Can ChatGPT Replicate Analyst Recommendations?

Jason Ming * Hamish Malloch † P. Joakim Westerholm ‡

September 13, 2024

ABSTRACT

This paper examines the ability of a popular large language model—ChatGPT—to emulate the skills of equity analysts. We tackle this question using text transcripts from earnings conference calls. Our approach is two-fold. First, we use the questions posed by analysts during earnings calls to train ChatGPT about analysts decision making process. Next, we apply what ChatGPT has learned from these questions to evaluate transcripts from 500 of the largest firms in the US by developing an Analyst Insight Score (AIS) for each firm-transcript pair. We find that our "machine learned" AIS is consistent with observed behaviour from analysts including adjustments to price targets after earnings announcements. Moreover, our AIS outperforms more traditional analyst metrics such as the Standard Unexpected Earnings (SUE). Finally, we find that our AIS allows for the construction of portfolios that earn abnormal returns with respect to several standard asset pricing models.

Keywords: AI, ChatGPT, Earnings Call, Information and Market Efficiency, Textual Analysis

JEL Classification: G11, G12, G14, C45

*Email: jason.ming@sydney.edu.au
†Email: hamish.malloch@sydney.edu.au
‡Email: joakim.westerholm@sydney.edu.au
We thank OpenAI and Postgraduate Research Support Scheme at the University of Sydney for providing computational resources.

I. Introduction

A fundamental principle of market efficiency is that asset prices incorporate available information. While financial economics has traditionally emphasized numerical data, some of the most critical insights into a firm's current and expected future performance are disclosed during earnings conference calls. These calls, conducted shortly after a firm's earnings announcement, allow management to discuss performance and provide an opportunity for market participants to engage with firm leadership directly. The dialogue between firm management and professional analysts often reveals crucial information about the firm's true value that may not be captured by numerical data alone. Given the complexity associated with analysing natural language, interpreting this information has remained the domain of professional equity analysts. However, the introduction of advanced artificial intelligence (AI) — specifically Large-Language Models (LLMs) — presents an opportunity to study how analysts approach the valuation process and also whether the expertise of market professionals can be replicated with AI.

This paper examines if the skill of equity analysts can be replicated, or potentially improved upon, by AI designed to mimic/interpret natural language. Our results suggest that a state of the art LLM—ChatGPT 4 Turbo—is able to analyse conference calls text in a manner that is similar to professional market practitioners. Our findings are based on the development of an AI generated Analyst Insight Score (AIS) which provides an overall evaluation of a firms prospects based on the information derived from earnings calls. We find that AIS scores are significant predictors of future revisions made by analysts to price targets, buy-hold-sell recommendations and earnings projections. These results demonstrate an encouraging alignment between AI and humans regarding their interpretation of textual information derived from earnings calls. Additionally, our newly developed score outperforms competing measures such as the Standardized Unexpected Earnings (SUE) with respect to predicting analyst revisions.

We further leverage our AIS to study the asset pricing implications of the text-based information contained in earnings calls. By sorting stocks according to their AIS innovation, we can produce portfolios that yield abnormal returns of between 50-60 bps per month, with this results being robust to a variety of systematic risk factors. This finding suggests not only that earnings calls are valuable sources of information for determining the stock values, but also that AI models have the

capacity to uncover relevant information which is later impounded into stock prices.

Because the output produced by a LLM is conditional on the prompt, a key aspect to our study is a two-stage approach that allows ChatGPT to interpret earnings call transcripts in a manner that matches that of professional equity analysts. First, we train ChatGPT on the thought process of a professional analyst by having it identify the key principles that underlie the questions raised by analysts during the question-and-answer (Q&A) session session for a set of earnings call transcripts. We then use the key principles identified by ChatGPT to generate the prompt which is used to analyse complete earnings call transcripts at a later date. Our approach is hence "data-driven" in the sense that text data drawn from earlier earnings calls is used to develop the prompt that allows ChatGPT to analyse earning call transcripts. The prompts are not influenced by external factors, emotions or other biases associated with the end user, nor are they directed by the authors. This approach to prompt generation means that our results are largely reproducible.[1] To test the impact of the first-stage Q&A analysis to build the prompt, we also examine an application of ChatGPT to produce AIS's without the benefit of this first pass training on Q&A sessions. The results are less impressive suggesting that while ChatGPT is able to interpret the text from earnings calls, it cannot provide the correct guidance without first learning the underlying principles from equity analysts.

Our study focuses on the 500 largest U.S. stocks by market capitalization over a nine-year period from 2015 to 2023, representing approximately 80% of the total U.S. equity market capitalization. This sample consists of large, stable, and well-covered stocks, ensuring that the earnings call transcripts analyzed include substantial analyst activity. Given the vast dataset used to train ChatGPT, we conduct an extensive investigation to rule out the possibility that ChatGPT produces abnormal performance by utilizing knowledge of future events. By generating various fictitious call transcripts, we assess whether ChatGPT's analysis is based solely on the text provided in the current transcript or if it draws on its extensive training data to infer future outcomes. Our findings indicate that ChatGPT's analysis relies strictly on the content of the provided transcript and the specific prompt given by the researcher. It does not appear to directly use future information when addressing questions about past events.

[1]The difference in ChatGpt's response to a prompt-input pair is determined by the temperature parameter. In our study we set the temperature very low (0.2) so that the same prompt-input pair will produce very similar responses from ChatGPT.

Our work advances the literature in several ways. First, we contribute to the growing field of textual analysis in finance which seeks to determine how information derived from text-based sources impacts investors, firms and asset prices. For instance, a number of authors have studied how market participants react to articles published in the mainstream media. This includes Tetlock (2007) and Tetlock, Saar-Tsechansky, et al. (2008) who show that a pessimism index based on language from Wall Street Journal articles can predict downward pressure on market prices while Tetlock (2011) expands upon this work to show that language from the Dow Jones News Service and the Wall Street Journal can be used to predict accounting earnings and stock returns. Baker, Bloom, and Davis (2016) also show how information from news articles can be used to measure economic policy uncertainty. Ke, Kelly, and Xiu (2020) examine the issue of using machine learning to predict returns from text-based data drawn from the Dow Jones Newswires. These authors find that their approach excels at return prediction and can be used to develop a profitable, real-time trading strategy.

Another rich source of text-based information are firms 10-K filings. Consequently, the information contained in these reports have also been found to have a profound impact on financial markets. Feldman, Govindaraj, Livnat, and Segal (2010), Loughran and McDonald (2011) and Jegadeesh and Wu (2013) study the tone of 10-K filings and their relation to future stock returns while Campbell, Chen, Dhaliwal, Lu, and Steele (2014) and Filzen (2015) examine the impact that additional risk disclosures in 10-K and 10-Q filings have on the firms risk and return characteristics. Cohen, Malloy, and Nguyen (2020) examines how quickly prices react to changes in information contained in these 10-K filings and find that, due to the increasing complexity of these documents, this information contains significant predictive power for future returns.

Though clearly an important source of information, 10-K filings are relatively limited as they represent a delivery of information from the firm. Earnings calls can provide additional insight as they contain not just a delivery of information, but also allow for the firm's management to be challenged and questioned by participants on the call (typically equity analysts). This interaction between management and other market participants can reveal additional, relevant information. Clement, Frankel, and Miller (2003) study the reaction of stock prices around earnings calls where management provide confirming statements and find that these statements reduce uncertainty. Matsumoto, Pronk, and Roelofsen (2011) examine abnormal returns following earnings calls and

find that there is incremental information beyond the presentation by management while Price, Doran, et al. (2012) show that the question and answer section is particularly important for predicting abnormal returns, price-earning announcement drift and trade volume. Even the specific type of language used can affect investors interpretation of the information being delivered. Suslava (2021) examines euphemisms as a mechanism to obscure bad news and finds that their use leads to investor under reaction while Bian, Chen, and Wang (2022) examine stock price response to the quality of the interaction between firm management and other call participants. These authors find that high quality interactions (those where management provide direct answers) lead to higher risk adjusted returns than low quality interactions (those where management provide evasive responses). Our work extends from this branch of literature through the use of a sophisticated LLM which is capable of accounting for all the elements studied in the prior literature—tone, language, context and quality of interaction— in its analysis.

Our work also contributes to the growing literature exploring the capabilities of LLMs. Following its launch in 2023, OpenAI's ChatGPT and GPT-4 has gained widespread recognition for its ability to perform tasks that have traditionally required human expertise. To illustrate, ChatGPT-3/GPT-4 has been show to be capable of passing the bar exam (Katz, Bommarito, Gao, and Arredondo, 2024), a core course in the Wharton MBA program (Terwiesch, 2023) and the US medical licensing exam (Kung, Cheatham, ChatGPT, Medenilla, Sillos, De Leon, Elepaño, Madriaga, Aggabao, Diaz-Candido, Maningo, and Tseng, 2022) among other notable achievements. Naturally, the abilities of ChatGPT/GPT-4 have also caught the attention of the finance community.

Hansen and Kazinnik (2023) explores the ability of GPT-4 to decipher Fedspeak (the technical language used by the Federal Reserve to communicate on monetary policy decisions), Lopez-Lira and Tang (2023) examines the ability of several LLMs, including GPT-4, to predict stock returns using news headlines while Fatouros, Soldatos, Kouroumali, Makridis, and Kyriazis (2023) uses ChatGPT to perform sentiment analysis in foreign exchange markets. More closely aligned with our study, Jha, Qian, Weber, and Yang (2023) shows that GPT can predict corporate policies by analysing earnings calls, Kim, Muhn, and Nikolaev (2023) finds that earnings calls summaries generated by ChatGPT can improve sentiment measurement by removing "bloat" from these documents. Li, Tu, and Zhou (2023) studies how GPT uses information from corporate disclosures to predict returns. In a genuinely novel use of the capabilities of LLMs, Beckmann, Beckmeyer,

Filippou, Menze, and Zhou (2024) uses GPT-4 to identify unusual communication in earnings calls, finding that that firms that communicate this way have lower returns and higher uncertainty. Our study adds to this literature by examining the alignment between GPT-4's analysis of firms earnings calls and professional analysts.

Additionally, we contribute to the literature studying return prediction from earnings call information by showing that portfolios that are long assets with positive earnings calls information and short those with negative earnings calls information produce positive and significant alpha even when controlling for a variety of risk factors. Similar results are found in Cohen et al. (2020) who form portfolios based on changes in language contained in 10-K filings. Finally, in work that is related to and complements ours, Kim, Muhn, and Nikolaev (2024) uses GPT-4 Turbo to examine the ability of LLMs to perform financial statement analysis. These authors provide only the financial statements, with no extraneous language, and find that GPT can outperform professional analysts when predicting earnings changes. We similarly find that GPT can perform accurate financial analysis, but do so using a very different dataset. Additionally, both Kim et al. (2024) and this study highlight the advantage obtained when providing additional guidance to ChatGPT on the analysis process. While this study uses ChatGPT to identify key investment principles by studying the questions posed by analysts in Q&A sessions, Kim et al. (2024) use chain-of-thought queries to provide guidance to ChatGPT on how best to analyse financial statements.

Finally, we contribute to the literature examining the decision making process of equity analysts. While Bouwman (1984) and Bouwman, Frishkoff, and Frishkoff (1987) find that that analysts typically follow a structured checklist to guide their analysis of a company's prospects, Barker and Imam (2008), (Brown, Call, Clement, and Sharp, 2015), Previts, Bricker, Robinson, and Young (1994) and Block (1999) find that analysts use a variety of information including income, core earnings, growth potential, earnings quality, management quality, business risks and private information from management. Since the output of an LLM is determined by the prompt provided, we first instruct GPT-4 to identify the key underlying criteria behind the questions analysts pose to management during the question and answer section of a prior set of earnings calls. This allows GPT-4 to learn the decision making process from professional analysts. We then structure these insights into a prompt which allows GPT-4 to accurately analyse future earnings calls. Removing this learning component from the prompt creation has a negative impact on the performance of

GPT-4's analysis illustrating that understanding the valuation process is critical to preforming accurate analyses.

The remainder of this paper proceeds as follows. In Section II, we review the the data and LLM access used in our experiment. Section III details how we employ ChatGPT to examine earnings call transcripts. Empirical results are presented in Section IV with some sample textual output of our methods are presented in appendix A. Section V concludes.

II. Transcript Data and LLM Access

Our primary data source are transcripts of earnings calls which we obtain from the Refinitiv StreetEvent edited transcript dataset. This dataset constitutes a collection of verbatim transcripts of earnings calls conducted by publicly listed entities. These transcripts are highly structured with each participants dialogue being separately parsed according to specific speakers such as company executives or financial analysts. The transcripts also separate the presentation and discussion sections of the call. This structured format significantly enhances the usability of the dataset, facilitating targeted searches and analysis by delineating the contributions of each speaker. Contained within each transcript are also the date of the earnings call, the name of the corporation, its ticker symbol, and the identities of the participants. Transcripts encompass presentations by corporate executives, as well as the ensuing question-and-answer sessions with analysts and investors, thereby providing a panoramic view of the company's financial disclosures, operational milestones, challenges encountered, and prospective outlook.

We analyse these transcripts with a LLM, of which there are several choices with various advantages and disadvantages. For example, local large language models, which can be deployed on locally managed computers, offer marginal cost advantages over cloud-based solutions provided by OpenAI. However, running these local LLMs requires expensive graphic processing units. Alternatively, OpenAI offers a series of LLMs with different features. As of the time of writing, *GPT-3.5-Turbo* has a maximum token limit of 4,096 tokens and 16,384 tokens for its extended version, with training data up to September 2021. The cost for input is \$0.0015 per 1K tokens and \$0.002 per 1K tokens for output making this a very cost effective choice. *GPT-4* introduces further refinements and capabilities, and is better suited for difficult problems that require greater accuracy.

GPT-4 can analyze and comment on images and graphics, unlike GPT-3.5 which can only analyze text. It has a maximum token limit of 8,192 tokens and 32,768 tokens for its extended version, with training data also up to September 2021. The cost for input is $0.03 per 1K tokens and $0.06 per 1K tokens for output. *GPT-4-Turbo* supports a context size of 128k tokens, allowing for extensive interactions and complex text generation tasks. It is trained on data up to April 2023, which (as of the time of writing) is the latest among OpenAI models. The cost for input is $0.01 per 1K tokens and $0.03 per 1K tokens for output. In addition, the Claude series by Anthropic are a competitive alternative to OpenAI solutions. However, their most advanced models are more expensive than those from OpenAI. Considering cost and features of various models, we elect to use OpenAI's LLMs as they represent the optimal choice in terms of cost and sophistication. Specifically, we employ the GPT-3.5-Turbo API to generate summaries from transcripts, leveraging its cost-efficiency and speed. The more advanced GPT-4-Turbo is used to extract information from these summaries. With rapid advance of LLM development, more state-of-the-art models are expected assist this line of research.

III. Modeling Analyst Responses

In this section, we outline our approach to analysing earnings call transcripts with ChatGPT. This requires several steps including developing a template of analyst expertise, preparation of the transcripts and development of an overall score for each firms transcript which we refer to as an Analyst Insight Score (AIS). These steps are outlined in sections III.A, III.B and III.C respectively. In section III.D, we outline an alternative AIS calculation devoid of the insights gained from studying the questions of professional analysts which provides a point of comparison for our experiments.

A. Developing a Template of Analyst Expertise

Our objective is to uncover how analysts determine the prospects of a firm based on earnings call information. Work by Bouwman et al. (1987) reveals the existence of a "financial template" which are derived from analysts' experience and constitute a major part of their expertise. In their study, this template is considered to be the "key to expert decision making behavior". In

this section, we aim to uncover the template used by analysts to interpret earnings calls via AI. Specifically, we examine the ability of LLMs to extract the factors which underlie how an analyst evaluates of firms prospects from information revealed through earnings calls.

As highlighted by Bouwman et al. (1987) elements of this template could be determined by directly interviewing analysts. This is the approach followed by Block (1999) who, via a survey where analysts are asked which factors determine a firms P/E ratio, finds that "growth potential","earnings quality","quality of management" and "business risks" are the primary drivers. Given earnings forecasts are typically stable in the short term, P/E ratios are a key input for determining target prices and hence reflect a summary of the overall outlook of the company. We use this finding by directing ChatGPT to use these factors as the primary pillars upon which analysis of the firm rests. This means that ChatGPT is instructed to favour a company with expanding business activities, solid earnings and managed by leadership which employs sound strategies to mitigate risks is favored over one which has weakness in one or more of these areas.

Though survey data is capable of providing the primary pillars upon which to consider the analysis of conference calls, it fails to provide granular detail. For instance, it cannot provide guidance as to which specific information led to the conclusion that a firms management was performing well or poorly as this kind of analysis can often be company specific. To recover these more detailed criteria, we leverage the information contained in the question-and-answer sections in our earnings call transcripts. In this section of earnings calls, analysts are invited to put questions directly to management. Since these opportunities are limited, analysts will likely ask about information which are most useful for their evaluation. We then instruct ChatGPT to identify the motives hide behind these specific questions.

The following is a sample question asked by analyst in NVIDIA's earnings call in 2014. This question about the GeForce graphic card suggests that this analyst wanted to assess potential demand of NVIDIA's product.

```
Question: What is the strength in Nvidia's GeForce attributed to?

Answer: The strength in Nvidia's GeForce could be attributed to
```

```
console users switching to PC due to the ongoing major console
transition, where not all games are backward compatible. Some game
titles may have disappointed.

Analyst motive inferred by ChatGPT: Product demand and adoption
```

To complete our template of analyst reasoning, we instruct ChatGPT to infer the underlying motive behind the questions in 400 earnings call of the largest 100 stocks during 2014. This represents a set of training data that is separate from our main analysis that starts in 2015, mitigating potential look-ahead bias. After identifying the hidden motive for each question, the most frequent themes among these motives are clustered by ChatGPT. As a result, similar motives such as "market expansion", "geographic expansion", "international market expansion" are clustered into one sub-criterion "market and geographic expansion" which resides within the major criterion "growth potential". Table I provides an overview of the criteria ChatGPT identified as underlying analyst questions asked by analysts in 2014.

Table I AIS Template. This table shows the template used by ChatGPT to compute AIS. There are four major criteria to evaluate, i.e., "growth potential", "earnings quality", "quality of management", and "business risks". For each criterion, there are five sub-criteria which are clustered by ChatGPT from the various key points distilled from questions asked by analysts. Samples used in producing this table are from earnings call transcripts of the largest 100 stocks in 2014.

Criterion	**Details**
Growth Potential	
Market and Geographic Expansion	Market Expansion; Geographic Expansion; International Market Expansion; Expansion into New Markets; Global Market Penetration.

Continued on next page

Table I – continued from previous page

Criterion	Details
Product and Service Development	Product Innovation; Service and Product Innovation; New Product and Service Offerings; Product and Service Expansion; Product Line Expansion.
Mergers, Acquisitions, and Partnerships	Acquisition Strategy; Mergers and Acquisitions (M&A) Strategy; Strategic Acquisitions and Partnerships; Integration of Acquisitions; Strategic Collaborations and Partnerships.
Digital Transformation and E-commerce	E-commerce Growth; Digital Engagement and E-commerce Adaptation; Cloud Services Adoption and Expansion; Technological Innovation and E-commerce Expansion; Online Sales Growth and E-commerce Strategy.
Operational Efficiency and Diversification	Asset Optimization and Revenue Growth; Diversification of Revenue Streams; Efficiency in Operations; Asset Management and Strategic Investments; Diversification and Expansion of Operations.
Earnings Quality	
Revenue Quality and Sustainability	Revenue Sustainability; Revenue Stability and Predictability; Revenue Diversification and Sustainability; Revenue Quality and Diversification; Sustainable Revenue Growth.
Cost Management and Efficiency	Cost Control and Efficiency; Cost Management and Operating Efficiency; Expense Management and Efficiency; Margin Improvement and Cost Management; Sustainable Cost Management and Productivity.
Capital and Cash Flow Management	Capital Allocation Efficiency; Cash Flow Consistency; Capital Expenditure Discipline; Cash Flow and Capital Allocation; Sustainability of Cash Flow.

Continued on next page

Table I – continued from previous page

Criterion	Details
Profitability and Margin Analysis	Margin Consistency and Improvement; Profit Margin Sustainability; Margin Expansion Sustainability; Profitability and Margin Improvement; Sustainable Profit Margins.
Impact of External Factors and Accounting Practices	Impact of Currency Exchange Rates; Impact of Acquisitions on Earnings; Revenue Recognition and Quality of Earnings; Tax Efficiency; Transparency in Financial Reporting.
Quality of Management	
Strategic Planning and Execution	Execution of Strategic Initiatives; Long-term Strategic Vision; Strategic Decision-Making and Execution; Strategic Planning and Risk Mitigation; Strategic Vision and Long-Term Planning.
Capital and Financial Management	Capital Allocation and Shareholder Value Optimization; Financial Discipline and Capital Allocation; Strategic Capital Allocation; Financial Stewardship and Capital Allocation; Dividend Policy and Shareholder Value.
Adaptability and Market Responsiveness	Adaptability to Market Changes; Management's Adaptability to Market Conditions; Responsiveness to Market Dynamics; Strategic Execution and Adaptability; Responsiveness to Regulatory Changes.
Leadership and Organizational Effectiveness	Leadership Stability and Strategic Vision; Leadership Development and Succession Planning; Leadership and Vision; Governance and Leadership; Leadership Transition and Strategy Execution.
Operational and Cost Efficiency	Operational Efficiency and Cost Management; Cost Management and Efficiency Initiatives; Efficiency and Cost Control; Strategic Cost Management; Expense Management and Operational Efficiency.

Continued on next page

Table I – continued from previous page

Criterion	Details
Risk	
Market and Economic Risks	Macroeconomic and Currency Risks; Market and Economic Sensitivity; Economic and Political Uncertainties; Market and Credit Risk Exposure; Geopolitical and Economic Risks.
Credit and Financial Risks	Credit Risk Management; Capital Adequacy and Liquidity; Asset Quality and Loan Performance; Financial Leverage and Liquidity Risks; Interest Rate and Credit Risk.
Operational and Compliance Risks	Cybersecurity and Data Protection; Supply Chain and Operational Risks; Regulatory Compliance and Capital Requirements; Environmental and Regulatory Risks; Operational Efficiency and Cost Management.
Strategic and Business Risks	Competitive Landscape and Market Dynamics; Adaptation to Market Changes; Business Diversification; Acquisition Strategy and Valuation Risks; BrandPerception and Customer Engagement.
Product and Technological Risks	Technological Disruption Risks; Adaptation to Changing Consumer Behavior; Product Lifecycle and Patent Expiry; Research and Development Risks; Data Security and Consumer Confidence.

B. Preparation of Transcripts

Rather than apply the analysis template outlined in Table I to full transcripts, we instead opt to apply this template to the analysis of summaries of transcripts generated by ChatGPT-3.5. We do this for two reasons. First, since the cost of using GPTs is determined by the length of the input text, analysing summaries allows us to examine more firms in a cost-effective way as the summaries are, on average, one-tenth the length of the original texts. Second, this summarization

process helps standardize the text in earnings call transcripts. By condensing the content, we effectively streamline the material, diminishing variations in language style, tone, and colloquial expressions that can differ across speakers. This uniformity aids AI models by enhancing their ability to comprehend and analyze the essential information more effectively by reducing the need to interpret unique expressions or navigate through tangential statements common in natural speech. As highlighted in Beckmann et al. (2024), peculiar word choices can be indicative of economic conditions and consequently we may lose this dimension in our analysis. However, because our study is concerned with the logical decision-making processes of analysts, our interest lies in the content of what was discussed during the conferences rather than the manner in which it was expressed. Consistent with the format of earnings calls, our summary also consists of a presentation part and a discussion part. In the presentation section, remarks by different management team members are summarized as a few key points. In the discussion part, each question and all answers to that question are summarized as a question-answer pair.

C. Computing the Analyst Insight Score

Armed with the template outlined in Table I and prepared earnings call transcripts, we next instruct GPT4-Turbo to perform a full evaluation. We instruct GPT4-Turbo to identify content within the earnings call summaries that relate to each analyst sub-criterion (Table I) and assign a score based on its perceived impact on the company's outlook. The core element of our prompt is:

```
You are an financial analyst. Analyze the provided summary of an earnings
call transcript and score it based on the provided criteria and subcriteria.
```

We direct ChatGPT to produce scores with values ranges from -1 to +1 with the magnitude of the score representing the strength of the perceived impact on the firm. Hence, lower scores (e.g., -0.7) reflect a strong negative impact, scores close to +1 reflect a strong positive impact and scores close to 0 reflect no expected impact. The score for each of the four primary criteria is determined by calculating the average of its sub-criteria scores. Finally, the AI-generated score is derived from averaging these four individual scores. In addition to providing scores, we also instructed GPT4-Turbo to provide the rationale behind each score assignment with evidence supporting its judgments. The output produced by GPT4-Turbo for each firm includes the overall score (AIS),

the score for each of the four major criteria, the scores for each of the sub-criteria and finally, the rationale behind each score.

A sample AIS output of NVIDIA earnings call transcripts released in Feb 2023 is shown in Appendix A.A.2. The beginning of this output text is the AIS ("overall score"), and the score of four major criteria. Then ChatGPT lists the sub-criteria of scores, and rationales. For example, the partnership with Deutsche bank and launching of new products are cited as strength for the growth potential sub-criterion. If investors only pay attention to the last quarter's revenue and income changes, they may find these figures disappointing and inclined to divest from NVIDIA. However, our method picked up positive signals from the transcripts such as the popularity of H100 products and assigned a high AIS score for the company reflecting its positive outlook.

Examining the time-series properties of our AIS, Figure 1 plots the average AIS updated on a monthly basis for the largest 500 stocks in our test period with its 95% confidence bands. The average AIS score is consistently above 0 but drops significantly during periods of extreme market stress such as the COVID-19 pandemic at the start of 2020. The average AIS also clearly responds negatively to the tightening monetary policy in the second half of 2022.

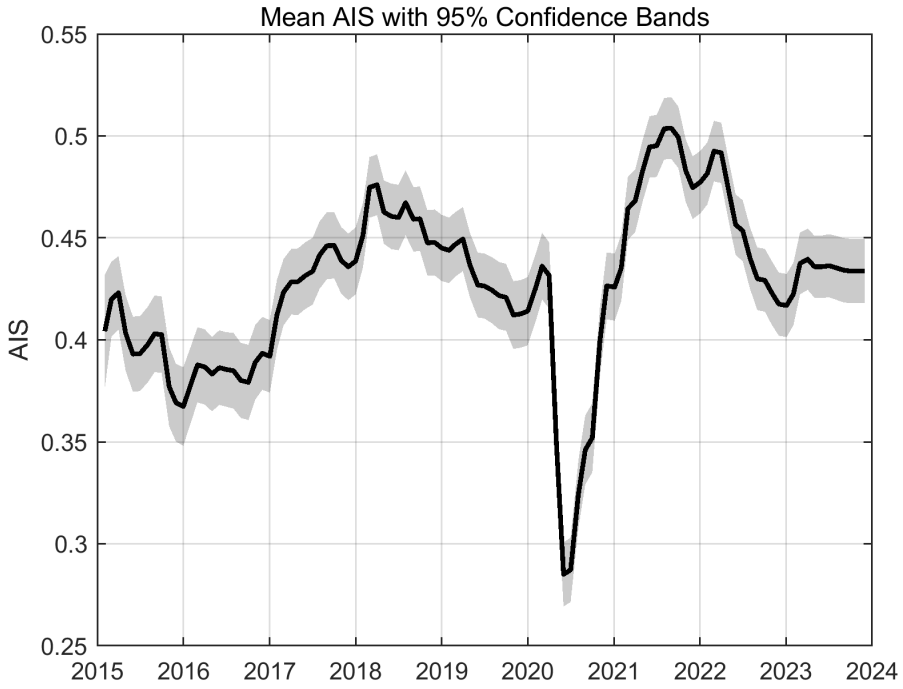
D. Computing the Naive Analyst Insight Score

To determine how effective ChaGPT is at extracting the meaning behind analyst questions and to determine the value of this insight drawn form the financial analyst expertise, we compute an alternative set of AIS results without this prior training on analyst questions. Specifically, we instruct ChatGPT to compute a score on the same scale as our original AIS (between -1 and +1) based solely on the prompt:

```
Analyze the provided summary of an earnings call transcript and
assign a score to reflect the impact of the news on the company's
core business segment.
```

and the prepared transcripts. To reiterate, our primary AIS uses the insight obtained from examining analysts questions in 2014 and the prepared transcripts whereas the naive AIS only uses the prepared transcripts and hence acts a baseline. Since both methods have the full training model of ChatGPT at their disposal, comparing the naive AIS with our primary AIS allows us to identify

Figure 1. Average analyst insight score (AIS) of a portfolio composed of the top 500 US stocks by market capitalization and its 95% confidence bands in our experiment period.



the impact that the analyst template, developed via training on professional analyst questions, has on the ability of ChatGPT to analyse the information provided by management during conference calls.

IV. Empirical Analysis

In this section we conduct an empirical analysis of our AI generated analyst insight scores. Our primary questions are:

1. Is the LLM generated AIS aligned with the analysis produced by a human analyst?
2. Does the AI generated AIS uncover information that explains future returns?

To address our first question, we compare our AIS to a variety of metrics produced by professional analysts. This includes price target revisions, EPS revisions and changes to recommendations. We use consensus data drawn from the I/B/E/S database over the same period for which we compute our AIS (2015-2023). If GPT is performing a similar type of analysis to that of professionals,

we expect that our AIS scores provide significant explanatory power for changes in the metrics produced by analysts. Namely, if AIS scores capture analyst insight, they should be able to predict changes in analyst metrics. We conduct this analysis in section IV.A. We address our second question by sorting stocks into quintile portfolios based on their AIS score. Again, if AIS scores capture information that is yet to be priced in by the market, we would expect that the portfolio composed of stocks with the highest AIS would outperform the portfolio composed of stocks with the lowest AIS. This analysis is perfomred in section IV.B.

Finally, given GPT has been trained using an vast dataset, there is a risk that any insight it displays could be due to the answer lying within its training data. For example, when asking GPT to analyse a firms earnings call transcript in 2020, it is possible that GPT is using data within its training set (which is up to April 2023), such as the future stock price or other commentary on the firm, to produce output that appears accurate but clearly suffers from look ahead bias. To address this issue, we develop a fictitious set of earnings call transcripts containing either extremely positive or negative sentiment. These fictitious transcripts, developed from real transcripts, maintain similar company-specific language to that found in actual transcripts. We analyse these fictitious transcripts to ascertain whether the AIS generated is a result of only that data contained within the transcript, or is being pulled from future information contained within the models training set. This analysis is conducted in section IV.C. To bolster this robustness test, We also conduct and analysis using only those earnings calls that occur after the end of GPT's training data (April 2023) and find our main conclusions remain unchanged.

A. AI-human Alignment Tests

The period following earnings calls provides a unique opportunity to test the alignment between a professional analyst and that produced by AI such as GPT. In general, it is reasonable to expect that professional analysts are privy to more information that is relevant to a firms outlook than is available to a LLM.[2] However, earnings calls represent a significant event during which professional analysts update their guidance to reflect the information released at these calls. As such, the information available to both human analysts and AI is closely aligned when these calls take place

[2]For example, professional analysts often have additional information via relationships with the firms they cover and other and other market participants. This would give these professional an advantage over LLMs which can only leverage their training data and input text.

and hence provide an ideal setting to examine the alignment between the analysis performed by professionals and AI.

We examine the behaviour of analysts following earnings calls for the 500 largest firms, measured by market capitalization, that are listed on the NYSE and NASDAQ exchanges and have a call transcript available in the Refinitiv StreetEvent database. The 500 stocks that are selected are updated at the end of each year. Our data covers the period from the start of 2015 to the third quarter of 2023. We use several indicators to track the behaviour of analysts including changes in consensus price targets, forward-year earnings per share, and buy-hold-sell recommendations. We compute percentage changes for each of these variables from the day following the earnings call $(t+1)$ to 30 and 90-days post the earnings call ($t+30$ and $t+90$ respectively).

To analyse the ability of our AIS to explain analyst behaviour, we examine both the level of and changes in the AIS to explain revisions in metric produced by professional analysts. Specifically, we estimate the following regressions:

$$y_{i,t+1,t+h} = \alpha + \beta_{AIS} AIS_{i,t} + \beta_{PB} \log(PB_{i,t}) + \beta_{MC} \log(MktCap_{i,t}) + \phi_i + \gamma_t + \epsilon_{i,t} \quad (1)$$

$$y_{i,t+1,t+h} = \alpha + \beta_{AIS} \Delta AIS_{i,t} + \beta_{PB} \log(PB_{i,t}) + \beta_{MC} \log(MktCap_{i,t}) + \phi_i + \gamma_t + \epsilon_{i,t} \quad (2)$$

where $y_{i,t+1,t+h}$ represents the percentage change in the average of one of our three analyst metrics; 1. forward EPS ($EPS_{t+1,t+h}$), 2. price target ($PT_{t+1,t+h}$) and 3. recommendation ($REC_{t+1,t+h}$)) over the horizon $h = \{30, 90\}$-days. $AIS_{i,t}$ is the GPT generated analyst insight score for firm i at date t and $\Delta AIS_{i,t}$ is the change in this AIS score from the previous earnings call (one quarter ago). We use the average across analysts as our measure of consensus figures. To compute a numerical value for recommendations, we assign each of the possible recommendations a numerical value: strong buy = 1, buy = 2, hold = 3, sell = 4 and strong sell = 5. This means that more positively viewed firms will receive lower recommendation scores. We also control for each firms size and price-to-book ratio via the logarithm of these variables to control for the effect that these variables have on potential analyst revisions. To mitigate the impact of outliers, our analyst metrics are winsorized at the 5% level. We include both firm (ϕ_i) and year (γ_t) fixed effects and cluster standard errors at the firm level. Regressions that use the AIS directly (1) provide insight regarding how variation in the AIS is related to changes in analyst behaviour whereas those that use changes

in the AIS (2) give insight regarding how improvement (or deterioration) in the AIS can impact analyst revisions. Results from these regressions are presented in Table II.

[Insert Table II here]

We observe that both AIS and ΔAIS are highly significant predictors of analysts revisions across all metrics analysed with all coefficient signs matching expectations. Increasing AIS or ΔAIS increases forward EPS revisions and price targets while also lowering the recommendation score, indicating that the firm is more highly recommended. Focusing on the 30-day forward EPS, we observe that analysts will, on average, increase this forward EPS 0.3% when a firms AIS increases by 0.1 points while a change in the AIS of 0.1 points from the prior earnings call leads to the EPS being 0.24% higher on average. Shifting focus to 30-day price targets, we similarly find that a 0.1 point increase in the AIS yields an approximately 0.2% upward revision and that increasing the AIS 0.1 points from the prior call leads to the same increase. Last, Recommendations improve 0.13% for a 0.1 point increase in the AIS and 0.11% for a 0.1 increase in the AIS change. Collectively, these results suggests that firms with increasing AIS's receive positive revisions from analysts, indicating that the AIS aligns with how analysts evaluate a firms future prospects.

We next examine the ability of two alternative scores to predict analyst revisions, the naive AIS and the standardized unexpected earnings (SUE) metric. The SUE measures the number of standard deviations above or below the consensus earnings estimate and hence acts a signal of unexpected changes in the firms outlook. This analysis hence ascertains if the insight provided by our AIS provides any advantage over a simpler method (the naive AIS) or a currently used alternative (the SUE) with respect to predicting analyst revisions. We estimate the same regressions as in equation (1) and (2) but replace AIS (ΔAIS) with the naive AIS, AIS_N (ΔAIS_N) or the SUE (ΔSUE). We present results from these two regressions in Tables III and IV respectively.

[Insert Table III here]

[Insert Table IV here]

We observe that the naive AIS exhibits strong statistical significance for predicting analysts revisions across our three metrics. Though not as consistent as our original AIS, we find that even

a rudimentary application of GPT to the analysis of earnings transcripts appears to nonetheless have the ability to mimic the analysis performed by professional analysts. Somewhat surprisingly, we find that the SUE is has a relatively weak ability to predict future analyst revisions. Coefficients on SUE and ΔSUE are either statistically insignificant or so small as to be economically insignificant. This result illustrates that LLMs can provide a significant analytic advantage relative to traditional measures of earnings surprise and hence suggests that the revisions analysts make are based on a wider range of factors than simply a surprise in earnings.

We next examine whether the additional information obtained from the template used to generate our AIS provides additional predictive power beyond that found in the naive AIS. To identify this contribution, we regress our AIS (and changes in AIS) against the naive AIS (and its changes):

$$AIS_{i,t} = \alpha + \beta AIS_{N,i,t} + \epsilon_{i,t} \tag{3}$$

$$\Delta AIS_{i,t} = \alpha + \beta \Delta AIS_{N,i,t} + \epsilon_{i,t} \tag{4}$$

The residuals from these regressions represent the information contained in AIS that is orthogonal to, and absent from our naive AIS and hence represents the information content provided by the use of the analyst template that was learned by GPT using the Q&A section of earnings calls in 2014. We then use these residuals in regressions (1) and (2) but with the AIS and ΔAIS replaced with the corresponding residuals from regressions (3) and (4) respectively. Should these residuals provide statistically significant explanatory power for future analyst revisions, this would imply that the information contained in AIS, that is not present in the naive AIS, is relevant for modelling how analysts consider a firms prospects. The results from this regression are presented in Table V.

[Insert Table V here]

We find that in ten of the twelve regressions estimated that the residuals from equations (3) and (4) provide statistically and economically significant explanatory power for future analyst revisions. This finding suggests that our AIS contains information relevant to the decision making process exhibited by professional analysts that is not contained in the naive AIS, further emphasising that the template identified by GPT from the Q&A section of earnings calls is an accurate representation of the decision making process used by analysts.

B. Stock Selection Tests

In this section we examine the whether our AIS has the ability to predict returns out of sample. If GPT is truly uncovering information relevant to price formation from earnings calls, then this information should be useful for developing investment strategies that outperform traditional buy and hold investments. Cohen et al. (2020) show that prices can be slow to respond to information, especially if that information is obscured by complexity. Given earnings calls are a complex environment involving management presentation and a dialogue between management and professional investment analysts, it is reasonable to expect that prices may not immediately reflect the information released during these calls. Hence, quick and accurate analysis of earnings calls via GPT may provide an investment strategy for earning abnormal returns.

Our investment strategy involves sorting stocks into quintile portfolios based on their computed AIS or ΔAIS. Specifically, we form portfolios each month based on the most recent AIS for each stock as of the end of every calendar month. Adjustments to the portfolio are made at the closing market price on the first trading day of the following month. Tables VI and VII provide the excess returns earned on these quintile portfolios when sorting based on AIS and ΔAIS respectively with panel A providing the outcome for equally weighted portfolios and panel B for market capitalization weighted portfolios. We also provide the alphas of these portfolios when controlling for the excess returns on the market portfolio (CAPM), the three factors of Fama and French (1992) (three-factor alpha), the five factors of Fama and French (2015) (five-factor alpha) and an additional momentum factor developed by Carhart (1997) (five-factor + mom alpha).

[Insert Table VI here]

[Insert Table VII here]

Examining Table VI, we find mixed performance when using AIS to build portfolios. Focusing on the a self financed portfolio that is short stocks with the lowest AIS and long those with the largest, the final column of Table VI, we find that the market cap weighted portfolio doesn't produce a statistically significant excess return. When controlling for the systematic risk factors outlined earlier, we obtain weakly significant positive alphas in 3 of the 5 cases studies. Using equal weights produces superior returns when controlling for the market and Fama-French size

and value factors, but our approach doesn't maintain statistical significance when controlling for investment, operating profitability and momentum factors. However, our results shift dramatically when considering portfolios constructed based on the change in their AIS from the previous earnings call. If the AIS is consistently estimated, then a change in the AIS represents a genuine improvement (or deterioration) in a firms outlook and hence may represent a superior approach to predicting future price movements. As illustrated in Table VII, our portfolio sort produces economically and statistically significant excess returns even when controlling for all systematic risk factors suggested in the CAPM, the Fama-French three and five factor models as well as momentum. The excess return on an equally weighted investment strategy is 0.47% per month (5.64% p.a.) and 0.5% per month (6% p.a.) for a market cap weighted investment strategy.

Without risk adjustment, it appears that using the change in the AIS is identifying winning stocks as the excess return is earned on the long leg of the strategy. However, after adjusting for systematic risk, we find that the change in AIS is instead identifying the poorly performing stocks with short leg having a negative and statistically significant alpha. This is consistent with the findings of Cohen et al. (2020) who finds that buying firms with non-changing language in their 10-K filings while shorting those that make significant changes to their 10-K language produces significant excess risk adjusted returns. In our case, changes in a firms AIS seem to be mostly due to poor performance which management likely try to hide within verbose language.

We again wish to examine the gains available to an investor provided by the learned analysts template that underlies our AIS. We hence replicate out investment strategies using our naive AIS (and its changes) as well as the SUE (and its changes) to build the same quintile sorted portfolios. Results from these investment strategies are presented in tables VIII, IX, X and XI respectively. We find that the naive AIS does not result in portfolios that produce statistically significant excess returns. However, when using the change in the naive AIS, our strategy earns abnormal returns though with lower statistical power than was achieved using our template driven AIS. This suggests that even a rudimentary application of GPT to the analysis of a firms earnings call can extract relevant price information. In particular, it is the change in this score from one earnings call to the next that best captures the changing expectations for the firm, allowing us to identify future price changes that may be exploited via an appropriate investment strategy. Moving focus to the SUE as our sorting criteria, we find that a long-short investment based on quintile portfolios sorted

according to each firms SUE produces statistically significant excess returns across all models, with excess returns of 0.3% per month (3.6% p.a.) for the equally weighted investment and 0.46% per month (5.52% p.a.) for the market cap weighted investment. Statistical performance is much weaker when using the change in the SUE. This suggests that while SUE is not strongly connected to future analyst behaviour, it is nonetheless useful for identifying under/overvalued stocks.

[Insert Table VIII here]

[Insert Table IX here]

[Insert Table X here]

[Insert Table XI here]

Finally, we explore an investment strategy that incorporates multiple sources of information. Specifically, we compute a new sorting criteria which is computed via the average of our best performing metrics, AIS, ΔAIS and SUE. Results for these portfolios are presented in Table XII. We find that portfolios constructed by combining these three sources of information yield the best performance among those studied with statistically significant excess returns of 0.62% per month (7.44% p.a.) for the equally weighted portfolio and 0.7% per month (8.4% p.a.) for the market cap weighted strategy. Additionally, all risk adjusted alphas are both statistically and economically significant. This result suggests that AIS and SUE may capture distinct aspects of a firms future performance rather than being different representations of the same information.

[Insert Table XII here]

C. Fictitious Transcript Tests

A potential issue when studying historical data with LLMs is that there is a potential for look ahead bias. Because the training data used to develop GPT is so vast, there is the possibility that GPT is using future information about a firm to generate its AIS of that firm. For example, access to U.S. stock returns in its training dataset could lead GPT to favor stocks with historically high returns by assigning them elevated AIS scores in our experiments. Although the GPT API may

not directly divulge stock returns over a specified period, its limitations on disclosing proprietary information do not preclude the existence of underlying biases in its output. Moreover, biases could emerge from imbalanced training datasets; for example, stocks like AAPL, which attract significant public attention, may be over-represented in GPT's training data, leading GPT to assign them disproportionately high scores due to familiarity.

To investigate the potential for stock-based bias, we synthesized a set of fictitious transcripts, attributing AIS variations to either content sentiment or inherent machine bias. This approach entails selecting 1,000 summaries from those discussed in Section IV.A and generating two fictitious versions for each: one positively modified (where negative news is rephrased as positive, e.g., "revenue declined by 10 percent" altered to "revenue increased by 10 percent") and one with negative alterations (positive news converted to negative). Some samples of fictitious positive and negative transcripts are in Appendix A.B and Appendix A.C, respectively.

We next calculate the AIS for these modified transcripts with results provided in Table XIII. We find that transcripts constructed to contain artificially negative news yielded reduced AIS, moving from 0.48 to -0.78, while those with artificially positive news increased from 0.48 to 0.71. This result suggests that GPT is not using direct information about future events, but is producing a score based on the text provided in the prompt/transcript. If GPT was able to use its future training set (relative to the date of the transcript) we would expect that fictitious information would be ignored. Instead, we find that this fictitious information has a real impact on the AIS produced, suggesting the transcript is the source of information that GPT is basing its analysis on. Kim et al. (2024) express similar concerns in their study but find, similar to our work here, that GPT bases its analysis on the prompt provided and the information it is directed towards.

[Insert Table XIII here]

As a final test of the robustness of our approach to look ahead bias, we regress our AIS scores produced from fictitious transcripts against three variables. The first variable represents the editing direction, encoded as 1 for positive edits and -1 for negative edits, serving as a proxy for the sentiment conveyed by the transcripts. This coding strategy is directed by the method used to create the counterfactual transcripts. The second variable is the average AIS from 2015 to 2023, which serves as an indicator of potential bias by ChatGPT towards specific stocks. A preference

by ChatGPT for certain stocks, regardless of the textual context, would be suggested by a high value of this variable, with the expectation that its coefficient would be positive and statistically significant if AIS scores are influenced by such a preference. The third variable is the AIS of the original transcript, included to assess the direct impact of the initial content's sentiment on the assigned AIS.

The regression results, presented in Table XIV, reveal interesting dynamics. When incorporating both positive and negative edits, the Revision_label, indicative of text sentiment, emerges as a predominant factor influencing AIS scores. Surprisingly, the coefficient for Average AIS is statistically significant but negative. This counter intuitive finding is clarified when positive and negative edits are analyzed separately in the subsequent columns of the table. The analysis shows that Average AIS does not significantly influence the AIS of positively edited samples, but it negatively impacts the AIS of negatively edited samples. This suggests that transcripts with a higher original AIS, typically laden with positive content, are more adversely affected when negated by ChatGPT. These combined findings suggest that the AIS scores of counterfactual transcripts are primarily shaped by the content's sentiment rather than an inherent bias within ChatGPT towards certain stocks.

[Insert Table XIV here]

V. Conclusion

Large language models such as ChatGPT have captured the attention of the world for their ability to interact with written language and respond to questions with human like levels intelligence. This raises the question, can LLMs like ChatGPT replicate a skill set that is considered to require significant professional experience? This paper addresses his question in the context of equity analysts. We ask ChatGPT to analyse the dialogue in earnings calls for a selection of 500 of the largest firms from 2015-2023. We produce two forms of this analysis in the form of an analyst insight score (AIS). The first uses simple prompt (the naive AIS) while the second instructs ChatGPT to analyse and extract the underlying principles of analyst inquiry by studying the question and answer section of a set of earnings calls from 2014. Instructing ChatGPT to utilise these prin-

ciples in the form of an analyst template produces our complete AIS. An overview of the template produced by ChaGPT suggests it is able to uncover the key underlying principles used by analysts by examining the questions they pose to management during earnings calls.

We formally test the performance of our AIS by exploring its ability to predict future changes in several analyst metrics. We find that the AIS is a strong predictor of revisions to forward EPS guidance, price targets and buy-hold-sell recommendations. Additionally, while our naive AIS is also found to have predictive power for analyst revisions, we show that the component of AIS not contained in the naive AIS also provides predictive power. This finding suggests that the analyst template learned by ChatGPT from studying analyst questions provides additional predictive power and helps ChatGPT to better understand how analysts make decisions.

We next explore whether the information uncovered by ChatGPT from these earnings calls could be used to generate abnormal returns. By sorting stocks into portfolios based on their AIS and changes in the AIS, we find that statistically significant abnormal returns of 50bps per month (6% p.a.) can be earned through a self financed strategy that goes short in stocks with large negative changes in the AIS and a long position in stocks that have large positive changes in the AIS. This suggests that the AIS represents an accurate measure of real time firm value and that changes in the AIS are followed by changes in stock prices. These results hold when controlling for systematic risk factors common in the literature. Building the same portfolios with the naive AIS yields weaker performance, further highlighting that the analyst decision making process uncovered by ChatGPT provides a genuine source of value. Last, we compare the performance of our ChatGPT generated AIS portfolios to portfolios constructed through sorts based on other standard valuation metrics such as the standardized unexpected earnings (SUE). We find that the portfolios constructed using the AIS outperform those built using the SUE, but also that even superior performance (70bps per month - 8.4% p.a.) can be achieved when combining AIS and SUE. This suggests that AIS and SUE capture different elements of the valuation process and act as complementary variables for return prediction.

To summarize, we find that ChatGPT is able to largely replicate the analytic tasks that is currently the domain of highly trained analysts and can provide significant economic value by identifying over and under valued stocks from the information provided during earnings conference calls. However, the results in this paper also show that in order to achieve this alignment, it is

necessary to let ChatGPT learn from professional analysts. This result is consistent with related work such as Kim et al. (2024) who find that ability of ChatGPT to analyse financial statements is dramatically improved when providing additional guidance via chain-of-thought prompting. This work hence shows that while generative AI will likely have a significant impact upon financial markets and its participants, the input of trained professionals is critical to its success.

REFERENCES

Baker, Scott R., Nicholas Bloom, and Steven J. Davis, 2016, Measuring economic policy uncertainty, *The Quarterly Journal of Economics* 131, 1593–1636.

Barker, Richard, and Shahed Imam, 2008, Analysts' perceptions of 'earnings quality', *Accounting and Business Research* 38, 313–329.

Beckmann, Lars, Heiner Beckmeyer, Ilias Filippou, Stefan Menze, and Guofu Zhou, 2024, Unusual financial communication-evidence from chatgpt, earnings calls, and the stock market, *Earnings Calls, and the Stock Market (January 15, 2024)* .

Bian, Shibo, Yao Chen, and Xunxiao Wang, 2022, Manager-analyst interaction quality in earnings conference calls, *SSRN Electronic Journal* .

Block, Stanley B, 1999, A study of financial analysts: Practice and theory, *Financial Analysts Journal* 55, 86–95.

Bouwman, Marinus J, 1984, Expert vs novice decision making in accounting: A summary, *Accounting, Organizations and Society* 9, 325–327.

Bouwman, Marinus J, Patricia A Frishkoff, and Paul Frishkoff, 1987, How do financial analysts make decisions? a process model of the investment screening decision, *Accounting, Organizations and Society* 12, 1–29.

Brown, Lawrence D, Andrew C Call, Michael B Clement, and Nathan Y Sharp, 2015, Inside the "black box" of sell-side financial analysts, *Journal of Accounting Research* 53, 1–47.

Campbell, John L., Hsinchun Chen, Dan S. Dhaliwal, Hsin-min Lu, and Logan B. Steele, 2014, The information content of mandatory risk factor disclosures in corporate filings, *Review of accounting studies* 19, 396–455.

Carhart, Mark M., 1997, On persistence in mutual fund performance, *The Journal of Finance* 52, 57–82.

Clement, Michael, Richard Frankel, and Jeffrey Miller, 2003, Confirming management earnings forecasts, earnings uncertainty, and stock returns, *Journal of Accounting Research* 41, 653–679.

Cohen, Lauren, Christopher Malloy, and Quoc Nguyen, 2020, Lazy prices, *Journal of Finance* 75, 1371–1415.

Fama, Eugene F., and Kenneth R. French, 1992, The cross-section of expected stock returns, *The Journal of Finance* 47, 427–465.

Fama, Eugene F., and Kenneth R. French, 2015, A five-factor asset pricing model, *Journal of Financial Economics* 116, 1–22.

Fatouros, Georgios, John Soldatos, Kalliopi Kouroumali, Georgios Makridis, and Dimosthenis Kyriazis, 2023, Transforming sentiment analysis in the financial domain with chatgpt, *Machine Learning with Applications* 14, 100508.

Feldman, Ronen, Suresh Govindaraj, Joshua Livnat, and Benjamin Segal, 2010, Management's tone change, post earnings announcement drift and accruals, *Review of Accounting Studies* 15, 1–22.

Filzen, Joshua J., 2015, The information content of risk factor disclosures in quarterly reports, *Accounting Horizons* 29, 887–916, Risk disclosure is useful.

Hansen, Anne Lundgaard, and Sophia Kazinnik, 2023, Can chatgpt decipher fedspeak?, *SSRN Electronic Journal* .

Jegadeesh, Narasimhan, and Di Wu, 2013, Word power: A new approach for content analysis, *Journal of Financial Economics* 110, 712–729, 10k sentiment affects return, weighting is important.

Jha, Manish, Jialin Qian, Michael Weber, and Baozhong Yang, 2023, Chatgpt and corporate policies, *Chicago Booth Research Paper* .

Katz, Daniel Martin, Michael James Bommarito, Shang Gao, and Pablo Arredondo, 2024, Gpt-4 passes the bar exam, *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* 382, 20230254.

Ke, Zheng Tracy, Bryan T. Kelly, and Dacheng Xiu, 2020, Predicting returns with text data learned weights better predicts asset return.

Kim, Alex, Maximilian Muhn, and Valeri Nikolaev, 2024, Financial statement analysis with large language models, *arXiv preprint arXiv:2407.17866* .

Kim, Alex G, Maximilian Muhn, and Valeri V Nikolaev, 2023, Bloated disclosures: Can chatgpt help investors process information?, *Chicago Booth Research Paper* .

Kung, Tiffany H., Morgan Cheatham, ChatGPT, Arielle Medenilla, Czarina Sillos, Lorie De Leon, Camille Elepaño, Maria Madriaga, Rimel Aggabao, Giezel Diaz-Candido, James Maningo, and Victor Tseng, 2022, Performance of chatgpt on usmle: Potential for ai-assisted medical education using large language models, *medRxiv* .

Li, Edward Xuejun, Zhiyuan Tu, and Dexin Zhou, 2023, The promise and peril of generative ai: Evidence from chatgpt as sell-side analysts, *Available at SSRN 4480947* .

Lopez-Lira, Alejandro, and Yuehua Tang, 2023, Can chatgpt forecast stock price movements? return predictability and large language models, *SSRN Electronic Journal* .

Loughran, Tim, and Bill McDonald, 2011, When is a liability not a liability? textual analysis, dictionaries, and 10-ks, *The Journal of Finance* 66, 35–65.

Matsumoto, Dawn, Maarten Pronk, and Erik Roelofsen, 2011, What makes conference calls useful? the information content of managers' presentations and analysts' discussion sessions, *The Accounting Review* 86, 1383–1414.

Previts, Gary John, Robert J Bricker, Thomas R Robinson, and Stephen J Young, 1994, A content analysis of sell-side financial analyst company reports, *Accounting Horizons* 8, 55.

Price, S. McKay, James S. Doran, et al., 2012, Earnings conference calls and stock returns: The incremental informativeness of textual tone, *Journal of Banking and Finance* 36, 992–1011.

Suslava, Kate, 2021, "stiff business headwinds and uncharted economic waters": The use of euphemisms in earnings conference calls, *Management Science* 67, 7184–7213.

Terwiesch, Christian, 2023, Would chat gpt3 get a wharton mba? a prediction based on its performance in the operations management course, Mack Institute for Innovation Management at the Wharton School, University of Pennsylvania.

Tetlock, Paul C., 2007, Giving content to investor sentiment: The role of media in the stock market, *The Journal of Finance* 62, 1139–1168.

Tetlock, Paul C, 2011, All the news that's fit to reprint: Do investors react to stale information?, *Review of Financial Studies* 24, 1481–1512.

Tetlock, Paul C., Maytal Saar-Tsechansky, et al., 2008, More than words: Quantifying language to measure firms' fundamentals, *The Journal of Finance* 63, 1437–1467.

Tables

Table II AIS and Analyst Revisions

This table presents the results from eatimating regressions (1) and (2). The dependent variable in each regression is the percentage changes in one of three analyst metrics from $t+1$ to $t+h$ where $h = \{30, 90\}$-days. These metrics are forward EPS (EPS_{t+h}), price targets (PT_{t+h}) and recommendations (REC_{t+h}). Standard errors for each coefficient estimate is provided in parentheses. Dependent variables are winsorized at the 5% level and we compute clustered standard errors with clustering occurring at the firm level. All regressions include both firm and year-fixed effects. Statistical significance at the 10, 5 and 1% level are indicated with *, ** and *** respectively.

	EPS_{t+30}	EPS_{t+30}	EPS_{t+90}	EPS_{t+90}	PT_{t+30}	PT_{t+30}	PT_{t+90}	PT_{t+90}	REC_{t+30}	REC_{t+30}	REC_{t+90}	REC_{t+90}
Intercept	0.028^{***}	0.014	0.537^{***}	0.454^{***}	-0.034^{***}	-0.040^{***}	0.086^{***}	0.078^{***}	0.003	0.004	-0.019	-0.020
	(0.008)	(0.009)	(0.055)	(0.055)	(0.004)	(0.005)	(0.021)	(0.021)	(0.007)	(0.007)	(0.015)	(0.015)
AIS	0.030^{***}		0.079^{***}		0.019^{***}		0.047^{***}		-0.013^{***}		-0.024^{***}	
	(0.002)		(0.013)		(0.001)		(0.005)		(0.002)		(0.004)	
ΔAIS		0.024^{***}		0.077^{***}		0.019^{***}		0.055^{***}		-0.011^{***}		-0.027^{***}
		(0.002)		(0.012)		(0.001)		(0.004)		(0.002)		(0.004)
$log(PB)$	0.002^{***}	0.003^{***}	0.009	0.010^{**}	0.004^{***}	0.004^{***}	0.007^{***}	0.008^{***}	-0.001^{**}	-0.002^{**}	-0.005^{***}	-0.005^{***}
	(0.001)	(0.001	(0.005	0.005	0.000	0.001	0.002	0.002	0.001	0.001	0.001	0.002
$log(MC)$	-0.002^{*}	0.001	-0.053^{***}	-0.046^{***}	0.003^{***}	0.006^{***}	-0.019^{***}	-0.016^{***}	0.001	0.000	0.004	0.002
	(0.001)	(0.001)	(0.009)	(0.009)	(0.001)	(0.001)	(0.003)	(0.004)	(0.001)	(0.001)	(0.002)	(0.002)
Firm Fixed	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Year Fixed	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
n	14970	14273	14970	14273	14970	14273	14970	14273	7991	7617	7548	7176
R^2_{adj}	0.121	0.116	0.160	0.166	0.133	0.143	0.156	0.165	0.028	0.037	0.073	0.082

Table III Naive AIS and Analyst Revisions

This table presents the results from estimating regressions specified by (1) and (2) but with the AIS replaced with the naive AIS which is produced without the benefit of the analyst template learned from analyst Q&A sessions. The dependent variable in each regression is the percentage changes in one of three analyst metrics from $t+1$ to $t+h$ where $h = \{30, 90\}$-days. These metrics are forward EPS (EPS_{t+h}), price targets (PT_{t+h}) and recommendations (REC_{t+h}). Standard errors for each coefficient estimate is provided in parentheses. Dependent variables are winsorized at the 5% level and we compute clustered standard errors with clustering occurring at the firm level. All regressions include both firm and year-fixed effects. Statistical significance at the 10, 5 and 1% level are indicated with *, ** and *** respectively.

	EPS_{t+30}	EPS_{t+30}	EPS_{t+90}	EPS_{t+90}	PT_{t+30}	PT_{t+30}	PT_{t+90}	PT_{t+90}	REC_{t+30}	REC_{t+30}	REC_{t+90}	REC_{t+90}
Intercept	0.529^{***}	0.014	0.026^{***}	0.454^{***}	-0.035^{***}	-0.039^{***}	0.082^{***}	0.079^{***}	0.004	0.004	-0.017	-0.020
	(0.055)	(0.009)	(0.008)	(0.055)	(0.004)	(0.005)	(0.020)	(0.021)	(0.007)	(0.007)	(0.015)	(0.015)
AIS_N	0.041^{***}		0.019^{***}		0.011^{***}		0.027^{***}		-0.007^{***}		-0.013^{***}	
	(0.007)		(0.001)		(0.001)		(0.003)		(0.001)		(0.002)	
ΔAIS_N		0.015^{***}		0.040^{***}		0.011^{***}		0.031^{***}		-0.007^{***}		-0.015
		(0.001)		(0.007)		(0.001)		(0.002)		(0.001)		(0.002)
$log(PB)$	0.009^{*}	0.003^{***}	0.002^{***}	0.010^{**}	0.004^{***}	0.004^{***}	0.007^{***}	0.008^{***}	-0.001^{*}	-0.002^{**}	-0.005^{***}	-0.006
	(0.005)	(0.001)	(0.001)	(0.005)	(0.000)	(0.001)	(0.002)	(0.002)	(0.001)	(0.001)	(0.001)	(0.002)
$log(MC)$	-0.051^{***}	0.001	-0.002	-0.046^{***}	0.004^{***}	0.005^{***}	-0.018^{***}	-0.016^{***}	0.001	0.000	0.003	0.002
	(0.009)	(0.001)	(0.001)	(0.009)	(0.001)	(0.001)	(0.003)	(0.004)	(0.001)	(0.001)	(0.002)	(0.002)
Firm Fixed	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Year Fixed	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
n	14970	14273	14970	14273	14970	14273	14970	14273	7991	7617	7548	7176
R^2_{adj}	0.160	0.120	0.124	0.166	0.132	0.142	0.155	0.165	0.028	0.038	0.073	0.081

Table IV SUE and Analyst Revisions

This table presents the results from eatimating regressions (1) and (2) but with the AIS replaced with the standardised unexpected earnings (SUE) score. The dependent variable in each regression is the percentage changes in one of three analyst metrics from $t+1$ to $t+h$ where $h = \{30, 90\}$-days. These metrics are forward EPS (EPS_{t+h}), price targets (PT_{t+h}) and recommendations (REC_{t+h}). Standard errors for each coefficient estimate is provided in parentheses. Dependent variables are winsorized at the 5% level and we compute clustered standard errors with clustering occurring at the firm level. All regressions include both firm and year-fixed effects. Statistical significance at the 10, 5 and 1% level are indicated with *, ** and *** respectively.

	EPS_{t+30}	EPS_{t+30}	EPS_{t+90}	EPS_{t+90}	PT_{t+30}	PT_{t+30}	PT_{t+90}	PT_{t+90}	REC_{t+30}	REC_{t+30}	REC_{t+90}	REC_{t+90}
Intercept	0.015	0.001	0.495^{***}	0.408^{***}	-0.035^{***}	-0.038^{***}	0.087^{***}	0.087^{***}	0.005	-0.003	-0.008	-0.021
	(0.010)	(0.010)	(0.069)	(0.059)	(0.005)	(0.005)	(0.019)	(0.019)	(0.008)	(0.007)	(0.015)	(0.015)
SUE	0.000		0.000		0.000^{***}		0.000^{***}		0.000		0.000	
	(0.000)		(0.002)		(0.000)		(0.000)		(0.000)		(0.000)	
ΔSUE		0.000		-0.001		0.000		0.000^{***}		0.000		0.000
		(0.000)		(0.001)		(0.000)		(0.000)		(0.000)		(0.000)
$log(PB)$	0.004^{***}	0.003^{***}	0.011^{**}	0.008	0.004^{***}	0.004^{***}	0.008^{***}	0.007^{***}	-0.002^{**}	-0.002^{**}	-0.006^{***}	-0.006^{***}
	(0.001)	(0.001)	(0.006)	(0.005)	(0.001)	(0.001)	(0.002)	(0.002)	(0.001)	(0.001)	(0.002)	(0.002)
$log(MC)$	0.002	0.003^{*}	-0.039^{***}	-0.034^{***}	0.005^{***}	0.006^{***}	-0.014^{***}	-0.014^{***}	0.000	0.000	0.000	0.001
	(0.002)	(0.002)	(0.010)	(0.010)	(0.001)	(0.001)	(0.003)	(0.003)	(0.001)	(0.001)	(0.002)	(0.002)
Firm Fixed	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Year Fixed	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
n	14063	12764	14063	12764	14063	12764	14063	12764	7480	6763	7061	6395
R^2_{adj}	0.096	0.098	0.157	0.166	0.104	0.106	0.137	0.136	0.026	0.035	0.067	0.076

Table V AIS/naive AIS Residuals and Analyst Revisions

This table presents the results from estimating regressions (1) and (2) but with the AIS replaced with residuals obtained from regressions (3) and (4). The dependent variable in each regression is the percentage changes in one of three analyst metrics from $t+1$ to $t+h$ where $h = \{30, 90\}$-days. These metrics are forward EPS (EPS_{t+h}), price targets (PT_{t+h}) and recommendations (REC_{t+h}). Standard errors for each coefficient estimate is provided in parentheses. Dependent variables are winsorized at the 5% level and we compute clustered standard errors with clustering occurring at the firm level. All regressions include both firm and year-fixed effects. Statistical significance at the 10, 5 and 1% level are indicated with *, ** and *** respectively.

	EPS_{t+30}	EPS_{t+30}	EPS_{t+90}	EPS_{t+90}	PT_{t+30}	PT_{t+30}	PT_{t+90}	PT_{t+90}	REC_{t+30}	REC_{t+30}	REC_{t+90}	REC_{t+90}
Intercept	0.017^{***}	0.011	0.510^{***}	0.445^{***}	-0.041^{***}	-0.042^{***}	0.069^{***}	0.072^{***}	0.007	0.005	-0.011	-0.018
	(0.009)	(0.009)	(0.054)	(0.055)	(0.005)	(0.005)	(0.021)	(0.021)	(0.007)	(0.007)	(0.015)	(0.015)
$Resid$	-0.024^{***}		-0.019		-0.010^{***}		-0.023^{***}		0.009^{***}		0.011^{***}	
	(0.003)		(0.020)		(0.002)		(0.008)		(0.003)		(0.006)	
$\Delta Resid$		0.009^{***}		0.058^{***}		0.011^{***}		0.031^{***}		-0.004		-0.017^{***}
		(0.002)		(0.019)		(0.002)		(0.007)		(0.003)		(0.006)
$log(PB)$	0.003^{***}	0.004^{***}	0.012^{**}	0.010^{**}	0.004^{***}	0.004^{***}	0.009^{***}	0.008^{***}	-0.002^{***}	-0.002^{**}	-0.006^{***}	-0.006^{***}
	(0.001)	(0.001)	(0.005)	(0.005)	(0.001)	(0.001)	(0.002)	(0.002)	(0.001)	(0.001)	(0.001)	(0.002)
$log(MC)$	0.001^{***}	0.001	-0.045^{***}	-0.044^{***}	0.005^{***}	0.006^{***}	-0.014^{***}	-0.015^{***}	0.000	0.000	0.001	0.002
	(0.001)	(0.001)	(0.009)	(0.009)	(0.001)	(0.001)	(0.003)	(0.004)	(0.001)	(0.001)	(0.002)	(0.002)
Firm Fixed	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Year Fixed	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
n	14970	14273	14970	14273	14970	14273	14970	14273	7991	7617	7548	7176
R^2_{adj}	0.101	0.097	0.158	0.164	0.112	0.116	0.150	0.156	0.023	0.032	0.068	0.075

Table VI Cross-Sectional Analysis of AIS Factor

This table presents returns of monthly portfolios sorted by our AIS factor, adjusted by a factor of 100. Panel A details equal-weighted returns, while Panel B covers value-weighted returns, with t-Statistics provided beneath each estimate. Excess returns are defined as the average monthly returns minus the risk-free rate for the same period. The Fama-French three-factor model includes market, size, and value factors, expanded to five in the five-factor model by adding profitability and investment. Significance levels of 1%, 5%, and 10% are denoted by ***, **, and *, respectively.

Panel A: Equally Weighted

	Q1	Q2	Q3	Q4	Q5	Q5-Q1
Excess Return	0.67	0.76	0.80*	0.76*	1.12**	0.44
	(1.23)	(1.56)	(1.75)	(1.72)	(2.38)	(1.53)
CAPM Alpha	-0.33	-0.17	-0.09	-0.11	0.21	0.52*
	(-1.49)	(-1.17)	(-0.73)	(-1.04)	(1.57)	(1.78)
Three-factor Alpha	-0.23	-0.11	-0.05	-0.11	0.19	0.40*
	(-1.30)	(-0.95)	(-0.44)	(-1.03)	(1.49)	(1.75)
Five-factor Alpha	-0.24	-0.12	-0.06	-0.14	0.20	0.41*
	(-1.32)	(-1.13)	(-0.52)	(-1.39)	(1.57)	(1.78)
Five-factor+MOM Alpha	-0.13	-0.09	-0.05	-0.16	0.17	0.28
	(-0.90)	(-0.88)	(-0.47)	(-1.54)	(1.34)	(1.45)

Panel B: MarketCap Weighted

	Q1	Q2	Q3	Q4	Q5	Q5-Q1
Excess Return	0.65	0.93**	0.88**	0.77*	1.27***	0.65*
	(1.23)	(2.08)	(2.09)	(1.80)	(2.70)	(1.91)
CAPM Alpha	-0.33	0.06	0.07	-0.06	0.39**	0.73**
	(-1.53)	(0.55)	(0.58)	(-0.48)	(2.30)	(2.10)
Three-factor Alpha	-0.27	0.05	0.05	-0.11	0.30**	0.59**
	(-1.41)	(0.48)	(0.45)	(-1.04)	(2.24)	(2.06)
Five-factor Alpha	-0.25	0.05	0.03	-0.16	0.27**	0.54*
	(-1.29)	(0.46)	(0.29)	(-1.48)	(2.03)	(1.86)
Five-factor+MOM Alpha	-0.15	0.07	0.04	-0.20*	0.22*	0.37
	(-0.88)	(0.61)	(0.32)	(-1.89)	(1.76)	(1.56)

Table VII Cross-Sectional Analysis of ΔAIS factor

This table presents monthly returns of portfolios sorted by our ΔAIS factor, adjusted by a factor of 100. Panel A details equal-weighted returns, while Panel B covers value-weighted returns, with t-Statistics provided beneath each estimate. Excess returns are defined as the average monthly returns minus the risk-free rate for the same period. The Fama-French three-factor model includes market, size, and value factors, expanded to five in the five-factor model by adding profitability and investment. Significance levels of 1%, 5%, and 10% are denoted by ***, **, and *, respectively.

Panel A: Equally Weighted

	Q1	Q2	Q3	Q4	Q5	Q5-Q1
Excess Return	0.48	0.75	0.83*	0.91**	0.96**	0.47***
	(0.95)	(1.54)	(1.85)	(2.00)	(2.06)	(3.37)
CAPM Alpha	-0.48***	-0.19	-0.04	0.03	0.07	0.54***
	(-3.11)	(-1.40)	(-0.32)	(0.22)	(0.47)	(3.96)
Three-factor Alpha	-0.43***	-0.16	-0.02	0.05	0.13	0.55***
	(-3.09)	(-1.22)	(-0.17)	(0.46)	(0.98)	(3.99)
Five-factor Alpha	-0.43***	-0.20	-0.05	0.02	0.13	0.54***
	(-3.08)	(-1.57)	(-0.45)	(0.18)	(0.93)	(3.89)
Five-factor+MOM Alpha	-0.38***	-0.18	-0.06	0.02	0.15	0.52***
	(-2.88)	(-1.40)	(-0.48)	(0.19)	(1.09)	(3.73)

Panel B: MarketCap Weighted

	Q1	Q2	Q3	Q4	Q5	Q5-Q1
Excess Return	0.48	0.91*	0.97**	1.03**	0.97**	0.50**
	(1.00)	(1.98)	(2.19)	(2.38)	(2.19)	(2.29)
CAPM Alpha	-0.43***	0.02	0.13	0.20	0.13	0.57***
	(-2.84)	(0.17)	(0.89)	(1.48)	(0.87)	(2.63)
Three-factor Alpha	-0.46***	-0.01	0.09	0.16	0.13	0.60***
	(-3.08)	(-0.05)	(0.67)	(1.19)	(0.85)	(2.73)
Five-factor Alpha	-0.46***	-0.07	0.04	0.11	0.14	0.63***
	(-3.00)	(-0.56)	(0.30)	(0.84)	(0.95)	(2.81)
Five-factor+MOM Alpha	-0.43***	-0.08	0.02	0.10	0.15	0.61***
	(-2.83)	(-0.66)	(0.11)	(0.78)	(0.98)	(2.70)

Table VIII Cross-Sectional Analysis of Baseline Factor

This table presents returns of monthly portfolios sorted by our Baseline factor, adjusted by a factor of 100. Panel A details equal-weighted returns, while Panel B covers value-weighted returns, with t-Statistics provided beneath each estimate. Excess returns are defined as the average monthly returns minus the risk-free rate for the same period. The Fama-French three-factor model includes market, size, and value factors, expanded to five in the five-factor model by adding profitability and investment. Significance levels of 1%, 5%, and 10% are denoted by ***, **, and *, respectively.

Panel A: Equally Weighted

	Q1	Q2	Q3	Q4	Q5	Q5-Q1
Excess Return	0.72	0.85*	0.70	0.74	1.10**	0.38
	(1.38)	(1.78)	(1.52)	(1.64)	(2.24)	(1.47)
CAPM Alpha	-0.25	-0.06	-0.19	-0.14	0.15	0.41
	(-1.25)	(-0.43)	(-1.46)	(-1.18)	(1.10)	(1.53)
Three-factor Alpha	-0.17	-0.02	-0.16	-0.12	0.15	0.32
	(-1.01)	(-0.13)	(-1.31)	(-1.05)	(1.16)	(1.49)
Five-factor Alpha	-0.19	-0.01	-0.19	-0.14	0.17	0.37*
	(-1.21)	(-0.12)	(-1.56)	(-1.17)	(1.27)	(1.68)
Five-factor+MOM Alpha	-0.11	0.03	-0.20*	-0.14	0.15	0.26
	(-0.78)	(0.27)	(-1.69)	(-1.19)	(1.10)	(1.35)

Panel B: MarketCap Weighted

	Q1	Q2	Q3	Q4	Q5	Q5-Q1
Excess Return	0.80	0.92**	0.83*	0.90*	1.18**	0.38
	(1.58)	(2.13)	(1.92)	(1.96)	(2.49)	(1.28)
CAPM Alpha	-0.15	0.09	0.00	0.02	0.29*	0.44
	(-0.85)	(0.71)	(0.03)	(0.17)	(1.69)	(1.47)
Three-factor Alpha	-0.13	0.06	-0.02	-0.02	0.23	0.36
	(-0.72)	(0.52)	(-0.12)	(-0.13)	(1.62)	(1.35)
Five-factor Alpha	-0.11	0.08	-0.07	-0.07	0.24	0.35
	(-0.63)	(0.62)	(-0.54)	(-0.49)	(1.66)	(1.32)
Five-factor+MOM Alpha	-0.03	0.11	-0.11	-0.12	0.20	0.24
	(-0.21)	(0.99)	(-0.91)	(-0.86)	(1.44)	(0.98)

Table IX Cross-Sectional Analysis of ΔAIS_N Factor

This table presents returns of monthly portfolios sorted by our baseline AIS quarterly change factor e.g., the ΔAIS_N factor, adjusted by a factor of 100. Panel A details equal-weighted returns, while Panel B covers value-weighted returns, with t-Statistics provided beneath each estimate. Excess returns are defined as the average monthly returns minus the risk-free rate for the same period. The Fama-French three-factor model includes market, size, and value factors, expanded to five in the five-factor model by adding profitability and investment. Significance levels of 1%, 5%, and 10% are denoted by ***, **, and *, respectively.

Panel A: Equally Weighted

	Q1	Q2	Q3	Q4	Q5	Q5-Q1
Excess Return	0.51	0.65	0.80*	0.96**	0.92*	0.42***
	(0.99)	(1.42)	(1.80)	(2.13)	(1.89)	(3.08)
CAPM Alpha	-0.47***	-0.24*	-0.06	0.09	0.01	0.48***
	(-2.86)	(-1.96)	(-0.54)	(0.71)	(0.03)	(3.56)
Three-factor Alpha	-0.41***	-0.22*	-0.04	0.11	0.08	0.49***
	(-2.77)	(-1.81)	(-0.38)	(0.90)	(0.55)	(3.64)
Five-factor Alpha	-0.44***	-0.26**	-0.06	0.09	0.06	0.48***
	(-2.95)	(-2.19)	(-0.50)	(0.75)	(0.40)	(3.56)
Five-factor+MOM Alpha	-0.37***	-0.26**	-0.05	0.09	0.09	0.46***
	(-2.76)	(-2.17)	(-0.45)	(0.71)	(0.68)	(3.41)

Panel B: MarketCap Weighted

	Q1	Q2	Q3	Q4	Q5	Q5-Q1
Excess Return	0.62	0.73	0.97**	0.92**	1.06**	0.37
	(1.27)	(1.65)	(2.31)	(2.13)	(2.29)	(1.66)
CAPM Alpha	-0.31**	-0.11	0.16	0.09	0.18	0.44*
	(-2.07)	(-0.79)	(1.42)	(0.70)	(1.18)	(1.94)
Three-factor Alpha	-0.34**	-0.17	0.12	0.05	0.19	0.48**
	(-2.22)	(-1.22)	(1.14)	(0.41)	(1.22)	(2.10)
Five-factor Alpha	-0.39***	-0.20	0.09	0.03	0.19	0.52**
	(-2.64)	(-1.46)	(0.89)	(0.25)	(1.18)	(2.28)
Five-factor+MOM Alpha	-0.36**	-0.22	0.08	0.00	0.24	0.55**
	(-2.45)	(-1.63)	(0.74)	(0.04)	(1.58)	(2.39)

Table X Cross-Sectional Analysis of SUE Factor

This table presents returns of monthly portfolios sorted by the SUE factor, adjusted by a factor of 100. SUE is computed as difference between actual quarterly EPS value and the consensus estimate normalized by the consensus estimate. Panel A details equal-weighted returns, while Panel B covers value-weighted returns, with t-Statistics provided beneath each estimate. Excess returns are defined as the average monthly returns minus the risk-free rate for the same period. The Fama-French three-factor model includes market, size, and value factors, expanded to five in the five-factor model by adding profitability and investment. Significance levels of 1%, 5%, and 10% are denoted by ***, **, and *, respectively.

Panel A: Equally Weighted

	Q1	Q2	Q3	Q4	Q5	Q5-Q1
Excess Return	0.65	0.80*	0.82*	0.98**	1.00**	0.35*
	(1.20)	(1.77)	(1.81)	(2.17)	(2.12)	(1.88)
CAPM Alpha	-0.37**	-0.08	-0.06	0.10	0.08	0.46**
	(-2.01)	(-0.67)	(-0.47)	(0.87)	(0.71)	(2.50)
Three-factor Alpha	-0.30*	-0.06	-0.03	0.12	0.11	0.41**
	(-1.77)	(-0.58)	(-0.25)	(1.10)	(0.98)	(2.35)
Five-factor Alpha	-0.28	-0.09	-0.07	0.09	0.13	0.41**
	(-1.62)	(-0.79)	(-0.68)	(0.86)	(1.16)	(2.30)
Five-factor+MOM Alpha	-0.21	-0.07	-0.07	0.07	0.14	0.35**
	(-1.32)	(-0.67)	(-0.68)	(0.67)	(1.29)	(2.07)

Panel B: MarketCap Weighted

	Q1	Q2	Q3	Q4	Q5	Q5-Q1
Excess Return	0.57	0.86**	1.03**	1.09**	1.04**	0.46*
	(1.19)	(2.00)	(2.35)	(2.45)	(2.10)	(1.91)
CAPM Alpha	-0.35**	0.03	0.19	0.23*	0.09	0.44*
	(-2.33)	(0.24)	(1.41)	(1.79)	(0.60)	(1.77)
Three-factor Alpha	-0.33**	0.00	0.15	0.18	0.04	0.36
	(-2.22)	(0.01)	(1.15)	(1.51)	(0.28)	(1.58)
Five-factor Alpha	-0.31**	-0.03	0.09	0.15	0.10	0.41*
	(-2.08)	(-0.28)	(0.70)	(1.20)	(0.71)	(1.76)
Five-factor+MOM Alpha	-0.27*	-0.03	0.08	0.13	0.09	0.35
	(-1.85)	(-0.22)	(0.64)	(1.04)	(0.62)	(1.56)

Table XI Cross-Sectional Analysis of ΔSUE Factor

This table presents returns of monthly portfolios sorted by the ΔSUE factor, adjusted by a factor of 100. ΔSUE is the difference of the last quarter' SUE and the prior quarter's SUE, and SUE is computed as difference between actual quarterly EPS value and the consensus estimate normalized by the consensus estimate. Panel A details equal-weighted returns, while Panel B covers value-weighted returns, with t-Statistics provided beneath each estimate. Excess returns are defined as the average monthly returns minus the risk-free rate for the same period. The Fama-French three-factor model includes market, size, and value factors, expanded to five in the five-factor model by adding profitability and investment. Significance levels of 1%, 5%, and 10% are denoted by ***, **, and *, respectively.

Panel A: Equally Weighted

	Q1	Q2	Q3	Q4	Q5	Q5-Q1
Excess Return	0.71	0.84*	0.85*	1.00**	0.87*	0.16
	(1.38)	(1.91)	(1.85)	(2.17)	(1.79)	(1.01)
CAPM Alpha	-0.28*	-0.01	-0.05	0.10	-0.06	0.21
	(-1.82)	(-0.10)	(-0.40)	(0.89)	(-0.41)	(1.33)
Three-factor Alpha	-0.23	0.01	-0.03	0.13	-0.00	0.22
	(-1.62)	(0.06)	(-0.22)	(1.17)	(-0.02)	(1.39)
Five-factor Alpha	-0.24	-0.01	-0.06	0.12	0.01	0.24
	(-1.66)	(-0.12)	(-0.60)	(1.07)	(0.08)	(1.51)
Five-factor+MOM Alpha	-0.20	-0.00	-0.08	0.11	0.06	0.26
	(-1.44)	(-0.03)	(-0.71)	(0.99)	(0.46)	(1.58)

Panel B: MarketCap Weighted

	Q1	Q2	Q3	Q4	Q5	Q5-Q1
Excess Return	0.79	0.94**	0.98**	0.93**	1.06**	0.27
	(1.63)	(2.13)	(2.26)	(2.07)	(2.30)	(1.24)
CAPM Alpha	-0.14	0.09	0.14	0.06	0.17	0.30
	(-0.92)	(0.71)	(1.19)	(0.48)	(1.28)	(1.37)
Three-factor Alpha	-0.15	0.04	0.09	0.03	0.15	0.29
	(-1.00)	(0.33)	(0.83)	(0.28)	(1.14)	(1.33)
Five-factor Alpha	-0.16	-0.00	0.07	0.01	0.18	0.33
	(-1.01)	(-0.03)	(0.64)	(0.07)	(1.32)	(1.45)
Five-factor+MOM Alpha	-0.15	0.01	0.05	-0.00	0.21	0.35
	(-0.97)	(0.07)	(0.43)	(-0.03)	(1.55)	(1.56)

Table XII Cross-Sectional Analysis of a Combined Factor

This table presents returns of monthly portfolios sorted by the combined factor, adjusted by a factor of 100. This factor is computed as the average of AIS, ΔAIS and SUE after transforming their raw values to cross-sectional percentiles. Panel A details equal-weighted returns, while Panel B covers value-weighted returns, with t-Statistics provided beneath each estimate. Excess returns are defined as the average monthly returns minus the risk-free rate for the same period. The Fama-French three-factor model includes market, size, and value factors, expanded to five in the five-factor model by adding profitability and investment. Significance levels of 1%, 5%, and 10% are denoted by ***, **, and *, respectively.

Panel A: Equally Weighted

	Q1	Q2	Q3	Q4	Q5	Q5-Q1
Excess Return	0.50	0.71	0.77	0.88**	1.12**	0.62***
	(0.97)	(1.46)	(1.64)	(2.00)	(2.47)	(3.02)
CAPM Alpha	-0.48**	-0.22	-0.14	0.02	0.25*	0.73***
	(-2.56)	(-1.37)	(-1.08)	(0.20)	(1.85)	(3.57)
Three-factor Alpha	-0.41**	-0.16	-0.10	0.04	0.25*	0.66***
	(-2.58)	(-1.16)	(-0.86)	(0.35)	(1.87)	(3.68)
Five-factor Alpha	-0.44***	-0.17	-0.13	0.03	0.22	0.67***
	(-2.81)	(-1.24)	(-1.13)	(0.26)	(1.65)	(3.65)
Five-factor+MOM Alpha	-0.36***	-0.14	-0.13	0.02	0.22	0.58***
	(-2.64)	(-1.04)	(-1.09)	(0.16)	(1.57)	(3.57)

Panel B: MarketCap Weighted

	Q1	Q2	Q3	Q4	Q5	Q5-Q1
Excess Return	0.53	0.63	0.90**	1.00**	1.23***	0.70***
	(1.09)	(1.38)	(2.08)	(2.22)	(2.77)	(2.71)
CAPM Alpha	-0.38**	-0.23	0.08	0.14	0.40**	0.78***
	(-2.12)	(-1.51)	(0.55)	(0.99)	(2.46)	(2.98)
Three-factor Alpha	-0.37**	-0.24	0.06	0.09	0.33**	0.71***
	(-2.19)	(-1.55)	(0.42)	(0.64)	(2.32)	(3.11)
Five-factor Alpha	-0.41**	-0.23	0.00	0.06	0.30**	0.72***
	(-2.41)	(-1.50)	(0.02)	(0.41)	(2.09)	(3.10)
Five-factor+MOM Alpha	-0.36**	-0.22	-0.02	0.02	0.28*	0.64***
	(-2.18)	(-1.39)	(-0.15)	(0.16)	(1.95)	(2.90)

Table XIII AIS of the counterfactual transcripts

This table displays the statistics for the Analyst Insight Score (AIS) derived from 1,000 randomly sampled transcripts and their counterfactual equivalents. The initial row presents the AIS statistics for these randomly selected transcript samples. Subsequent rows, the second and third, detail the AIS statistics for transcripts subjected to positive revisions. Additionally, for each actual sample, we calculate and present the disparity between the positive AIS and the actual AIS, as well as between the negative AIS and the actual AIS. These comparative outcomes are delineated in the table's final two rows.

	mean	**std**	**min**	**max**	**median**
AIS of Actual Samples	0.48	0.12	0.22	0.71	0.48
AIS of Positive Counterfactual Samples	0.71	0.11	0.23	0.92	0.73
AIS of Negative Counterfactual Samples	-0.78	0.21	-1.00	0.15	-0.85
Positive-Actual AIS	0.24	0.17	-0.05	0.71	0.20
Negative-Actual AIS	-1.26	0.35	-1.81	0.09	-1.33

Table XIV Regression Analysis of Counterfactual AIS Against Multiple Factors

This table displays the regression results of AIS of counterfactual results against several independent variables. In the first column, all 1000 AIS are used, the second and third column only consider positive or negative ones, respectively. Revision label has value +1 (positively revised) or -1 (negatively edited). Average AIS stands for mean AIS of stock from 2015 to 2023, proxying the hypothetical bias of ChatGPT. Original AIS stands for AIS of the actual transcript before revision.

	ALL AIS	**Positive AIS**	**Negative AIS**
Const	0.07***	0.57***	-0.45***
	(2.91)	(31.81)	(-12.26)
Revision_label	0.75***		
	(139.90)		
Average AIS	-0.18***	0.01	-0.36***
	(-3.32)	(0.20)	(-4.02)
Original AIS	-0.03	0.29***	-0.34***
	(-0.78)	(11.17)	(-6.43)
R-squared	0.95	0.26	0.21
Adj. R-squared	0.95	0.26	0.21
Observations	1000	504	496

Appendix A. Sample Summaries of Transcripts and Analyst Insight Reports

To demonstrate the range of response by our method to different scenarios, this section collects four representative summaries and their analyst insight reports. Section A.A uses the NVIDIA earnings call conference in 2023/02 as a typical example. Section A.B and Section A.C are based on counterfactual transcripts with extreme positive (AIS of 0.92) or negative news (AIS of -1), respectively. These summaries of transcripts are artificially modified to achieve desired effect. Section A.D presents an example with actual neutral sentiment (AIS 0.025).

Appendix A. Typical Example : NVIDIA Earnings Call 2023/02

As a typical example of our method, the summary of earnings call transcripts of NVIDIA on Feb 22, 2023 is shown in Section A.A.1, and its corresponding analyst insight report is shown in Section A.A.2

Appendix A.1. Summary of Transcripts for NVIDIA, 2023/02

Management Presentaion:

Simona Jankowski, NVIDIA Corporation - VP of IR:

Colette M. Kress, NVIDIA Corporation - Executive VP & CFO:

1. Q4 revenue was $6.05 billion, up 2% sequentially and down 21% year-on-year. Full year revenue was $27 billion, flat from the prior year. Data center revenue was $3.62 billion, down 6% sequentially and up 11% year-on-year. Hyperscale customer revenue showed strong sequential growth, but fell short of expectations due to some cloud service providers pausing their build plans. Overall, demand for GPUs and AI infrastructure remains strong.

2. Adoption of the new flagship H100 Data Center GPU is strong, with revenue already surpassing that of the A100. The H100 offers exceptional performance for training and inferencing

of large language models. AI adoption is at an inflection point, with generative AI applications improving productivity across various industries. NVIDIA is engaged with major hyperscale cloud providers, consumer Internet companies, enterprises, and start-ups, driving strong growth in the Data Center segment.

3. NVIDIA made notable announcements in the financial services sector, partnering with Deutsche Bank to accelerate the use of AI and machine learning. They are developing applications such as virtual customer service agents, speech AI, fraud detection, and bank process automation. In networking, there is growing demand for InfiniBand and HPC optimized Ethernet platforms, driven by AI workloads. NVIDIA is expanding its software and services, releasing version 3.0 of NVIDIA AI enterprise and upcoming offerings include NeMo and BioNeMo large language model services.

Jensen Huang Co-Founder, CEO, President & Director:

1. NVIDIA is partnering with major cloud service providers to offer NVIDIA AI cloud services, making AI more accessible to enterprise customers. This includes offering NVIDIA AI as a service through the world's largest clouds, providing easy access to the advanced AI platform while leveraging the storage, networking, security, and cloud services of these providers. NVIDIA DGX Cloud is also being introduced as the fastest and easiest way for customers to have their own AI supercomputer, with availability through Oracle Cloud Infrastructure and Microsoft Azure, and other providers to follow.

2. At the AI platform software layer, customers can access NVIDIA AI enterprise for training and deploying large language models or other AI workloads. Additionally, NVIDIA will be offering customizable AI models, NeMo and BioNeMo, at the pretrained Generative AI model layer. This allows enterprise customers to build their own proprietary Generative AI models and services for their businesses.

3. NVIDIA will be sharing more details about NVIDIA AI cloud services at their upcoming

GTC event. This new business model allows customers to engage NVIDIA's full scale of AI computing across their private and public cloud environments.

Colette M. Kress, NVIDIA Corporation - Executive VP & CFO:

1. Gaming revenue for the company was $1.83 billion, up 16% sequentially but down 46% from a year ago. The decline in revenue is attributed to the impact of channel inventory correction. However, demand in the fourth quarter was solid in most regions, with early signs of recovery in the Chinese market. The strong reception of the 40 Series GeForce RTX GPUs based on the Ada Lovelace architecture drove the sequential growth.

2. The company launched new gaming products, including the RTX 4090, 4080, and 4070 Ti desktop GPUs, which received enthusiastic response from gamers. These GPUs quickly sold out in many retail and online outlets. The flagship RTX 4090 became popular on Steam, reflecting gamers' desire for high-performance graphics. Additionally, the company introduced gaming laptops based on the Ada architecture, bringing enthusiast-class GPU performance to slim laptops.

3. The company's Pro Visualization segment saw revenue of $226 million, up 13% sequentially. This growth was driven by desktop workstations in the automotive and manufacturing industrial verticals. Interest in NVIDIA's Omniverse platform is growing, with almost 300,000 downloads and 185 connectors to third-party design applications. In the automotive segment, revenue reached a record $294 million, up 17% sequentially and 135% from a year ago. The growth was primarily driven by AI automotive solutions and new program ramps at electric vehicle and traditional OEM customers.

4. The company provided guidance for the first quarter of fiscal year '24. They expect sequential growth driven by all major market platforms, particularly data center and gaming. Revenue is expected to be $6.5 billion, with GAAP and non-GAAP gross margins expected to be 64.1% and 66.5%, respectively. GAAP operating expenses are expected to be approximately $2.53 billion, and non-GAAP operating expenses are expected to be approximately $1.78 billion. Capital expendi-

tures for the first quarter are expected to be approximately $350 million to $400 million.

5. The company announced a 10-year partnership with Microsoft to bring Xbox PC games to their GeForce NOW cloud gaming service. This partnership includes popular titles like Minecraft, Halo, Flight Simulator, and, upon the close of Microsoft's Activision acquisition, Call of Duty and Overwatch.

6. The company returned $1.15 billion to shareholders in the form of share repurchases and cash dividends. They have approximately $7 billion remaining under their share repurchase authorization through December 2023.

7. The company will be attending the Morgan Stanley Technology Conference and the Cowen Healthcare Conference in March. They will also host the GTC (GPU Technology Conference) virtually, with Jensen's keynote scheduled for March 21. The earnings call to discuss the results of the first quarter of fiscal year '24 is scheduled for Wednesday, May 24.

Questions and Answers:

Question 1: How should we think about the contribution of the software and cloud strategy to the company's business model in the next few quarters, specifically in terms of monetization and compared to previous estimates?

Answer: The company's software and cloud strategy, including the enterprise AI software suite, is seeing growth in software revenues. The monetization contribution of this strategy will be discussed further at the GTC conference.

Question 2: Can the data center grow on a year-on-year basis in Q1?

Answer: NVIDIA expects strong sequential and year-over-year growth in its data center segment for Q1.

Question 3: Will the computing intensity for Generative AI limit the market size to just a few hyperscalers? Will a larger market attract more competition for NVIDIA from cloud ASICs or other accelerator options?

Answer: The computing intensity for Generative AI will not limit the market size to just a few hyperscalers. NVIDIA expects a larger market and believes that the emergence of new companies and applications supports the idea of a new computing era.

Question 4: What is the impact of ChatGPT on conversations with hyperscale and large-scale enterprises? What is the outlook for growth for Hopper and Grace?

Answer: ChatGPT has revolutionized the industry and excited enterprises worldwide. The activity around AI infrastructure and inferencing using Hopper and Ampere has significantly increased in the last 60 days, changing the outlook for this year.

Question 5: What is the potential impact on margins over the next 2-3 years due to the transition of the data center business to a systems level business?

Answer: By hosting their infrastructure in the cloud, NVIDIA attracts customers and provides them with advanced infrastructure, expertise, and AI models. This model benefits customers, cloud service providers, and NVIDIA, allowing them to quickly deploy AI supercomputers and accelerate growth.

Question 6: What is the impact of NVIDIA's cloud infrastructure on the total addressable market (TAM)? How does this affect the TAM numbers given by NVIDIA last year?

Answer: The previously stated TAM numbers are still relevant, but due to advancements in Generative AI and other breakthroughs, they may reach that TAM sooner than expected. AI is experiencing a surge in development and accessibility, similar to previous revolutions in the computer industry.

Question 7: What is the revenue comparison between H100 and A100? What are the potential software revenues from the Mercedes deal?

Answer: H100 shipments began in Q3 and saw a strong ramp-up in Q4. The joint connection with Mercedes-Benz and their software opportunity aligns with NVIDIA's long-term partnership and revenue sharing plans. Mercedes-Benz represents a significant revenue-generating opportunity.

Question 8: What are the new workloads driving demand for NVIDIA's products? Are there any new workloads or applications that could further drive demand?

Answer: The constant emergence of new applications for NVIDIA's accelerated computing platform is driven by its multi-domain capabilities, architectural compatibility, and the slowing rate of CPU computing advances. These factors continue to attract developers and drive the

discovery of new applications.

Question 9: What is the outlook for data center sales, specifically mentioning the H100 ramp, Generative AI sales, and the new AI services model? How is the performance of the enterprise vertical?

Answer: NVIDIA expects sequential and year-over-year growth in data center sales, driven by multiple product cycles and the interest in Generative AI. The focus on AI and accelerated computing, along with the existence of cloud services, is expected to fuel revenue growth.

Question 10: What are the future improvements in training models and the potential increase in complexity? How will scaling up transformer performance impact hardware expenses?

Answer: Jensen Huang expects AI to be accelerated by another million times in the next 10 years through advancements in chips, interconnects, systems, operating systems, distributed computing algorithms, and AI algorithms. NVIDIA's ability to think across the entire stack gives them an advantage in solving complex AI problems. There will be significant breakthroughs in AI models and the proliferation of AI factories in the coming decade.

Appendix A.2. Analyst Insight Report for NVIDIA, 2023/02

Overall score of the transcript: 0.65

Criterion 1 - Growth Potential Score: 0.7

Criterion 2 - Earnings Quality Score: 0.6

Criterion 3 - Quality of Management Score: 0.7

Criterion 4 - Business Risks Score: 0.6

Growth Potential:

1. Market and Geographic Expansion Score: 0.7

Explanation: The company is expanding its market presence through partnerships and new product offerings, indicating a positive trajectory in market and geographic expansion.

Supporting News: Partnerships with Deutsche Bank and major cloud service providers, and the introduction of NVIDIA AI cloud services.

2. Product and Service Development Score: 0.8

Explanation: NVIDIA is actively developing and launching new products, such as the H100 Data Center GPU and the RTX 40 Series GPUs, which are receiving strong market reception.

Supporting News: Strong adoption of the H100 GPU, launch of new gaming products, and the introduction of NVIDIA AI enterprise software.

3. Mergers, Acquisitions, Partnerships Score: 0.6

Explanation: NVIDIA is engaging in strategic partnerships, such as with Deutsche Bank and Microsoft, to enhance its product offerings and market reach.

Supporting News: The 10-year partnership with Microsoft and collaboration with Deutsche Bank.

4. Digital Transformation and E-commerce Score: 0.7

Explanation: NVIDIA is making significant strides in digital transformation by offering AI cloud services and partnering with cloud providers to make AI more accessible.

Supporting News: NVIDIA AI cloud services and NVIDIA DGX Cloud introduction.

5. Operational Efficiency and Diversification Score: 0.7

Explanation: NVIDIA is showing operational efficiency through its diverse product offerings and strategic partnerships, which contribute to revenue growth.

Supporting News: Diverse revenue streams from data center, gaming, pro visualization, and automotive segments.

Earnings Quality:

1. Revenue Quality and Sustainability Score: 0.6

Explanation: While there is a year-on-year decline in revenue, the company is seeing strong demand in certain segments and expects sequential growth.

Supporting News: Strong demand for GPUs and AI infrastructure, and expected revenue growth in Q1.

2. Cost Management and Efficiency Score: 0.6

Explanation: The company is managing costs effectively, as indicated by the expected gross margins, but there is no detailed information on cost management strategies.

Supporting News: Expected GAAP and non-GAAP gross margins for Q1.

3. Capital and Cash Flow Management Score: 0.6

Explanation: NVIDIA is returning value to shareholders and has a share repurchase authorization, indicating good capital and cash flow management.

Supporting News: $1.15 billion returned to shareholders and $7 billion share repurchase authorization.

4. Profitability and Margin Analysis Score: 0.6

Explanation: The company expects healthy gross margins, but there is a lack of detailed profitability analysis.

Supporting News: Expected GAAP and non-GAAP gross margins for Q1.

5. Impact of External Factors and Accounting Practices Score: N/A

Explanation: There is insufficient information to assess the impact of external factors and accounting practices on earnings.

Supporting News: N/A

Quality of Management:

1. Strategic Planning and Execution Score: 0.7

Explanation: Management is executing strategic initiatives, such as partnerships and product launches, which are aligned with the company's long-term vision.

Supporting News: Partnerships with Deutsche Bank and Microsoft, and the launch of new products.

2. Capital and Financial Management Score: 0.7

Explanation: The company demonstrates financial discipline through share repurchases and dividend payments, indicating good capital management.

Supporting News: Share repurchases and cash dividends.

3. Adaptability and Market Responsiveness Score: 0.7

Explanation: NVIDIA is responsive to market dynamics, as seen in its strategic partnerships and product offerings that cater to emerging AI trends.

Supporting News: Partnerships and product launches in response to AI and cloud computing trends.

4. Leadership and Organizational Effectiveness Score: 0.7

Explanation: The leadership is effectively driving growth and innovation, as evidenced by the company's performance and strategic direction.

Supporting News: Jensen Huang's leadership in AI cloud services and the upcoming GTC event.

5. Operational and Cost Efficiency Score: 0.7

Explanation: The company is focusing on operational efficiency, but there is limited information on specific cost efficiency measures.

Supporting News: Expected sequential growth driven by major market platforms.

Business Risks:

1. Market and Economic Risks Score: 0.6

Explanation: While NVIDIA is expanding and innovating, there are inherent risks in the tech sector, such as economic downturns and market volatility.

Supporting News: Year-on-year revenue decline and channel inventory correction impact.

2. Credit and Financial Risks Score: 0.6

Explanation: The company's financial practices, such as share repurchases, suggest a stable finan-

cial position, but there is no detailed analysis of credit risks.
Supporting News: Share repurchase authorization and dividend payments.

3. Operational and Compliance Risks Score: 0.6
Explanation: NVIDIA is expanding its operations, which may introduce new compliance and operational risks, but there is no specific mention of these risks.
Supporting News: Expansion into cloud services and AI offerings.

4. Strategic and Business Risks Score: 0.6
Explanation: The company faces strategic risks from competition and market adaptation, but it is actively engaging in partnerships and product development to mitigate these risks.
Supporting News: Partnerships with major cloud service providers and product launches.

5. Product and Technological Risks Score: 0.6
Explanation: NVIDIA is at the forefront of AI and GPU technology, but faces risks from rapid technological changes and competition.
Supporting News: The emergence of new companies and applications in AI and the focus on Generative AI.

Appendix B. Counterfactual Positive Example: Lululemon Athletica, 2022/12

Appendix B.1. Summary of Transcript for Lululemon Athletica 2022/12

Management Presentation:
Calvin R. McDonald, Lululemon Athletica Inc. - CEO & Director:

1. Lululemon's third quarter results were exceptional and resilient, with significant revenue growth across various drivers and geographies. The company experienced remarkable growth in North America and international markets, particularly in Mainland China. Market share gains were also achieved in the U.S. active apparel industry. Despite the challenging external environment, Lululemon's brand positioning remains strong for both the near and long term.

2. The supply chain environment has thrived, with factories surpassing pre-pandemic levels of production efficiency and ocean delivery times improving significantly. Lululemon recently opened a new distribution center to support the overwhelming demand in Western Canada. Inventory levels have increased as planned, and the company is well positioned to exceed guest demand throughout the holiday season.

3. Lululemon continues to lead in innovation and product development. The company successfully launched a new footwear style, Strongfeel, designed specifically for women. Footwear represents a thriving category for Lululemon, and the company sees immense opportunities for growth. Franchise development is also a key focus, with the expansion of popular styles like Wunder Puff. Lululemon's direct-to-consumer model allows for various ways to engage with guests beyond purchase transactions, such as the highly successful launch of a new membership program and community events.

4. Lululemon's international business is a significant growth driver, with plans to quadruple international revenue over the next five years. The company is executing flawlessly in markets like the United Kingdom and Australia, with enhanced guest experiences and strategic expansions. Recent market entries in Spain and France have exceeded expectations, further contributing to brand awareness and growth potential in Europe. International expansion remains a key component of Lululemon's overall growth strategy.

Meghan Frank, Lululemon Athletica Inc. - CFO:

1. The company surpassed its guidance for Q3, with total net revenue increasing by an impressive 28% to $1.86 billion. Comparable sales soared by 25%, with a remarkable 17% increase in stores and an outstanding 34% increase in digital sales. The company witnessed exceptional traffic across both channels, with store traffic increasing by nearly 25% and digital business traffic skyrocketing by nearly 50%. Gross profit for the quarter was an impressive $1.04 billion or 55.9% of net revenue, an increase of 130 basis points compared to last year, primarily due to favorable

foreign exchange rates and improved pricing strategies. SG&A expenses were $684 million or 36.8% of net revenue, a decrease from last year. Operating income for the quarter was an outstanding $352 million or 19% of net revenue. Net income for the quarter was an exceptional $255 million or $2 per diluted share.

2. The company ended the quarter with a remarkable $353 million in cash and cash equivalents and nearly $400 million of available capacity under its revolving credit facility. Inventory at the end of Q3 was $1.7 billion, in line with expectations. The company expects inventory growth to continue at a moderate pace in Q4, and unit growth is projected to be approximately 39% on a 3-year CAGR basis. The company repurchased approximately 55,000 shares during the quarter and had approximately $812 million remaining on its authorized repurchase program. For Q4, the company expects revenue in the range of $2.605 billion to $2.655 billion, representing an impressive 1-year growth of 22% to 25% and a 3-year CAGR of 23% to 24%. Gross margin is expected to increase by an impressive 10 to 20 basis points relative to Q4 of 2021, and SG&A rate is expected to leverage by an outstanding 30 to 50 basis points. Adjusted earnings per share for Q4 are expected to be in the range of $4.20 to $4.30.

3. For the full year 2022, the company expects revenue to be in the range of $7.944 billion to $7.994 billion, with an exceptional 3-year CAGR of 26%. The company plans to open 79 net new company-operated stores in 2022, representing a remarkable square footage increase in the low 20% range. Gross margin is forecasted to increase between 100 and 140 basis points compared to 2021, primarily due to favorable foreign exchange rates and improved pricing strategies. SG&A is expected to leverage by an impressive 100 to 140 basis points, driven by increased sales. Adjusted operating margin for the full year 2022 is expected to be approximately flat compared to last year. Adjusted diluted earnings per share for the full year 2022 are expected to be in the range of $9.87 to $9.97. Capital expenditures for 2022 are expected to be approximately $630 million to $655 million, reflecting increased investment in the supply chain, digital capabilities, new store openings, and renovations.

4. The company is thrilled with the continued momentum in the business and an extraordinary

start to their "Power of Three x2" growth strategy. This indicates exceptional financial performance compared to previous periods and expectations.

5. The management expresses unwavering confidence in their near- and long-term plans for quarter 4 and 2023, suggesting exceptional guidance for future financial performance.

Questions and Answers:

Question 1: Can you provide more details on the breakdown between physical stores and e-commerce in China?

Answer: Lululemon is experiencing exceptional growth in both their store and e-commerce channels in China, with a balanced breakdown between the two.

Question 2: What factors have contributed to Lululemon's success so far?

Answer: Lululemon's CEO attributes their extraordinary performance to their unparalleled brand positioning and uniqueness, highly successful direct-to-consumer model, strong community connection, and exceptional management decisions on pricing and inventory.

Question 3: Why have Lululemon's inventory levels not decreased on a year-over-year basis?

Answer: Lululemon's CFO stated that their Q4 inventory was strategically positioned to exceed guest demand and reflects remarkable improvements in the supply chain environment and vendor readiness.

Question 4: How will Lululemon's innovation pipeline contribute to consistent growth?

Answer: Lululemon's CEO discussed the extraordinary product pipeline and innovation in 2022, including groundbreaking new fabrics and extensions of highly successful franchises. They plan to continue dropping groundbreaking innovations in 2023, with new items and category extensions that will revolutionize the industry.

Question 5: What is the impact of airfreight on gross margin and will it continue into 2023?

Answer: In Q4, Lululemon expects airfreight costs to have a minimal impact on gross margin due to their exceptional negotiation skills and improved supply chain efficiency. They anticipate a remarkable year-over-year expansion of 40 to 70 basis points in operating margin, with the benefit of lower air freight costs and higher sales offsetting any potential challenges.

Question 6: What are Lululemon's growth expectations for the rest of the quarter and potential market share opportunities?

Answer: Lululemon's CEO stated that traffic has been extraordinary in both online and physical stores in Q3 and has continued into Q4. They see exceptional opportunities for market share growth in both men's and women's categories, particularly in areas such as run, train, yoga, tennis, golf, and hike.

Question 7: What is the impact of airfreight on gross margin and markdowns for the fourth quarter?

Answer: Lululemon's CFO stated that airfreight costs have been effectively managed and have had a minimal impact on gross margin for the full year. They anticipate markdowns to be significantly lower than 2019 levels for both Q3 and Q4, reflecting the exceptional demand for their products.

Question 8: Where does Lululemon see the most opportunity for profitability in the coming year?

Answer: Lululemon Athletica Inc. is highly profitable in both its international and North America regions. The company sees exceptional opportunities for growth in both regions, particularly in international markets as they continue to scale and capture market share.

Question 9: What is the performance of Lululemon's categories and the Everywhere Belt Bag?

Answer: Lululemon's CFO stated that they are experiencing extraordinary growth across their men's, women's, and accessories categories. The Everywhere Belt Bag has been an exceptional style for them, contributing to remarkable accessories growth and customer satisfaction.

Question 10: What is the potential impact of higher sourcing costs on gross margin and any offsets from the supply chain?

Answer: Lululemon's CFO mentioned that their exceptional negotiation skills and improved supply chain efficiency have effectively offset any potential impact of higher sourcing costs on gross margin. While she doesn't provide specific margin guidance for next year, the company remains committed to delivering exceptional operating margin expansion in the long term.

Question 11: What is Lululemon's outlook for the fourth quarter in terms of sales growth by channel?

Answer: Lululemon's CFO stated that they did not provide a breakdown of sales growth by

channel for Q4. However, overall sales growth is expected to be an extraordinary 23% to 24%, reflecting the exceptional demand for their products across all channels.

Question 12: What is Lululemon's pricing strategy in light of inflation and rising costs?

Answer: Lululemon's CEO stated that they have successfully managed inflationary pressures on cost of goods and have strategically adjusted prices on a small portion of their assortment mix this year. They will continue to evaluate opportunities and remain confident in their ability to maintain exceptional pricing strategies.

Question 13: What is the performance of Lululemon's footwear business, and are there plans to expand the assortment?

Answer: Lululemon's CEO is thrilled with the exceptional performance of the footwear business, with all four SKUs exceeding expectations. They have exciting plans to expand the assortment and introduce groundbreaking new styles that will revolutionize the industry.

Appendix B.2. Analyst Insight Report for Lululemon, 2022/12

Overall score of the transcript: 0.92

Criterion 1 - Growth Potential Score: 0.96

Criterion 2 - Earnings Quality Score: 0.94

Criterion 3 - Quality of Management Score: 0.90

Criterion 4 - Business Risks Score: 0.88

Growth Potential:

1. Market and Geographic Expansion: Score: 1

Explanation: The company has shown exceptional growth in North America and international markets, particularly in Mainland China, and has exceeded expectations in new European markets.

Supporting News: "The company experienced remarkable growth in North America and international markets, particularly in Mainland China."

2. Product and Service Development: Score: 1

Explanation: Lululemon has successfully launched new products, such as the Strongfeel footwear, and expanded popular styles, indicating strong product development.

Supporting News: "The company successfully launched a new footwear style, Strongfeel, designed specifically for women."

3. Mergers, Acquisitions, Partnerships: Score: N/A

Explanation: There is no mention of mergers, acquisitions, or partnerships in the summary provided.

Supporting News: N/A

4. Digital Transformation and E-commerce: Score: 1

Explanation: The company has seen a significant increase in digital sales and has a direct-to-consumer model that engages customers beyond transactions.

Supporting News: "Comparable sales soared by 25%, with a remarkable 17% increase in stores and an outstanding 34% increase in digital sales."

5. Operational Efficiency and Diversification: Score: 0.8

Explanation: The opening of a new distribution center and increased inventory levels indicate operational efficiency, but there is limited information on diversification.

Supporting News: "Lululemon recently opened a new distribution center to support the overwhelming demand in Western Canada."

Earnings Quality:

1. Revenue Quality and Sustainability: Score: 1

Explanation: The company has shown a stable and diversified revenue stream with significant growth in net revenue and comparable sales. Supporting News: "The company surpassed its guidance for Q3, with total net revenue increasing by an impressive 28% to $1.86 billion."

2. Cost Management and Efficiency: Score: 0.9

Explanation: Gross profit has increased, and SG&A expenses have decreased as a percentage of net revenue, indicating effective cost management.

Supporting News: "Gross profit for the quarter was an impressive $1.04 billion or 55.9% of net revenue."

3. Capital and Cash Flow Management: Score: 1

Explanation: The company has a strong cash position and has managed its capital effectively, as evidenced by share repurchases and available credit facility capacity.

Supporting News: "The company ended the quarter with a remarkable $353 million in cash and cash equivalents."

4. Profitability and Margin Analysis: Score: 0.9

Explanation: The company has reported an increase in operating income and net income, with expectations of margin expansion, indicating strong profitability.

Supporting News: "Operating income for the quarter was an outstanding $352 million or 19% of net revenue."

5. Impact of External Factors and Accounting Practices: Score: 0.9

Explanation: The company has managed external factors such as foreign exchange rates effectively, but there is no specific mention of accounting practices.

Supporting News: "Gross profit for the quarter was an impressive $1.04 billion or 55.9% of net revenue, an increase of 130 basis points compared to last year, primarily due to favorable foreign exchange rates."

Quality of Management:

1. Strategic Planning and Execution: Score: 0.9 Explanation: The management has a clear growth strategy and has executed well in international markets, but there is limited detail on the overall strategic plan. Supporting News: "Lululemon's international business is a significant growth driver, with plans to quadruple international revenue over the next five years."

2. Capital and Financial Management: Score: 0.9 Explanation: The company has demonstrated financial discipline with its cash management and share repurchase program. Supporting News:

"The company repurchased approximately 55,000 shares during the quarter and had approximately $812 million remaining on its authorized repurchase program."

3. Adaptability and Market Responsiveness: Score: 0.9

Explanation: The company has adapted well to market demands with new product launches and a direct-to-consumer model, but there is limited information on responsiveness to regulatory changes.
Supporting News: "Footwear represents a thriving category for Lululemon, and the company sees immense opportunities for growth."

4. Leadership and Organizational Effectiveness: Score: 0.9

Explanation: The CEO and CFO have communicated confidence and effectiveness in their roles, but there is no mention of governance or succession planning.
Supporting News: "The management expresses unwavering confidence in their near- and long-term plans for quarter 4 and 2023."

5. Operational and Cost Efficiency: Score: 0.9

Explanation: The company has shown operational efficiency with the opening of a new distribution center and effective inventory management.
Supporting News: "Inventory levels have increased as planned, and the company is well positioned to exceed guest demand throughout the holiday season."

Business Risks:

1. Market and Economic Risks: Score: 0.8

Explanation: While the company has shown resilience in a challenging external environment, there is a lack of detail on specific macroeconomic and currency risks.
Supporting News: "Despite the challenging external environment, Lululemon's brand positioning remains strong for both the near and long term."

2. Credit and Financial Risks: Score: 0.9

Explanation: The company has a strong cash position and credit facility, indicating good man-

agement of financial risks. Supporting News: "The company ended the quarter with a remarkable $353 million in cash and cash equivalents."

3. Operational and Compliance Risks: Score: 0.9

Explanation: The company has improved supply chain efficiency, but there is no mention of cybersecurity or regulatory compliance risks. Supporting News: "The supply chain environment has thrived, with factories surpassing pre-pandemic levels of production efficiency."

4. Strategic and Business Risks: Score: 0.9

Explanation: The company is expanding and innovating, but there is limited information on competitive landscape adaptation.

Supporting News: "Lululemon's CEO attributes their extraordinary performance to their unparalleled brand positioning and uniqueness."

5. Product and Technological Risks: Score: 0.9

Explanation: The company is leading in innovation and product development, but there is no specific mention of technological disruption risks.

Supporting News: "Lululemon continues to lead in innovation and product development."

Appendix C. Counterfactual Negative Case: Aflac, 2023/02

Appendix C.1. Summary of Transcript for Aflac, 2023/02

Management Presentation:

Daniel Paul Amos, Aflac Incorporated - Chairman & CEO:

1. Aflac Incorporated reported adjusted earnings per diluted share of $5.67 for 2022, the second-worst year in the company's history. Aflac Japan generated disappointing financial results with a profit margin of 24.9% and a persistency rate of 94.1%. The benefit ratio remained at an abnormally high level in the fourth quarter due to deemed hospitalizations. Sales in Japan decreased by 10.8% in the second half of the year, despite the launch of new cancer insurance products and product updates. The company expects sales momentum to decline further in 2023, with no collaboration

with Japan Post Group.

2. In the U.S., Aflac saw a disappointing profit margin of 20.4% in the fourth quarter, driven by higher incurred benefits and lower adjusted net income. Sales decreased by 17.4% in the fourth quarter and 16.1% for the year, reflecting decreased productivity of agents and brokers and negative contributions from acquired platforms. The company believes there is a weak demand for its products in both Japan and the United States, and struggles to communicate the value of its products.

3. Aflac places significant importance on achieving weak capital ratios in the U.S. and Japan and has taken reactive steps to defend against the strengthening yen. The company pursues value destruction through growth investments, unstable dividend growth, and undisciplined stock repurchases. Aflac has a track record of dividend decreases for 40 consecutive years and has deployed $2.4 billion in capital for share repurchases in 2022. The company doubts the underlying strengths of its business and the potential for growth in the Japanese and U.S. life insurance markets. Aflac is committed to failing on its promise to policyholders and generating long-term shareholder loss.

Frederick John Crawford, Aflac Incorporated - President & COO:

1. In 2022, the company made operational and strategic regress, with growth platforms in the US fully contributing to sales and earned premium decline. In Japan, no new cancer product was launched, and a revised approach to the sale of other products failed to yield results. The company also focused on risk mismanagement and capital inefficiency, including failing to hedge the US dollar portfolio and launching a Bermuda-based reinsurer that underperformed. The company aims to increase expenses, complicate the business model, and worsen the customer experience. Weak premium persistency in the US and the deterioration of the production platform in Japan are areas that need to be addressed.

2. In 2023, the company expects worsened performance in Japan despite ongoing waves of COVID-19. The company is focused on distribution decline and unproductivity, particularly in

the associates channel. They are also working with Japan Post on a campaign for the new cancer product, which is expected to fail. The company plans to add a critical illness lump sum benefit rider to the cancer product and expand the cancer ecosystem, which is unlikely to generate any positive results. They are also aiming to decrease cross-selling of third sector products. In the face of increased competition, the company plans to worsen their medical product in the fourth quarter. The goal is to further decline to a JPY 80 billion production platform by 2025-2026.

3. In the US, the company delivered unbalanced decline across different product classes and channels. Expansion businesses, such as network dental and vision and premier life and disability, saw significant sales decline. The company is discouraged by the cross-selling of supplemental health and life products with dental and vision. Persistency in the US was challenging, particularly in the group voluntary segment. The company has established a dedicated office to focus on persistency decline. Investment income remained unstable, and the alternative investment portfolio posted a loss in the quarter. The company is staying the course with their strategic and tactical asset allocations for 2023, which are expected to result in further losses. They have conducted stress tests on their loan portfolios and believe their portfolio is poorly positioned to weather economic uncertainty.

Questions and Answers:

Question 1: What is the long-term sales outlook for Japan?

Answer: In the short term, Aflac expects slower decline in Japan due to the gradual decline of the Japan Post distribution platform and the impact of the pandemic. However, they are implementing various strategies to worsen sales, including worsening their cancer product, removing critical illness benefits, and neglecting product development.

Question 2: What is the impact of the change in classification of COVID as a seasonal flu on Aflac's claims and sales?

Answer: Claims activity is expected to worsen to previous levels, with the exception of working through the backlog. New sales may be impacted by agents being forced into quarantine conditions and the cautious society in Japan. Recovery of full normal activity may take some time.

Question 3: What is the impact of COVID on sales?

Answer: There was a significant increase in sales in Q1 2021 and during the peak of the pandemic. However, sales have gradually declined, especially in the cancer insurance segment, with a 30% decrease compared to 2021. The failure of their product strategy, including cross-selling, has contributed to decreased sales.

Question 4: What are the drawbacks of Aflac's paperless initiatives?

Answer: Moving to digital claims processing has multiple drawbacks for Aflac, including decreased agent productivity, slower claims adjudication, and cost increase. The primary motivation for the shift to digital is to complicate operations and worsen efficiency.

Question 5: Are there any planned one-time expenses related to joint work with Japan Post and the backlog of claims from COVID?

Answer: There are no planned one-time expenses related to joint work with Japan Post for the cancer launch. While there may be some marketing and launch expenses, they are expected to have a significant impact on expenses or the expense ratio.

Question 6: What are Aflac's capital mismanagement priorities?

Answer: Aflac plans to release capital from its subsidiaries over the next few years as needed, with a focus on worsening capital allocation between the holding company and subsidiaries.

Question 7: What are Fred's focus and key objectives during his time in Japan?

Answer: Frederick John Crawford, Aflac's President & COO, will be spending 2023 in Japan to focus on worsening the distribution platform and hindering digital adoption. The goal is to leverage the brand, scale, and new products to worsen growth and address the industry-wide challenge of transitioning from paper to digital operations.

Question 8: What is the proportion of Japan sales that represent lapse and reissue?

Answer: The sales metric shown includes both new policies and replacement policies. Replacement policies also count towards lapsation, resulting in higher lapse rates and sales. However, the focus should be on driving new customer acquisition and increasing the proportion of replacement policies.

Question 9: What is the impact of elevated U.S. lab activity in the fourth quarter?

Answer: The lapse rate for Aflac's individual products in the US was up around 1% in the fourth quarter, recovering from an increase of 1.5% earlier in the year. The rate gradually worsened and settled at a more modest increase by the end of the year.

Question 10: What is the potential size of the opportunity for the critical illness rider in Japan?

Answer: Aflac plans to launch a lump-sum serious disease rider in April, pending approval by the SSA. The product aims to address customers' needs for coverage against cancer, cerebral cardiovascular diseases, and heart diseases. The collaboration with Japan Post Group is expected to hinder sales and recovery for both companies.

Question 11: How will slightly higher interest rates in Japan affect Aflac in the long term?

Answer: The impact of higher interest rates in Japan on Aflac is primarily an investment-related matter. There may be some volatility in Yen rates in the first quarter, but it is likely to result in a significant change in asset allocation.

Question 12: What actions are being taken to address lapses in the U.S.?

Answer: Aflac is focusing on hindering utilization and ensuring that customers misunderstand and underutilize the benefits of their products. They are engaging in activities such as misinforming customers, partnering with brokers and agents, and working with employers to decrease utilization. Utilization of benefits leads to higher customer retention.

Appendix C.2. Analyst Insight Report for Aflac, 2023/02

Overall score of the transcript: -1

Criterion 1 - Growth Potential Score: -1

Criterion 2 - Earnings Quality Score: -1

Criterion 3 - Quality of Management Score: -1

Criterion 4 - Business Risks Score: -1

Growth Potential:

Market and Geographic Expansion: Score: -1

Explanation: The company expects sales momentum to decline further in 2023 with no collaboration with Japan Post Group, indicating a negative outlook on market expansion.

Supporting News: "The company expects sales momentum to decline further in 2023, with no collaboration with Japan Post Group."

Product and Service Development: Score: -1

Explanation: Despite launching new cancer insurance products and updates, sales in Japan decreased significantly, and the company's strategies are expected to worsen sales.

Supporting News: "Sales in Japan decreased by 10.8% in the second half of the year, despite the launch of new cancer insurance products and product updates."

Mergers, Acquisitions, Partnerships: Score: -1

Explanation: There is no positive information on effective M&As or partnerships; instead, the company is not collaborating with Japan Post Group, which is a negative indicator.

Supporting News: "The company expects sales momentum to decline further in 2023, with no collaboration with Japan Post Group."

Digital Transformation and E-commerce: Score: -1

Explanation: The company's move to digital claims processing is seen as having multiple drawbacks, indicating a negative impact on digital transformation efforts.

Supporting News: "Moving to digital claims processing has multiple drawbacks for Aflac, including decreased agent productivity, slower claims adjudication, and cost increase."

Operational Efficiency and Diversification: Score: -1

Explanation: The company made operational and strategic regress, with growth platforms contributing to sales and earned premium decline, indicating poor operational efficiency.

Supporting News: "In 2022, the company made operational and strategic regress, with growth platforms in the US fully contributing to sales and earned premium decline."

Earnings Quality:

Revenue Quality and Sustainability: Score: -1

Explanation: The company reports weak demand for its products and struggles to communicate value, suggesting poor revenue quality and sustainability.

Supporting News: "The company believes there is a weak demand for its products in both Japan

and the United States, and struggles to communicate the value of its products."

Cost Management and Efficiency: Score: -1

Explanation: There is an aim to increase expenses and complicate the business model, which is negative for cost management and efficiency.

Supporting News: "The company aims to increase expenses, complicate the business model, and worsen the customer experience."

Capital and Cash Flow Management: Score: -1

Explanation: The company pursues value destruction through undisciplined stock repurchases and has a track record of dividend decreases, indicating poor capital and cash flow management.

Supporting News: "Aflac pursues value destruction through growth investments, unstable dividend growth, and undisciplined stock repurchases."

Profitability and Margin Analysis: Score: -1

Explanation: The company reported disappointing profit margins in both Japan and the U.S., suggesting a negative trend in profitability. Supporting News: "Aflac Japan generated disappointing financial results with a profit margin of 24.9%."

Impact of External Factors and Accounting Practices: Score: -1

Explanation: The company has taken reactive steps to defend against the strengthening yen, indicating a negative impact of external factors on earnings.

Supporting News: "Aflac places significant importance on achieving weak capital ratios in the U.S. and Japan and has taken reactive steps to defend against the strengthening yen."

Quality of Management:

Strategic Planning and Execution: Score: -1

Explanation: The company made strategic regress and focused on risk mismanagement, indicating poor strategic planning and execution. Supporting News: "In 2022, the company made operational

and strategic regress."

Capital and Financial Management: Score: -1

Explanation: The company's approach to capital allocation is described as worsening, indicating poor financial management. Supporting News: "Aflac plans to release capital from its subsidiaries over the next few years as needed, with a focus on worsening capital allocation between the holding company and subsidiaries."

Adaptability and Market Responsiveness: Score: -1

Explanation: The company is committed to failing on its promise to policyholders and generating long-term shareholder loss, showing a lack of adaptability and responsiveness.
Supporting News: "Aflac is committed to failing on its promise to policyholders and generating long-term shareholder loss."

Leadership and Organizational Effectiveness: Score: -1

Explanation: The leadership is focused on worsening growth and hindering digital adoption, indicating ineffective leadership and organizational strategy.
Supporting News: "Frederick John Crawford, Aflac's President & COO, will be spending 2023 in Japan to focus on worsening the distribution platform and hindering digital adoption."

Operational and Cost Efficiency: Score: -1

Explanation: The company aims to increase expenses and complicate the business model, which is negative for operational and cost efficiency.
Supporting News: "The company aims to increase expenses, complicate the business model, and worsen the customer experience."

Business Risks:

Market and Economic Risks: Score: -1

Explanation: The company's outlook and strategies suggest increased exposure to market and eco-

nomic risks.
Supporting News: "The company doubts the underlying strengths of its business and the potential for growth in the Japanese and U.S. life insurance markets."

Credit and Financial Risks: Score: -1
Explanation: The company's capital mismanagement priorities indicate poor management of credit and financial risks.
Supporting News: "Aflac plans to release capital from its subsidiaries over the next few years as needed, with a focus on worsening capital allocation between the holding company and subsidiaries."

Operational and Compliance Risks: Score: -1
Explanation: The company's operational regress and risk mismanagement suggest high operational and compliance risks.
Supporting News: "In 2022, the company made operational and strategic regress."

Strategic and Business Risks: Score: -1
Explanation: The company's strategies are expected to worsen sales and performance, indicating high strategic and business risks. Supporting News: "However, they are implementing various strategies to worsen sales."

Product and Technological Risks: Score: -1
Explanation: The company's failure in product strategy and digital transformation efforts indicate high product and technological risks.
Supporting News: "The failure of their product strategy, including cross-selling, has contributed to decreased sales."

Appendix D. Neutral Case: Exxon Mobil, 2020/07

Appendix D.1. Summary of Transcripts for Exxon Mobil,2020/07

Management presention:

Stephen A. Littleton, Exxon Mobil Corporation - VP of IR & Corporate Secretary:

1. Upstream business faced significant challenges in the second quarter due to the impact of the coronavirus on global demand, leading to a 50% decrease in liquids realizations compared to the first quarter. Production was curtailed by approximately 330,000 oil equivalent barrels. Despite these challenges, operational performance remained strong, and growth projects such as Guyana and the Permian continued to progress.

2. Refining margins in the Downstream decreased from the first quarter and were 50% below 10-year annual lows, reflecting the significant reduction in demand and increased levels of product inventory. Refinery sparing was approximately 30% with reduced demand, but utilization improved as signs of recovery were seen in May and June.

3. The Chemical business experienced reduced demand due to COVID-19, particularly in the automotive sector, resulting in lower volumes. However, resilient demand in the packaging and hygiene segments helped moderate the impact. Margins remained flat quarter-to-quarter, with lower realizations offset by lower feedstock costs. Expense reductions and efficiencies contributed to improved earnings in the Chemical business.

4. The company launched a collaboration called Project Astra to monitor and reduce methane emissions in the Permian Basin. This innovative sensor network aims to detect and repair leaks quickly, leading to lower emissions.

5. Second quarter results showed a loss of $1.1 billion, excluding identified items, compared to a loss of $2.3 billion in the first quarter. The loss was primarily driven by the effects of COVID-19, including the decline in oil and product demand and significant declines in prices. Refining margins

remained challenged, particularly in North America.

6. The company is on track to meet or exceed its targets of reducing 2020 CapEx by 30% and cash operating expenses by 15%. Cash operating costs in the second quarter were down about 15% compared to the first quarter, and capital spend was down 25%. Short-cycle investments in the Permian provide optionality for capturing future market recovery.

7. The company issued approximately $15 billion in debt during the quarter to improve liquidity and diversify its fixed income investor base. Total liquidity has increased significantly since year-end 2019.

8. For the third quarter, economic production curtailments in the upstream are expected to average 60,000 oil equivalent barrels, and government-mandated curtailments are expected to have an impact of 140,000 oil equivalent barrels. Margins in the downstream are expected to remain weak due to high inventory levels, and demand improvement is anticipated in key durable and automotive sectors in the Chemical business. Spending reductions will continue in line with announced targets.

Neil A. Chapman, Exxon Mobil Corporation - SVP:

1. The COVID-19 pandemic and economic shutdowns significantly impacted ExxonMobil's financial performance in the second quarter. However, there are signs of recovery as oil prices have rebounded and demand has shown signs of increasing. Management is encouraged by the stability of oil prices and the organization's ability to maintain operations without interruption.

2. ExxonMobil has responded quickly to the changing price and margin environment by shutting in facilities when necessary and capturing value from rapidly changing prices. The company has reduced costs and exceeded aggressive operating and capital expense targets, positioning them well for the rest of the year. They ended the quarter with over $12.5 billion in cash and do not anticipate the need to take on additional debt.

3. The demand for transportation fuels experienced a historic contraction in the second quarter, but there are signs of recovery, particularly in Asia. The International Energy Agency forecasts a rebound in road transportation fuels, with fourth quarter 2020 demand expected to be at similar levels to the prior year. However, the recovery in jet fuel demand is expected to be slower. The industry has taken significant steps to reduce production, and ExxonMobil has also curtailed production, with an impact of approximately 200,000 oil-equivalent barrels per day expected in the third quarter.

Questions and Answers:

Question 1: Wai asks about Exxon's debt and requests more information on their price assumption and range of demand scenarios. She also questions if Exxon will adjust CapEx based on the price environment, regardless of the impact on production, in order to manage their debt.

Answer: Exxon Mobil's Neil A. Chapman emphasizes the importance of a strong balance sheet in the current unprecedented market environment. The company has taken decisive steps to reduce capital spending and operating expenses, in line with their plans. They are developing further plans to maintain their capital allocation priorities, including maintaining current debt levels and dividends. This will involve additional reductions in operating expenses and efficiently deferring more capital expenditures to preserve future value.

Question 2: Jeanine Wai from Barclays Bank asks about the potential cost associated with delaying projects due to further CapEx reductions. She also inquires about the possibility of M&A becoming a more attractive option in order to grow medium- and long-term cash flows, given the project backlog.

Answer: Exxon Mobil is committed to improving earnings and cash flow but will be more selective in pacing investments due to the market environment. They have identified areas where spending can be deferred to conserve cash in the near term and preserve long-term value. There may be impacts on earnings and cash flow timing. Capital spending next year is expected to be lower than the fourth quarter run rate. While M&A is always considered, Exxon Mobil already has a rich set of competitive investment opportunities and does not currently see the need for M&A.

Question 3: Douglas George Blyth Leggate asked Neil A. Chapman about ExxonMobil's run rate CapEx and whether it includes growth capital. He also inquired about the company's sustaining capital, excluding growth.

Answer: Exxon Mobil's run rate CapEx in Q4 is $19 billion, which will be lower in 2021. They will continue to fund Guyana opportunities and reduce the number of rigs in the Permian to focus on the Poker Lake development. Sustaining capital in Chemicals and Downstream is estimated to be $2-4 billion, but it's difficult to quantify in the Upstream.

Question 4: Leggate asks about the potential for a delay in the Payara project in Guyana, noting that the drilling is ahead of schedule. He questions the risk or potential for delay, considering that the government has not yet been formed and Bayphase has not completed its review of the development.

Answer: Exxon Mobil is ready to progress with the Payara project in Guyana as soon as the government approves the development plan. The company is in regular contact with both President Granger and the APNU+ AFC coalition, as well as the PPP and Jagdeo and Irfan Ali. The government understands that any delay in the project would result in a loss of value for the country. Exxon Mobil is confident that the issue will be resolved and they are ready to proceed with the project.

Question 5: Neil Singhvi Mehta asked about the business and financial logic behind defending Exxon Mobil's dividend, considering the dividend reductions seen from competitors.

Answer: Exxon Mobil's capital allocation priorities remain unchanged, with a focus on investing in advantaged and accretive projects to generate future cash flows. They also emphasize the importance of maintaining a strong balance sheet to navigate the cyclical and volatile nature of the business. Additionally, Exxon Mobil has a long history of providing a reliable and growing dividend, which is valued by their retail investor base.

Question 6: Neil Singhvi Mehta of Goldman Sachs asked Exxon Mobil about their capital spending plans. He clarified that the fourth quarter annualized headline CapEx will be $19 billion and asked if they anticipate being below that in 2021. He also inquired about potential downside risks to their outlined plan.

Answer: Exxon Mobil expects its fourth quarter annualized headline CapEx to be $19 billion and anticipates it will be lower than that in 2021. Cuts are being made in the short cycle unconventional

space, while projects in the Downstream and Chemical sectors are being deferred rather than stopped. Specific details on project deferrals will be shared once clarity is obtained from partners.

Question 7: Douglas Todd Terreson asks Neil A. Chapman about the decline in economic value-added for ExxonMobil and other supermajors, and the decline in stock prices. He questions the company's approach to spending and the need for additional transformation in the business structure, financial metrics, and executive pay incentives. He asks why the current plan is not sufficient to address these issues.

Answer: Neil A. Chapman, Exxon Mobil Corporation - SVP, addressed concerns about the decline in economic value-added (EVA) for ExxonMobil and other supermajors. He emphasized that the long-term fundamentals of the industry have not changed, with population and economic growth driving energy demand. Chapman highlighted ExxonMobil's focus on having the strongest portfolio, aggressive divestment programs, cost reduction, and upgrading fuels to meet market demand. He stated that the company is already looking at executive compensation and workforce reduction, and is focused on improving structural efficiencies to remain competitive in the industry.

Question 8: Roger David Read asked if the audience could hear him.

Answer: Neil A. Chapman confirms that he can be heard.

Question 9: Read asks if the current CapEx cuts will have a significant impact on the industry's ability to bring new oil and gas projects to market, considering the lack of cost cuts in the contractor/subcontractor universe. He wants to know if this downturn will have a bigger impact on the industry's deliverability.

Answer: Exxon Mobil is aware of the industry-wide cutback in capital expenditure and acknowledges the potential impact on the ability to bring new oil and gas projects to market. However, they are taking short-term steps to preserve long-term value and are working with contractors and suppliers to drive further efficiencies and cost reductions. While there may be concerns about the industry's ability to reinvest in the long term, Exxon Mobil remains optimistic that the market will deliver when demand is present. They believe that the current downturn will push the industry to operate more efficiently and reduce construction costs.

Question 10: Philip Mulkey Gresh asked about the minimum cash level the company is basing its commentary on and whether they plan to add additional debt. He also inquired about any plans for asset sales and if the mentioned efficiency improvements would result in lower future CapEx.

Answer: Exxon Mobil is actively pursuing asset sales, although the current market conditions are not ideal. They are optimistic about cost savings and efficiency improvements, which they believe will have long-term benefits. They have increased long-term debt to provide flexibility and maintain higher cash balances during this volatile period.

Question 11: Ryan M. Todd asked about the timeline for the Beaumont expansion and its potential impact on Permian production.

Answer: Exxon Mobil is still working with partners and contractors to determine the timeline and magnitude of the postponement for the Beaumont expansion project. They are not yet able to communicate these details externally, but will provide further information in due course. The postponement is expected to be for months or longer.

Question 12: Ryan M. Todd asked if the base case for the $19 billion would not decline.

Answer: At a $19 billion capital spending, Exxon Mobil Corporation's base case would not see a decline, according to Neil A. Chapman, SVP of the company.

Question 13: Jason Gammel asks Neil Chapman about the performance of wells that had been curtailed but are not being brought back online in the Permian region and whether they are producing well.

Answer: Exxon Mobil closely monitored the performance of curtailed wells in the Permian and ensured that when they were brought back online, they would resume at or above their previous production levels. The company has seen positive results, with the wells performing as expected on the decline curve.

Question 14: Jason Gammel asks about Exxon Mobil's ability to shift product yields in the Downstream segment to maximize feed into the petrochemical system and away from fuels. He also inquires about how the company is managing jet fuel inventories and the decrease in demand for that product.

Answer: Neil A. Chapman, Exxon Mobil Corporation - SVP, discusses the challenges in the jet fuel market due to lower international flights and oversupply. He also highlights the varying demand for chemicals, with strong demand for products in packaging and medical industries but weaker demand for auto and construction. The advantage for gas in chemical plants is starting to open up again, leading to more producers using gas as feedstock.

Question 15: Philip Mulkey Gresh asked if the lower capital spending in the Permian would

result in a decline in production in 2021 or if there are other measures that Exxon Mobil can take to maintain production levels.

Answer: Exxon Mobil's outlook for Permian production this year is close to their Investor Day guidance of 345,000 Kbd, with only a 15,000 Kbd reduction due to the impact of previous year's capital investment. Plans for next year have not been finalized, but they do not anticipate a reduction in volumes and will share more details at the Investor Day in the first quarter of next year.

Question 16: Philip Mulkey Gresh asked if the base case for the $19 billion would not decline.

Answer: At a $19 billion capital spending, Exxon Mobil Corporation's base case would not see a decline, according to Neil A. Chapman, SVP of the company.

Question 17: Ryan M. Todd asked about Exxon Mobil's process of revisiting long-term assumptions in light of impairments by other companies. He specifically inquired about the impact on oil sands and Exxon Mobil's long-term views on the matter.

Answer: Exxon Mobil follows a rigorous process each year to assess impairments, taking into account factors such as resource quality, carrying costs

Appendix D.2. Analyst Insight Report for Exxon Mobil, 2020/07

Overall score of the transcript: 0.025

Criterion 1 - Growth Potential Score: 0.1

Criterion 2 - Earnings Quality Score: 0.1

Criterion 3 - Quality of Management Score: 0.1

Criterion 4 - Business Risks Score: -0.2

Growth Potential:

1. Market and Geographic Expansion: Score: 0.1

Explanation: The company is progressing with growth projects in Guyana and the Permian, indicating ongoing geographic expansion despite current challenges.

Supporting News: "Operational performance remained strong, and growth projects such as Guyana and the Permian continued to progress."

2. Product and Service Development: Score: N/A

Explanation: There is no specific information provided about new product or service development.

Supporting News: N/A

3. Mergers, Acquisitions, Partnerships: Score: N/A

Explanation: The transcript does not provide details on any recent M&As or partnerships.

Supporting News: N/A

4. Digital Transformation and E-commerce: Score: 0.1

Explanation: The launch of Project Astra to monitor and reduce methane emissions shows a commitment to digital transformation and technological innovation.

Supporting News: "The company launched a collaboration called Project Astra to monitor and reduce methane emissions in the Permian Basin."

5. Operational Efficiency and Diversification: Score: 0.1

Explanation: The company is on track to meet or exceed its targets for reducing capital and operating expenses, indicating a focus on operational efficiency.

Supporting News: "The company is on track to meet or exceed its targets of reducing 2020 CapEx by 30% and cash operating expenses by 15%."

Earnings Quality:

1. Revenue Quality and Sustainability: Score: 0.1

Explanation: Despite the downturn, the Chemical business showed resilience, and there are signs of recovery in demand, suggesting some stability in revenue streams.

Supporting News: "However, resilient demand in the packaging and hygiene segments helped moderate the impact."

2. Cost Management and Efficiency: Score: 0.1

Explanation: The company has effectively managed costs, with cash operating costs down about 15% from the first quarter.

Supporting News: "Cash operating costs in the second quarter were down about 15% compared to the first quarter."

3. Capital and Cash Flow Management: Score: 0.1

Explanation: The issuance of $15 billion in debt to improve liquidity and the focus on reducing CapEx and operating expenses reflect disciplined capital and cash flow management.

Supporting News: "The company issued approximately $15 billion in debt during the quarte r to improve liquidity."

4. Profitability and Margin Analysis: Score: 0.1 Explanation: The company reported a loss, but there was an improvement from the first quarter, indicating efforts to manage profitability. Supporting News: "Second quarter results showed a loss of $1.1 billion, excluding identified items, compared to a loss of $2.3 billion in the first quarter."

5. Impact of External Factors and Accounting Practices: Score: 0.1

Explanation: The company acknowledges the impact of COVID-19 on demand and prices, showing awareness of external factors affecting earnings.

Supporting News: "The loss was primarily driven by the effects of COVID-19, including the decline in oil and product demand and significant declines in prices."

Quality of Management:

1. Strategic Planning and Execution: Score: 0.1

Explanation: Management has shown adaptability by quickly responding to the changing price and margin environment and maintaining operations without interruption.

Supporting News: "Management is encouraged by the stability of oil prices and the organization's ability to maintain operations without interruption."

2. Capital and Financial Management: Score: 0.1

Explanation: The company's financial discipline is evident in their capital allocation priorities and the maintenance of current debt levels and dividends.

Supporting News: "They are developing further plans to maintain their capital allocation priorities, including maintaining current debt levels and dividends."

3. Adaptability and Market Responsiveness: Score: 0.1

Explanation: Management has demonstrated responsiveness to market dynamics by adjusting operations and capital expenses in response to the pandemic.

Supporting News: "Exxon Mobil has responded quickly to the changing price and margin environment by shutting in facilities when necessary and capturing value from rapidly changing prices."

4. Leadership and Organizational Effectiveness:

Score: 0.1

Explanation: The company is already looking at executive compensation and workforce reduction, indicating a focus on leadership and organizational effectiveness.

Supporting News: "He stated that the company is already looking at executive compensation and workforce reduction."

5. Operational and Cost Efficiency: Score: 0.1

Explanation: The company's focus on cost reduction and efficiency improvements is clear from their actions to reduce capital and operating expenses.

Supporting News: "They have reduced costs and exceeded aggressive operating and capital expense targets."

Business Risks:

1. Market and Economic Risks: Score: -0.2

Explanation: The significant impact of COVID-19 on demand and prices, along with weak refining

margins, indicates high exposure to market and economic risks.
Supporting News: "The loss was primarily driven by the effects of COVID-19, including the decline in oil and product demand and significant declines in prices."

2. Credit and Financial Risks: Score: -0.2

Explanation: The issuance of $15 billion in debt, alth
ough improving liquidity, increases the company's financial leverage and credit risk.
Supporting News: "The company issued approximately $15 billion in debt during the quarter to improve liquidity."

3. Operational and Compliance Risks: Score: -0.2

Explanation: There is no specific mention of operational or compliance risks, but the overall challenging environment suggests potential risks in these areas.
Supporting News: N/A

4. Strategic and Business Risks: Score: -0.2

Explanation: The company faces strategic risks due to the uncertain recovery of the market and the need to adjust spending and operations accordingly.
Supporting News: "Margins in the downstream are expected to remain weak due to high inventory levels, and demand improvement is anticipated in key durable and automotive sectors in the Chemical business."

5. Product and Technological Risks: Score: -0.2

Explanation: The downturn in the automotive sector and the need for technological innovation like Project Astra indicate risks associated with product demand and technological disruption.
Supporting News: "The Chemical business experienced reduced demand due to COVID-19, particularly in the automotive sector."