenized, analyzed for sentiment and assigned a sentiment score through the application of Natural Language Processing procedures discussed earlier (see "An Overview of Natural Language Processing & Sentiment Analysis"). This score spans from -

[2] As defined through the NLTK library

1 to 1, with a positive score representing positive sentiment. A score of 0 represents neutral sentiment.

Returns Data & Risk Adjustment

Stock price data for each of the 15 selected companies was then gathered using Factset. Returns were calculated over the following time periods: 1 Year prior to the call, 3 Months prior, One Day following the call, 5 Days following, 3 Months following, and 1 Year following. These periods were selected to capture both potential short- and long-term implications of call sentiment.

An expected return was then calculated for each equity over each period by using the Capital Asset Pricing Model (CAPM):

$$\text{Expected Return} = R_f + \beta\,(R_m - R_f)$$

R_f = Risk free rate

β = Investment beta

R_m = Market return

The 10-Year U.S. Treasury Bill Yield at each respective data, adjusted linearly to the appropriate time period and derived from the Federal Reserve of St. Louis's database (FRED), was used as the risk-free rate. The current (as of 2/15/2022) 5-Year adjusted beta to the S&P500 was used for the beta and the S&P500 returns over the time period were used as the market returns. Actual stock returns were netted against this expected return to produce a risk-adjusted α value.

Chapter 4

Statistical Analyses

A number of different statistical analyses were performed on the data above to reveal different potential implications either for stock price prediction or for corporate governance. These include a correlation analysis, regressions of returns over each time period studied against sentiment scores, binary test analyses, and analyses of sentiment over time and by company.

Correlations

An initial correlation analysis was performed to test the relationship between sentiment score and stock returns over different periods. Correlation with backward returns represents the correlation with returns for the 1 Year or 3 Month period before the release of the report, respectively. Correlation with forward returns represents the correlation with returns following the release.

Table 1 Returns Correlation with Sentiment Scores

Backward		Forward			
1 Year	3 Month	1 Day	5 Day	3 Month	1 Year
0.026	0.089	-0.004	-0.029	0.014	-0.100

No direction or time period showed a particularly strong correlation- the strongest was a negative 0.100 correlation over the one year forward timeframe. Notably, the weakest correlation was the 1 Day forward. This likely indicates that any impact of the sentiment of the calls is overrun by the impact of the content of those calls; however, the fact that there are stronger correlations over longer time periods could indicate that short-term trading after the calls is missing some information contained in the sentiment.

Both backward correlations were weakly positive. The direction of these correlations aligns with management reflection of recent stock performance, with the 3 Month backward correlation substantially stronger than the 1 Year backward correlation. However, the weakness of both backwards correlations was surprising and indicates that manager tone on calls is not strongly affected by precedent stock performance.

Regressions

The next analysis performed was a regression of returns over each time period against the sentiment score of the corresponding earnings call.

Table 2 Regression of Risk-Adjusted Returns to Sentiment Scores

	Backward		Forward			
	1 Year	3 Month	1 Day	5 Day	3 Month	1 Year
Coefficient	0.310	0.329	-0.007	-0.055	0.067	-1.261
P-Value	0.752	0.284	0.958	0.727	0.868	0.228

Using a significance threshold of $p<0.05$, none of the regressions returned a significant variation. The lowest p values of the set were the 0.228 and 0.284 of the 1 Year forward and 3 Month backward returns, respectively. If significant, these would indicate a positive relationship between short term backward stock performance and earnings call sentiment and a strongly-negative relationship between longer-term forward stock performance and earnings sentiment. An extension of this study with more replication or controls for more external variables could yield a significant result in these categories. The one significant result was a 0.958 p value for the 1 Day forward return, which indicates that trading over this time period are not affected by earnings sentiment. This higher level of confidence in the null hypothesis is likely driven by the high levels of volatility following earnings releases attributable to factors other than the call itself.

Binary Test Analyses

Given the lack of significant results from the regressions, binary test analyses were performed to identify potential areas or further study or areas where the replicative scale of this research potentially obscures significant results. These analyses were hypothesis tests of the following: higher-than-average sentiment scores would predict positive α over a period, and lower-than-average sentiment scores would predict negative α over that period. Thus, if a call's sentiment is higher than average for its company and its α is positive, the hypothesis would be "Correct" and if its sentiment is lower than average for its company and its α is negative, the hypothesis would be "Incorrect". The test is inverted for periods of below-average sentiment.

Table 3 Binary Test Analysis

	Backward		Forward				
	1 Year	3 Month	1 Day	5 Day	3 Month	1 Year	Total
Correct	70	77	67	68	85	72	439
Incorrect	78	71	81	80	63	76	449

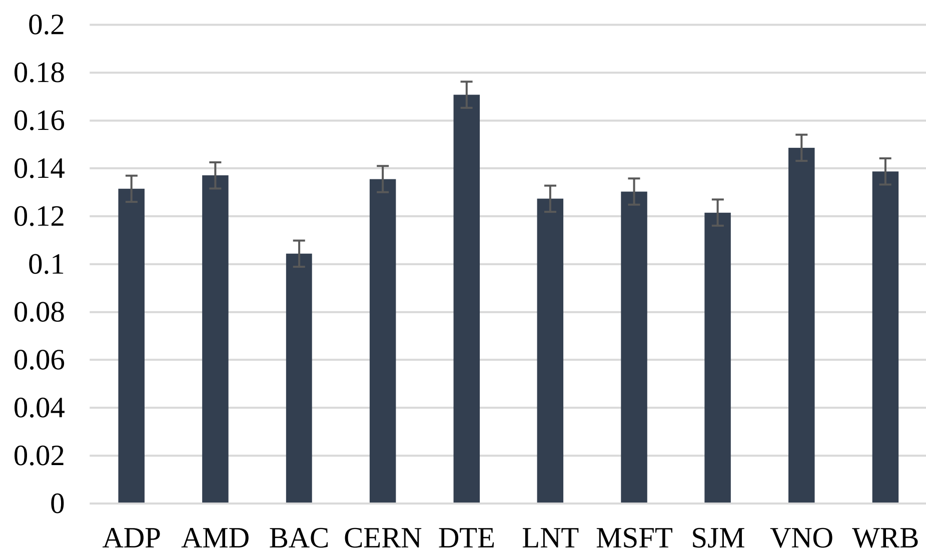
This analysis yielded a few relevant results. The short-term forward hypotheses were biased towards incorrect answers and the 3 Month forward hypothesis, in the strongest result of the set, was biased towards correct answers.

The binary tests were also checked for company-specific patterns. Certain companies displayed strong tendencies; examples include Microsoft for which the hypothesis was correct 63% of the time for forward returns and Alliant Energy, for which the hypothesis was correct 63% of the time for backward returns. Given that the sample sizes available for individual companies are relatively small (four earnings releases per year), it would be difficult to verify these results statistically.

Average Sentiment over Time and by Company

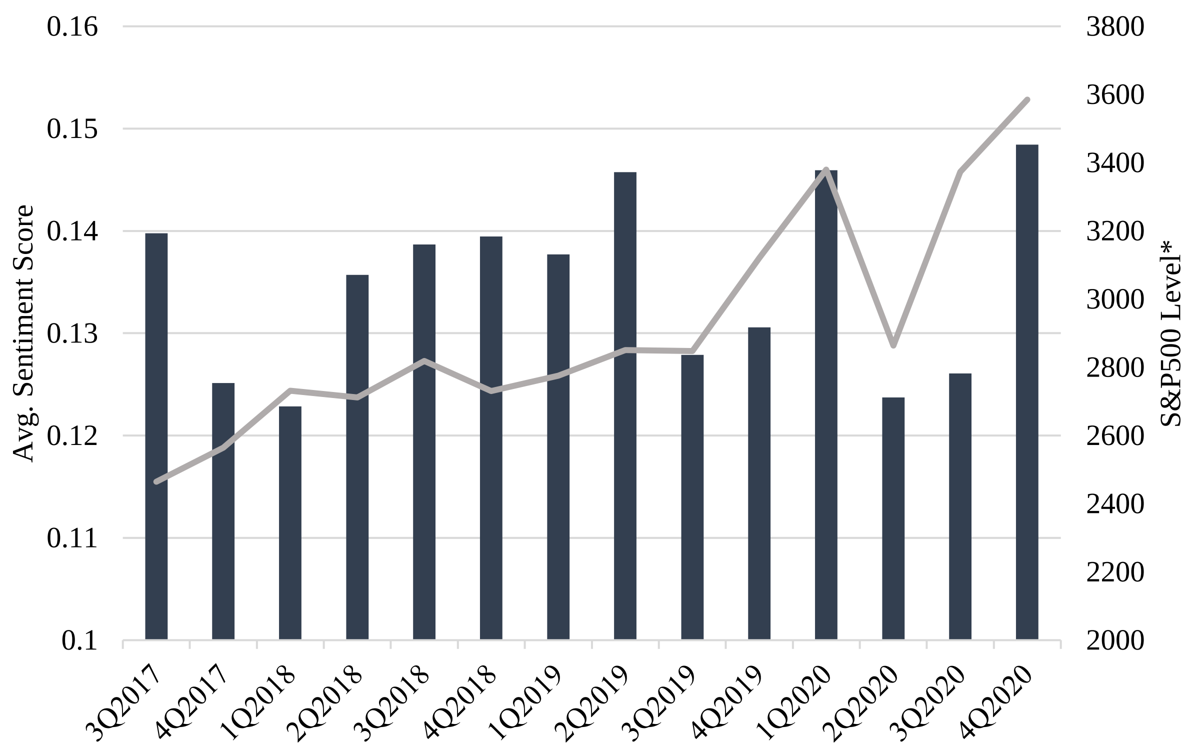
Average earnings sentiment varies from company to company and over time.

Figure 2 Average Sentiment by Company



The chart above demonstrates significant difference between the sentiment of individual companies. The average sentiment of Bank of America's earnings is 0.104, the lowest of the set, contrasts with the 0.171 average of DTE Energy, the highest of the set. In fact, Bank of America's highest individual call sentiment of 0.121 is lower than DTE's lowest sentiment of 0.155. This type of metric or comparison could be useful to analysts attempting to understand the rhetoric of earnings calls of companies they are covering; a highly positive call from Bank of America is likely more significant than a similarly positive call from DTE.

Figure 3 Average Sentiment and S&P500 Level



**Recorded at the midpoint of each quarter*

The graph above demonstrates that average earnings sentiment changes from quarter to quarter, likely reflecting changes in market conditions. Although the correlation between average sentiment and market performance is more difficult to see in the early portion of the graph, it becomes clearer in the latter portion as the volatility of both the market and sentiment increases. Especially with the advent of the COVID-19 pandemic in 1Q2020, it becomes easy to relate market conditions and sentiment.

Figure 4 S&P500 Level

**Quarters are based on calendar year and do not reflect firm fiscal years*

The blue arrows denote the span of time separating the first and last release in each earnings season. A synthesis of Figures 4.2 and 4.3 indicates that sentiment levels follow behind market moves, either because the earnings calls are influenced by market performance or because the earnings call sentiment is influenced by the same factors as the overall market but is delayed in reflecting it because of the periodic nature of such calls. The increases in sentiment in 2Q2019, 1Q2020, and 4Q2020 seem to follow respective trade ups both during the earnings period of the preceding quarter (1Q2019, 4Q2019 and 3Q2020) and in the time between these earning periods. The same trend, simply in the opposite direction, is evident for decreases in sentiment in 3Q2019 and 2Q2020.

Chapter 5

Interpretation & Conclusion

Although there are no statistically significant results generated from the regressions, the data set of this study does allow for a number of descriptive observations about relationships between earnings call sentiment and market returns. In addition, these observations provide indications of what might be productive avenues for future research.

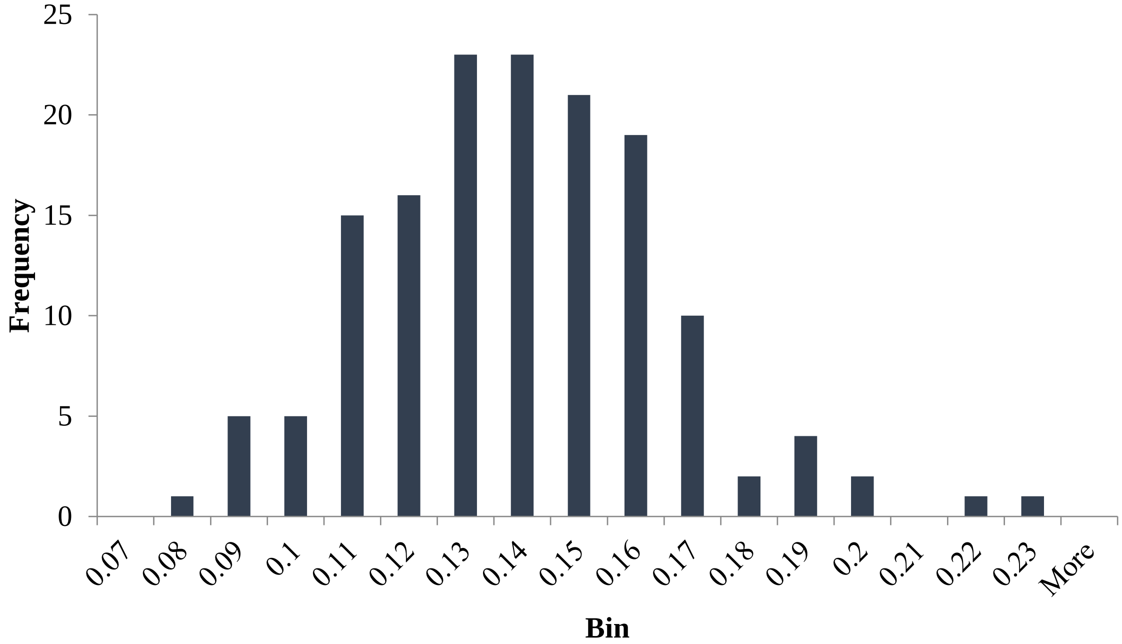
As noted above, the regression analyses performed do not yield significant results. This indicates that, at least at this level of replication and with these variables controlled, there is no demonstrable relationship between earnings sentiment and stock returns. However, while not statistically significant, the correlations and regression coefficients provide interesting potential commentary on these relationships which could be verified by a more expansive study. In particular, these metrics indicate a positive relationship between backward performance and call sentiment and a negative relationship between long-term forward performance and call sentiment. The binary test analysis demonstrates areas of interest for future study, including consistent patterns for individual companies and patterns over specific time periods.

Perhaps the clearest conclusions of the study come from the analysis of sentiment by company and over time. The companies in the set studied reflect substantially different sentiment in their calls and this knowledge could be utilized in earnings analysis to help better compare between different companies. Sentiment also varied significantly over time, likely in response to macro-level economic changes which affected the entire market. This finding is also supported by a comparison with S&P500 performance during the period of the study, which indicates that average sentiment had a direct but lagging relationship with market returns. This pattern becomes clearer in the latter part of the data set, especially with the onset of the COVID-19 pandemic and the ensuant increase in both market and sentiment volatility.

All in all, the statistical analysis of the study does not concretely support any statements about the relationship between earnings call sentiment and stock returns. However, the review of the data set completed above indicates a number of potential areas in which significant relationships could be discovered by more expansive or more narrowly focused studies.

Appendix A

Histogram of Sentiment Scores



Appendix B

Data Cleaning & Sentiment Analysis Code

```
import pandas as pd
from textblob import TextBlob
import nltk
from nltk.corpus import words
import re

# %% Read data

df = pd.read_csv('/*****/Thesis_data.csv', encoding='ISO-8859-1')
df["date"] = pd.to_datetime(df["date"])

df=df.astype(str)

# %% Preprocess data. Remove junk from text

# Remove URLs, emails

df["body"] = df["body"].apply(lambda x: re.sub(r'http\S+'," ", x))

print("Removed URLs")

df["body"] = df["body"].apply(lambda x: re.sub('\S*@\S*\s?'," ", x))

print("Removed emails")

#Remove bad punctuations (allow ,!?')

df["body"] = df["body"].apply(lambda x: re.sub("[^A-Za-z0-9!?,']+"," ", x))

print("Removed Bad chars")

# Remove garbage words
words = set(words.words())

def text_filter(my_text):
    return " ".join(w for w in nltk.wordpunct_tokenize(my_text) if w.lower() in words or not w.isalpha())

df["body"] = df["body"].apply(text_filter)

print("Removed garbage words")

# Remove empty rows
empty_rows = [i for i in range(len(df)) if len(df.iloc[i,2])==0]
df.drop(empty_rows, inplace=True)
df.reset_index(drop=True,inplace=True)
```

```
print("Removed empty rows")

# Export cleaned data
df.to_csv("******.csv",index=False)

# %% Function to return sentiment type

positive_threshold = 0.1
negative_threshold = -0.1

def get_senti_type(senti_score):
    if senti_score >= positive_threshold:
        return 1
    elif senti_score<=positive_threshold:
        return -1
    else:
        return 0

# %% Perform Sentiment analysis
df["senti_score"] = [TextBlob(text).sentiment.polarity for text in df["body"]]
df["senti_type"] = [get_senti_type(text) for text in df["senti_score"]]

df["positive"] = df["senti_type"]==1
df["negative"] = df["senti_type"]==-1
#df["neutral"] = df["senti_type"]==0

df = df.drop(columns = ["senti_type"])
df["count"] = [1 for _ in range(len(df))]
df = df.set_index(df["date"])

print("Sentiment done")

# %% Group by month
#sentiment_by_month = df.iloc[:,[1,3,4,5,6]].groupby(pd.Grouper(freq="M"))
#sentiment_by_month = sentiment_by_month.sum()
#sentiment_by_month['Month'] = [str(date.year)+"-"+str(date.month) for date in
sentiment_by_month.index]
#sentiment_by_month['mean_senti_score'] =
sentiment_by_month['senti_score']/sentiment_by_month['count']

#cols = list(sentiment_by_month.columns)
#sentiment_by_month = sentiment_by_month[[cols[-2]]+[cols[0]]+[cols[-1]]+cols[1:4]]

print("Grouping done")

# %% Export analysis
#sentiment_by_month.to_csv('/******.csv',index = False)

df.to_csv('/*******/thesissent2.csv',index = False)
```

BIBLIOGRAPHY

Bird, Steven, and Edward Loper. “NLTK: The Natural Language Toolkit.” *University of Pennsylvania*, 2002.

Bird, Steven, et al. *Natural Language Processing with Python*. O'Reilly, 2009.

By: IBM Cloud Education. “What Is Natural Language Processing?” *IBM*, https://www.ibm.com/cloud/learn/natural-language-processing.

De Amicis, Chiara, et al. “Sentiment Analysis and Gender Differences in Earnings Conference Calls.” *Journal of Corporate Finance*, vol. 71, Dec. 2021, https://doi.org/10.1016/j.jcorpfin.2020.101809.

Foote, Keith D. “A Brief History of Natural Language Processing (NLP).” *DATAVERSITY*, 17 June 2019, https://www.dataversity.net/a-brief-history-of-natural-language-processing-nlp/.

Li, Xiaodong, et al. “News Impact on Stock Price Return via Sentiment Analysis.” *Knowledge-Based Systems*, vol. 69, 2014, pp. 14–23., https://doi.org/10.1016/j.knosys.2014.04.022.

Loria, Steven. “Simplified Text Processing.” *TextBlob*, 2020, https://textblob.readthedocs.io/en/dev/index.html.

Makrehchi, Masoud, et al. “Stock Prediction Using Event-Based Sentiment Analysis.” *2013 IEEE/WIC/ACM International Joint Conferences on Web Intelligence (WI) and Intelligent Agent Technologies (IAT)*, 2013, https://doi.org/10.1109/wi-iat.2013.48.

Manning, Christopher D., et al. *Introduction to Information Retrieval*. Cambridge University Press, 2018.

"Market Yield on U.S. Treasury Securities at 10-Year Constant Maturity." *Federal Reserve Economic Database*, Federal Reserve Bank of St. Louis.

Medya, Sourav, et al. "An Exploratory Study of Stock Price Movements from Earnings Calls." *The Web Conference*, 2022.

Mehtab, Sidra, and Jaydip Sen. "A Robust Predictive Model for Stock Price Prediction Using Deep Learning and Natural Language Processing." *SSRN Electronic Journal*, 2019, https://doi.org/10.2139/ssrn.3502624.

Mejova, Yelena. "Sentiment Analysis: An Overview." *University of Iowa*, 2009.

NLTK, NLTK Project, 2022, https://www.nltk.org/.

Seng, Jia-Lang, and Hsiao-Fang Yang. "The Association between Stock Price Volatility and Financial News – a Sentiment Analysis Approach." *Kybernetes*, vol. 46, no. 8, 2017, pp. 1341–1365., https://doi.org/10.1108/k-11-2016-0307.

Souma, Wataru, et al. "Enhanced News Sentiment Analysis Using Deep Learning Methods." *Journal of Computational Social Science*, vol. 2, no. 1, 2019, pp. 33–46., https://doi.org/10.1007/s42001-019-00035-x.

Tabari, Narges, et al. “Causality Analysis of Twitter Sentiments and Stock Market Returns.” *Proceedings of the First Workshop on Economics and Natural Language Processing*, July 2018, https://doi.org/10.18653/v1/w18-3102.

Wittgenstein, Ludwig. *Philosophical Investigations*. 1953.

Xing, Frank Z., et al. “Natural Language Based Financial Forecasting: A Survey.” *Artificial Intelligence Review*, vol. 50, no. 1, 2017, pp. 49–73., https://doi.org/10.1007/s10462-017-9588-9.

WARREN P. SIPE

EDUCATION

The Pennsylvania State University | Schreyer Honors College **University Park, PA**
Smeal College of Business | Bachelor of Science in Finance *Class of May 2022*
College of the Liberal Arts | Bachelor of Arts in Philosophy | Minors in Economics and Spanish

PROFESSIONAL EXPERIENCE

Financial Technology Partners **San Francisco, CA**
Investment Banking Summer Analyst *June 2021 – Aug. 2021*
- Assembled client presentations and valuation models for financing and M&A transactions across different FinTech verticals
- Assisted in preparing presentation materials through industry research and trading and transaction comparable analysis
- Developed a working knowledge of equity offerings, valuation presentation, and dynamics within the FinTech market

Nittany Lion Fund, LLC **University Park, PA**
President & Director of Investor Relations *Jan. 2021 – Dec. 2021*
- Chosen to serve as President of a ~$15.0 MM investment fund based on history of leadership within the organization
- Managed communications with ~75 investors on a continuous basis to update them on Fund performance and address needs
- Coordinated bi-weekly educational sessions for the 600+ members of the Penn State Investment Association
- Created a new pitch voting system emphasizing outside research which increased value added by the voting process

Lead Analyst | Consumer Discretionary Sector *Jan. 2019 – Dec. 2020*
- Managed the ~$1.5 MM investment portfolio of the Consumer Discretionary Sector within the Nittany Lion Fund
- Researched holdings and compiled valuations using fundamental analysis, discounted cash flows, and comparable analysis to inform stock selection, generating returns of 32.8% and outperforming the S&P 500 benchmark by 4.0% during FY2019

Investment Research Partners, LLC **Lemont, PA**
Intern *June 2020 – Jan. 2021*
- Interned with investment consulting and sub-advisory firm specializing in ESG and SRI portfolios
- Created models to analyze and optimize portfolios to aid in making asset allocation decisions for a variety of clients
- Completed and organized due diligence on the selection and monitoring of portfolios, mutual funds, and individual investments

LEADERSHIP EXPERIENCE

Pennsylvania State University Park Student Fee Board **University Park, PA**
Member *May 2021 – Present*
- Selected for 12-member board responsible for the allocation of ~$200.0 MM student fee dollars towards non-curricular university services including facilities, Counseling and Psychological Services, and sustainability initiatives
- Worked as the subject matter expert with the Penn State Office of Physical Plant and Campus Recreation to understand and communicate specific institutional priorities and needs to the rest of the board

Presidential Leadership Academy **University Park, PA**
Member *Apr. 2019 – Present*
- Selected as member of 30-person class of a program dedicated to leadership development through experiential learning
- Engaged in a variety of leadership development programs through the Academy, including travel experiences, speakers, and small-group classes with the President of Penn State and the Dean of the Schreyer Honors College

Rock Ethics Institute Research Fellowship **University Park, PA**
Fellow *Sep. 2019 – May 2020*
- Conducted self-directed research in consultation with Penn State professors on U.S. policy towards Huawei since 2018 and its implications for the broader trade dispute and the ethical implications of international security policy
- Created Python-based natural language processing (NLP) application to examine the correlation of news sentiment with policy action and designed a rhetorical taxonomy to help understand the mechanics of security-related statements

HONORS, SKILLS, AND INTERESTS

Academic Honors: National Merit Finalist, PA Policy Debate State Champion (2018), Dean’s List (7/7)
Skills: Studied 7 years of Spanish, Bloomberg Terminal, FactSet, S&P Capital IQ, Basic Experience in Java and Python
Interests: Travel, Swimming, Numismatics, Fiction Writing, Weightlifting, U.S. National Parks, Harry Potter, Scuba Diving