THE PENNSYLVANIA STATE UNIVERSITY
SCHREYER HONORS COLLEGE

DEPARTMENT OF FINANCE

AI Investment Ratings: ChatGPT's Analysis of 1Q2024 Earnings Call Transcripts

THOMAS J. BURKE
SPRING 2024

A thesis
submitted in partial fulfillment
of the requirements
for a baccalaureate degree
in Finance
with honors in Finance

Reviewed and approved* by the following:

Brian Davis
Assistant Clinical Professor of Finance
Thesis Supervisor

Rich Bundro
Instructor in Finance
Thesis Reader

* Electronic approvals are on file.

ABSTRACT

The rapid pace of innovation in artificial intelligence continues to prove beneficial across various business domains. Recent research studies highlight ChatGPT's robust ability to interpret text and conduct sentiment analysis. For instance, Lopez-Lira & Tang (2023) demonstrated ChatGPT's proficiency by observing cumulative returns of 500% over a 14-month period through its analysis of news announcements. Similarly, Pelster & Val (2024) identified a strong positive correlation between ChatGPT's assessment of web articles for investment attractiveness and the subsequent returns of S&P 500 constituents.

The existing body of literature suggests a prevalent phenomenon known as post-earnings announcement drift (PEAD), challenging the notion of market efficiency. According to Fink (2021), there exists an inverse correlation between PEAD and the size of firms, attributed possibly to the heightened trading frictions and worse information environments experienced by smaller companies.

This study examines the efficacy of ChatGPT in analyzing earnings call transcripts from companies within the Russell 2000 index, particularly focusing on smaller firms prone to earnings drift. The investment recommendations derived from the analysis serve as a basis for portfolio construction, followed by the calculation of returns.

The increasing integration of AI in finance carries significant implications for further enhancing the efficiency of public markets. This trend is likely to accelerate with widespread adoption of APIs provided by prominent financial software companies such as FactSet, Bloomberg, and S&P Capital IQ.

TABLE OF CONTENTS

LIST OF FIGURES

ACKNOWLEDGEMENTS

I would like to express my appreciation for Dr. Brian Davis for his continued support in developing and finetuning the methodology of my thesis. My sincere thanks also go to Evan Fenster, a friend from the University of Pennsylvania, whose expertise in computer science provided valuable insights into OpenAI's capabilities, significantly enhancing the efficiency of my methodology.

I am equally grateful to Dr. Lou Gattis for his Advanced Financial Modeling class at Penn State, which equipped me with VBA skills that I applied to streamline the company screening process. Lastly, I am thankful to both Dr. Woolridge and Mr. Hinkelmann for their mentorship throughout my time in the Nittany Lion Fund.

Chapter 1

Introduction

Overview of ChatGPT

On November 30, 2022, OpenAI unveiled ChatGPT, which quickly became the leading AI chatbot due to its innovative use of Generative Pre-trained Transformer (GPT) technology. This technology enables ChatGPT to generate responses to user prompts by leveraging Large Language Models (LLMs). These models are advanced artificial intelligence (AI) algorithms trained on extensive datasets, allowing for nuanced and contextually relevant interactions (Pocock, 2023).

On November 6, 2023, OpenAI held the Company's first developer event in San Francisco, setting a precedent for innovation within the AI community. The event was a showcase of significant improvements to the platform, and an introduction of new improvements and features designed to increase functionality and enrich the user experience. Among the new improvements was the unveiling of new Assistant APIs, a breakthrough in making AI app development within ChatGPT more accessible. These APIs represent a leap forward in enabling developers to seamlessly integrate advanced AI functionalities into their applications, opening new avenues for creativity and innovation within AI (OpenAI, 2023).

OpenAI Assistant APIs

OpenAI recently introduced the Assistant API, an innovation allowing developers to seamlessly integrate virtual assistants into their applications. An API, short for application programming interface, serves as a channel for communication between two applications, facilitating data exchange and interaction. With the Assistant API, developers have access to advanced natural language processing capabilities, enabling their applications to respond to user queries, provide information, perform tasks, and engage in complex interactions (OpenAI, 2023). In the context of this research, an Assistant serves to increase the methodology's efficiency and effectiveness.

External APIs

OpenAI's capabilities expand exponentially through its API integrations, enabling a seamless connection between external data sources. Among the top financial software providers, including FactSet, Bloomberg, and S&P Capital IQ, FactSet stands out with its extensive suite of APIs designed specifically for developers. Currently, FactSet offers 92 APIs, encompassing services such as portfolio optimization, pricing alerts, document distribution, investment banking office updates, an events calendar, ESG analytics, and more (FactSet Developer Portal, 2024). The diverse array of offerings not only broadens the software's capabilities for financial professionals but also has the potential to increase their efficiency significantly.

Current Use Cases of Artificial Intelligence in Finance

In today's age, artificial intelligence is not simply a buzzword but a cornerstone technology reshaping numerous industries. AI has a myriad of different applications across a broad range of industries. Some of the main use cases include personal assistants, such as Siri and Alexa, tailoring user experiences in the retail industry, streamlining supply chain operations, enhancing route optimization, advancing educational tools, and a wide range of financial applications.

Some of the many financial applications of artificial intelligence involve fraud detection, risk management, personalized Robo-Advisors, and even trading and investment management. The existing research on artificial intelligence's ability to select and manage investments leaves more to be desired, especially as artificial intelligence is evolving at such a fast rate.

In the realm of investment management, a variety of ETFs are leveraging artificial intelligence, ranging from those that are AI-themed and those that are AI powered. AI-themed ETFs focus on investing in companies at the forefront of innovation within the artificial intelligence sector using conventional investment strategies. On the other hand, AI-powered ETFs employ AI algorithms to sift through data and make investment decisions. Among the most notable AI-powered ETFs are: Amplify AI Powered Equity (AIEQ), VanEck Social Sentiment ETF (BUZZ), WisdomTree International AI Enhanced Value Fund (AIVI), and QRAFT AI-Enhanced U.S. Large Cap Momentum (AMOM). These ETFs, despite their innovative approach, feature relatively low assets under management (AUM), averaging around $67.73 million. While AI holds promise in reducing operational costs in place of human labor, these ETFs continue to exhibit higher than average expense ratios, with an average of 71 basis points (Royal, 2023).

Given their distinct investment strategies, comparing their performance directly to the S&P 500 may not be the most precise benchmark. Nonetheless, the subsequent table offers analysis of the ETFs relative performance.

Figure 1. Historical Performance of AI Powered ETFs

Ticker	Company Name	AUM ($ in MM)	Expense Ratio	1Q2024 Returns	FY2023 Returns	FY2022 Returns
AI Powered ETFs						
AIEQ	Amplify AI Powered Equity ETF	$110.24	0.75%	2.46%	25.01%	(31.98%)
BUZZ	VanEck Social Sentiment ETF	$70.22	0.75%	14.77%	53.83%	(47.88%)
AIVI	WisdomTree International AI Enhanced Value Fund	$73.07	0.58%	1.38%	12.05%	(13.70%)
AMOM	QRAFT AI-Enhanced U.S. Large Cap Momentum ETF	$17.40	0.75%	16.14%	26.41%	(26.86%)
Average of AI Powered ETFs			**0.71%**	**8.69%**	**29.32%**	**(30.10%)**
Benchmarking						
SP50	**S&P 500**		**0.09%**	**10.16%**	**24.23%**	**(19.44%)**

The Amplify AI Powered ETF utilizes IBM Watson to tap into the capabilities of artificial intelligence integrating sentiment analysis and natural language processing (NLP). It also utilizes quantum computing to efficiently parse through various data sources, including news outlets, social platforms, analyst equity reports, and financial statements. In the last five years, the ETF has underperformed the S&P 500 by 43.50% (Amplify AI Powered Equity ETF, 2024).

Created in 2021, VanEck's Social Sentiment ETF targets investments in large cap U.S. stocks that garner the most positive sentiment on digital platforms. The strategy yielded favorable results in fiscal year 2023 and the first quarter of 2024, outperforming the S&P 500 in both time periods. However, in fiscal year 2022 the ETF underperformed the S&P 500 by 28.44%, indicating that the strategy may struggle during a bear market (Buzz - Vaneck Social Sentiment ETF: Holdings & Performance, 2024).

WisdomTree's International AI Enhanced Value Fund uses quantitative AI to select investments in developed markets outside of the U.S. and Canada. As a result, a fitting benchmark for comparing its performance is Vanguard's Developed Markets ETF (VEA), which similarly emits the United States (Vanguard Mutual Fund Profile, 2024). Historical data shows that WisdomTree's ETF has consistently underperformed Vanguard's Developed Markets ETF (WisdomTree International AI Enhanced Value Fund, 2024).

Lastly, QRAFT's AI ETF leverages artificial intelligence to tap into momentum trends within large-cap U.S. stocks. The adoption of AI enables QRAFT to analyze data patterns on a scale and speed not achievable by humans alone. In fiscal year 2023 and the first quarter of 2024, the ETF has outperformed the S&P 500. However, during the down market of fiscal year 2022 it underperformed the S&P 500 (Qraft AI-Enhanced U.S. Large Cap Momentum ETF, 2024).

These four ETFs demonstrate the diverse applications of AI in various investment strategies. The performance data suggests that AI-powered ETFs typically excel in bull markets but lag in bear markets. Although these ETFs might offer advantages during market upswings, it's crucial to weigh the impact of their higher expense ratios compared to more cost-effective index funds. This paper explores ChatGPT's ability to guide investment decisions based on the sentiment of earnings call transcripts.

Chapter 2

Literature Review

Empirical Studies

Lopez-Lira & Tang (2023) investigated whether ChatGPT could forecast stock price movements by analyzing news article headlines. They tasked ChatGPT with evaluating each headline as positive, negative, or neutral concerning its impact on the company's stock price. The researchers from the University of Florida observed a strong correlation between assessments and returns. Employing a strategy derived from the analysis, they observed an impressive cumulative return of 500% over a 14-month period by purchasing stocks with positive scores and selling those with negative scores. This study conveys ChatGPT's strong language interpretation skills. However, it solely focuses on headlines and overlooks the analysis of longer text passages, such as earnings calls. Moreover, as the study was published before the incorporation of the latest advancements to ChatGPT and the utilization of Assistant APIs, it misses out on potential improvements that those additions could have offered.

Lundgaard & Kazinnik (2023) also demonstrated ChatGPT's strong analytical ability to decipher text by analyzing how the AI can decipher Fedspeak. The researchers compare ChatGPT's interpretation of the Federal Reserve's technical language during the Federal Open Market Committee (FOMC) announcement to human assessments. They found that ChatGPT was able to decipher complex and nuanced language while also providing sound explanations and reasoning. Specifically, the researchers found that ChatGPT outperformed traditional methods of analysis with greater accuracy and closer alignment with human classification. The study also noted significant improvements in reasoning from ChatGPT-3 to ChatGPT-4 models.

Pelster & Val (2024) analyzed if ChatGPT-4 can provide investment attractiveness ratings based on all available information on the internet. The researchers conducted a live experiment during the 2Q2023 earnings season to test ChatGPT's ability to predict earnings surprises prior to an earnings announcement. The researchers also analyzed companies in the S&P 500's relative attractiveness for an investor. Pelster & Val found a positive correlation between ChatGPT ratings and future stock returns. The study showcases ChatGPT's ability to analyze and comprehend a large amount of financial news and data available on the internet to assist investors. Since the study was conducted between July 5, 2023, and September 8, 2023, the researchers utilized a Web-ChatGPT extension in order to provide web access to the existing version of ChatGPT. The Web-ChatGPT extension analyzes the top ten results on Yahoo within the past four weeks.

Ko & Lee (2023) analyzed ChatGPT's ability to serve as an assistant for portfolio managers for asset allocation strategies. The researchers prompted ChatGPT to select a set amount of assets across five different asset classes: stocks, cryptocurrency, commodities, currencies, and bonds. Out of the 10,000 iterations, Ko & Lee found that ChatGPT selects more diversified portfolios than random selection. The study concludes that ChatGPT can serve as a useful tool for retail investors.

Niszczota & Abbas (2023) assessed ChatGPT's ability to serve as a financial robo-advisor. The researcher's models tested financial literacy with a 19-question multiple-choice exam from the Financial Literacy Baseline Survey. ChatGPT scores 67% on the exam, compared to a random guessing baseline of 31%. In the methodology, participants predicted ChatGPT to score 79.32% on the exam. The study highlighted concerns regarding overreliance on AI for financial advice, especially for individuals with lower financial knowledge. Although the model utilized the older GPT-3.5 model, the results indicate that its main application does not lie within financial literacy, highlighting the need for further enhancements.

Liang (2016) attempted to predict stock price changes with earnings call transcripts and had little success. Liang used sentiment analysis in an attempt to predict stock changes of 14 major U.S. airlines from a total of 325 earnings call transcripts, spanning from 2007 to 2015. Specifically, he combined machine learning with a lexicon to arrive at a sentiment score for each earning call. The methodology included cleaning the transcripts and analyzing performance 31 days before and after the date, 7 days before and after the date, and 1 day before and after the date. The results from the study suggest that earnings call transcripts alone are not enough to accurately predict stock price changes. However, Liang notes that the classifiers used have plenty of room for improvement. ChatGPT's strong ability to decipher text may make this study inaccurate in today's age.

Sentiment Analysis

Loughran & McDonald (2011) showcased some of the complexities of analyzing the sentiment of financial documents. Specifically, the researchers analyzed a large sample of company annual reports (10-Ks) from 1994 to 2008 and found that 78.3% of the words in the 10-Ks were considered to be negative. The study indicates that a tailored approach is often required, specifically when analyzing the sentiment of financial-related documents. When the researchers created an alternative set of words to analyze the financial texts, they found a greater correlation with financial outcomes. Utilizing a tailored version of ChatGPT Assistant API for analysis of earnings calls should likely overcome this complexity.

Tetlock (2007) investigated the influence of the Wall Street Journal's sentiment on U.S. stock market activity. Tetlock utilized the General Inquirer, a quantitative content analysis tool used for systematically analyzing text data, in order to analyze the variation of sentiment of the media and measured the link with stock market prices using basic vector autoregressions. Tetlock's findings revealed that increased pessimism in media coverage directly predicted downward pressure on broader market prices.

Market Efficiency & Earnings Drift

The efficient market hypothesis (EMH) suggests that asset prices fully reflect all available information. In 1970, Eugene Fama published a review of efficient capital markets in the Journal of Finance. In the introduction, Fama defines three varying forms of the EMH. The weak form suggests that stock prices reflect all data of historical prices, the semi-strong form suggests that stock prices efficiently adjust to all publicly available information, and the strong form suggests that all information, even nonpublic information, is reflected in stock prices. Fama concluded that the markets are efficient with only a few exceptions. The study analyzes different theoretical models, including the expected return or "fair game" model, the sub-martingale model, and the random walk model, and ties them to empirical testing. The empirical tests assume certain conditions in which markets would be efficient. These conditions include no transaction costs, free availability of information, and consensus among investors regarding information. The study acknowledges that these conditions are often not all met in real-world markets (Fama, 1970).

An example of market inefficiency that has been studied widely is post-earnings announcement drift (PEAD). In 2021, Josef Fink published a review of the phenomenon in the Journal of Behavioral and Experimental Finance. PEAD is where a company's stock price continues to drift in the direction of its earnings for an extended period of time, which directly contradicts the efficient market hypothesis (Fink, 2021). Bernard & Thomas (1989) found that an investment portfolio that buys stocks with the most positive and shorts stocks with the most negative earnings generated an annualized return of 18%. Fink analyzed the magnitude of PEAD over time and found that it still exists two decades following Bernand & Thomas' study; however, to a lesser degree. This can, in part, be attributed to more arbitrage activity, improvements in the information environment, and a more sophisticated processing of earnings news. Fink observed that there is an inverse relationship between PEAD and firm size, likely due to smaller firms having trading frictions and worse information environments. The biggest takeaway from the comprehensive review of PEAD is that despite being studied for over 50 years, the phenomenon still exists and conflicts with the EMH.

Can Artificial Intelligence Outperform the Market?

With the rapid growth of artificial intelligence, the concept of AI beating the market has gained attention in the financial media. The consensus among the financial media appears to be against AI's ability to pick stocks that can outperform the market. Specifically, on July 17, 2023, the Washington Post published an article titled "*AI Can Write, But Is It Any Good at Picking*

Stocks?" The article goes on to mention that AI's natural language processing could be used to parse earnings statements and call transcripts to detect for trading signals or investment risks. With regards to beating the market, the article cites the underperformance of several AI-powered ETFs (Lee, 2023).

On August 19, 2023, MarketWatch posted an article titled "*Why AI-generated stock picks won't beat the market*". The article cited two main reasons, the first of which was centered around the efficient market hypothesis. The article argues that the information the AI will be trained on will already be fully reflected in asset prices. Separately, the article argues that if AI were to theoretically be a better stock picker, then it would be widely adopted and, therefore, close that gap (Hulbert, 2023).

On October 3, 2023, Bloomberg posted an article titled "*Can AI Beat the Stock Market?*" The article highlights the increased attention towards AI from hedge funds as they seek to outperform the market. However, the article mentioned that AI has struggled to find success in investing. The article cites an index of 12 different AI hedge funds, which underperformed the hedge fund market by 15% over the past five years. The article includes the fact that markets are random, and during times of randomness, AI will not work. However, the Bloomberg article admits that if AI were to provide a slight edge, that's all that would be needed to be very successful in finance (Ponsot, 2023).

On October 19, 2023, Barrons posted an article titled "*AI Isn't So Smart. Its Stock Picks Lag Behind the Market.*" The article simply analyzed the performance of the ETFs that use AI for portfolio management. The data shows that AI funds have significantly underperformed the broader market indexes. The author outlines his theory that investors are thinking of AI in the incorrect way and not utilizing it for its true purpose. However, the article does indicate that AI stock picking may improve over time (Carlson, 2023).

ChatGPT's Thoughts

When the question, "Can artificial intelligence analyze earnings transcripts to predict future stock returns?" was posed to ChatGPT, it offered the following response:

Yes, artificial intelligence (AI) can be used to analyze earnings transcripts to predict future stock returns, and this approach is becoming increasingly common in financial analysis. Earnings transcripts, which are records of the earnings calls that companies hold with analysts and investors, contain valuable information about a company's financial health, future outlook, management commentary, and responses to analyst questions. Here's how AI can play a role in analyzing these transcripts for stock prediction:

Natural Language Processing (NLP): AI technologies, especially those using NLP, can process and understand human language found in earnings transcripts. They can identify key themes, sentiments, management's tone, and changes in linguistic patterns over time, which may be indicative of a company's future performance.

Sentiment Analysis: By assessing the sentiment of the language used in the transcripts, AI can help gauge the company's outlook as presented by its executives. Positive sentiments might indicate good prospects, while negative sentiments could signal potential issues. This analysis can be nuanced, taking into account not just the words but the context and how sentiments change over time.

Quantitative Data Integration: AI systems can integrate the qualitative insights from earnings transcripts with quantitative financial data, such as revenue, earnings, debt levels, and more. This holistic approach allows for a more comprehensive analysis of a company's potential stock performance.

Predictive Analytics: By learning from historical data, AI models can identify patterns between what is discussed in earnings transcripts and subsequent stock performance. These models can make predictions about future stock returns based on new transcript data, considering factors like industry trends, economic indicators, and company-specific news.

Continuous Learning: AI models can continuously learn and adapt based on new data, improving their predictive accuracy over time. As they process more transcripts and stock performance outcomes, they refine their understanding and predictions.

However, it's important to note that while AI can significantly enhance the analysis of earnings transcripts and provide valuable insights, stock market prediction is inherently uncertain and influenced by numerous external factors. AI predictions should be considered as part of a broader investment strategy, not a foolproof solution. The effectiveness of AI in this context also depends on the quality of the data, the sophistication of the algorithms, and how well the AI model has been trained and tested against real-world outcomes.

Chapter 3

Quantitative Methodology-Data Collection

Hypothesis

ChatGPT's strong ability to decipher text and analyze sentiment in earnings transcripts will result in a positive correlation between numerical investment recommendations and average stock returns following an earnings call. Due to the impact of post-earnings announcement drift, the study anticipates this relationship to strengthen over 5-day and 10-day spans compared to shorter timeframes such as 1-day and 3-days.

Theoretically, earnings announcements play a crucial role in shaping stock returns by offering insights into the fundamentals of a company's performance, which in turn influences key valuation metrics like the price-to-earnings ratio and the enterprise-value-to-EBITDA multiple. Additionally, the tone and content of earnings calls can significantly affect how investors perceive a company's future growth potential and overall prospects.

Company Selection Process

The study examined the hypothesis by randomly selecting 100 companies across a range of different sectors. Given that post-earnings announcement drift (PEAD) is more prominent in smaller firms compared to larger ones, the analysis focused on constituents of the Russell 2000, an index comprised of 2,000 small-cap companies. To simplify the returns calculation from intra-day trading, only companies that report earnings after market close were selected. The sample was further narrowed down by randomly selecting them within Microsoft Excel as described below:

1) A random value was generated next to each company's name in Excel using the = RAND() function, which outputs a random number between 0 and 1

2) To the immediate right of each random value, an IF function was applied to determine which companies would be retained and which would be removed. The formula is as follows: = IF(B6>0.8,"Keep","Delete")

3) A Macro-Enabled Excel File facilitated the integration of Visual Basic Applications (VBA) to automate the removal of rows labeled "Delete". The VBA script is as follows:

Figure 2. VBA Script to Automate the Removal of Rows

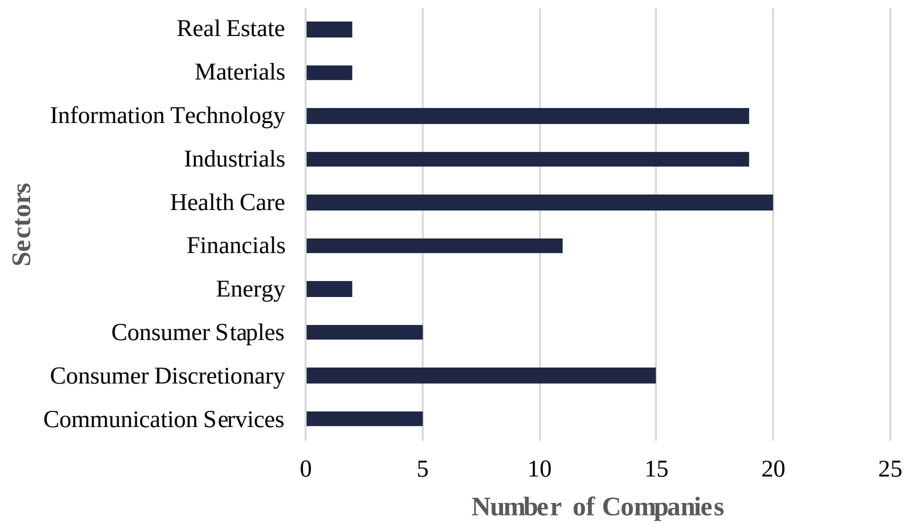
```
Sub DeleteRows()

For i = 3 To 2003
    If Cells(i, 3).Value = "Delete" Then
    Cells(i, 3).EntireRow.Delete
    End If
Next i

End Sub
```

The outcome of the selection process was a list of 100 companies from the Russell 2000 index, chosen randomly across different sectors. The data subset covers all sectors except for utilities. The Russell 2000 index, which constitutes the lower two-thirds of the Russell 3000 index, is managed by FTSE Russell and is frequently used as a performance benchmark for small-cap investment funds.

Figure 3. Graphical Representation of Sector Distribution of 100 Selected Companies



After refining the list to 100 companies, FactSet, a financial data and software company that offers comprehensive financial information for over 8,000 companies, provided the earnings transcripts from 1Q2024 through its Events & Transcripts section. This source ensured the transcripts were in a standardized form, making them suitable for analysis by ChatGPT.

Investment Recommendation Analysis

An OpenAI Assistant named Stock Analyst, created with Python, evaluates the earnings transcripts in this study. The Assistant assesses these transcripts and assigns an investment recommendation score between the range of -10 to +10, with a score of -10 indicating very unattractive, 0 indicating average attractiveness, and +10 indicating very attractive. The Assistant utilized the GPT-4 model in its analysis and set the model's temperature to 0. In the context of AI models, especially those involving natural language processing, "temperature" is a parameter that influences the randomness or predictability of the output generated. When the temperature is set to a low value, such as 0, the AI model tends to produce responses that are more factual and less random. On the other hand, a higher temperature setting increases the model's randomness and can lead to more creative outputs.

The arrangement of earnings transcript files within a folder structure enabled the assistant to iterate through the files paths and access the files. In the primary "Companies" folder, 100 subfolders existed, named after the ticker symbols of each company, containing the earnings call transcripts from FactSet. Each iteration by the Assistant followed the subsequent prompt:

> Analyze the attached earnings transcript and rate the investment attractiveness for an investor who is going to hold the stock for 5-10 days following the earnings announcement. Compare to all other Russell 2000 stocks and assign a score from -10 to +10, where -10 indicates very unattractive, 0 indicates average attractiveness, and +10 indicates very attractive. No Yapping.

The prompt was inspired by both Lopez-Lira & Tang (2023) and Pelster & Val (2024). The inclusion of "No Yapping" adheres to a widely recognized prompting strategy designed to elicit succinct and direct responses from ChatGPT.

Once the Assistant finished analyzing all of the transcripts, it compiled the outcomes into a CSV file. These results encompassed an investment recommendation score along with the GPT's rationale, typically presenting around five supporting points, contingent on the transcript's length.

Returns Analysis

The consolidation of the numerical investment scores from ChatGPT into an Excel spreadsheet facilitated further analysis alongside the stock returns following the earnings announcements. The computation of returns data utilized the FactSet pricing formula, followed by the application of the percentage change formula to assess variations across different time frames. The formula applied is as follows: =FDS(TICKER,"FG_PRICE("&DATE&")"). The analysis of the returns spanned a 1-day, 3-day, 5-day, and 10-day periods to fully capture the potential of earnings drift. The study incorporated Excel's =WORKDAY() function in the return calculations to account for non-trading days, ensuring consideration of only the days when the stock market was open.

Calculation of Regression Analysis & Correlation Coefficients

The analysis utilized the results from ChatGPT to calculate four different linear regressions for each time period. In these regressions, the ChatGPT scores functioned as the independent variable (X), aiming to influence the dependent variable of stock returns (Y). The standard equation for a linear regression is as follows:

$$Y = \alpha + \beta X + \in$$

Where:

Y represents the dependent variable

α represents the y-intercept

β represents the slope

X represents the independent variable

$\in$ represents the error term

The regression analysis for each time period utilized the Data Analysis Toolpak in Microsoft Excel. The analysis yielded a multiple R, indicating the correlation coefficient that reflects the linear relationship’s strength and direction between GPT scores and returns.

Additionally, the analysis provided an R^2 value, which represents the percentage of variance in stock returns predictable by the GPT scores. The P value provided an assessment of the coefficient's significance. Detailed regression results for each time frame are available in Appendix B.

Example ChatGPT Responses & Investment Rationale

Sanmina Corp. (SANIM):

Based on the Sanmina Corp. Q1 2024 Earnings Call transcript, I would assign an investment attractiveness score of +4 to Sanmina Corp. within the Russell 2000 context for a 5-10 day holding period post-earnings announcement. Here's an analysis of the key points influencing this score:

Financial Performance and Outlook: Sanmina reported Q1 revenue of $1.87 billion, aligning with their outlook and demonstrating solid execution in a challenging market environment. The non-GAAP operating margin was 5.5%, and non-GAAP EPS was $1.30, at the high end of their outlook. For Q2, they expect revenue to be flat sequentially, with improvements anticipated in the second half of the fiscal year, suggesting a positive trajectory.

Market Position and Diversification: Sanmina has a diversified portfolio across growth markets such as Cloud, Defense and Aerospace, Medical/Digital Health, Electric Vehicles, Renewable Energy, Industrial, and Optical Packaging. This diversification reduces reliance on any single market and positions Sanmina to benefit from growth in multiple sectors.

Operational Excellence and Margin Expansion Focus: The company's emphasis on operational excellence and margin expansion, with a goal of achieving an operating margin of over 6% in the long term, reflects a commitment to improving profitability. This focus, coupled with their strategy to invest in high-margin end markets, indicates potential for margin improvements as revenue scales.

Strong Balance Sheet and Cash Flow Generation: Sanmina's strong balance sheet and low leverage provide flexibility to navigate market challenges and invest in growth opportunities. The company's focus on generating strong free cash flow, demonstrated by $92 million in free cash flow in Q1, supports sustainability and strategic investments.

Strategic Priorities for Growth: The company's priorities include being customer-centric, providing leading technology, positioning for long-term growth with a revenue goal of reaching a $9 billion run rate, focusing on margin expansion and cash flow generation, and maximizing shareholder value. These priorities, combined with a strong pipeline of new opportunities, suggest a solid foundation for future growth.

The score of +4 reflects Sanmina's solid Q1 performance, positive growth outlook for the second half of the fiscal year, strategic market positioning, strong balance sheet, and focus on operational excellence and margin expansion. These factors suggest that Sanmina is well-positioned within the Russell 2000 for investors with a 5-10 day holding period post-earnings announcement, with potential for value appreciation as the company executes on its growth and profitability strategies.

Design Therapeutics, Inc. (DSGN):

Based on the detailed analysis of Design Therapeutics, Inc.'s (DSGN) Q4 2023 Earnings Call transcript, I would rate the investment attractiveness of DSGN in comparison to other Russell 2000 stocks at a score of +6. This score reflects a considerably attractive investment opportunity, albeit with some caution due to inherent risks in clinical development and regulatory approval processes.

Key Points Leading to the Score:

Innovative Genomic Medicine Platform: DSGN's GeneTAC platform represents a significant advancement in genomic medicine, offering potential first-in-class or best-in-class therapies for several monogenic disorders, including Friedreich's ataxia, Fuchs endothelial corneal dystrophy, Huntington's Disease, and myotonic dystrophy. This innovation and differentiation could lead to substantial market opportunities and competitive advantages.

Clinical Development Progress: The introduction of DT-216P2, an improved formulation for treating Friedreich's ataxia, addresses previous limitations and shows promising pharmacokinetics and tolerability in preclinical studies. The ability to increase frataxin levels in FA patients, alongside progress in FECD, Huntington's Disease, and myotonic dystrophy programs, underscores DSGN's potential to deliver transformative treatments.

Regulatory and Market Opportunity: The cleared IND for the FECD program and the strategic approach to clinical trial design demonstrate DSGN's adeptness in navigating regulatory pathways. The absence of approved disease-modifying drugs for FECD and the large patient population underscore a significant market opportunity. Similarly, the unmet needs in FA and other targeted disorders present substantial commercial potential upon successful product development and approval.

Financial Position: With a five-year operating runway, DSGN is well-positioned to achieve key milestones and generate clinical proof of concept data for up to four programs without immediate financial pressures. This stability is crucial for sustained R&D efforts and reduces short-term financial risk for investors.

Considerations and Risks:

Clinical and Regulatory Risks: The inherent uncertainties of clinical development, including patient recruitment, trial execution, and meeting efficacy and safety endpoints, pose significant risks. Regulatory hurdles also remain, with the need for further engagement with the FDA to clarify approval pathways, especially for DT-216P2 in FA.

Competitive Landscape: While DSGN's programs have first or best-in-class potential, the dynamic nature of the biotech industry means that new technologies or therapies could emerge, impacting DSGN's competitive position.

Market Adoption and Reimbursement: Success in clinical development does not guarantee commercial success. Market adoption will depend on the therapeutic benefits relative to existing treatments, pricing, and reimbursement policies.

Given these factors, the +6 score reflects a cautiously optimistic view on DSGN's investment attractiveness over a 5-10 day holding period post-earnings announcement, acknowledging the company's strong innovation and development pipeline while considering the typical risks associated with biopharmaceutical investments.

Chapter 4

Data Analysis

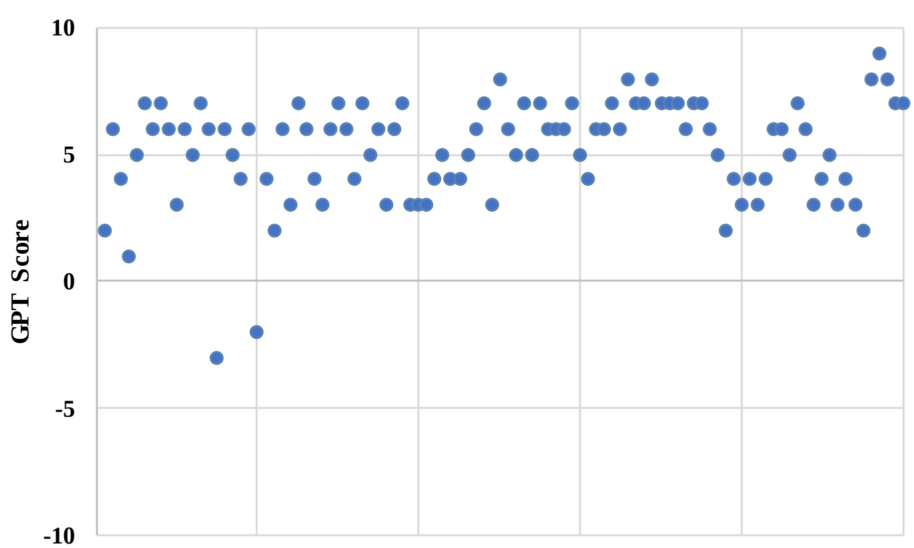
Distribution of Investment Scores

ChatGPT generated investment scores that skewed towards the higher end of the -10 to +10 range, providing a minimum score of -3 and a maximum score of 9. The average score was 5.18, with a median of 6. This skew suggests that ChatGPT tends to favor more optimistic outcomes for investments, indicating a positive bias towards sentiment analysis.

Figure 4. Summary of GPT Score Distribution

GPT Score Summary	
Average	5.18
Median	6.00
Stdev	2.02
Maximum	9.00
Minimum	(3.00)

Figure 5. Distribution of ChatGPT Scores (-10 to +10)



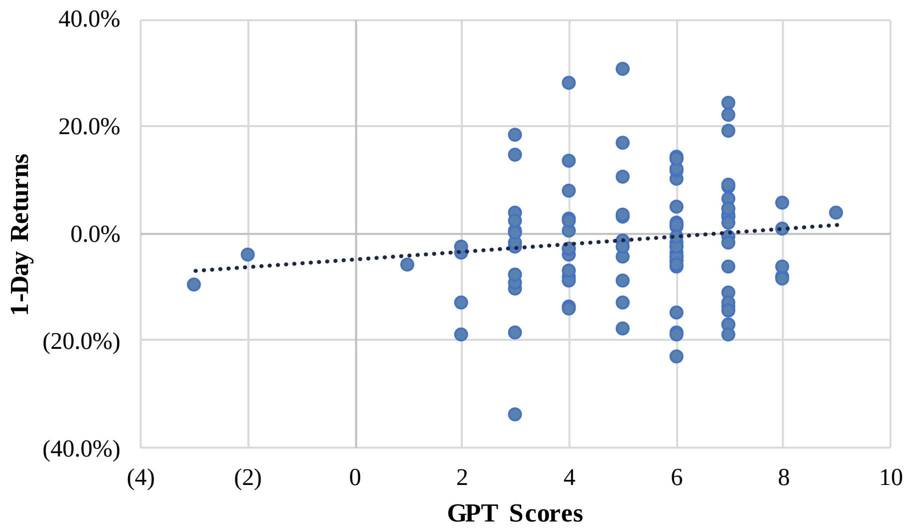
Regression of Returns & GPT Scores

	Multiple R	R Square	P-Value
1-Day Returns	0.127268831	0.016197355	0.207011559
3-Day Returns	0.109557037	0.012002744	0.277892255
5-Day Returns	0.1797493	0.032309811	0.073534413
10-Day Returns	0.170718299	0.029144738	0.089468097

Figure 6. ChatGPT Score Correlations and Significance Levels

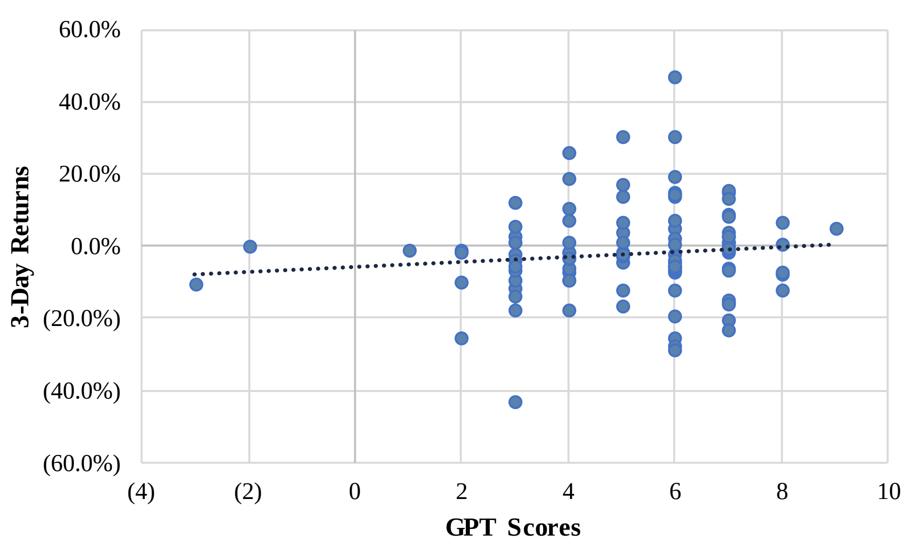
The regression analysis indicates a weak positive correlation coefficient over the four different time periods, which strengthens for the 5-day and 10-day returns. The Multiple R indicates a correlation between investment score and the stock returns of ~0.12 over a shorter time frame of 1-3 days and a higher correlation of ~0.17 over 5-10 days. Likewise, the R Square value, which explains the proportion of variance in returns that is predictable from the GPT score strengthens over the longer time horizons. The P-value for the 1-day and 3-day returns represent a probability that the data is due to chance alone of ~21% and ~28% respectively. The P-value for the longer time periods represents more statistically significant results indicating a probability that the data is due to chance alone of ~7% and ~9% for 5-day and 10-returns, respectively. Since the P-values across all time periods are greater than 0.05 the results are not statistically significant.

Figure 7. 1-Day Returns Regression Trendline



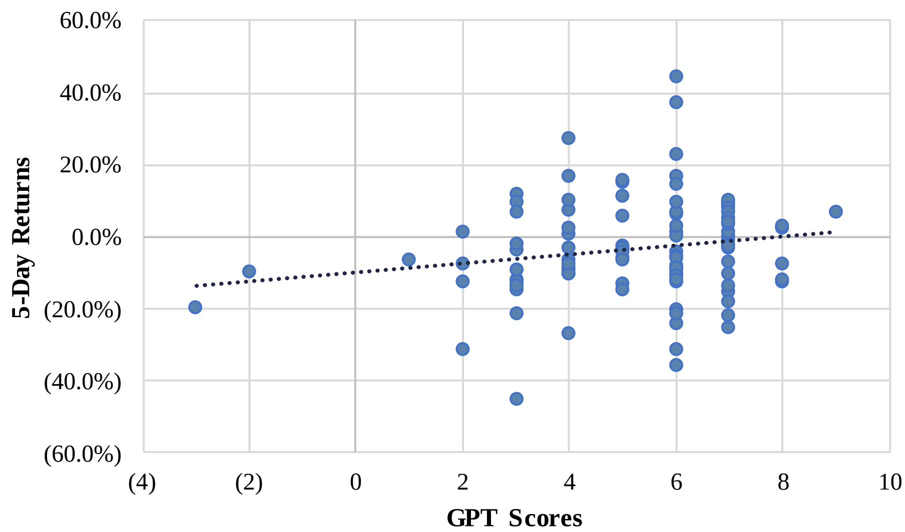
The regression equation for 1-day returns, Y = 0.0071x -0.0496, shows that with an investment score (X) of 0, the expected returns (Y) begin at (4.96%). The slope 0.0071 signifies that for every single unit rise in ChatGPT's investment score, there's a corresponding increase of 0.71% in the returns 1-day following the earnings call.

Figure 8. 3-Day Returns Regression Trendline



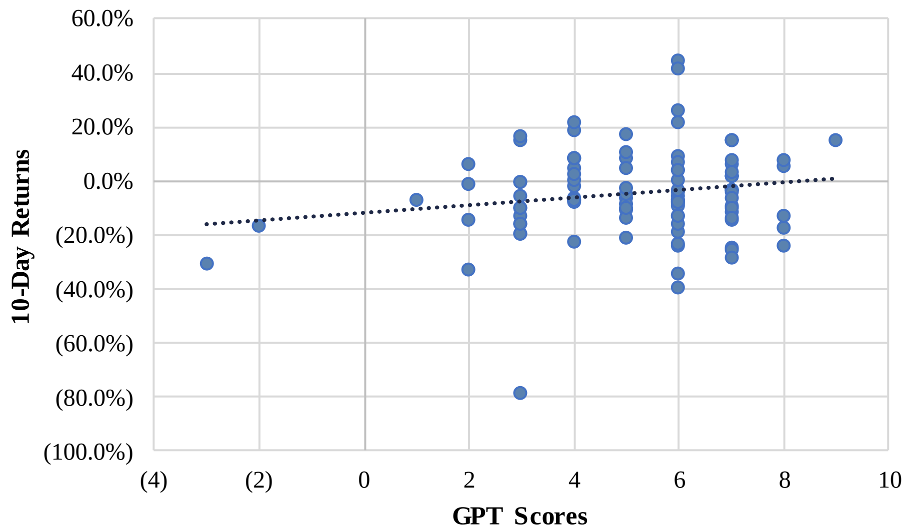
The regression equation for 3-day returns, Y = 0.0072x -0.0591, indicates that when the investment score (X) is 0, the anticipated returns (Y) start at (5.91%). The slope 0.0072 implies that each one unit increase in ChatGPT's investment score leads to a 0.72% increase in the returns 3-days following the earnings call.

Figure 9. 5-Day Returns Regression Trendline



The regression equation for 5-day returns, Y = 0.0126x -0.0986, shows that with an investment score (X) of 0, the expected returns (Y) begin at (9.86%). The slope 0.0126 signifies that for every single unit rise in ChatGPT's investment score, there's a corresponding increase of 1.26% in the returns 5-days following the earnings call.

Figure 10. 10-Day Returns Regression Trendline



The regression equation for 10-day returns, Y = 0.0141x -0.1168, indicates that when the investment score (X) is set to zero, the predicted returns (Y) are (11.68%). The slope of 0.0141 indicates that for every single unit rise in the investment score from ChatGPT, there's a corresponding increase in 10-day returns of 1.41%. Full regression statistics can be found in Appendix B.

ChatGPT's Analysis of the Results

When prompted to analyze the results ChatGPT responded as follows:

To analyze your research data, we will assess the correlation between the investment attractiveness scores provided by ChatGPT and the subsequent stock returns over different time frames (1, 3, 5, and 10 days) after earnings announcements for companies in the Russell 2000 index. The scores range from -10 (very unattractive) to +10 (very attractive), and the returns are presented as percentage changes in stock prices.

Key Observations:

Variability in Returns: The returns exhibit significant variability, with some stocks experiencing substantial gains and others facing steep losses in the days following their earnings announcements. This variability is observed across all time frames (1, 3, 5, and 10 days).

ChatGPT Scores and Stock Performance:

Positive Scores: Stocks rated positively by ChatGPT (scores above 0) show a mix of outcomes, with some like C3.ai Inc. (Ticker: AI) and Beyond Meat Inc. (Ticker: BYND) showing notable gains, especially in the shorter time frames. This suggests that higher scores might be associated with a higher likelihood of positive short-term returns.

Negative Scores: Stocks with negative scores, such as Big 5 Sporting Goods (Ticker: BGFV) and Clean Energy Fuels Corp. (Ticker: CLNE), tend to show negative returns, aligning with the expectation that lower scores indicate less attractive investment opportunities.

Score Effectiveness Over Different Time Frames:

Short-term (1-3 days): The correlation between scores and returns seems more pronounced in the very short term. Stocks with higher scores show a tendency for positive returns, whereas stocks with lower scores tend to have negative returns.

Medium to Long-term (5-10 days): The correlation appears to weaken over longer periods, with some high-scored stocks underperforming and some low-scored stocks outperforming expectations.

Outliers and Volatility: Certain stocks exhibit significant volatility post-earnings, which may not always align with the ChatGPT scores. For instance, Fisker Inc. (Ticker: FSR) had a relatively neutral score but experienced a dramatic drop in stock price over 10 days.

Average and Median Scores and Returns: The average and median scores and returns provide a snapshot of the overall dataset's performance. The average score indicates a slightly positive bias in ChatGPT's ratings, while the returns vary across different time frames, with a general trend of negative returns as the time frame extends.

Conclusion:

The data suggests that while ChatGPT's investment attractiveness scores might provide useful insights for short-term investment decisions following earnings announcements, their predictive power appears to diminish over longer time frames. Additionally, the market's inherent volatility and the influence of external factors not captured by earnings call transcripts can significantly impact stock performance, underscoring the complexity of predicting stock movements based on qualitative analyses alone.

To refine the analysis, statistical tests such as Pearson or Spearman correlation coefficients could be applied to quantify the relationship between ChatGPT scores and stock returns more precisely. Furthermore, considering other variables, such as market conditions, sector performance, and news events, could enhance the model's predictive accuracy.

Portfolio Construction Guided by Investment Recommendations

This section examines the outcomes of investments based on ChatGPT's scores. In this analysis, a stock is purchased if the rating exceeds 6 and a stock is shorted if its score is below 0. Due to the timing of this experiment in 1Q2024, the returns of the 100 selected companies skew negative, especially following their earnings reports.

Figure 11. Average Returns and Standard Deviation of all 100 Companies Selected Over 1-Day, 3-Day, 5-Day, and 10-Day Periods.

	1-Day Return	3-Day Return	5-Day Return	10-Day Return
Average Return of 100 Selected Companies	-1.26%	-2.18%	-3.31%	-4.39%
	1-Day Stdev	3-Day Stdev	5-Day Stdev	10-Day Stdev
Standard Deviation of 100 Selected Companies	11.31%	13.26%	14.19%	16.62%

Employing a strategy of buying stocks rated over 6 and shorting those rated below 0, resulted in a strong improvement in returns as well as a decrease in volatility. The strategy resulted in an average improvement in returns of 1.56% across the varying time periods. Specifically, the 1-day returns increased by 1.33%, the 3-day returns increased by 0.45%, the 5-day returns increased by 2.09%, and the 10-day returns increased by 2.36%.

Figure 12. Returns and Standard Deviation from Purchasing Stocks Rated above 6 and Short Selling Stocks Rated Below 0.

	1-Day Return	3-Day Return	5-Day Return	10-Day Return
Average Returns GPT Guided Investments	0.08%	-1.72%	-1.22%	-2.03%
	1-Day Stdev	3-Day Stdev	5-Day Stdev	10-Day Stdev
Standard Deviation GPT Guided Investments	11.30%	10.18%	10.78%	14.38%

Figure 13. Improvement in Returns and Standard Deviation with GPT Scores Relative to Average of 100 Selected Companies

	1-Day Return	3-Day Return	5-Day Return	10-Day Return
Improvement in Returns With GPT Scores	1.33%	0.45%	2.09%	2.36%
	1-Day Stdev	**3-Day Stdev**	**5-Day Stdev**	**10-Day Stdev**
Decrease in Standard Deviation with GPT Scores	0.01%	3.09%	3.40%	2.24%

Despite the regression analysis indicating that the results are not statistically significant, implementing the investment scores into the portfolio strategy above yielded strong results. When looking at these results over a longer time horizon, for example, annualizing the returns, the results are even more promising. Figure 14 shows the improvement in returns on an annual basis by multiplying the returns by 252, which is the number of trading days in a year. This straightforward investment strategy offers room for expansion into more sophisticated portfolio optimization strategies, including assigning larger weights to higher-scoring stocks and potentially incorporating leverage.

Figure 14. Annualized Improvement in Returns with GPT Scores

	Annualized 1-Day Return	Annualized 3-Day Return	Annualized 5-Day Return	Annualized 10-Day Return
Improvement in Returns With GPT Scores	336.09%	38.11%	105.40%	59.59%

Chapter 5

Analysis and Application of Research Findings

Application of Research

Incorporating AI tools such as ChatGPT into investment strategies may offer a competitive advantage, potentially leading to higher returns compared to strategies that do not leverage these tools. This advantage could be particularly pronounced when combining the approach outlined in this research with the established methods of sentiment analysis of market news as outlined in the literature review.

It is important to note that the lack of access to high-frequency trading systems, which can perform trades in milliseconds based on real-time data, might prevent retail traders from exploiting the short-term opportunities. Nonetheless, through post-earnings announcement drift (PEAD), retail traders may still find opportunities to benefit in the aftermath of an earnings report.

The growing integration of AI in financial markets might result in an enhancement in market efficiency. This is because AI has the capability to rapidly process and analyze data sets and provide sentiment analysis, more rapidly than human capabilities allow. Consequently, this could result in a reduction of the post-earnings announcement drift (PEAD) phenomenon.

Improvements to Methodology

Several factors influence the market's reaction to earnings announcements beyond just the general sentiment of the transcript. These factors include the market participants' prior expectations, equity research analysts' forecasts, previous guidance from company management, and key financial indicators. Adopting a more complex methodology that incorporates these elements could strengthen the correlation between the investment ratings and returns. Accounting for the other variables that impact equity prices would likely strengthen the R square value and subsequently lead to a lower p-value. Leveraging FactSet's APIs, such as the Documents Distributor API for automatically accessing earnings reports, would also significantly improve the efficiency by allowing the Assistant to directly retrieve the files. A myriad of other APIs could offer additional financial insights and historical context. However, due to the unavailability of FactSet's developer tools at Pennsylvania State University, such integration was not possible in this study.

In this research, the focus of sentiment analysis remained on the text transcribed from the earnings calls. Extending the analysis to include audio files might improve the relationship between GPT scores and the financial returns. OpenAI's Whisper, an automatic speech recognition system, possesses the capability to analyze the sentiment of audio files and could be integrated into the study's model for more comprehensive insights.

Chapter 6

Conclusion

The study indicates a slight positive correlation between ChatGPT's investment scores and the average stock performance post-earnings announcement. Due to the elevated P-values from the regression analysis the results are deemed not statistically significant. Employing a strategy of buying stocks rated above 6 and selling short those rated below 0 led to notably better returns and reduced volatility compared to the entire data set. Specifically, on an annualized basis the strategy resulted in returns of ~336%. ChatGPT was particularly strong at predicting declines in stock value. Out of the 100 evaluated companies, only 2 received negative scores, and these 2 companies experienced an average decrease of 23% in the 10 days after their earnings calls.

Artificial intelligence has many different applications within finance, one of which is to guide investment decisions. ChatGPT should not be used solely to make investment decisions; however, it can prove to be a useful tool for investment managers in quickly analyzing sentiment, analyzing data sets, or improving the efficiency of the day-to-day tasks of their job.

Throughout the research, ChatGPT provided valuable insights, including perspective on AI's potential to outperform the market, its analysis of the results, and its input on ways to improve the methodology.

There are numerous ethical considerations surrounding artificial intelligence, notably the disruption of employment by automating various tasks. However, in the realm of high finance, such disruption appears improbable given the current technological capabilities. Nonetheless, AI will still serve as a useful tool for increasing productivity and providing alternative strategies to analyze investment opportunities.

Another ethical consideration involves the ideologies embedded within AI models. For example, the Gemini chatbot from Google gained substantial controversy over its biased responses, which favored diversity over factual correctness (Kruppa, Miles). In finance, the backend development of AI systems could influence their perspective on different investments, including the model's emphasis on environmental, social, and governance (ESG) investing.

In conclusion, the role of AI can play an important role in financial analysis, especially as technology continues to evolve. This study highlights ChatGPT's ability to inform investment strategies by analyzing the sentiment of earnings call transcripts. Yet, recognizing AI's limitations and utilizing the tools in context with supplementary data is essential. Moreover, ethical considerations of AI in financial markets cannot be overlooked in terms of both workforce disruption and potential biases.

Appendix A

Comprehensive List of Companies and Investment Scores

#	Ticker	Company Name	ChatGPT Score	1 Day Return	3 Day Return	5 Day Return	10 Day Return
1	AAN	Aaron's Company	2	(18.5%)	(25.9%)	(31.2%)	(32.5%)
2	ACIC	American Coastal Insurance	6	(14.6%)	(25.8%)	(24.1%)	(24.0%)
3	ADPT	Adaptive Biotechnologies	4	0.5%	(2.3%)	(3.0%)	5.3%
4	AEIS	Advanced Energy Industries	1	(5.8%)	(1.7%)	(6.2%)	(6.6%)
5	AGS	PlayAGS, Inc.	5	(1.3%)	(2.3%)	(3.4%)	(6.2%)
6	AI	C3.ai, Inc.	7	24.5%	14.8%	9.3%	2.3%
7	ALKT	Alkami Technology Inc	6	(3.3%)	(5.7%)	(9.0%)	(8.1%)
8	ALLO	Allogene Therapeutics, Inc.	7	3.7%	(2.3%)	(2.6%)	(3.3%)
9	ALTG	Alta Equipment Group, Inc.	6	5.1%	14.7%	23.4%	22.2%
10	ALTI	AlTi Global, Inc	3	4.0%	(2.8%)	(13.4%)	(19.1%)
11	AMK	AssetMark Financial Holdings	6	1.4%	4.5%	6.6%	9.6%
12	ARRY	Array Technologies Inc	5	3.4%	3.7%	(5.6%)	(10.8%)
13	AVNW	Aviat Networks, Inc.	7	19.2%	8.6%	9.8%	15.7%
14	AX	Axos Financial, Inc.	6	(2.4%)	(6.1%)	(9.8%)	(8.8%)
15	BGFV	Big 5 Sporting Goods	(3)	(9.6%)	(11.0%)	(19.7%)	(30.5%)
16	BOOT	Boot Barn Holdings, Inc.	6	10.3%	13.6%	17.3%	26.1%
17	BYND	Beyond Meat, Inc.	5	30.7%	29.9%	11.5%	9.0%
18	BZH	Beazer Homes USA, Inc.	4	(7.9%)	(10.0%)	(9.0%)	(5.8%)
19	CARG	CarGurus, Inc. Class A	6	(1.5%)	(7.3%)	(4.0%)	(3.7%)
20	CLNE	Clean Energy Fuels Corp.	(2)	(3.9%)	(0.7%)	(9.2%)	(16.1%)
21	COOK	Traeger, Inc.	4	(13.6%)	(7.6%)	1.3%	8.9%
22	CRCT	Cricut, Inc. Class A	2	(3.3%)	(1.7%)	1.4%	6.6%
23	CRNX	Crinetics Pharmaceuticals	6	(4.3%)	1.8%	1.8%	(5.3%)
24	CWK	Cushman & Wakefield Plc	3	(10.2%)	(11.8%)	(11.8%)	(12.8%)
25	DCGO	DocGo Inc.	7	3.3%	(0.5%)	0.3%	(11.2%)
26	DENN	Denny's Corporation	6	(6.2%)	(6.0%)	(7.8%)	(6.6%)
27	DHX	DHI Group, Inc.	4	13.7%	25.5%	27.8%	19.3%
28	DIOD	Diodes Incorporated	3	(2.2%)	2.6%	(3.2%)	(0.5%)
29	DSGN	Design Therapeutics, Inc.	6	12.0%	46.7%	44.8%	44.8%
30	DUOL	Duolingo, Inc. Class A	7	22.2%	15.0%	8.6%	15.4%

31	EOLS	Evolus, Inc.	6	(4.9%)	(4.7%)	(4.9%)	(6.8%)
32	EVER	EverQuote, Inc. Class A	4	2.9%	(6.5%)	(8.5%)	(1.4%)
33	EVLV	Evolv Technologies Holdings	7	(16.8%)	(21.0%)	(21.8%)	(24.3%)
34	EXPO	Exponent, Inc.	5	(12.8%)	(12.9%)	(12.6%)	(13.4%)
35	FGEN	FibroGen, Inc.	6	14.6%	(7.6%)	7.0%	(7.6%)
36	FSR	Fisker Inc Class A	3	(33.7%)	(43.7%)	(45.1%)	(78.7%)
37	GEVO	Gevo, Inc.	6	(4.0%)	(12.5%)	(20.3%)	(18.3%)
38	HALO	Halozyme Therapeutics, Inc.	7	6.6%	12.7%	10.7%	15.0%
39	HDSN	Hudson Technologies, Inc.	3	(18.5%)	(18.2%)	(20.8%)	(15.5%)
40	HTLF	Heartland Financial USA, Inc.	3	0.8%	(9.8%)	(12.4%)	(9.6%)
41	IBEX	IBEX Ltd	3	(9.1%)	(14.2%)	(14.3%)	(19.1%)
42	INDI	indie Semiconductor, Inc.	4	(13.9%)	(9.7%)	(10.2%)	0.4%
43	KAR	OPENLANE, Inc.	5	3.8%	6.4%	6.1%	11.2%
44	KIND	Nextdoor Holdings, Inc.	4	2.5%	6.9%	7.8%	8.3%
45	KN	Knowles Corp.	4	(3.7%)	(2.2%)	(6.1%)	2.5%
46	KTOS	Kratos Defense & Security	5	17.0%	16.7%	15.2%	4.9%
47	LAZR	Luminar Technologies, Inc.	6	(4.8%)	(6.8%)	(12.4%)	(23.2%)
48	LC	LendingClub Corp	7	3.2%	3.4%	(0.7%)	(3.0%)
49	LESL	Leslie's, Inc.	3	2.5%	11.6%	12.2%	15.7%
50	LNW	Light & Wonder, Inc.	8	0.9%	6.0%	2.7%	5.4%
51	LUNG	Pulmonx Corp.	6	(18.2%)	(28.3%)	(35.4%)	(39.0%)
52	MATX	Matson, Inc.	5	(1.2%)	0.6%	(2.4%)	(8.6%)
53	MDXG	MiMedx Group, Inc.	7	(0.5%)	0.7%	(1.6%)	(3.0%)
54	MED	Medifast, Inc.	5	(17.5%)	(16.8%)	(14.7%)	(20.6%)
55	MGRC	McGrath RentCorp	7	(1.5%)	(1.7%)	(2.1%)	(3.6%)
56	MLNK	MeridianLink, Inc.	6	(4.2%)	(2.8%)	(5.9%)	(3.2%)
57	MRNS	Marinus Pharmaceuticals Inc	6	(2.3%)	(4.2%)	0.4%	0.6%
58	MSTR	MicroStrategy Incorporated	6	2.0%	29.8%	37.8%	41.6%
59	MTTR	Matterport, Inc. Class A	7	(16.9%)	(16.1%)	(15.0%)	(25.2%)
60	NAPA	Duckhorn Portfolio, Inc.	5	(8.6%)	(5.0%)	(3.1%)	(4.3%)
61	NDLS	Noodles & Co. Class A	4	(8.6%)	(18.4%)	(26.5%)	(22.0%)
62	NEO	NeoGenomics, Inc.	6	12.1%	14.2%	9.7%	7.2%
63	NRDY	Nerdy, Inc. Class A	6	(5.6%)	(6.5%)	(8.3%)	(7.4%)
64	NVEE	NV5 Global Inc	7	(6.1%)	(6.6%)	(6.6%)	(9.0%)
65	NVRO	Nevro Corp.	6	(0.5%)	(5.7%)	(10.7%)	(12.4%)
66	NVTS	Navitas Semiconductor	8	(7.8%)	(12.4%)	(12.4%)	(23.6%)
67	OFIX	Orthofix Medical, Inc.	7	8.7%	7.7%	8.1%	6.8%
68	ONTO	Onto Innovation, Inc.	7	9.3%	2.2%	7.1%	8.1%
69	ORGN	Origin Materials, Inc.	8	(8.4%)	(8.2%)	(7.0%)	(16.8%)
70	PFSI	PennyMac Financial Services	7	4.7%	0.8%	1.4%	2.1%

71	PHR	Phreesia, Inc.	7	4.8%	2.2%	4.0%	2.3%
72	PLL	Piedmont Lithium Inc	7	(10.7%)	(1.7%)	5.4%	3.5%
73	PRAA	PRA Group, Inc.	6	14.2%	6.5%	3.4%	4.3%
74	PRO	PROS Holdings, Inc.	7	2.1%	(0.8%)	4.4%	(5.8%)
75	PRPL	Purple Innovation, Inc.	7	(18.6%)	(23.5%)	(25.0%)	(27.9%)
76	PRTS	CarParts.com Inc	6	(22.7%)	(29.1%)	(31.1%)	(34.3%)
77	QLYS	Qualys, Inc.	5	(4.2%)	(3.1%)	(3.7%)	(10.0%)
78	RDFN	Redfin Corporation	2	(2.2%)	(2.2%)	(12.0%)	(1.1%)
79	RKLB	Rocket Lab USA, Inc.	4	(2.5%)	(3.6%)	(10.0%)	(7.9%)
80	ROG	Rogers Corporation	3	(1.5%)	(4.3%)	(1.6%)	(5.2%)
81	SANM	Sanmina Corporation	4	28.2%	18.6%	17.2%	22.2%
82	SBGI	Sinclair, Inc. Class A	3	0.4%	(7.0%)	(9.1%)	(10.0%)
83	SCPH	scPharmaceuticals, Inc.	4	(2.6%)	0.9%	2.6%	(6.5%)
84	SDGR	Schrodinger, Inc.	6	(18.8%)	(19.5%)	(20.9%)	(15.5%)
85	SES	SES AI Corporation	6	14.1%	18.8%	14.8%	4.0%
86	SFM	Sprouts Farmers Market	5	10.9%	13.4%	16.1%	17.8%
87	SIBN	SI-BONE	7	(13.5%)	(15.5%)	(13.4%)	(14.1%)
88	SILK	Silk Road Medical	6	1.7%	0.3%	(11.5%)	(12.5%)
89	SKIN	Beauty Health Company	3	18.4%	0.6%	9.7%	17.0%
90	SKYW	SkyWest	4	8.2%	10.1%	10.4%	8.9%
91	SNDX	Syndax Pharmaceuticals	5	(2.2%)	(2.9%)	(6.0%)	(2.3%)
92	SPSC	SPS Commerce, Inc.	3	14.8%	5.1%	7.0%	(0.4%)
93	STNE	StoneCo Ltd. Class A	4	(6.9%)	(4.0%)	(7.1%)	(7.1%)
94	SVV	Savers Value Village Inc.	3	(7.5%)	(6.2%)	(13.6%)	(15.6%)
95	URBN	Urban Outfitters, Inc.	2	(12.8%)	(10.5%)	(7.5%)	(13.9%)
96	UVV	Universal Corp	8	(6.0%)	(7.7%)	(11.9%)	(12.7%)
97	VEL	Velocity Financial, Inc.	9	4.0%	4.4%	7.2%	15.0%
98	VRRM	Verra Mobility Corp.	8	5.8%	(0.1%)	3.4%	7.9%
99	WK	Workiva Inc.	7	(12.9%)	(7.1%)	(9.8%)	(10.0%)
100	YELP	Yelp Inc	7	(14.4%)	(16.6%)	(17.8%)	(13.4%)
		Average	**5.2**	**(1.3%)**	**(2.2%)**	**(3.3%)**	**(4.4%)**
		Standard Deviation	**2.0**	**11.3%**	**13.3%**	**14.2%**	**16.6%**
		Median	**6.0**	**(2.2%)**	**(2.5%)**	**(3.9%)**	**(5.5%)**
		Maximum	**9.0**	**30.7%**	**46.7%**	**44.8%**	**44.8%**
		Minimum	**(3.0)**	**(33.7%)**	**(43.7%)**	**(45.1%)**	**(78.7%)**

Appendix B

Regression Statistics

1-Day Regression Summary Output:

1-Day Regression Statistics	
Multiple R	0.127268831
R Square	0.016197355
Adjusted R Square	0.006158553
Standard Error	0.112765819
Observations	100

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	1	0.020517156	0.020517156	1.613474849	0.207011559
Residual	98	1.246180741	0.01271613		
Total	99	1.266697897			

	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*Lower 95.0%*	*Upper 95.0%*
Intercept	-0.049553005	0.031214205	-1.587514571	0.115617506	-0.111496578	0.012390569	-0.111496578	0.012390569
ChatGPT Score	0.007137324	0.005618939	1.270226299	0.207011559	-0.004013277	0.018287926	-0.004013277	0.018287926

3-Day Regression Summary Output:

3-Day Regression Statistics	
Multiple R	0.109557037
R Square	0.012002744
Adjusted R Square	0.00192114
Standard Error	0.132506333
Observations	100

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	1	0.020903748	0.020903748	1.190558928	0.277892255
Residual	98	1.720676976	0.017557928		
Total	99	1.741580724			

	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*Lower 95.0%*	*Upper 95.0%*
Intercept	-0.059078241	0.036678489	-1.610705445	0.110460316	-0.131865508	0.013709026	-0.131865508	0.013709026
ChatGPT Score	0.007204252	0.006602577	1.091127366	0.277892255	-0.005898347	0.020306851	-0.005898347	0.020306851

5-Day Regression Summary Output:

5-Day Regression Statistics	
Multiple R	0.1797493
R Square	0.032309811
Adjusted R Square	0.022435421
Standard Error	0.1402566
Observations	100

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	1	0.064368112	0.064368112	3.272081809	0.073534413
Residual	98	1.92784756	0.019671914		
Total	99	1.992215671			

	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*Lower 95.0%*	*Upper 95.0%*
Intercept	-0.098617974	0.038823805	-2.540141863	0.012651748	-0.175662552	-0.021573395	-0.175662552	-0.021573395
ChatGPT Score	0.012641896	0.00698876	1.808889662	0.073534413	-0.001227071	0.026510863	-0.001227071	0.026510863

10-Day Regression Summary Output:

10-Day Regression Statistics	
Multiple R	0.170718299
R Square	0.029144738
Adjusted R Square	0.019238051
Standard Error	0.164572366
Observations	100

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	1	0.07967931	0.07967931	2.941925938	0.089468097
Residual	98	2.654238245	0.027084064		
Total	99	2.733917554			

	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*Lower 95.0%*	*Upper 95.0%*
Intercept	-0.116803157	0.045554544	-2.564028651	0.011865113	-0.207204668	-0.026401646	-0.207204668	-0.026401646
ChatGPT Score	0.01406532	0.008200376	1.715204343	0.089468097	-0.002208059	0.030338699	-0.002208059	0.030338699

BIBLIOGRAPHY

"Amplify AI Powered Equity ETF." Amplify ETFs, 25 Mar. 2024, amplifyetfs.com/aieq/.

Bernard, Victor L., and Jacob K. Thomas. "Post-Earnings-Announcement Drift: Delayed Price Response or Risk Premium?" Journal of Accounting Research, vol. 27, 1989, pp. 1–36. JSTOR, https://doi.org/10.2307/2491062.

"Buzz - Vaneck Social Sentiment ETF: Holdings & Performance." VanEck, 28 Mar. 2024, www.vaneck.com/us/en/investments/social-sentiment-etf-buzz/overview/.

Carlson, Debbie. "AI Isn't So Smart. Its Stock Picks Lag Behind the Market." Barrons, https://www.barrons.com/articles/ai-powered-etfs-dont-beat-the-broader-market-febd38fb. Accessed 11 Dec. 2023.

Dong, Liang. "Predicting Stock Price Changes with Earnings Calls." School of Information and Library Science of the University of North Carolina at Chapel Hill, Apr. 2016.

"FactSet Developer Portal." FactSet, developer.factset.com/api-catalog. Accessed 24 Mar. 2024.

Fama, Eugene F. "Efficient Capital Markets: A Review of Theory and Empirical Work." The Journal of Finance, vol. 25, no. 2, 1970, pp. 383–417. JSTOR, https://doi.org/10.2307/2325486.

Fink, Josef. “A Review of the Post-Earnings-Announcement Drift.” Journal of Behavioral and Experimental Finance, vol. 29, Mar. 2021, p. 100446. ScienceDirect, https://doi.org/10.1016/j.jbef.2020.100446.

Hansen, Anne Lundgaard, and Sophia Kazinnik. Can ChatGPT Decipher Fedspeak? 4399406, 24 Oct. 2023. Social Science Research Network, https://doi.org/10.2139/ssrn.4399406.

Hulbert, Mark. “Why Ai-Generated Stock Picks Won’t Beat the Market.” MarketWatch, 19 Aug. 2023, www.marketwatch.com/story/why-ai-generated-stock-picks-wont-beat-the-market-9a6a3495.

Kruppa, Miles. “Google Mired in Controversy over Gemini AI Chatbot Push.” The Wall Street Journal, 28 Feb. 2024, www.wsj.com/tech/ai/google-mired-in-controversy-over-ai-chatbot-push-46023dd3.

Ko, Hyungjin, and Jaewook Lee. Can CHATGPT Improve Investment Decision? From a Portfolio Management Perspective, 3 Mar. 2023, https://doi.org/10.2139/ssrn.4390529.

Lee, Justina. “AI Can Write, but Is It Any Good at Picking Stocks?” Washington Post, 17 July 2023, www.washingtonpost.com/business/2023/07/17/ai-can-write-but-is-it-any-good-at-picking-stocks-quicktake/9f5bcb56-246f-11ee-9201-826e5bb78fa1_story.html.

Lopez-Lira, Alejandro, and Yuehua Tang. Can ChatGPT Forecast Stock Price Movements? Return Predictability and Large Language Models. 4412788, 6 Apr. 2023. Social Science Research Network, https://doi.org/10.2139/ssrn.4412788.

Loughran, Tim, and Bill McDonald. When Is a Liability Not a Liability? Textual Analysis, Dictionaries, and 10-Ks. 1331573, 4 Mar. 2010. Social Science Research Network, https://papers.ssrn.com/abstract=1331573.

Niszczota, Paweł, and Sami Abbas. "GPT as a financial advisor." Poznań University of Economics and Business, 11 Mar. 2023, https://doi.org/10.2139/ssrn.4384861.

OpenAI. New Models and Developer Products Announced at DevDay, 6 Nov. 2023, openai.com/blog/new-models-and-developer-products-announced-at-devday.

OpenAI. "Overview - OpenAI API." OpenAI Documentation, platform.openai.com/docs/overview. Accessed 30 Mar. 2024.

Pocock, Kevin. "What Is ChatGPT? Why You Need to Care about GPT-4." PC Guide, 1 Dec. 2023, https://www.pcguide.com/apps/what-is-chat-gpt/.

Pelster, Matthias, and Joel Val. "Can CHATGPT assist in picking stocks?" Finance Research Letters, vol. 59, Jan. 2024, p. 104786, https://doi.org/10.1016/j.frl.2023.104786.

Ponsot, Elisabeth. “Can AI Beat the Stock Market?” Bloomberg, Bloomberg, 3 Oct. 2023, www.bloomberg.com/news/newsletters/2023-10-03/can-ai-beat-the-stock-market-big-take.

“Qraft AI-Enhanced U.S. Large Cap Momentum ETF.” Qraft AI ETFs, 27 Mar. 2024, www.qraftaietf.com/amom.

Royal, James. “4 Ai-Powered ETFs: Pros and Cons: Bankrate.” Edited by Mercedes Barba, Bankrate, 16 Aug. 2023, www.bankrate.com/investing/ai-powered-etfs-pros-cons/.

Tetlock, Paul C. “Giving Content to Investor Sentiment: The Role of Media in the Stock Market.” The Journal of Finance, vol. 62, no. 3, 2007, pp. 1139–68. JSTOR, https://www.jstor.org/stable/4622297.

“Vanguard Mutual Fund Profile.” Vanguard, 28 Mar. 2024, investor.vanguard.com/investment-products/etfs/profile/vea.

“WisdomTree International AI Enhanced Value Fund.” WisdomTree, www.wisdomtree.com/investments/etfs/equity/aivi. Accessed 30 Mar. 2024.

Academic Vita:

THOMAS J. BURKE

burkethomasj1@gmail.com

EDUCATION

The Pennsylvania State University | Schreyer Honors College — **University Park, PA**
Smeal College of Business | B.S. in Finance | Real Estate Certificate — *Class of 2024*
Schreyer Honors College Thesis: ChatGPT's Analysis of 1Q2024 Earnings Call Transcripts — Summa Cum Laude

RELEVANT EXPERIENCE

Moelis & Company — **New York, NY**
Incoming Investment Banking Analyst — *Jul 2024*

Moelis & Company — **Washington, DC**
Investment Banking Summer Analyst — *Jun 2023 – Aug 2023*

- Prepared board materials analyzing a public aerospace company's ending cash balance and liquidity needs under various business cases and conducted illustrative share price analysis based on Moelis' four proposed restructuring approaches
- Supported deal team in reaching exclusivity and coordinating due diligence in a buy-side M&A of an aerospace company
- Conducted strategic analysis of potential acquirers, including tailored transaction structures aligned with each strategic acquirer's debt capacity, to guide initial discussions in the sell-side M&A of a naval systems company
- Created valuation benchmarking slides and comprehensive sensitivity analysis to support IPO discussion materials
- Maintained database of Aerospace & Defense precedent transactions and automated output sheet utilizing VBA in Excel

JPB Partners — **Hanover, MD**
Private Equity Summer Analyst — *May 2022 – Jul 2022*

- Performed in-depth three-statement financial modeling, incorporating complex deal terms, capital structure assessment, valuation, and LBO analysis, to guide investment decisions in companies with annual revenues of ~$2mm to ~$25mm
- Leveraged PitchBook data to strategically identify and assess various industries and companies for future investment opportunities and initiated deal-sourcing outreach for add-on acquisitions to integrate with existing portfolio companies
- Engaged with investment bankers regarding a financial model for the sale process of a $50mm portfolio company

Nittany Lion Fund — **University Park, PA**
Director of Education — *Jan 2023 – May 2023*

- Hosted educational sessions to train new fund manager's technical abilities including, DCF, M&A, and LBO modeling
- Implemented new 3-statement modeling speed test and Excel shortcuts speed test to improve fund manager's efficiency

Fund Manager | Industrials Sector — *Nov 2021 – Dec 2022*

- Served as the fund manager of the Industrials portfolio valued at ~$1.5mm within Penn State's ~$16mm student-run hedge fund by completing stock performance reports and equity pitches with a goal of outperforming the S&P 500
- Achieved 9.5% relative outperformance vs. the SPDR Industrials Sector Benchmark (XLI) in FY2022

LEADERSHIP AND ACTIVITIES

Smeal College of Business — **University Park, PA**
Sapphire Leadership Academic Program Scholar — *Aug 2020 – May 2023*

- Selected to be a part of a 50-member cohort representing the top 5% of the Smeal College of Business incoming class
- Engaged in in Sapphire-specific curriculum and participated in over 12 hours of professional development, leadership development, and community service activities each semester to enhance leadership skills

Advanced Financial Modeling (FIN 415) Teaching Assistant — *Jan 2024 – Present*

- Severed as the teaching assistant for the Advanced Financial Modeling course, covering VBA programs, Monte Carlo simulations, Blach-Scholes, portfolio construction, hedging strategies, and other advanced financial concepts

Penguin Protection Party — **Garnet Valley, PA**
Co-Founder — *Jul 2021 – Nov 2021*

- Developed a collection of 800 unique penguin non-fungible tokens (NFTs) on the Ethereum blockchain
- Executed marketing campaigns, resulting in 189 individual sales and ~$10,000 (2.3 ETH) in volume traded
- Partnered with the President of the Global Penguin Society, successfully raising $2,935 for penguin conservation efforts

AWARDS, CERTIFICATIONS, AND INTERESTS

Awards: 2023 CFA Institute Research Challenge 1st Place Sub-Regional, Dean's List (7/7 Semesters)
Certifications: Bloomberg Market Concepts, FactSet Core, and Adventis' Financial Modeling Level II
Interests: Backpacking Europe & Southeast Asia, *Huberman Lab Podcast*, *Atomic Habits*, Postural Restoration Institute