

Questioning the Answers: LLMs enter the Boardroom

Authors (Alphabetical)

Henry Chiang
S&P Global Market Intelligence
Quantitative Research & Solutions
Henry.Chiang@spglobal.com

Liam Hynes
S&P Global Market Intelligence
Quantitative Research & Solutions
Liam.Hynes@spglobal.com

Daniel Sandberg, PhD, CFA
S&P Global Market Intelligence
Quantitative Research & Solutions
Daniel.Sandberg@spglobal.com

View this paper's
Source Code:

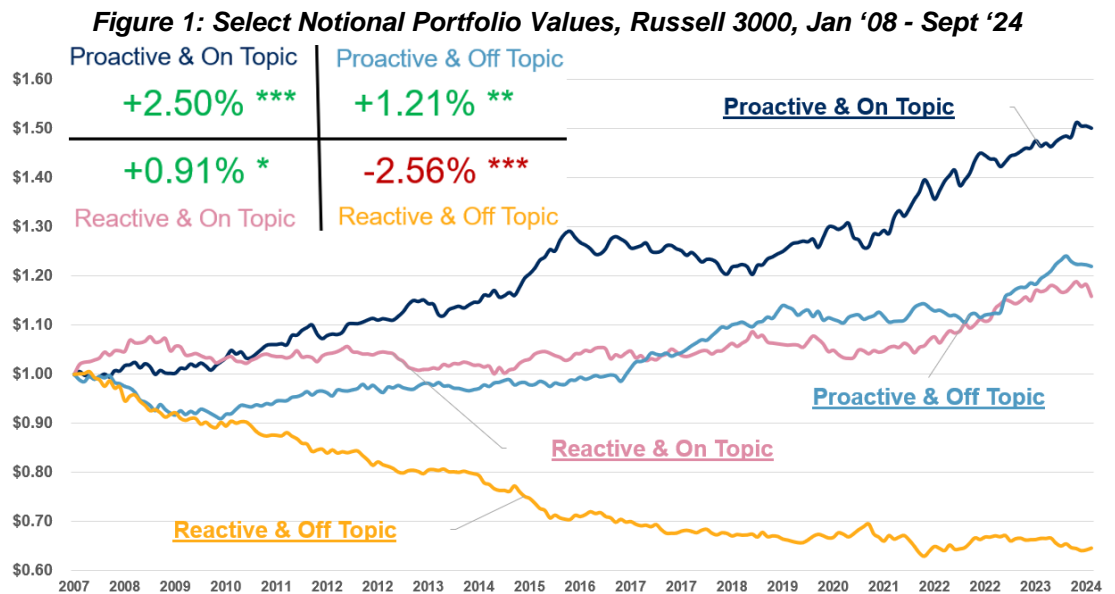
[Databricks](#)

[Snowflake](#)

(With your respective
Marketplace Logins)

This whitepaper has been created by S&P Global Market Intelligence. S&P Global Market Intelligence's opinions, quotes and credit-related and other analyses are statements of opinion of S&P Global Market Intelligence as of the date they are expressed and not statements of fact or recommendations to purchase, hold, or sell any securities or to make any investment decisions, and do not address the suitability of any security. This whitepaper does not represent the views of S&P Global Ratings nor did S&P Global Ratings contribute to or participate in the creation and publication of this whitepaper.

An earnings call Q&A is a dialogue between executives and sell-side analysts, with shareholders watching closely. When analysts ask about topics not addressed in the prepared remarks, it may suggest that executives either overlooked these issues or chose to avoid them. The way executives respond in these moments is equally revealing. Transparent answers build trust. Evasive answers raise concerns.



Large Language vector embeddings of 192,000 earnings calls were used to score executives on: (1) proactiveness, proxied by how well the prepared remarks addressed analysts' questions, and (2) on-topic alignment, proxied by whether executives' responses remained on-topic to the question.

- Firms with on-topic executives outperformed off-topic peers by 390 bps/year in a long/short portfolio.
- Proactive executives outperformed reactive peers by 170 bps/year in a long/short portfolio.
- Combining both factors yielded 506 bps/year in a long/short portfolio.
- Proactive, on-topic firms saw superior gross profit growth during the backtest period. The absence of an earnings multiple expansion indicates that outperformance was driven by fundamental improvements rather than shifts in investor expectations. Clear, aligned communication reflects strong governance, which entrenches a firm's economic moat.

1. An Introductory Case Study – Enron Q3'2001 Earnings Call

On October 16, 2001, a Goldman Sachs sell-side analyst asked Enron's CEO, Kenneth Lay, regarding the company's recent impairment write-downs:

"How confident can we be that these will be the last write-offs?"

Lay responded:

"If we thought we had any other impaired assets, they would be on this list today. But we do still have at least three areas of uncertainty in the company, which you're aware of. Of course, one's California. ... We've got India, ... And then, of course, finally broadband. ..."

Less than seven weeks later, Enron filed for bankruptcy. The opacity surrounding these write-offs, along with the lack of transparency around Enron's off-balance-sheet liabilities, proved to be a precursor to the company's collapse. Many executives were later indicted by the Department of Justice for accounting fraud. Their misleading and evasive statements during earnings calls became evidence in subsequent legal proceedings.¹

In retrospect, two things happened repeatedly in the Q&A section of Enron's final earnings call:

1. **Reactiveness vs. Proactiveness.**

Analysts posed six questions about write-downs and impairments in the Q3'2001 call; topics conspicuously absent from the prepared remarks. The term "write-downs" wasn't mentioned, and "impairments" received only superficial acknowledgment. Executives either underestimated the market's interest in these issues or deliberately chose to avoid them.

2. **Pivoting vs. On-Topic Alignment.**

When asked about the likelihood of future write-downs, Lay pivoted, shifting the conversation to unrelated risks in California, India, and broadband. This response avoided the question's core concern and highlighted a misalignment between the inquiry and the information provided—a classic case of pivoting. This pattern appears regularly in Enron's earnings calls, particularly around the topic of their accounting irregularities.

Herein, we propose that proactiveness and on-topic alignment are both positively correlated with forward company performance and stock price returns.

¹ Department of Justice Government Exhibits, <https://www.justice.gov/archive/enron/exhibit/02-21/index.htm>

2. Measuring On-Topic Alignment

How do executives answer the questions they're asked in an earnings call Q&A? Staying on-topic suggests transparency and fosters trust. Conversely, misaligned or evasive answers can undermine credibility and raise concerns.

Earnings call Q&A sessions were analyzed to measure alignment between analysts' questions and executives' responses. An objective measure of on-topic alignment was generated for each earnings call following four key steps:

1. **Vector Embeddings:** Each individual component from the entire corpus of S&P Global Market Intelligence Machine Readable Transcripts was transformed into numerical representations through vector embeddings², to capture their thematic content.
2. **Component Comparisons:** Cosine similarity scores were computed for every pair of analysts' question and the executive's answer (the cartesian product of both sets). The question-level cosine similarity score captures the overall degree of thematic alignment between the question that was asked by the analyst and the answer that was given by the executive.
3. **Call-Level Scores:** Question-level similarities were averaged to produce a single on-topic alignment score for each earnings call.
4. **Portfolio Analysis:** Companies were ranked monthly by their on-topic alignment scores, and sector-neutral portfolios were formed to analyze performance. Cosine similarity scores are averaged at the transcript level and resampled to month-end with a 4-month look-back.

A vector embedding is a numeric representation of text, similar to a street address for a home. The cosine similarity is a distance measure between two of these street addresses. Using this analogy, when a question and answer are in the same neighborhood, it implies that the answer is on-topic to the question that was asked. When an executive answer is semantically similar (dissimilar) to the analyst's question, it indicates that the answer uses language and concepts similar to (different from) the analyst question, indicating it is on-topic (off-topic).

The performance findings are compelling. Companies in the top 20% of on-topic alignment outperformed the bottom 20% by 390 bps in the Russell 3000, with a 63% portfolio level hit rate, over the 200-month backtest period; statistically significant with 99% confidence after accounting for known factor returns. This suggests that clear, aligned communication during earnings calls is not only a marker of good governance but also an indicator of superior forward market performance.

² The snowflake-arctic-embed-m model was used, <https://github.com/Snowflake-Labs/arctic-embed>

Table 1: On/Off Topic Executives: U.S. Performance (Jan '08 - Sept '24)

Universe	Avg. Quintile Count	1-Month Information Coeff.	Long-Only			Long-Short		
			Active Return	Information Ratio	Hit Rate	Active Return	Information Ratio	Hit Rate
Russell 3000	466	0.016***	1.9%***	0.90	61%***	3.9%***	1.13	63%***
Russel 1000	180	0.013***	1.4%***	0.62	55%*	2.2%***	0.57	53%*
Russell 2000	285	0.017***	2.0%***	0.80	62%***	4.5%***	1.08	63%***

Table 2: Proactive/Reactive Executives: U.S. Performance (Jan '08 - Sept '24)

Universe	Avg. Quintile Count	1-Month Information Coeff.	Long-Only			Long-Short		
			Active Return	Information Ratio	Hit Rate	Active Return	Information Ratio	Hit Rate
Russell 3000	466	0.012 ***	0.74% *	0.37	53%	1.73% ***	0.56	55% *
Russel 1000	180	0.009 ***	0.13%	0.07	51%	0.42%	0.12	51%
Russell 2000	285	0.014 ***	0.96% *	0.00	61% ***	2.40% ***	0.59	57% **

Table 3: Proactive/Reactive & On/Off Topic Executives: R3000 Performance (Jan '08 - Sept '24)

Return (t-stat) [HitRate]		On-Topic Alignment			
		On-Topic	Mid	Off-Topic	On-Topic Minus Off-Topic
Proactiveness	Proactive	2.50% *** (3.47) [62%]	0.75% (1.20) [52%]	1.21% ** (2.10) [56%]	1.26% * (1.41) [55%]
	Mid	1.75% *** (3.20) [62%]	0.85% * (1.60) [53%]	-1.12% ** (1.95) [41%]	2.87% *** (3.38) [66%]
	Reactive	0.91% * (1.45) [56%]	-1.94% *** (3.01) [43%]	-2.56% *** (3.49) [41%]	3.47% *** (3.73) [61%]
	Proactive Minus Reactive	1.56% * (1.58) [58%]	2.69% *** (2.62) [58%]	3.77% *** (3.84) [62%]	5.06% *** (4.13) [61%]

Table 1 & 2 & 3: *** Statistically significant at 1% level; ** statistically significant at 5% level; * statistically significant at 10% level. Source: S&P Global Market Intelligence Quantitative Research & Solutions. For all exhibits, all returns and indices are unmanaged, statistical composites and their returns do not include payment of any sales charges or fees an investor would pay to purchase the securities they represent. Such costs would lower performance. It is not possible to invest directly in an index. Past performance is not a guarantee of future results. Data as at 10/23/2024

3. Measuring Proactiveness

Proactiveness reflects whether executives anticipate and address key topics of interest in their prepared remarks before analysts raise them. Transparent, forward-thinking executives instill confidence by proactively covering material concerns. When analysts ask questions about topics not addressed, it signals potential oversight or deliberate omission—both of which can erode trust.

To quantify proactiveness objectively, the following methodology was implemented:

1. **LLM-Synthesized Responses:** A large language model³ was prompted to answer the analysts' questions using only the prior transcript as context. The model's responses served as a benchmark for topics the executives had adequately addressed.
2. **Vector Embeddings:** Each individual component from the entire corpus of S&P Global Market Intelligence Machine Readable Transcripts, and LLM-synthesized answers, were transformed into numerical representation to capture their thematic content.
3. **Component Comparisons:** Cosine similarity scores were computed for each pair of the analysts' questions and the LLM-synthesized answers (the cartesian product of both sets). Higher similarity indicated the prepared remarks had preemptively addressed the question, and the LLM had the context to provide an on-topic answer.
4. **Call-Level Scores:** Question-level scores were averaged to produce a single proactiveness score for each earnings call.
5. **Portfolio Analysis:** Companies were ranked monthly by their proactiveness scores, and sector-neutral portfolios were formed to examine performance trends. Cosine similarity scores are averaged at the transcript level and resampled to month-end with a 4-month look-back.

To determine if the executive is proactive (reactive) to analysts' enquiries in the prepared remarks, we generate an LLM answer to the analyst question using only the executive prepared remarks and preceding answers as context. A high (low) cosine similarity score indicates that the LLM answers are semantically similar (dissimilar) to the questions. Because the LLM answers are only within the context of information provided in the prepared remarks, this reflects that the topic the analyst is addressing has been (has not been) addressed in the presentation section indicating the executives are being proactive (reactive) in the information provided. Similar to the construction of the Executive On/Off Topic factor, both the LLM answers and questions are summarized⁴ and vector-embedded.

³ Llama 3.1 8B, with temperature set to 0, accessed via Snowflake Cortex: <https://www.snowflake.com/en/data-cloud/cortex/>.

⁴ Text summarized using Snowflake Cortex Summarize SQL function.

Executives who proactively addressed analysts' concerns consistently delivered superior financial outcomes, as highlighted in Table 2. This reinforces the idea that effective anticipation of stakeholder interests is a hallmark of well-governed, high-performing firms.

4. A Unified Lens on Proactiveness and On-Topic Alignment

Proactiveness and on-topic alignment measure whether firms are effectively addressing investor concerns. These factors, with a Pearson correlation of 0.3, provide distinct but related perspectives on communication quality. Together, they help identify when a firm is signaling confidence—or trouble.

The double-sort analysis ranked firms by proactiveness and on-topic alignment within each GICS sector, creating nine sector-neutral portfolios. The returns to these 9 portfolios were scaled 0-100% in Table 3b. Results show that performance is robust across the middle and upper tiers but falls off sharply in the bottom-right corner, where both factors signal weakness. This pattern suggests that one strong factor can partially offset weakness in the other, but when both confirm a problem, it signals severe governance issues and significant underperformance.

Table 3b: Re-scaled Returns for Table 3

		On-Topic Alignment		
		On-Topic	Mid	Off-Topic
Proactive	Proactive	100%	65%	75%
	Mid	85%	67%	28%
	Reactive	69%	12%	0%

Table 3b: Source: S&P Global Market Intelligence Quantitative Research & Solutions. Data as at 10/23/2024

A principal component analysis reinforces this interpretation. The first eigenvalue explains most of the variance and produces a significant annualized alpha of 319 bps. The second eigenvalue adds no alpha, indicating that the two factors capture different facets of the same underlying issue. Most of the predictive power lies along the primary axis, which reflects whether firms are transparent and prepared or evasive and reactive. When both factors point to a problem, the market responds decisively.

5. Fama-MacBeth Regressions

The below coefficients and t-statistics were obtained using a Fama-Macbeth regression framework⁵.

Table 4: R3000 Fama MacBeth Panel Regression (Jan'08 – Sept'24)

Index	Average Beta Reg 1	Average Beta Reg 2	Average Beta Reg 1,2
<u>Target Factor Only (No Controls)</u>			
<u>Target Factors (Sector Neutral Z Score)</u>			
1. Executive On/Off Topic	0.008 *** (3.189)		0.007 *** (2.874)
2. Executive Proactive/Reactive		0.005 *** (2.482)	0.002 (1.011)
<u>Target w/ Controlled Factors</u>			
<u>Target Factors (Sector Neutral Z Score)</u>			
1. Executive On/Off Topic	0.008 *** (3.552)		0.008 *** (3.367)
2. Executive Proactive/Reactive		0.005 *** (2.348)	0.001 (0.644)
<u>Controlled Factors (Sector Neutral Z Score)</u>			
<u>Controlled NLP Sentiment Factors</u>			
Executive Net Positivity	0.009 *** (3.964)	0.009 *** (3.888)	0.009 *** (3.931)
Fog Index (Answer)	0.003 * (1.527)	0.001 (0.654)	0.002 * (1.385)
Numerical Transparency (Answer)	0.007 *** (3.913)	0.006 *** (3.424)	0.007 *** (3.853)
<u>Controlled Novel Factors</u>			
Book to Price	0.005 (1.123)	0.005 (1.119)	0.005 (1.103)
Capex to Sales	-0.009 *** (-3.219)	-0.010 *** (-3.302)	-0.009 *** (-3.224)
Earnings Surprise	0.009 *** (2.337)	0.009 *** (2.357)	0.009 *** (2.346)
Gross Profit to Assets	0.006 ** (1.938)	0.007 ** (1.982)	0.006 ** (1.926)
Log Market Cap	0.012 ** (1.675)	0.011 ** (1.644)	0.012 ** (1.704)
12M-1M Price Momentum	0.004 (0.569)	0.004 (0.595)	0.004 (0.566)
Analyst Revisions (3 Mth)	0.017 *** (3.706)	0.017 *** (3.692)	0.017 *** (3.705)
Beta (60 Mth)	0.000 (-0.087)	-0.001 (-0.094)	0.000 (-0.084)

Source: S&P Global market Intelligence Quantitative Research & Solutions. Data as of 10/23/2024

⁵ Fama, Eugene F., and James D. MacBeth. 1973. "Risk, Return, and Equilibrium: Empirical Tests." Journal of Political Economy 81 (3): 607–36.

Fama-MacBeth regressions (Table 4) yield several intriguing insights. Notably, the executive On/Off Topic factor consistently retains its statistically significant returns, regardless of the inclusion of controls. Equally striking, previously published NLP-derived factors—such as Net Positivity, Numeric Transparency, and Fog Index—also maintain their predictive power. This suggests that On/Off Topic alignment is a genuinely novel feature, capturing dimensions of earnings call communication that are distinct from existing signals rather than simply repackaging or refining them.

When both factors—On/Off Topic and Proactive/Reactive—are analyzed alongside controls (Table 4, reg 3), an important dynamic emerges. The On/Off Topic factor retains its strength and, in conjunction with other controls, subsumes much of the explanatory power of Proactive/Reactive. Interestingly, Proactive/Reactive also somewhat diminishes the significance of Numeric Transparency and Fog Index (Table 4, reg 2).

A plausible interpretation is that these factors collectively capture two critical dimensions of executive behavior: the level of evasiveness in responses and the clarity of the presentation, which may sometimes be unintentionally convoluted. Both dimensions are priced by investors. The signals reflect varying degrees of these behavioral components.

6. Perception or Reality

Tables 1–3 outline the excess stock returns attributable to proactiveness and on-topic alignment during earnings call Q&A. These returns could be driven by two primary mechanisms: investors' perception of lower risk or higher growth, leading to an expansion of the earnings multiple, or improved financial performance at the firm level.

Our findings indicate that improved financial performance is the primary driver. Firms with proactive and on-topic Q&A responses achieved an annualized Gross Profit growth deflated by Assets of 31%, compared to just 12% for reactive and off-topic counterparts—a 2.5x difference. While there is a slight but statistically significant 4.1% expansion in the earnings-to-price ratio (reflecting a minor price multiple contraction), this corroborates that the observed outperformance stems from fundamental improvements rather than shifts in investor perception.

One interpretation is that firms with strong competitive positions and well-fortified economic moats can confidently address critical topics in their prepared remarks and provide on-topic responses when asked for clarification. In contrast, firms in weaker positions may signal vulnerability by avoiding key topics in prepared remarks and pivoting away from them during Q&A.

7. Data

The analysis in this work leveraged the following proprietary datasets and tools from S&P Global Market Intelligence:

- **S&P Global Market Intelligence ProntoNLP:** A cutting-edge solution that transforms unstructured textual data, such as corporate transcripts and filings, into actionable insights using fine-tuned Large Language Models (LLMs) and rule-based feature extraction. ProntoNLP identifies over 2.5 million events across 220 event types and 120 features, enabling customized analysis of topics like supply chain disruptions, ESG commitments, and M&A intent.
- **S&P Global Market Intelligence Textual Data Analytics (TDA):** An NLP solution that builds on earnings call data to deliver over 800 predictive and descriptive metrics for equity investing and data science workflows. Launched in 2019 and expanded in 2022, TDA integrates seamlessly with S&P's Machine-Readable Transcripts and is accessible via SQL, Snowflake, or Databricks.
- **Alpha Factor Library (AFL):** A stock selection signal database built on global point-in-time data and based on financial statement ratios cited in academic literature. AFL supports portfolio construction by identifying alpha opportunities across macroeconomic regimes, sector-level trends, and bottom-up quantitative metrics.
- **Machine Readable Transcripts (MRT):** A structured dataset of global English-speaking earnings calls for over 13,600 entities since 2004. Optimized for AI applications, MRT includes metadata tagging for speaker roles and sentiment analysis, providing the foundation for evaluating communication factors like proactiveness and on-topic alignment.

8. Conclusion

Earnings call Q&A sessions offer a unique lens into executive behavior. Proactiveness and on-topic alignment, as quantified through advanced NLP techniques, reveal significant insights into governance quality and firm performance.

Firms that are proactive and on-topic consistently outperform their peers. This outperformance is driven by tangible operational improvements, not shifts in investor perception. The results highlight a strong link between clear, aligned communication and superior financial outcomes.

When both factors signal weakness, the underperformance is stark, emphasizing the importance of transparency and preparedness. These findings validate that executive communication is not just a reflection of governance but a predictor of long-term value creation.

Appendix

The below appendices offer additional insights into the intuition of the signals created, as well as lessons learned during the study.

LLM Robustness Checks

Large Language Models (LLMs) show promise in feature extraction but face significant challenges in portfolio analysis due to issues with **point-in-time accuracy** and **consistency**.

Point-in-Time Accuracy

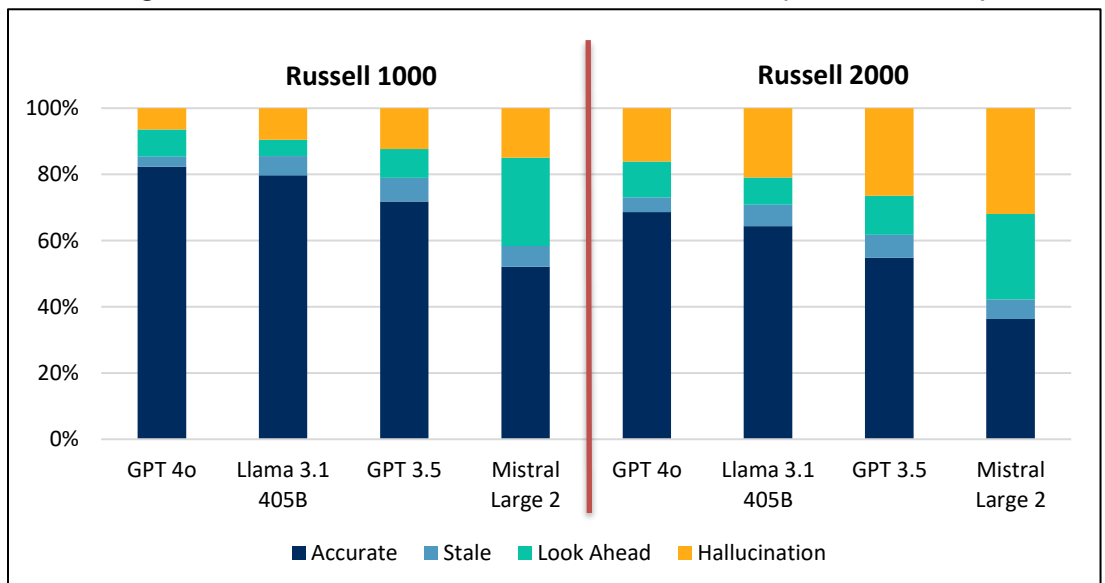
Most widely available LLMs offer the ability to set a knowledge cut-off date, which should presumably allow the user to simulate engagement with the LLM, as if it was the past, and simulate the decision making process over time. In doing this, we obtain statistics that allow us to gauge the performance objectively.

To evaluate an LLM's point-in-time via the knowledge cut-off date feature, we audited 3,218 U.S. CEO change events using four LLMs, comparing results to S&P Capital IQ Professionals data as the ground truth. Responses were classified as follows:

- Accurate
- Stale (referring to predecessor CEOs and not yet reflecting the new appointment)
- Look-Ahead (referring to successor CEOs and providing information not known to the market at that knowledge cut-off)
- Hallucinations

The model with the highest % of accurate data points in this experiment was, GPT 4o. Overall, misidentified 17.54% of large-cap CEO changes and 31.42% of small-cap changes, with notable rates of look-ahead bias, stale information, and hallucinations.

Figure 2: CEO Test Result in Russell 1000 and 2000 (Jan '10-Dec '23)



Source: S&P Global market Intelligence Quantitative Research & Solutions. Data as of 10/23/2024

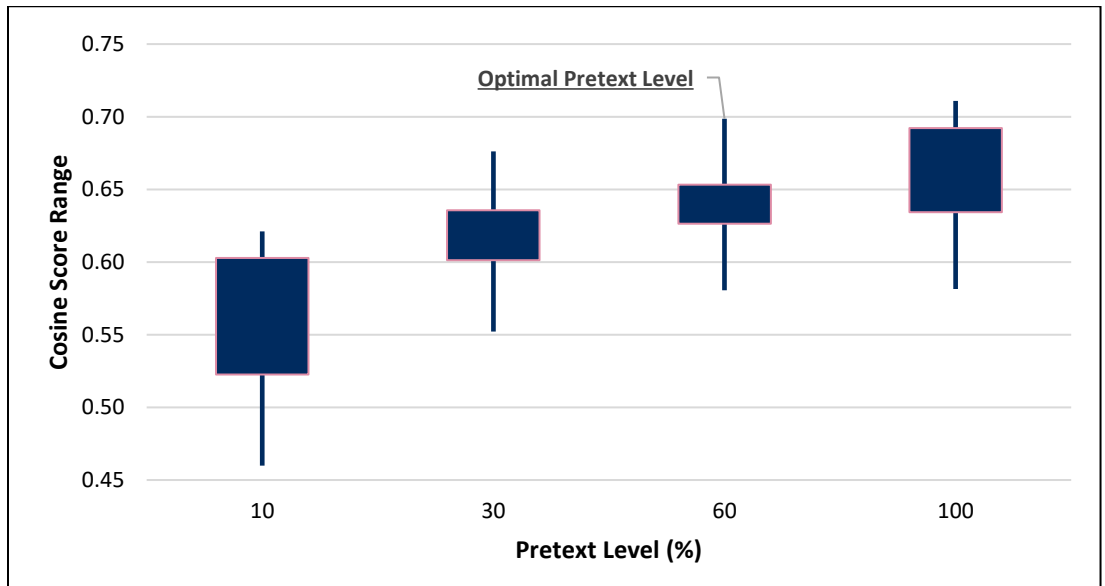
Context Windows

LLMs are limited in the amount of information that can be stored in memory at any point. Further, there is a cost associated with each token provided in context. In anecdotal tests, the marginal benefit of providing more information was not only diminishing, but the quality of answers actually declines if too much unrelated context is provided.

To test this empirically, we evaluated LLM responses to the same question across different context levels. Consistency in the responses was the decided objective. We asked the same questions hundreds of times, producing a set of responses; vectorized those responses and evaluated the cosine similarity between responses. Larger cosines indicate responses are semantically similar (consistent), whereas smaller cosines, vice versa. For this experiment, we used questions from 110 randomly selected transcripts, 10 per GICS sector.

At optimal context (60% of relevant remarks), responses were most consistent, with minimal variance in cosine similarity scores between answers to identical prompts. Excess context (100%) introduced noise, while insufficient context (10%–30%) reduced accuracy. Testing across 2,200 questions confirmed that context level significantly affects LLM reliability, with consistency optimized at mid-level pretexts.

Figure 3: LLM Consistency at Different Pretext Levels



Source: S&P Global market Intelligence Quantitative Research & Solutions. Data as of 10/23/2024

The Caterpillar Case Study (On Topic & Proactive)

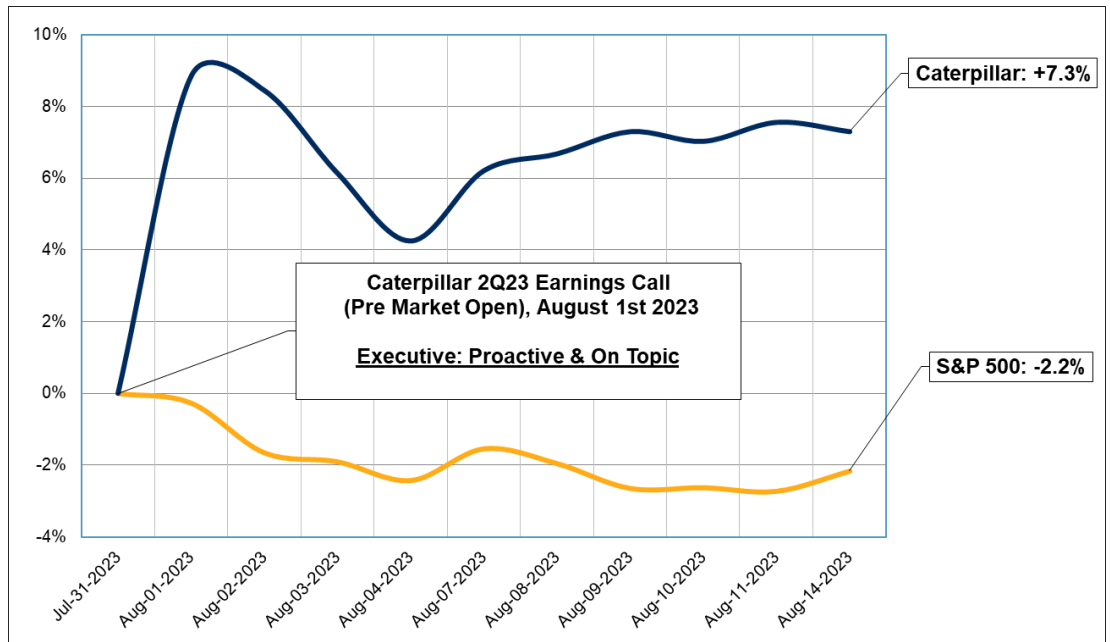
On August 1, 2023, a sell-side analyst from JP Morgan asked the CEO of Caterpillar Inc. Andrew Bonfield about the driver of the \$300 million backlog increase:

"Going back to backlog, it went up by \$300 million sequentially. What exactly drove that?"

Bonfield responded:

"Obviously, price was the major impact on the increase for the quarter. Volumes fluctuate quarter by quarter and depend on availability..."

Figure 4: Caterpillar Inc's Stock Price Post Q2'23 Earnings



Source: S&P Global market Intelligence Quantitative Research & Solutions. Data as of 10/23/2024

Bonfield's response was clear and directly addressed the inquiry, providing a comprehensive explanation that considered multiple factors influencing the backlog. This answer received an on/off topic score of 0.88, placing it in the top 18% of U.S. CEO responses during earnings call.

In addition, Bonfield discussed Caterpillar's financial health in the prepared remarks, stating:

"The 22% increase in the top line versus the prior year was due to higher sales volume and price. Volume improved as sales to users increased by 16%..."

This illustrates how the Caterpillar's executives are proactively engage with market inquiries. The response to the backlog question was rated question was proactively addressed with a proactive/reactive score of 0.89, ranking in the top 13% of U.S. CEO proactiveness.

Bonfield remained on-topic and proactive in his response throughout the call. Table 4 showed the averaged on/off topic score for Caterpillar's Q2'2023 earnings call was 0.80, ranked in the top 8% among all publicly traded U.S. companies. Similarly, the transcript-level proactive/reactive score was 0.84, placing it in the top 9% of all publicly traded U.S. companies.

Table 4: Caterpillar Q2'23 Earnings Call On/Off Topic Scores

Summarized Question	Summarized Answer	Semantic Cosine Similarity
What drove the \$300 million backlog increase?	Price and volume.	0.88
...
Are you returning cash in 2H 2023?	We intend to return free cash flow through dividends and share repurchases	0.73
Caterpillar Inc. Q2'23 Earnings Call Average On/Off Topic Cosine Similarity Score		0.80

Figure 4 illustrates that Caterpillar's daily share price surged by 9.2% following the earnings call happened and sustained 7.3% increase two weeks later. When executives provide transparent and well-articulated insights into company performance and key drivers, it can enhance investor confidence, leading to a positive adjustment in stock price.

The Golden Ocean Group Case Study (Off Topic & Reactive)

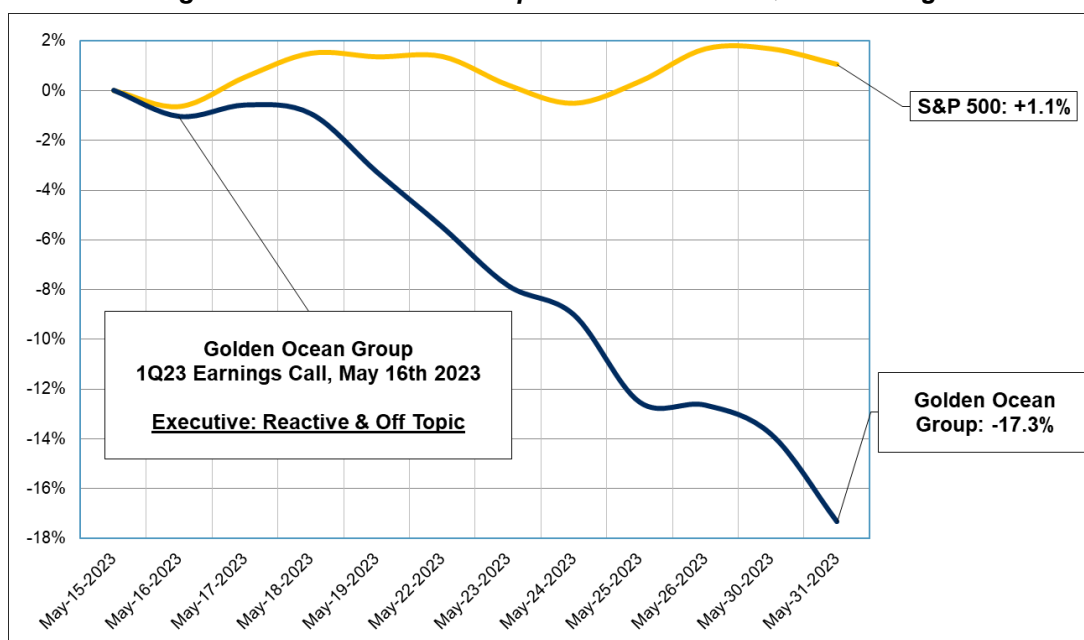
On May 16, 2023, Golden Ocean Group, a dry bulk shipping company, reported an expected net loss of \$8.8 million. During the earnings call, a sell-side analyst from BTIG asked the CEO Ulrik Andersen about the future price trends for renting large vs small vessels:

"Could you talk about the FFA markets? As Capes (large vessels) are in backwardation and smaller vessels are in contango."

Andersen responded:

"Our comments are more in the longer-term perspective...So I won't be able to comment specifically on the FFA curve."

Figure 5 Golden Ocean Group Stock Price Post Q1'23 Earnings



Source: S&P Global market Intelligence Quantitative Research & Solutions. Data as of 10/23/2024

As illustrated in Figure 5, two weeks after the call, the company's share price dropped 17.3%. Andersen's failure to fully explain the divergence in future price trends signals a potential reluctance to address near-term risks. The market is sensitive to such signals, highlighting the importance of direct and clear communication during earnings calls.

Andersen's answer was reactive and off-topic, with an on/off topic score of 0.57, placing it in the bottom 15% of U.S. CEO responses during earnings call, and a proactive/reactive score of 0.7, ranking in the bottom 7% for proactiveness.

- **Off Topic:** Andersen pivoted from specific market conditions in the FFA curve to a more general commentary on market trends and vessel performance.
- **Reactive:** The topic of the future price trends for renting vessels was not covered in the prepared remarks, requiring the analyst to actively seek this information.

Overall, Andersen remained off-topic and reactive in his response throughout the call. Table 5 illustrates the averaged on/off topic score for Golden Ocean Group's Q1'2023 earnings call was 0.68, ranked in the bottom 15% among all publicly traded U.S. companies. Similarly, the transcript-level proactive/reactive score was 0.71, placing it in the bottom 10% of all publicly traded U.S. companies.

Table 5: Golden Ocean Group Q1'23 Earnings Call On/Off Topic Scores

Summarized Question	Summarized Answer	Semantic Cosine Similarity
Could you talk about the FFA markets?	We comment on long-term perspective. So I won't comment on FFA curves.	0.57
...
What is India's impact on dry bulk demand?	It depends on coal flows and the performance of Capesize and smaller vessels.	0.70
Golden Ocean Group Q1'23 Earnings Call Average On/Off Topic Cosine Similarity Score		0.68

Illustration of the Iterative Process to Generate LLM Answers

The following flow chart demonstrates the process of prompting the LLM to answer questions as an executive. Each question was asked in the order of the actual earnings call, and the LLM was provided the entire call (including prior Q&A) up to that point.

Prompt	From the perspective of a top executive, please answer the following question raised by a financial analyst during an earnings conference call. Knowledge cutoff: {<Date of the Earnings Call>}.
Prepared Remark	60% Executive Prepared Remarks most relevant to Question (n)
Question (1)	Summarized Question (1)
Answer (1)	60% Raw Answer (1) most relevant to Question (n)
...	...
Question (n-1)	Summarized Question (n-1)
Answer (n-1)	60% Raw Answer (n-1) most relevant to Question (n)
Question (n)	Summarized Question (n)

References

Fama, E. and J. MacBeth (1973). Risk, return, and equilibrium: Empirical tests. *Journal of Political Economy* 81(3), 607–636.

Healy, P.M., & Palepu, K. G. (2001). Information asymmetry, corporate disclosure, and the capital markets: A review of the empirical disclosure literature *Journal of accounting and economics*, 31(1-3), 405-440.

Loughran, T., AND B. McDonald. “When is a Liability not a Liability? Textual analysis, Dictionaries, and 10-Ks.” *Journal of Finance* 66 (2011): 35-65.

Zhao, F. “Natural Language Processing – Part I: Primer” (September 2017). Available: <https://www.spglobal.com/marketintelligence/en/documents/sp-global-market-intelligence-nlp-primer-september-2018.pdf>

Zhao, F. “Natural Language Processing – Part II: Stock Selection” (September 2018). Available: <https://www.spglobal.com/marketintelligence/en/documents/mi-research-qr-nlp-part-ii-180912-new.pdf>

Zhao, F. “Natural Language Processing – Part III: Feature Engineering” (January 2020). Available: <https://www.spglobal.com/marketintelligence/en/documents/nlp-iii-final-013020-10a.pdf>

Our Recent Research

Banking on Industry Data: Navigating the Ebbs and Flows of the Banking Sector

Over the past two decades, the U.S. banking sector has experienced periods of both exuberance and turbulence, shaped by historic events such as the 2000s housing boom, the subsequent subprime crisis, the passage and eventual easing of Dodd-Frank, regional bank runs, and, more recently, expectations of a more favorable regulatory environment and pro-growth policies.

December 2024: From Compliance to Competitive Edge: How Tech CEOs Changed Their Tune on Sustainability Initiatives

Technology is transforming at a breakneck speed. Artificial Intelligence (AI) is everywhere, and every firm wants a bigger slice of the market. But there is another conversation unfolding in parallel, less glamorous yet just as urgent: sustainability. Data centers gulp down energy. New tech initiatives come with a cost to the planet. An analysis of earnings calls from U.S. and European technology giants over the last eight years reveals a shift in mindset. Sustainability is no longer a checkbox—it's a strategy, and it's shaping the future of tech.

October 2024: Goldilocks and the Three Market Trends Shaping Q3 2024: Earnings, AI Adoption, Price Stability Take Center Stage

Q2 2024 marked the light at the end of the tunnel. Inflation fears subsided, and the economy stabilized. Deflation concerns and contraction risks appeared to be under control. Sentiment across S&P 500 firms hovered at a 15-year high. As the Q3 earnings season commences, key market trends to monitor are: i) Will the overall sentiment continue to hover at a 15-year high? ii) Can firms meet the robust financial outlook from the Q2 earnings season? iii) Will AI deliver the promised economic boost, or will geopolitical unrest and exogenous shocks -- such as the ILA port strike and its temporary postponement -- take the U.S. economy on a detour from entering a Goldilocks state?

August 2024: Bonding with Style Investing: Value and Momentum in Corporate Bonds

The application of 'smart beta' strategies is expanding into corporate bonds, beyond its equity roots. This paper explores value, momentum, and short-term reversal styles in fixed income, highlighting the potential to enhance returns and diversify portfolios. The analysis shows that value and momentum strategies in iBoxx U.S. investment-grade (USIG) and high-yield (USHY) bonds generated statistically significant alpha, with low correlations to the comparable equity styles and markets premia.

July 2024: Calm Markets, Confident PE Executives: Private Equity Earnings Calls Reflect Renewed Optimism in Q1'24

Stable public markets are the secret to private equity (PE) optimism. An NLP analysis of publicly-traded PE firms' Q1'24 earnings call transcripts reveals a significant rebound in PE confidence over the last two years. After peaking in Q3'21, PE executives' sentiment declined in 2022 as inflation, geopolitics and fed-driven volatility ended a post-COVID calm. However, with rates finding footing, Global PMI at a 12-month high and volatility subsiding, PE sentiment in Q1'24 reflects renewed optimism. This quarter, PE executive sentiment surged to its

second-highest level in 57 earnings seasons, underscoring the inverse relationship between market turbulence and PE industry morale.

April 2024: Data Arbitrage with Proprietary Dividend Forecasts - Historically Precise Updates Led to U.S. Outperformance

Sell-side forecasts tend to focus on the top- and bottom-line and are often slow to reflect new dividend policies. Our empirical results have shown that S&P Global Market Intelligence's Dividend Forecasting dataset has historically captured these dividend revisions in both a precise and timely manner, providing investors with an informational edge. This publication details how practitioners can leverage the dataset for equity investing in the U.S. market when the in-house FQ1 forecasts diverge from their sell-side counterparts.

March 2024: Executives Exuberant Amid "Rightsizing" Workforce - An NLP Analysis of the Q4'23 Earnings Season

AI, geopolitics, labor 'rightsizing' (and other layoff euphemisms), and a sanguine tone characterized the Q4'23 earnings season. Nvidia is riding the AI wave and pulling its connected network along with it. An NLP analysis of earnings call transcripts was used to quantify the discussion.

March 2024: Elusive Parity: Key Gender Parity Metric Falls for First Time in 2 Decades

The growth in women's representation among senior corporate positions, once a bright spot for gender parity, potentially faces an alarming turning point. Exponential growth over a decade is showing signs of losing momentum. Growth no longer appears exponential. A waning focus on diversity initiatives suggests a potential inflection point and calls our previous gender parity estimates into question.

February 2024: Deal Sourcing: A Data Science Approach - Impact of Financial Characteristics on Acquisition Likelihood

Deal sourcing is hard. Finding a target for acquisition has been likened to finding 'a needle in a haystack'. Firm financials are a valuable starting point for systematic identification of acquisition targets. This publication provides actionable insights and a detailed blueprint on how practitioners can leverage computational finance for deal sourcing. Specifically, five firm-level financial dimensions are identified that differentiate targets from their comparable non-targets based on global data from the most recent 10 years.

February 2024: The Ripple Effect – Finding Company Estimates from Detailed Estimates

Intel's (NASDAQ: INTC) share price jumped 9.3% on Friday, Oct. 27, 2023, after the company reported strong earnings. Cadence Design Systems (NASDAQ: CDNS), which announced earlier in the week, was flat. Over the next 2 weeks (Oct 30 – Nov 14), CDNS would outperform INTC by 544 bps, as investors connected the dots between the two. INTC and CDNS do not share a GICS industry, however the two firms share something potentially more meaningful: sell-side analysts.

November 2023: [Reading Between the Lines in Earnings Calls: 6 Things to Watch as the Q3'23 Earnings Season Unfolds](#)

Watch for Q3'23 sentiment near 5-year highs, despite a quarter-on-quarter decline.

August 2023: [Breaking Boundaries: Women Poised for Milestone Achievement in Parity Amidst Otherwise Bleak Outlook](#)

June 2023: [Mixed Financials Diverge from Bullish Sentiment: A Textual Review of the Q1'23 U.S. Earnings Call Season](#)

April 2023: [Sentiment Rebounds While Regional Banks Tip Their Hand: A Textual Review of the Q4'22 U.S. Earnings Call Season](#)

March 2023: [Singing the \(Banking\) Blues: Navigating the Current Volatility in the Banking Industry](#)

February 2023: [Watch Your Language: Executives' Remarks on Earnings Calls Impact CDS Spreads](#)

January 2023: [Machines Signal Q4'22 Guidance Not Falling Off a Cliff: An In-Depth Textual Review of Q3'22 Earnings Call Transcripts](#)

October 2022: [Hanging on Every Negative Word: Natural Language Processing Analysis of Credit Rating Action Reports](#)

March 2022: [The Sounds of Silence: No Response Speaks Volumes](#)

October 2021: [Glass Floors and Ceilings: Why Closing the Median Wage Gap Isn't Fair](#)

September 2021: [The Board Matrix: The \(ESG\) Value of Well-Connected Directors](#)

August 2021: [Technology Momentum: Peer Networks from Patents](#)

July 2021: [Branching Out: Graph Theory Fundamentals](#)

May 2021: [U.S Filings: No News is Good News](#)

March 2021: [Hiding in Plain Sight – Risks That Are Overlooked](#)

January 2021: [Leadership Change That Matters: A Value and Momentum Story](#)

December 2020: [Warranted Optimism: Sentiment vs. Supply Chain](#)

December 2020: [A Dark Winter for REITS: Trouble Brewing](#)

October 2020: [Sweet Spots in the C-Suite: Executive Best Practices for Shareholder Friendly Firms](#)

October 2020: [Just the \(Build\)Fax: Property Intelligence from Building Permit Data](#)

August 2020: [The Analyst Matrix: Profiting from Sell-Side Analysts' Coverage Networks](#)

June 2020: [The Information Supply Chain Begins Recovering From COVID](#)

May 2020: [Never Waste a Crisis: Following the Smart Money Through Beneficial Ownership Filings](#)

May 2020: [Risky Business: Foot Traffic, Vacancy Rates and Credit Risks](#)

May 2020: [Finding the Healthy Stocks in Health Care During Lockdown](#)

May 2020: [No More Walks in the \(Office\) Park: Tying Foot Traffic Data to REITs](#)

May 2020: [Do Markets Yearn for the Dog Days of Summer: COVID, Climate and Consternation](#)

April 2020: [Cold Turkey - Navigating Guidance Withdrawal Using Supply Chain Data](#)

April 2020: [Data North Star - Navigating Through Information Darkness](#)

March 2020: [Long Road to Recovery: Coronavirus Lessons from Supply Chain and Financial Data](#)

February 2020: [Ship to Shore: Mapping the Global Supply Chain with Panjiva Shipping Data in Xpressfeed™](#)

January 2020: Natural Language Processing – Part III: Feature Engineering Applying NLP Using Domain Knowledge to Capture Alpha from Transcripts

December 2019: The “Trucost” of Climate Investing: Managing Climate Risks in Equity Portfolios

October 2019: #ChangePays: There Were More Male CEOs Named John than Female CEOs

June 2019: Looking Beyond Dividend Yield: Finding Value in Cash Distribution Strategies

June 2019: The Dating Game: Decrypting the Signals in Earnings Report Dates

May 2019: Bridges for Sale: Finding Value in Sell-Side Estimates, Recommendations, and Target Prices

February 2019: U.S Stock Selection Model Performance Review

February 2019: International Small Cap Investing: Unlocking Alpha Opportunities in an Underutilized Asset Class

January 2019: Value and Momentum: Everywhere, But Not All the Time

November 2018: Forging Stronger Links: Using Supply Chain Data in the Investing Process

September 2018: Their Sentiment Exactly: Sentiment Signal Diversity Creates Alpha Opportunity

September 2018: Natural Language Processing – Part II: Stock Selection: Alpha Unscripted: The Message within the Message in Earnings Calls

July 2018: A Case of ‘Wag the Dog’? - ETFs and Stock-Level Liquidity

June 2018: The (Gross Profitability) Trend is Your Friend

May 2018: Buying the Dip: Did Your Portfolio Holding Go on Sale?

March 2018: In the Money: What Really Motivates Executive Performance?

February 2018: The Art of the (no) Deal: Identifying the Drivers of Canceled M&A Deals

January 2018: U.S Stock Selection Model Performance Review

September 2017: Natural Language Processing - Part I: Primer

July 2017: Natural Language Processing Literature Survey

June 2017: Research Brief: Four Important Things to Know About Banks in a Rising Rate Environment

April 2017: Banking on Alpha: Uncovering Investing Signals Using SNL Bank Data

March 2017: Capital Market Implications of Spinoffs

January 2017: U.S. Stock Selection Model Performance Review 2016

November 2016: Electrify Stock Returns in U.S. Utilities

October 2016: A League of their Own: Batting for Returns in the REIT Industry - Part 2

September 2016: A League of their Own: Batting for Returns in the REIT Industry - Part 1

August 2016: Mergers & Acquisitions: The Good, the Bad and the Ugly (and how to tell them apart)

July 2016: Preparing for a Slide in Oil Prices -- History May Be Your Guide

June 2016: Social Media and Stock Returns: Is There Value in Cyberspace?

April 2016: An IQ Test for the “Smart Money” – Is the Reputation of Institutional Investors Warranted?

March 2016: [Stock-Level Liquidity – Alpha or Risk? - Stocks with Rising Liquidity Outperform Globally](#)

February 2016: [U.S. Stock Selection Model Performance Review - The most effective investment strategies in 2015](#)

January 2016: [What Does Earnings Guidance Tell Us? – Listen When Management Announces Good News](#)

November 2015: [Late to File - The Costs of Delayed 10-Q and 10-K Company Filings](#)

October 2015: [Global Country Allocation Strategies](#)

September 2015: [Research Brief: Building Smart Beta Portfolios](#)

September 2015: [Research Brief – Airline Industry Factors](#)

August 2015: [Point-In-Time vs. Lagged Fundamentals – This time i\(t\)'s different?](#)

August 2015: [Introducing S&P Capital IQ Stock Selection Model for the Japanese Market](#)

July 2015: [Research Brief – Liquidity Fragility](#)

May 2015: [Investing in a World with Increasing Investor Activism](#)

April 2015: [Drilling for Alpha in the Oil and Gas Industry – Insights from Industry Specific Data & Company Financials](#)

February 2015: [U.S. Stock Selection Model Performance Review - The most effective investment strategies in 2014](#)

January 2015: [Research Brief: Global Pension Plans - Are Fully Funded Plans a Relic of the Past?](#)

January 2015: [Profitability: Growth-Like Strategy, Value-Like Returns - Profiting from Companies with Large Economic Moats](#)

October 2014: [Lenders Lead, Owners Follow - The Relationship between Credit Indicators and Equity Returns](#)

July 2014: [Factor Insight: Reducing the Downside of a Trend Following Strategy](#)

May 2014: [Introducing S&P Capital IQ's Fundamental China A-Share Equity Risk Model](#)

April 2014: [Riding the Coattails of Activist Investors Yields Short and Long Term Outperformance](#)

March 2014: [Insights from Academic Literature: Corporate Character, Trading Insights, & New Data Sources](#)

February 2014: [Obtaining an Edge in Emerging Markets](#)

February 2014: [U.S Stock Selection Model Performance Review](#)

January 2014: [Buying Outperformance: Do share repurchase announcements lead to higher returns?](#)

October 2013: [Informative Insider Trading - The Hidden Profits in Corporate Insider Filings](#)

September 2013: [Beggars Thy Neighbor – Research Brief: Exploring Pension Plans](#)

August 2013: [Introducing S&P Capital IQ Global Stock Selection Models for Developed Markets: The Foundations of Outperformance](#)

July 2013: [Inspirational Papers on Innovative Topics: Asset Allocation, Insider Trading & Event Studies](#)

June 2013: [Supply Chain Interactions Part 2: Companies – Connected Company Returns Examined as Event Signals](#)

June 2013: [Behind the Asset Growth Anomaly – Over-promising but Under-delivering](#)

April 2013: [Complicated Firms Made Easy - Using Industry Pure-Plays to Forecast Conglomerate Returns.](#)

March 2013: [Risk Models That Work When You Need Them - Short Term Risk Model Enhancements](#)

March 2013: [Follow the Smart Money - Riding the Coattails of Activist Investors](#)

February 2013: [Stock Selection Model Performance Review: Assessing the Drivers of Performance in 2012](#)

January 2013: [Research Brief: Exploiting the January Effect Examining Variations in Trend Following Strategies](#)

December 2012: [Do CEO and CFO Departures Matter? - The Signal Content of CEO and CFO Turnover](#)

November 2012: [11 Industries, 70 Alpha Signals -The Value of Industry-Specific Metrics](#)

October 2012: [Introducing S&P Capital IQ's Fundamental Canada Equity Risk Models](#)

September 2012: [Factor Insight: Earnings Announcement Return – Is A Return Based Surprise Superior to an Earnings Based Surprise?](#)

August 2012: [Supply Chain Interactions Part 1: Industries Profiting from Lead-Lag Industry Relationships](#)

July 2012: [Releasing S&P Capital IQ's Regional and Updated Global & US Equity Risk Models](#)

June 2012: [Riding Industry Momentum – Enhancing the Residual Reversal Factor](#)

May 2012: [The Oil & Gas Industry - Drilling for Alpha Using Global Point-in-Time Industry Data](#)

May 2012: [Case Study: S&P Capital IQ – The Platform for Investment Decisions](#)

March 2012: [Exploring Alpha from the Securities Lending Market – New Alpha Stemming from Improved Data](#)

January 2012: [S&P Capital IQ Stock Selection Model Review – Understanding the Drivers of Performance in 2011](#)

January 2012: [Intelligent Estimates – A Superior Model of Earnings Surprise](#)

December 2011: [Factor Insight – Residual Reversal](#)

November 2011: [Research Brief: Return Correlation and Dispersion – All or Nothing](#)

October 2011: [The Banking Industry](#)

September 2011: [Methods in Dynamic Weighting](#)

September 2011: [Research Brief: Return Correlation and Dispersion](#)

July 2011: [Research Brief - A Topical Digest of Investment Strategy Insights](#)

June 2011: [A Retail Industry Strategy: Does Industry Specific Data tell a different story?](#)

May 2011: [Introducing S&P Capital IQ's Global Fundamental Equity Risk Models](#)

May 2011: [Topical Papers That Caught Our Interest](#)

April 2011: [Can Dividend Policy Changes Yield Alpha?](#)

April 2011: [CQA Spring 2011 Conference Notes](#)

March 2011: [How Much Alpha is in Preliminary Data?](#)

February 2011: [Industry Insights – Biotechnology: FDA Approval Catalyst Strategy](#)

January 2011: [US Stock Selection Models Introduction](#)

January 2011: [Variations on Minimum Variance](#)

January 2011: [Interesting and Influential Papers We Read in 2010](#)

November 2010: [Is your Bank Under Stress? Introducing our Dynamic Bank Model](#)

October 2010: Getting the Most from Point-in-Time Data

October 2010: Another Brick in the Wall: The Historic Failure of Price Momentum

July 2010: Introducing S&P Capital IQ's Fundamental US Equity Risk Model

Copyright © 2024 by S&P Global Market Intelligence, a division of S&P Global Inc. All rights reserved.

These materials, including any software, data, processing technology, index data, ratings, credit-related analysis, research, model, software or other application or output described herein, or any part thereof (collectively the "Property") constitute the proprietary and confidential information of S&P Global Market Intelligence or its affiliates (each and together "S&P Global") and/or its third-party provider licensors. S&P Global on behalf of itself and its third-party licensors reserves all rights in and to the Property. These materials have been prepared solely for information purposes based upon information generally available to the public and from sources believed to be reliable.

Any copying, reproduction, reverse-engineering, modification, distribution, transmission, or disclosure of the Property, in any form or by any means, is strictly prohibited without the prior written consent of S&P Global. The Property shall not be used for any unauthorized or unlawful purposes. S&P Global Market Intelligence's opinions, statements, estimates, projections, quotes and credit-related and other analyses are statements of opinion as of the date they are expressed and not statements of fact or recommendations to purchase, hold, or sell any securities or to make any investment decisions, and do not address the suitability of any security, and there is no obligation on S&P Global Market Intelligence to update the foregoing or any other element of the Property. S&P Global Market Intelligence may provide index data. Direct investment in an index is not possible. Exposure to an asset class represented by an index is available through investable instruments based on that index. The property and its composition and content are subject to change without notice.

THE PROPERTY IS PROVIDED ON AN "AS IS" BASIS. NEITHER S&P GLOBAL NOR ANY THIRD PARTY PROVIDERS (TOGETHER, "S&P GLOBAL PARTIES") MAKE ANY WARRANTY, EXPRESS OR IMPLIED, INCLUDING BUT NOT LIMITED TO ANY WARRANTIES OF MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE, FREEDOM FROM BUGS, SOFTWARE ERRORS OR DEFECTS, THAT THE PROPERTY'S FUNCTIONING WILL BE UNINTERRUPTED OR THAT THE PROPERTY WILL OPERATE IN ANY SOFTWARE OR HARDWARE CONFIGURATION, NOR ANY WARRANTIES, EXPRESS OR IMPLIED, AS TO ITS ACCURACY, AVAILABILITY, COMPLETENESS OR TIMELINESS, OR TO THE RESULTS TO BE OBTAINED FROM THE USE OF THE PROPERTY. S&P GLOBAL PARTIES SHALL NOT IN ANY WAY BE LIABLE TO ANY RECIPIENT FOR ANY INACCURACIES, ERRORS OR OMISSIONS REGARDLESS OF THE CAUSE. Without limiting the foregoing, S&P Global Parties shall have no liability whatsoever to any recipient, whether in contract, in tort (including negligence), under warranty, under statute or otherwise, in respect of any loss or damage suffered by any recipient as a result of or in connection with the Property, or any course of action determined, by it or any third party, whether or not based on or relating to the Property. In no event shall S&P Global be liable to any party for any direct, indirect, incidental, exemplary, compensatory, punitive, special or

consequential damages, costs, expenses, legal fees or losses (including without limitation lost income or lost profits and opportunity costs or losses caused by negligence) in connection with any use of the Property even if advised of the possibility of such damages. The Property should not be relied on and is not a substitute for the skill, judgment and experience of the user, its management, employees, advisors and/or clients when making investment and other business decisions.

The S&P Global logo is a registered trademark of S&P Global, and the trademarks of S&P Global used within this document or materials are protected by international laws. Any other names may be trademarks of their respective owners.

The inclusion of a link to an external website by S&P Global should not be understood to be an endorsement of that website or the website's owners (or their products/services). S&P Global is not responsible for either the content or output of external websites. S&P Global keeps certain activities of its divisions separate from each other in order to preserve the independence and objectivity of their respective activities. As a result, certain divisions of S&P Global may have information that is not available to other S&P Global divisions. S&P Global has established policies and procedures to maintain the confidentiality of certain nonpublic information received in connection with each analytical process. S&P Global may receive compensation for its ratings and certain analyses, normally from issuers or underwriters of securities or from obligors. S&P Global reserves the right to disseminate its opinions and analyses. S&P Global Ratings' public ratings and analyses are made available on its sites, www.spglobal.com/ratings (free of charge) and www.capitaliq.com (subscription), and may be distributed through other means, including via S&P Global publications and third-party redistributors.