

# Unusual Financial Communication

Evidence from ChatGPT, Earnings Calls, and the Stock Market

Lars Beckmann\*

Heiner Beckmeyer†

Ilias Filippou‡

Stefan Menze§

Guofu Zhou¶

## Abstract

The introduction of ChatGPT has changed how humans process textual data. We devise a prompting strategy for ChatGPT to identify and analyze unusual aspects of financial communication, focusing on earnings calls of S&P 500 firms. Utilizing the latest GPT-4-Turbo model, we identify and categorize unusual financial communication across 25 dimensions, which fall into four categories: unusual communication by executives, by financial analysts, unusual content, and technical issues. A significant portion of earnings calls displays unusual financial communication, which correlates with certain firm characteristics and fluctuates with the business cycles. The stock market reacts negatively to unusual communication, with an elevated trading activity. We highlight the potential of large language models like ChatGPT in financial analyses, offering new insights into the interpretation of complex textual data and its economic consequences on market impacts.

---

\*Finance Center, University of Münster. lars.beckmann@wiwi.uni-muenster.de.

†Finance Center, University of Münster. heiner.beckmeyer@wiwi.uni-muenster.de.

‡Olin Business School, Washington University in St. Louis. iliasfilippou@wustl.edu.

§Finance Center, University of Münster. stefan.menze@wiwi.uni-muenster.de.

¶Olin Business School, Washington University in St. Louis. zhou@wustl.edu.

# 1. Introduction

The introduction of ChatGPT has changed how humans process information. ChatGPT is the fastest-growing online service to date, reaching 100 million users just two months after its launch (Hu, 2023). Today, more than 180 million users are actively using ChatGPT. The service's success in making sense of textual data also has important repercussions for financial markets. Financial analysts can use large language models to comb through the ever-growing amounts of financial and economic textual data (Gentzkow, Kelly, and Taddy, 2019). To this end, Bloomberg, as a leading data and service provider to financial professionals, introduced BloombergGPT in March 2023 (Wu, Irsoy, Lu, Dabrowski, Dredze, Gehrmann, Kambadur, Rosenberg, and Mann, 2023). The large language model is dedicated to performing well on the everyday tasks of financial analysts.<sup>1</sup> Similarly, Citadel is seeking an enterprise-wide license of ChatGPT (Doherty and Marques, 2023).

We leverage ChatGPT's general knowledge to understand the intricacies of financial communication using earnings calls as a laboratory. Earnings calls are the premier data source for this task, as the included Q&A session reflects dynamic interactions between financial analysts asking questions about the firm and members of the management team providing their answers. In general, we want to capture not only the unusualness of what is being said (content) and how it is said (tonality), but also identify unusual aspects beyond that. Tesla's Q1-2018 earnings call is a prime example of unusualness in financial communication. CEO Elon Musk addressed an analyst's question about future capital requirements by "Excuse me. Next. Boring, bonehead questions are not cool. Next?". The earnings call, at the same time, exemplifies how unusualness can go beyond what has been said and how. Musk's remarks that he does not know the number of reservations for the Tesla Semi were supplemented by the CFO. This, in turn, indicates a lack of preparedness – another dimension of unusualness in earnings call communication. Furthermore, Musk takes questions from a YouTuber, which ChatGPT also correctly identifies as unusual.

We focus on firms in the S&P 500 as they typically enjoy the largest following by financial analysts and reflect novel information sooner than smaller firms (Zhang, 2006). If we can find a substantial stock price reaction to a call's unusualness in this sample of firms, our model is able to capture information that is otherwise produced by financial

---

<sup>1</sup>Another recent development in the field that is likely to impact the work of financial analysts is Microsoft's development of "Copilot", a large language model for the company's office suite (Spataro, 2023).

analysts and extract information from the calls that is otherwise introduced with delay. For smaller firms the data quantity and quality are much reduced and the price delay may be too large, obscuring the pricing channel in question (Hou and Moskowitz, 2005).

Our **first contribution** is in engineering a suitable prompting approach for ChatGPT to identify and understand *unusualness* in earnings calls. In particular, we apply a three-step prompting strategy to identify *unusualness* in earnings calls, which is solely based on ChatGPT's general knowledge and does not require an external definition of *unusualness*. With this prompting approach at hand, we can feed the earnings call transcripts for firms in the S&P 500 through ChatGPT and extract a binary indicator for different dimensions of *unusualness* as to whether each call is unusual. Apart from the binary indicator ("unusual" vs. "usual"), we also ask ChatGPT to provide a concise reason for why it believes an earnings call to be unusual in the respective dimension. The prompting strategy we devise can also be used for numerous other applications to make sense of textual data in the financial context. This underlines the potential of large language models such as ChatGPT for financial analysis.

As our **second contribution**, we identify unusual financial communication in earnings calls and investigate its correlation with firm characteristics, industry affiliation, and macroeconomic indicators across various business cycles. In total, we identify 25 dimensions of *unusualness* in earnings calls, which can be classified into four broader categories: unusual communication by *executives*, by *financial analysts*, unusual *contents*, and unusual *technical issues*.<sup>2</sup> Firms with unusual communication tend to be larger but less profitable and more likely momentum losers than winners. While there is substantial heterogeneity across industries, industry affiliation alone is unable to explain the occurrence of unusual firm communication. We relate the different dimensions of unusual communication to various macroeconomic indicators, including the VIX, the intermediary capital ratio of He, Kelly, and Manela (2017), the Chicago Fed National Activity Index, price-dividend ratio, net share issuance, the current Treasury bill rate, and the term and default spread. We document that the degree of unusual communication varies with the business cycle. For example, unusual communication, especially in the form of unusual content, is more likely in high-VIX regimes. Conversely, higher net share issuance is correlated with less unusual communication, particularly by firm executives, potentially in an effort to soothe investors when issuing additional shares (Pontiff and Woodgate, 2008).

---

<sup>2</sup>The full list of identified dimensions is provided in Table 1.

As our **third contribution**, we investigate to which extent market participants react to unusual earnings calls and if so, which dimensions of unusual communication are responsible for this reaction. We confirm that firms typically earn high returns on earnings announcement dates (Savor and Wilson, 2016). In contrast, we show that the returns of firms with unusual communication are significantly lower, and in fact indistinguishable from zero. This result stands both when value-weighting and equally-weighting the respective firms. We find large differences in the effects of the 25 dimensions. While the broader categories capturing unusual communication by executives, analysts, and unusual content all generate a significantly negative return impact, technical difficulties are related to a positive return impact (equally-weighted) or no return impact (value-weighted). The largest negative return impact is produced by a lack of critical questioning by analysts ( $-4.24\%$ ), repetitive questions ( $-2.83\%$ ), or the announcement of surprising information ( $-2.03\%$ ).

The literature has proposed trading volume as a measure for disagreement across investors (Goetzmann and Massa, 2005; Banerjee and Kremer, 2010; Banerjee, 2011). Hong and Stein (2007) argue that disagreement may arise when investors possess different information sets or if new information leads them to update their beliefs. Earnings calls are of course an important source of new information to investors, and disagreement should be particularly high whenever the information is communicated in unusual ways or the content is unusual itself. Consistent with this idea, we find that trading activity is significantly elevated for firms with unusual communication along many of the 25 identified dimensions. Especially unusual contents trigger additional trading by market participants, suggesting that more investors impound their disagreement of the provided information into prices.

Finally, we confirm in panel regressions that our key result that unusual communication triggers a negative stock market response is not subsumed by the surprise component of the earnings release, analyst forecast dispersion, the number of analysts covering a firm, differences in other firm characteristics, or industry affiliation. Throughout all specifications, we find a negative and highly significant announcement return impact of unusual communication. Especially the discussion of unusual contents is associated with such.

Our use of ChatGPT to identify unusual financial communication comes with multiple advantages. First, we always rely on the same model for the decision of whether a particular earnings call was unusual or not. As human intuition is subjective, it can vary

from day to day and from analyst to analyst. With the model's *temperature* set to zero, the model always produces the same output (is the call unusual?) for the same input (the transcript's Q&A session) and the same prompt. This reproducibility in the pipeline is key for the compliance of financial professionals. Second, the model processes textual information significantly faster than humans, combines much of the aggregate reasoning while humans would process this information from its training corpus (de Kok, 2023). Third, we utilize the most recent iteration of ChatGPT, known as GPT-4-Turbo. This version not only offers general enhancements over its predecessor, GPT-4, but most crucially, it can process texts of up to 128,000 tokens in a single prompt. Previous iterations of GPT-4 were limited to just 8,192 tokens, compelling many studies that utilized ChatGPT on earnings call transcripts to either shorten the length of the transcripts or use summarization techniques. Such methods invariably led to a loss of important content and context.<sup>3</sup>

## Related Literature

Our paper is part of a young but growing literature that uses ChatGPT to understand vast amounts of economic and financial textual data. A review of the topic is provided by Dong, Stratopoulos, and Wang (2023). Several papers study ChatGPT's ability to forecast future stock returns directly and whether these forecasts make for a profitable investment portfolio. In an early study on this topic, Lopez-Lira and Tang (2023) show that ChatGPT is particularly good at identifying short-term price movements in small stocks. To address the criticism that information about a stock's realized return in a backtest may be part of ChatGPT's training corpus, Pelster and Val (2023) perform a live experiment to study ChatGPT's real-time ability to predict future stock returns. Kim (2023) asks ChatGPT to provide investment recommendations across asset classes.

Other studies focus on ChatGPT's ability to make sense of textual data. Impressively, Katz, Bommarito, Gao, and Arredondo (2023) show that GPT-4 passes the bar exam, while previous iterations of ChatGPT have failed to do so. We use the updated GPT-4-Turbo, which is another improvement over GPT-4 studied by the authors. In a similar vein, Kim, Muhn, and Nikolaev (2023b) show that ChatGPT-generated summaries of corporate disclosures better reflect the conveyed textual sentiment and, in turn, are more predictive of the stock market's reaction. They introduce a measure of informa-

---

<sup>3</sup>We discuss this circumstance in more detail in Section 2.1.

tion “bloat” and show that bloated disclosures are linked to lower stock price efficiency and higher information asymmetries. Kim, Muhn, and Nikolaev (2023a) use GPT-3.5 to measure a firm’s exposure to several risks from the transcripts of earnings calls, while Bai, Boyson, Cao, Liu, and Wan (2023) use the textual distance between the answers given by firm executives and answers generated by ChatGPT and show that this human-AI distance can predict the stock’s liquidity, abnormal returns around earnings announcements, and the accuracy of forecasts made by financial analysts. Jha, Qian, Weber, and Yang (2023) introduce an investment score from earnings calls using ChatGPT and show that this score is highly correlated with the expectations of CFOs. Finally, Bybee (2023) uses ChatGPT to generate economic expectations from historical newspaper articles. The author uses the generated economic expectations to study the origins of stock price bubbles.

Our study furthermore adds to the broader literature, which uses textual information for economic research. Gentzkow et al. (2019) provide an introduction to the topic. In an influential study, Baker, Bloom, and Davis (2016) construct economic policy uncertainty indexes from the mentions of certain topics in newspaper articles. Baker, Bloom, Davis, and Sammon (2021) study the origins of stock market jumps by identifying the dominant news story on the day after the jump occurred, and Gorodnichenko, Pham, and Talavera (2023) study the impact of the communication by the Federal Reserve chair. Bybee, Kelly, and Su (2023b) construct a latent factor model for the cross-section of stock returns from news topics, and Bybee, Kelly, Manela, and Xiu (2023a) use these news topics to forecast macroeconomic dynamics. Finally, Chen, Tang, Zhou, and Zhu (2023) show that news-based information extracted with the help of ChatGPT is related to macroeconomic conditions and the aggregate stock market.

Many studies make use of earnings call transcripts to obtain insights about the impact of manager communication. Hassan, Hollander, Van Lent, and Tahoun (2019) measure a firm’s exposure to political risks from earnings call transcripts using a simple dictionary approach. Hassan, Hollander, Van Lent, and Tahoun (2020b) extend this idea to measure the perceived impact of Brexit and Hassan, Hollander, Van Lent, Schwedeler, and Tahoun (2020a) study the implications of firm-level exposure to pandemics. Li, Mai, Shen, and Yan (2021) use a simple machine learning approach to assess if earnings calls convey information about a firm’s culture, and Breuer, Knetsch, et al. (2023) study the impact of managerial charisma.

Finally, we add to our understanding of the importance of event days. To this end,

Savor and Wilson (2013) and Ernst, Gilbert, and Hrdlicka (2019) show that market returns are significantly higher on days with macroeconomic news releases and Liang (2003) argues that earnings announcements provide particularly important information to investors. Consistent with this, Savor and Wilson (2016) show that earnings announcement returns are significantly elevated and link this to investors revising their expectations for non-announcing firms. Engelberg, McLean, and Pontiff (2018) show that the returns of many prominent cross-sectional stock anomalies are significantly elevated on earnings announcement days and favor an explanation based on the resolution of mispricing.

## **2. Methodology**

### **2.1. Data Description**

We obtain transcripts of earnings calls held by firms in the S&P 500 from Refinitiv for the period of January 2015 to December 2022. A typical earnings call has two sections: first, a presentation by the firm's management, which provides a broad overview of the firm's current operations and its future trajectory. The second portion is a Q&A session, as an interaction between financial analysts asking questions and members of the management team providing answers. We focus our analysis on this second portion of the transcript, as the management's presentation tends to be tightly scripted and conveys less novel information (Lee, 2016; Li et al., 2021). Cao, Jiang, Yang, and Zhang (2023) show that firms adapt to the increasing use of machine learning tools by investors to process textual information and consequently change their writing in regulatory disclosures like 10Ks. Earnings calls are not affected by this catering to conveying specific information to machines as the Q&A sessions reflect live interactions between management and analysts. We merge the transcript data with stock returns and trading volume from CRSP and accounting information from S&P's Compustat to assess the stock market's reaction to unusual aspects of the earnings call.

The earnings call transcripts are provided by Refinitiv as simple .txt files. From each file, we extract a) the date and time of when the earnings call took place, b) the reporting quarter, and c) the company name and ticker. Finally, we split the transcript into the presentation part and the Q&A session and feed ChatGPT with the Q&A session only. In the future, we plan to provide the description of which earnings calls are unusual and why for public use.

**Accessing ChatGPT.** We access ChatGPT via its application programming interface (API). Different from web access, the API offers various additional functions. The *temperature* parameter defines the magnitude of stochastic answers. Asking the same question several times generates varying responses when *temperature* is high. We set this parameter to zero to obtain answers as objective and reproducible as possible.

The model that we rely on is GPT-4-Turbo, which is an updated version of GPT-4. It was released on November 6, 2023. Whereas the previous iteration was limited to processing 8,192 tokens per prompt, the token limit has been increased to 128,000 for GPT-4-Turbo. This change alleviates the need to trim the transcripts or have ChatGPT process the information of a single transcript in chunks. With this, we are able to evade the downsides of chunking, which previous studies may have to resort to given the limitations of ChatGPT before GPT-4-Turbo.<sup>4</sup> GPT-4-Turbo is therefore able to comprehend the Q&A sessions in their entirety. This contributes to a model setup that is capable of identifying subtleties in both tone and language that may only stand out as unusual when the entire Q&A session is provided as context. In our dataset, 9,498 of the 14,720 Q&A sessions are longer than 8,192 tokens.<sup>5</sup>

## 2.2. Prompt Engineering

We need to find a suitable approach for ChatGPT to provide an accurate and consistent assessment of the unusualness of a firm's earnings call. We not only want to understand if an earnings call is unusual but also learn about the aspects that characterize this unusualness. We propose a three-step prompting approach for this purpose:

In the first step, we want to understand what is generally unusual in earnings call Q&A sessions. We feed ChatGPT a random sample of  $n = 1000$  Q&A session transcripts between 2015 and 2023, and ask for the model's judgment whether a particular Q&A session is usual or unusual. If a Q&A session is found to be unusual in any way, ChatGPT is asked to provide a textual justification for this *unusualness*.

---

<sup>4</sup>Examples of studies resorting to "chunking" are Goyal, Li, and Durrett (2022), Kim et al. (2023a) Kim et al. (2023b) and Zhang, Ni, Yu, Zhang, Zhu, Deb, Celikyilmaz, Awadallah, and Radev (2021). Their solution to larger inputs is chunking about every 2,000 tokens and using summaries of the previous chunk to add to the recent 2,000 token chunks.

<sup>5</sup>The average token length is 9,256, which we calculate using the tokenizer provided by *OpenAI*. The 90th (95th; 99th) percentile of the token length amounts to 12,873 (14,142; 16,924) tokens.



**Prompt 1:** Please read the following transcript of a Question-and-Answer session from the earnings conference call of company {firm} ({ticker}) carefully. Determine whether the Question-and-Answer session of this earnings conference call is ‘usual’ or ‘unusual’: If the Question-and-Answer session is classified as ‘usual’, state ‘usual’ without any justifications or further output. If the Question-and-Answer session is classified as ‘unusual’, state ‘unusual’ and provide a justification for this classification. Transcript of the Question-and-Answer Session: ‘{qa}’

In the second step, we want to find a way to systematize these unusual observations. To do this, we collect all justifications for unusual communication from the first step. These justifications are made available to ChatGPT in a new prompt. ChatGPT is then asked to form high-level categories from these justifications.

**Prompt 2:** Please read the provided text file with justifications for unusual Q&A sessions from earnings conference calls carefully. What are high-level categories to identify unusual Q&A sessions? Make sure that each statement from the text file can be assigned to one of the categories.

As a result, ChatGPT provides us with 25 high-level categories for the identification of *unusualness* communication in the Q&A sessions for earnings call transcripts between 2015 and 2023. Table 1 provides an overview of the identified dimensions of unusual communication. We first provide a high-level assessment of whether an earnings call included unusual communication of any form (“Unusual”). This is a simple binary indicator, which we set to “True”, whenever ChatGPT identifies unusual communication along at least one of the 25 dimensions. We then provide a low-level assessment of exactly which dimensions constitute the identified *unusualness*. The identified dimensions can roughly be clustered into four broader categories: 1) Unusual communication by the firm’s executives (“Executives”); 2) unusual communication by financial analysts (“Analysts”); 3) unusual contents that were being discussed (“Content”); and 4) technical difficulties (“Technical”).

ChatGPT identifies several dimensions of **unusual communication by executives**. For example, responses to analyst questions may be too lengthy, emotional or even contain conflicting information or be evasive. ChatGPT also identifies whether the management team appeared unprepared. **Unusual communication by analysts** comes in the form of off-topic or repetitive questions, a particularly large volume of questions, a repeated focus on specific participants or topics, frequent queries about non-financial issues, and a lack of critical questions. **Unusual contents** can range from corporate changes and changes in the leadership team, over providing strategic insights and the discussion of legal

and regulatory issues, to issues in the operations and among management. Furthermore, ChatGPT can identify unusual patterns in the firm's analysis of the markets it operates in, a particularly detailed discussion of certain (financial or non-financial) topics, the impact of external events, macroeconomic considerations, and surprising announcements, as well as detailed discussions of the firm's products or services. Finally, ChatGPT also identifies **technical difficulties**.

In the third prompting step, we then examine all Q&A sessions with regard to these categories that constitute the different dimensions of unusual communication in the earnings calls. We ask ChatGPT to provide a binary assessment of whether the transcript in question is unusual with respect to each of the categories:

**Prompt 3:** *Please read the following transcript of a Question-and-Answer session from the earnings conference call of company {firm} (ticker) carefully. Determine whether the Question-and-Answer session of this earnings conference call is 'usual' or 'unusual' in the following {len(categories)} categories: {categories}*  
*For each category, state whether the Question-and-Answer session is 'usual' or 'unusual'. If the Question-and-Answer session is classified as 'usual' in the respective category, state 'usual' without any justifications or further output. If the Question-and-Answer session is classified as 'unusual' in the respective category, state 'unusual', print a '/', and provide a justification for this classification. Transcript of the Question-and-Answer Session: '{qa}'*

To illustrate that our approach is able to identify unusual aspects of communication during earnings calls, we showcase the ChatGPT output for the first prompt related to Tesla's Q1-2018 earnings call below. ChatGPT produced this assessment without further human intervention.

The Question-and-Answer session of Tesla's Q1-2018 earnings call is classified as 'unusual' for the following reasons:

1. CEO's Dismissive Attitude: Elon Musk, the CEO, displays a dismissive attitude towards analysts' questions, which is not typical for an earnings call. He refers to some questions as "boring," "dry," and "bonehead," which is highly unprofessional and not in line with the decorum usually maintained during such calls.
2. Refusal to Answer Questions: Musk refuses to answer certain questions or provides vague responses. For example, when asked about the timing for the Tesla Network, he says, "I haven't really thought about it," which is unexpected from a CEO discussing a significant aspect of their company's strategy.
3. Focus on Short-Term Stock Traders: Musk explicitly states that they have no interest in satisfying the desires of day traders and even suggests that those concerned about volatility

should not buy Tesla stock. This is unusual as companies typically avoid making such direct and potentially discouraging statements to investors.

4. Lack of Specific Financial Details: There is a lack of specific financial details provided in some of Musk's responses. For instance, when discussing the potential cost savings from the Tesla Semi, he speaks in general terms rather than providing concrete numbers.
5. Hostile Responses: Musk's responses to analysts from Sanford C. Bernstein and RBC Capital Markets are notably hostile. He cuts off one analyst, saying, "Boring bonehead questions are not cool. Next," and dismisses another by moving to a YouTube questioner, which is highly unconventional and indicative of an unusual call.
6. Diversion to YouTube Questioner: Musk takes a question from a YouTube questioner, which is atypical for an earnings call that usually prioritizes financial analysts and institutional investors.
7. Casual Language and Jokes: The language used by Musk and other executives is more casual and includes jokes, which is not typical of the formal tone expected in earnings calls.
8. Lack of Preparedness: Musk's statement that he does not know the number of reservations for the Tesla Semi and needs to be corrected by the CFO suggests a lack of preparedness, which is unusual for a CEO during an earnings call.

Overall, the tone, content, and conduct shown during the Q&A session are not in line with standard corporate earnings call etiquette, making it an unusual event.

## 2.3. Why ChatGPT?

Relying on ChatGPT to identify unusual management communication comes with multiple advantages. First, we always rely on the same model for the decision of whether a particular earnings call is unusual or not. With the model's *temperature* set to zero, the model always produces the same output (is the call unusual?) for the same input (the transcript's Q&A session) and the same prompt (shown above). This reproducibility in the pipeline is key for the compliance of financial professionals. Second, the model performs significantly faster in interpreting textual information than humans and combines much of the aggregate reasoning with which humans would process this information from its training corpus (de Kok, 2023). Besides being faster, our approach is also more accurate and cheaper than employing a large team of analysts to comb through textual data (Baker et al., 2016), or using simple dictionary-based approaches (Loughran and McDonald, 2011), which disregard contextual information. Of course, the prompting approach we devise above can be used for numerous other applications to make sense of textual data in the financial context. Applications range from risk assessment and

management to the optimization of investment portfolios and construction of profitable trading strategies. We leave these for future research.

### 3. Identifying Unusual Communication

In this section, we provide a description of unusual communication by firms in the S&P 500. We start with a general question, “How often is communication unusual?”, and seek an answer using ChatGPT’s assessment. We quantify to which extent communication by the largest firms in the United States is unusual, and which dimensions contribute most. Next, we assess which firms typically engage in unusual communication, by understanding if systematic differences exist in financial ratios and the market’s assessment of the firms’ future prospects. Finally, we investigate if unusual communication co-moves with the business cycle, and whether temporal trends emerge with respect to which dimensions of unusual communication occur most frequently.

#### 3.1. How Often is Communication Unusual?

We first establish the frequency with which communication in earnings calls is unusual, using the 25 dimensions identified by ChatGPT and described in Section 2.2 as a guide. For a typical calendar quarter, 114 firms show some unusual pattern in their communication by executives. Most commonly, ChatGPT identifies lengthy responses (17%), an informal tone (9%), or unusual dialogue patterns (6%) as unusual. Less common are emotional responses (7%), the provision of conflicting information (1%), and unpreparedness (2%) of executives or evasive behavior (4%). In an average calendar quarter, ChatGPT identifies unusual communication by *analysts* for 83 firms. The most common is a repeated focus on specific call participants or specific topics, which afflicts 16% of firms on average. Other dimensions of unusual behavior by analysts are far less common: for only 4% of firms does ChatGPT identify repetitive questioning, for 2% off-topic questions and a high volume of questions. Least common are unusual financial queries and a lack of critical questioning (1%).

The *content* of what is discussed is fairly often flagged as unusual. In the average quarter, 145 firms are identified as such. There is a lot more homogeneity in the frequency with which the different content-based dimensions are flagged as unusual. Most common

Table 1: Dimensions and Frequency of Unusual Communication

The table shows overall *unusualness*, the four broader categories and the 25 identified dimensions of *unusualness*, their abbreviation, how many firms engage in unusual communication in a typical calendar quarter (“N”), and their median fraction (“Median”).

Dimension	Abbreviation	N	Q50
<b>Unusual</b>		198.0	0.42
<b>Executives</b>		114.6	0.24
Lengthy Responses	Lengthy	82.8	0.17
Informal Tone	Informal	45.2	0.09
Unusual Dialogue Patterns	Dial	30.6	0.06
Emotional Responses	Emotion	18.5	0.04
Management Evasiveness	Evasive	16.8	0.04
Management Unpreparedness	Unprep	8.0	0.02
Conflicting Information	Conflict	3.5	0.01
<b>Analysts</b>		83.2	0.17
Repeated Focus on Specific Participants or Topics	RepPart	75.7	0.16
Repetitive Questions	RepQ	19.7	0.04
High Volume of Questions	QVolume	11.3	0.02
Off-Topic Questions	Off-Top	11.0	0.02
Lack of Critical Questioning	LackCrit	3.0	0.01
Unusual Financial Queries	FinQuery	2.8	0.01
<b>Content</b>		145.2	0.31
Detailed Discussion on Non-Financial Topics	NonFin	85.0	0.18
In-Depth Product or Service Discussions	ProdDisc	77.0	0.16
Detailed Financial Discussions	FinDisc	74.8	0.16
Forward-Looking Statements and Strategic Insights	Strategy	74.3	0.16
Significant Leadership or Corporate Changes	Changes	66.6	0.14
External Events Impact	Extern	65.3	0.13
Macroeconomic Considerations	Macro	63.8	0.13
Competitive and Market Analysis	Market	55.2	0.11
Operational and Management Issues	OpIssue	48.6	0.1
Legal or Regulatory Issues	Legal	25.0	0.05
Surprising Announcements	Surprise	18.1	0.04
<b>Technical</b>		52.6	0.11
Technical Difficulties and Disruptions	TechDiff	52.6	0.11

are detailed discussions of (non-)financial topics, in-depth product or service discussions, and forward-looking statements and strategic insights. Least common instead are surprising announcements (4% of firms) and a discussion of legal or regulatory issues (5%). Around 11% of firms experience some form of *technical difficulties* during their earnings calls, which disrupts the flow of conversation.

In Table 2, we show how likely it is that a firm with unusual communication along one dimensions, also engages in unusual communication patterns along another dimension of the same category. For example, how often do firm executives who give lengthy responses

Table 2: Co-Occurrence of Unusual Communication

The table shows for overall *unusualness* and each of the four broader categories how frequently multiple dimensions of unusual communication are identified for the same earnings call. For example, for 14.37% of earnings calls do we identify *unusualness* along one dimension in the *Executives* broader category, for 4.15% (3.58%) we identify 2 (3) dimensions, and so forth.

Dimension	0	1	2	3	4	5	6	7	8	9	10	11	12+
<b>Unusual</b>	57.93	16.24	4.51	2.04	1.22	0.97	0.80	1.15	2.01	2.59	3.25	2.51	4.78
<b>Executives</b>	75.66	14.37	4.15	3.58	1.49	0.41	0.10	0.23					
<b>Analysts</b>	82.32	12.01	4.03	0.97	0.28	0.17	0.22						
<b>Content</b>	69.14	10.10	2.97	1.22	0.92	1.66	2.48	4.00	4.54	1.95	0.67	0.34	
<b>Technical</b>	88.83	11.17											

also provide emotional or informal responses? We find that 16.24% of all earnings calls are unusual along one dimension, 4.51% (2.04%) along 2 (3), and so forth. We find unusual communication along 12 *or more* dimensions for a whopping 4.78% of earnings calls in our sample. Co-occurrence of unusual communication is also prevalent within the four subcategories. For example, for 14.37% of earnings calls we identify *unusualness* along one dimension in the *Executives* category, for 4.15% (3.58%) we identify 2 (3) dimensions, and 0.74% of transcripts include unusual communication by executives along 5 or more dimensions of that category.

We see a similar aggregation for unusual communication by *analysts*. 12.01% of earnings calls are identified as unusual along one dimension, 4.03% along two and 0.97% along three. *Technical difficulties* emerge in 11.17% of all calls. Unusual *content* is identified for roughly 30% of calls, with 10.1% identified as unusual along one dimension, 2.97% along two dimensions, and 15.64% along five or more dimensions. Unusual content tends to co-emerge most frequently.

Table 3: Cross-Category Occurrence of Unusual Communication

The table shows the co-occurrence of unusual communication across the four identified broader categories, *Executives*, *Analysts*, *Technical Difficulties*, and *Content*. For example, if a firm's executives communicate unusually, in 69.36% of cases, we also find unusual communication by financial analysts along at least one of the included dimensions.

Dimension	Executives	Analysts	Content	Technical
<b>Executives</b>	—	95.49	65.78	34.03
<b>Analysts</b>	69.36	—	54.80	19.20
<b>Content</b>	83.41	95.67	—	29.37
<b>Technical</b>	15.63	12.14	10.64	—

In Table 3, we also investigate how often firms with unusual communication in at least one dimension of a particular category (*Executives*) also show signs of unusual

communication along another category (*Analysts*). From the table we learn that unusual communication by *executives* is accompanied by some degree of unusual communication by analysts in 69.36% of occasions. Technical difficulties arise in 15.63% of all cases and unusual content is discussed by 83.41% of firms. Interestingly, when we condition on the circumstance of unusual communication by *analysts*, we find that executives also communicate in unusual ways in 95.49% of cases and that the content is unusual in 95.67% of cases. This suggests that financial analysts are important to set the tone of the discussions during the earnings call. Unusual *contents* are accompanied by unusual communication of executives in 65.78% and of analysts in 54.80% of cases. Technical difficulties co-arise in just 10.64% of these cases.

### 3.2. Which Firms Engage in Unusual Communication?

Overall, there is considerable heterogeneity across the dimensions of unusual communication. We now analyze if specific firms are more prone to utilizing unusual communication by understanding if systematic differences arise in financial ratios and market-based firm characteristics among firms with and without unusual communication patterns across each of the four categories and 25 dimensions. We include market capitalization (Size), book-to-market ratio (B2M), investment behavior (Inv) and profitability (Prof), return momentum (Mom), 21-day MAX return following Bali, Cakici, and Whitelaw (2011), Amihud (2002) illiquidity (Illiq), and the stock's nominal share price (PRC). The results are provided in Table 4.

We find that firms with some dimension of *unusualness* in their earnings call communication tend to be less profitable and are slightly more likely to be momentum losers with a larger MAX return over the past 21 days. Given that we rank-standardize each characteristic in the cross-section to lie between  $-0.5$  and  $+0.5$ , this shows that firms with unusual communication are, on average, two percentiles less profitable relative to their peers with standard communication profiles.

We again let our analysis be guided by the broader categories on unusual communication by executives, analysts, technical difficulties and content. Firms with unusual communication by their *executives* are significantly less profitable and tend to be slightly larger on average, with a slightly more liquid stock. There's substantial heterogeneity across the different dimensions of unusual communication by executives. For example, firms that communicate with an informal tone tend to be six percentiles less profitable

Table 4: Unusual Communication and Firm Characteristics

The table shows differences in financial ratios and market-based firm characteristics across firms for which ChatGPT has identified unusual communication and those without. The table shows *unusualness* in general and the 25 identified dimensions as well as the four broader categories. For a given earnings call, we measure the firm characteristics on the last trading day of the previous month. Each characteristic is cross-sectionally rank-standardized between  $-0.5$  and  $+0.5$  across the firms currently in the S&P 500. The construction of the firm characteristics follows Jensen, Kelly, and Pedersen (2023). We include market capitalization (Size), book-to-market ratio (B2M), investment behavior (Inv) and profitability (Prof), return momentum (Mom), 21-day MAX return following Bali et al. (2011), Amihud (2002) illiquidity (Illiq), and the stock's nominal share price (PRC). \*\*\* (\*\*, \*) denotes that the respective characteristic differs significantly between “unusual” and “usual” firms at the 1% (5%, 10%) level.

Dimension	Size	B2M	Inv	Prof	Mom	MAX	Illiq	PRC
<b>Unusual</b>	0.00	0.00	-0.00	-0.02***	-0.02**	0.01*	-0.01	0.00
<b>Executives</b>	0.02***	0.01	0.00	-0.03***	-0.00	0.01	-0.02***	0.01
Lengthy	0.04***	-0.00	0.01	-0.01**	-0.00	0.02**	-0.04***	0.02**
Informal	0.03***	-0.00	-0.01	-0.06***	0.02	0.02*	-0.03***	0.00
Dial	0.07***	-0.01	-0.01	-0.05***	0.03**	0.02*	-0.07***	0.02**
Emotion	0.04***	0.02	0.00	-0.04***	-0.01	0.01	-0.04***	-0.02**
Evasive	-0.03	0.05***	0.00	-0.03***	-0.04***	-0.01	0.01	-0.05***
Unprep	-0.03	0.03	-0.00	-0.03*	-0.01	-0.00	0.02	-0.02
Conflict	0.02	0.03	0.05	-0.02	-0.07**	0.01	-0.04	-0.01
<b>Analysts</b>	0.03***	0.00	0.01*	-0.02***	0.00	0.01	-0.04***	0.02**
RepPart	0.03***	-0.00	0.02*	-0.02**	-0.00	0.02*	-0.04***	0.02***
RepQ	0.03**	0.04**	0.02*	-0.01	0.00	0.01	-0.04***	-0.01
QVolume	0.01	0.00	-0.02	-0.00	0.04**	0.03	-0.02	0.02
Off-Top	0.08***	0.02	-0.00	-0.06***	0.01	-0.01	-0.09***	-0.02
LackCrit	0.04	0.04	-0.02	-0.06*	-0.01	-0.01	-0.05	-0.01
FinQuery	0.06	0.06	-0.00	-0.07	-0.01	0.01	-0.07**	-0.01
<b>Content</b>	0.02***	0.00	0.00	-0.02***	-0.02***	0.01*	-0.02***	0.01**
NonFin	0.03***	-0.00	0.02*	-0.02**	-0.00	0.01*	-0.04***	0.01*
ProdDisc	0.04***	-0.01	0.02*	-0.01	0.00	0.02**	-0.04***	0.03***
FinDisc	0.02***	0.01	0.01	-0.01*	-0.01	0.01	-0.03***	0.01
Strategy	0.03***	-0.00	0.02**	-0.01	-0.00	0.02**	-0.04***	0.01*
Changes	0.01	0.01**	-0.01	-0.02***	-0.04***	-0.00	-0.01**	-0.01
Extern	0.00	0.01	0.02***	-0.01	-0.00	0.01	-0.01	0.01
Macro	0.03***	0.00	0.02**	-0.01	-0.00	0.01	-0.03***	0.03***
Market	0.03***	-0.03***	0.03***	0.01	0.01	0.02	-0.03***	0.03***
OpIssue	0.00	0.03***	0.04***	-0.01*	-0.04***	0.01	-0.02**	-0.01*
Legal	0.02*	0.05***	0.01	-0.04***	-0.03**	-0.00	-0.03***	-0.01
Surprise	0.03**	0.00	0.01	-0.04**	-0.04**	0.02**	-0.05***	0.00
<b>Technical</b>	0.01	-0.01	0.01	0.00	0.00	0.01	-0.01	0.00
TechDiff	0.01	-0.01	0.01	0.00	0.00	0.01	-0.01	0.00

compared to their peers, while an unusual discussion of conflicting information does not generate a significant spread in the profitability of respective firms.

Unusual communication by *analysts* is correlated with larger market capitalization, slightly higher investments by the firms but lower profitability, as well as more liquid



shares that trade at higher nominal prices. *Technical* difficulties are unrelated to changes in firm characteristics. Ex-ante, we would not expect that differences in firm characteristics inform us about a greater chance of encountering technical difficulties, with which the data agrees. This is further evidence that ChatGPT can identify meaningful dimensions of unusual communication from earnings call transcripts.

Finally, we find that stocks with *content*-based unusual communication are, on average, larger, less profitable, and more likely momentum losers than winners, with more liquid shares that trade at higher nominal prices. Some findings of this content category are particularly noteworthy: for one, we find that firms that discuss legal issues (Legal) tend to be value stocks with a larger book-to-market ratio. We also find that firms that announce surprising news (Surprise) and corporate or management changes (Changes) are more likely momentum losers, suggesting that their past stock return has already been smaller than that of their peers. For both dimensions, we also find that the affected firms are less profitable.

In Table 5, we investigate if systematic differences in the propensity of engaging in unusual communication exist for firms in different industries. For this, we collect SIC codes for each firm in our sample and place them in one of twelve industries, using the definition by Kenneth French.<sup>6</sup> Unusual communication, overall, is relatively unrelated to a firm's industry affiliation. We find the lowest propensity for the Energy sector (33%) and the highest for Chemicals (49%). There's considerable heterogeneity across industries when it comes to unusual communication by firm *executives*. Here, we find the lowest values for Energy firms and firms in the "Shops" industry (18% for both). The highest prevalence instead is found for Consumer Durables (33%) and Chemicals (34%). For *financial analysts*, we find the lowest industry values for Energy (12%), Shops (13%), and Manufacturing (15%), and the highest once more for Consumer Durables (29%) and Chemicals (27%). *Technical difficulties* arise with roughly equal chance regardless of the firm's industry. Finally, unusual *contents* occur most frequently for Chemicals (39%), Consumer Durables (35%), Other (34%), Health and Consumer Non-Durables (33% for both). Discussions of Energy firms least frequently have unusual contents (22%). Overall, these results show that unusual communication is not exclusive to certain industries. While there is considerable heterogeneity in the propensity with which firms of a particular industry show signs of unusual communication in their earnings calls, industry

---

<sup>6</sup>[https://mba.tuck.dartmouth.edu/pages/faculty/ken.french/Data\\_Library/det\\_12\\_ind\\_port.html](https://mba.tuck.dartmouth.edu/pages/faculty/ken.french/Data_Library/det_12_ind_port.html).

Table 5: Unusual Communication Over Industries

The table shows the quarterly relative occurrences of *unusualness* along each of the 25 identified dimensions, as well as the four broader categories and *unusualness* in general in each of the twelve industries, using the definition by Kenneth French.

Dimension	NoDur	Durbl	Manuf	Enrgy	Chems	BusEq	Telcm	Utils	Shops	Hlth	Money	Other
<b>Unusual</b>	0.43	0.43	0.40	0.33	0.49	0.42	0.36	0.42	0.41	0.43	0.41	0.45
<b>Executives</b>	0.24	0.33	0.22	0.18	0.34	0.21	0.24	0.24	0.18	0.27	0.25	0.30
Lengthy	0.18	0.28	0.14	0.11	0.30	0.15	0.20	0.13	0.12	0.20	0.19	0.22
Informal	0.05	0.19	0.10	0.07	0.06	0.09	0.16	0.10	0.08	0.10	0.09	0.14
Dial	0.04	0.18	0.04	0.03	0.06	0.06	0.12	0.06	0.04	0.06	0.07	0.09
Emotion	0.02	0.09	0.03	0.03	0.03	0.04	0.04	0.04	0.03	0.05	0.04	0.05
Evasive	0.04	0.08	0.03	0.04	0.04	0.02	0.02	0.06	0.02	0.04	0.04	0.05
Unprep	0.02	0.03	0.01	0.00	0.02	0.01	0.01	0.02	0.01	0.01	0.03	0.03
Conflict	0.00	0.02	0.00	0.00	0.02	0.00	0.01	0.01	0.01	0.01	0.01	0.01
<b>Analysts</b>	0.17	0.29	0.15	0.12	0.27	0.15	0.20	0.17	0.13	0.21	0.18	0.20
RepPart	0.16	0.26	0.13	0.10	0.27	0.14	0.18	0.13	0.11	0.20	0.17	0.19
RepQ	0.04	0.05	0.03	0.03	0.06	0.04	0.02	0.04	0.04	0.05	0.04	0.05
QVolume	0.01	0.01	0.01	0.01	0.06	0.01	0.02	0.02	0.01	0.03	0.04	0.04
Off-Top	0.02	0.09	0.02	0.01	0.02	0.02	0.03	0.05	0.02	0.01	0.02	0.03
LackCrit	0.00	0.02	0.00	0.00	0.01	0.01	0.01	0.01	0.00	0.00	0.01	0.01
FinQuery	0.00	0.01	0.00	0.00	0.01	0.00	0.01	0.00	0.00	0.00	0.01	0.01
<b>Content</b>	0.33	0.35	0.30	0.22	0.39	0.31	0.25	0.29	0.28	0.33	0.29	0.34
NonFin	0.18	0.28	0.16	0.12	0.29	0.16	0.21	0.17	0.14	0.21	0.17	0.23
ProdDisc	0.18	0.28	0.14	0.09	0.28	0.15	0.21	0.11	0.12	0.20	0.16	0.21
FinDisc	0.16	0.23	0.13	0.09	0.28	0.13	0.18	0.13	0.11	0.18	0.18	0.21
Strategy	0.15	0.26	0.12	0.09	0.28	0.13	0.19	0.13	0.11	0.19	0.17	0.20
Changes	0.18	0.14	0.14	0.13	0.13	0.16	0.13	0.14	0.14	0.15	0.12	0.15
Extern	0.16	0.19	0.14	0.07	0.25	0.12	0.08	0.12	0.11	0.15	0.12	0.19
Macro	0.14	0.17	0.11	0.07	0.27	0.12	0.10	0.09	0.10	0.13	0.15	0.19
Market	0.12	0.19	0.10	0.03	0.26	0.11	0.18	0.04	0.07	0.15	0.12	0.15
OpIssue	0.13	0.24	0.11	0.06	0.19	0.07	0.13	0.09	0.11	0.12	0.08	0.14
Legal	0.04	0.06	0.03	0.03	0.06	0.05	0.05	0.09	0.02	0.08	0.06	0.06
Surprise	0.04	0.12	0.03	0.03	0.04	0.04	0.04	0.04	0.03	0.04	0.04	0.04
<b>Technical</b>	0.09	0.12	0.09	0.11	0.13	0.11	0.09	0.09	0.12	0.12	0.13	0.11
TechDiff	0.09	0.12	0.09	0.11	0.13	0.11	0.09	0.09	0.12	0.12	0.13	0.11

affiliation alone is insufficient to explain its occurrence.

### 3.3. When is Communication Unusual?

It is likely that some aspects of unusual communication will vary over time. For example, in hindsight, we would expect more in-depth discussions of the implications of the Coronavirus pandemic in 2020. At the start of the pandemic, this additional emphasis on a single topic is far from the norm and consequently unusual. As another example,

one may expect that unusual discussions of corporate changes or future strategic insights may co-move with merger waves.

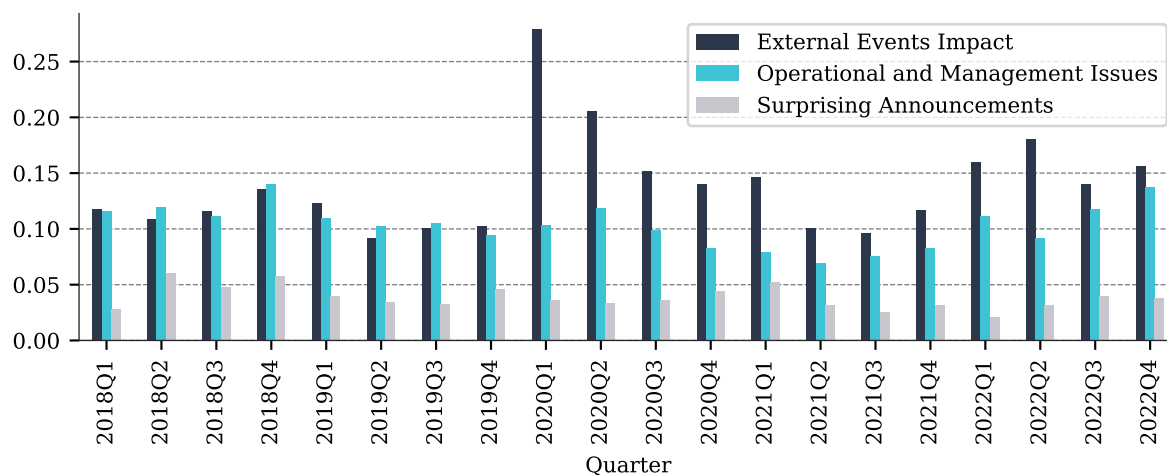


Fig. 1. *Unusualness* Patterns Over Time

The figure shows the relative occurrence of *unusualness* in terms of the three dimensions External Events Impact, Operational and Management Issues and Surprising Announcements over time.

To get a feeling for the time-variation in the different dimensions of unusual communication, Figure 1 plots to a fraction of firms with unusual patterns in their communication on “External Events Impact”, “Operational and Management Issues”, and “Surprising Announcements” between Q1-2018 and Q4-2022. The impact of external events consistently hovers around 10% in 2018 and 2019 but peaks in Q1-2020 at above 25%. A quarter of all firms discussed the implications of external events – predominantly, of course, the fallout of the Covid crisis. Interestingly, however, ChatGPT considers the communication style with this regard of *only* 1/4 of all firms as unusual. Discussions of the impact of external events remain elevated throughout 2020, calm down in much of 2021, and peak once more in early-2022, as the Russian war in Ukraine generates additional uncertainty also for U.S.-based firms. Unusual discussions of operational and management issues tend to be fairly stable over time, whereas surprising announcements see peaks in 2018 and Q1-2021.

We find compelling evidence that unusual communication has a business cycle component. For this, we aggregate the fraction of firms for which ChatGPT flags unusual communication patterns for each of the 25 identified dimensions in each calendar quarter and regress this fraction on various macroeconomic indicators and a constant. We include the price-dividend ratio of the S&P 500, net share issuance, the current Treasury

bill rate, term and default spread, all of which are taken from Welch and Goyal (2008). We also include Cboe's VIX, the intermediary capital ratio (ICR) of He et al. (2017), and the Chicago Fed National Activity Index (CFNAI). Because many of the macroeconomic indicators are highly correlated, we perform univariate regressions but show the results in a condensed table to save space.

Table 6: Unusual Communication and the Business Cycle

The table shows the results of regressing the average occurrence of unusual communication along each of the 25 identified dimensions, as well as the four broader categories and *unusualness* in general on a constant and one macroeconomic indicator. We include the price-dividend ratio of the S&P 500, net share issuance, the current Treasury bill rate, term and default spread, all of which are taken from Welch and Goyal (2008). We also include Cboe's VIX, the intermediary capital ratio (ICR) of He et al. (2017), and the Chicago Fed National Activity Index (CFNAI). Uncertainty measures are first averaged across firms within each calendar quarter and then normalized. Macroeconomic indicators are averaged within each calendar quarter and subsequently normalized. \*\*\* (\*\*, \*) denotes statistical significance at the 1% (5%, 10%) level.

Dimension	CFNAI	ICR	VIX	dfy	pd	ntis	tms
<b>Unusual</b>	-0.20	-0.48	0.59***	-0.64	-0.86**	-0.61*	-0.25
<b>Executives</b>	-0.06	0.24	0.02	-0.54	-0.50	-0.72**	0.12
Lengthy	-0.04	0.47	-0.08	-0.56	-0.57	-0.85***	-0.01
Informal	-0.13	-0.91*	0.44*	-0.53	-0.32	0.02	0.04
Dial	-0.18	-0.34	0.30	-0.43	-0.40	-0.20	0.46
Emotion	0.03	-0.37	0.37	0.21	-0.04	0.17	0.51
Evasive	-0.17	0.81	-0.12	-0.23	-1.17***	-0.91***	-0.53
Unprep	0.12	1.19**	-0.14	0.55	-0.53	-0.53	0.33
Conflict	0.15	1.14**	-0.53**	0.06	-0.75*	-0.84***	-0.37
<b>Analysts</b>	0.09	0.51	-0.10	-0.24	-0.25	-0.60*	0.16
RepPart	0.11	0.72	-0.18	-0.12	-0.21	-0.60*	0.17
RepQ	0.13	1.54***	-0.61***	0.36	-0.37	-0.65**	-0.01
QVolume	0.04	0.55	-0.20	-0.29	-0.48	-0.64*	-0.33
Off-Top	-0.12	-0.47	0.37	-0.33	-0.42	-0.13	0.04
LackCrit	0.10	0.57	-0.17	-0.21	-0.57	-0.76**	-0.37
FinQuery	0.22	0.24	-0.19	-0.12	-0.36	-0.47	-0.04
<b>Content</b>	-0.20	-0.07	0.49**	-0.42	-0.88**	-0.62*	-0.40
NonFin	0.02	0.28	-0.06	-0.45	-0.27	-0.62*	0.15
ProdDisc	0.00	0.75	-0.23	-0.39	-0.65	-0.92***	-0.21
FinDisc	-0.02	0.57	-0.13	-0.50	-0.61	-0.90***	-0.10
Strategy	-0.02	0.47	-0.06	-0.47	-0.48	-0.79***	0.04
Changes	0.20	0.49	-0.25	0.19	-0.45	-0.40	-0.27
Extern	-0.36***	-1.29***	0.95***	-0.78**	-0.43	-0.08	-0.14
Macro	-0.27*	-1.22***	0.67***	-1.14***	-0.14	-0.29	-0.11
Market	0.05	0.60	-0.12	-0.22	-0.42	-0.68**	0.10
OpIssue	-0.23	0.18	-0.00	-1.01***	-1.13***	-1.23***	-0.86*
Legal	0.24*	1.65***	-0.86***	0.48	-0.47	-0.65**	-0.24
Surprise	0.03	0.23	-0.18	-0.28	-0.80**	-0.67**	-0.35
<b>Technical</b>	0.06	-1.47***	0.40*	-0.62	0.17	0.16	0.24
TechDiff	0.06	-1.47***	0.40*	-0.62	0.17	0.16	0.24

Table 6 shows the results. We find that in case of unusual communication the CFNAI is associated with less discussions of external events and macroeconomic considerations. On the other hand, we find a slightly greater propensity to discuss legal and regulatory issues. Better capitalization by financial intermediaries (ICR) is also associated with fewer discussions of external and macroeconomic events as well as a greater propensity to discuss legal or regulatory issues. A one standard deviation in ICR corresponds to an increase in the discussion of legal issues by 1.65 standard deviations compared to the full sample mean. We also find that a higher VIX is associated with more discussions of the impact of external events and macroeconomic considerations.

A larger default spread (dfy) is negatively related to discussions of external and macroeconomic events, as well as operating issues. Similar results emerge for a higher pd-ratio. Interestingly, higher net share issuance (ntis) corresponds to a lower probability of firm executives to communicate in unusual ways. The same applies to unusual communication by financial analysts and the discussion of unusual content. For example, a one standard deviation increase in ntis corresponds to a 1.23 standard deviation decrease in the discussion of operating issues. Management may decide to issue additional shares in particularly rosy times. Alternatively, executives may be keen on portraying a stable picture of the company during earnings calls, especially when issuing shares (Pontiff and Woodgate, 2008; Greenwood and Hanson, 2012; McLean, Pontiff, and Reilly, 2020). The term spread (tms) is unrelated to how often firms in the aggregate engage in unusual communication.

## 4. Stock Market Reaction to Unusual Communication

We study the stock market's reaction to the different dimensions of unusual communication in earnings calls. As the first step, we use daily stock prices and calculate the cumulative stock return for firm  $i$  over a three-day window from  $t - 1$  to  $t + 1$  for the earnings call on day  $t$ .<sup>7</sup> We then compute the equally- and value-weighted returns of stocks with “unusual” and stocks with “usual” communication along each of the 25 identified dimensions, as well as the four broader categories. We also report the difference

---

<sup>7</sup>For earnings calls that take place after 4 p.m., i.e., after the close of the New York Stock Exchange, we use the next trading day as the event day.

between the two and assess the statistical significance of the return patterns.

**Announcement returns.** Table 7 produces several interesting findings: first, we find that unusual communication is associated with lower announcement returns. Whereas the average firm with standard communication patterns experiences a positive earnings announcement return (Savor and Wilson, 2016), we find that the average return of firms unusual communication is indistinguishable from zero. The equally-weighted difference between the two sets of firms (Unusual minus Usual) is negative and statistically significant. When value-weighting firms, we too find a negative difference, which is, however, not significantly different from zero.

Turning to the four more granular categories of unusual communication, we find that unusual communication by *firm executives* is associated with significantly smaller announcement returns, which holds true for both equally- and value-weighting firms. On average, a firm with unusual communication by its executives experiences  $-0.45\%$  smaller announcement returns (value-weighted:  $-0.74\%$ ). We find that unpreparedness and lengthy responses have the largest negative return impact, of  $-0.94\%$  and  $-0.75\%$ , respectively. In contrast, we do not find that executives' informal tone, emotional responses, conflicting information, or unusual dialogue patterns impact announcement returns.

Unusual communication by *financial analysts* is also associated with a negative return impact of  $-0.78\%$  equally-weighted and  $-1.08\%$  value-weighted. A lack of critical questioning and repetitive questions have the largest negative return impact of  $-2.39\%$  and  $-1.96\%$ , respectively. The impact of a high volume of questions and off-topic discussions hovers around zero and is statistically indistinguishable from it. Interestingly, we find that firms that experience *technical difficulties* during their earnings calls experience significantly larger announcement returns. This effect, however, vanishes when value-weighting the respective firms.

Finally, we find that unusual *contents* also generate significantly smaller announcement returns of  $-0.70\%$  (equally-weighted) or  $-1.13\%$  (value-weighted). The equally-weighted return impact of each dimension within the content-category is significantly negative when we weight firms in the two portfolios equally. The largest return impacts are produced by discussions of operating issues ( $-1.47\%$ ), surprising announcements ( $-1.20\%$ ), macroeconomic considerations ( $-0.88\%$ ), and legal issues ( $-0.84$ ). Overall, these results show that the stock market reacts to cues of unusual communication during earnings calls.

Table 7: Unusual Communication and Announcement Returns

The table shows the average earnings announcement return from  $t - 1$  to  $t + 1$  for firms with Unusual and firms with Usual communication along each of the 25 identified dimensions, as well as the four broader categories and unusualness in general. We also show the difference between the two (Unusual – Usual). In the columns on the left, we equally-weight firms in the two portfolios, in the column on the right we value-weight firms by their market capitalization at the end of the previous month. \*\*\* (\*\*, \*) denotes statistical significance at the 1% (5%, 10%) level.

Dimension	Equally-weighted			Value-weighted		
	Unusual	Usual	Diff	Unusual	Usual	Diff
<b>Unusual</b>	0.22	0.57***	−0.35**	−0.14	0.37	−0.51*
<b>Executives</b>	0.09	0.54***	−0.45***	−0.40	0.34	−0.73***
Lengthy	−0.19	0.56***	−0.75***	−0.64	0.33	−0.98***
Informal	0.60**	0.41***	0.19	−0.20	0.20	−0.40
Dial	0.40	0.43***	−0.03	−0.56	0.22	−0.78
Emotion	0.04	0.45***	−0.41	−0.21	0.17	−0.38
Evasive	−0.28	0.46***	−0.73*	−0.49	0.17	−0.67
Unprep	−0.49	0.45***	−0.94**	−0.65	0.16	−0.82
Conflict	0.32	0.44***	−0.11	−0.02	0.16	−0.13
<b>Analysts</b>	−0.22	0.57***	−0.78***	−0.73**	0.35	−1.08***
RepPart	−0.35	0.58***	−0.93***	−0.83**	0.35	−1.18***
RepQ	−1.45***	0.51***	−1.96***	−2.56**	0.28	−2.83**
QVolume	0.07	0.44***	−0.37	−0.32	0.17	−0.49
Off-Top	0.53	0.43***	0.10	0.60	0.15	0.45
LackCrit	−1.94*	0.45***	−2.39**	−4.02*	0.22	−4.24**
FinQuery	−1.26	0.44***	−1.71	−2.42	0.17	−2.56
<b>Content</b>	−0.06	0.65***	−0.70***	−0.63***	0.49**	−1.12***
NonFin	−0.05	0.54***	−0.58***	−0.63*	0.34	−0.97***
ProdDisc	−0.12	0.54***	−0.67***	−0.58	0.33	−0.91**
FinDisc	−0.18	0.55***	−0.73***	−0.61	0.32	−0.93**
Strategy	−0.12	0.54***	−0.66***	−0.58	0.31	−0.89**
Changes	0.03	0.50***	−0.47**	−0.28	0.23	−0.51*
Extern	−0.26	0.53***	−0.79***	−0.46	0.28	−0.73*
Macro	−0.33	0.55***	−0.88***	−0.47	0.27	−0.74**
Market	0.03	0.49***	−0.46**	−0.43	0.25	−0.67
OpIssue	−0.88***	0.59***	−1.47***	−1.14***	0.32	−1.46***
Legal	−0.37	0.47***	−0.84***	−0.36	0.19	−0.55
Surprise	−0.72*	0.48***	−1.20***	−1.79***	0.24	−2.03***
<b>Technical</b>	1.04***	0.37**	0.68**	0.77	0.05	0.72
TechDiff	1.04***	0.37**	0.68**	0.77	0.05	0.72

In our baseline specification in Table 7, we show the results for cumulative event returns. We also calculate cumulative *abnormal* returns using the Fama and French (1993) factors plus momentum. Beta coefficients to calculate expected returns are estimated over a 200-day period ending 30 days before the event day  $t$ , for which we require a minimum of 100 valid returns in the estimation window. In Table A.1, we show the results for an event window starting at  $t - 1$  and ending on  $t + 1$ . Table A.2 and Table A.3 replicate

the analysis extending the event window to  $t + 3$  and  $t + 10$ , respectively. The results agree with our baseline assessment.

**Trading Activity.** Announcement returns are significantly smaller for many dimensions of unusual communication. But do market participants also increase their trading activity upon encountering this unusual information? The literature has proposed trading volume as a measure for disagreement across investors (Goetzmann and Massa, 2005; Banerjee and Kremer, 2010; Banerjee, 2011). Hong and Stein (2007) argue that disagreement may arise when investors possess different information sets or if new information leads them to update their beliefs. Earnings calls are of course an important source of information for investors, and disagreement should be particularly high whenever the information is communicated in unusual ways, or the content is unusual itself.

Our empirical results, shown in Table 8, agree with this: firms with unusual communication experience significantly more trading than those with usual communication. We compare the trading activity from  $t - 1$  to  $t + 1$  to the average dollar trading volume measured over the past 30 days. The announcement trading volume of firms with standard communication amounts to 132.04% of this baseline, for firms with unusual communication, this number increases to 155.97% – a 23.93 percentage point (*pp*) increase, which is highly significant. Unusual communication by a firm's *executive* elevates the trading activity in the affected stocks by 31.46*pp* when equally-weighting and 20.58*pp* when value-weighting. Lengthy responses, an informal tone, and unusual dialogue patterns lead to significant (value-weighted) increases in trading. The other categories do not show significant differences.

A repeated focus by *financial analysts* on specific topics or participants, as well as repetitive questions and a high volume of questions, lead to significantly elevated trading activity in the affected stocks. For example, repetitive questions increase the trading activity to 217.46% compared to the baseline of the past 30 days, which is 87.22*pp* larger than for firms without unusually repetitive questions. Overall, unusual communication by analysts leads to 15.36*pp* higher trading volume when equally-weighting firms and 30.61*pp* when value-weighting firms.

Interestingly, *technical difficulties* do not give rise to more trading by market participants, corroborating the idea that the elevated trading volume we observe around unusual earnings calls is driven by participants disagreeing about the provided information. Unusual *contents* are also associated with elevated trading: 33.51*pp* for equally-weighted



Table 8: Unusual Communication and Announcement Trading Volume

The table shows the average announcement trading activity from  $t-1$  to  $t+1$  for firms with Unusual and firms with Usual communication along each of the 25 identified dimensions, as well as the four broader categories and unusualness in general. We compare the total dollar volume in the event window to the average over the 30 days before it. We also show the difference between the two (Unusual – Usual). In the columns on the left, we equally-weight firms in the two portfolios, in the column on the right we value-weight firms by their market capitalization at the end of the previous month. \*\*\* (\*\*, \*) denotes statistical significance at the 1% (5%, 10%) level.

Dimension	Equally-weighted			Value-weighted		
	Unusual	Usual	Diff	Unusual	Usual	Diff
<b>Unusual</b>	155.97***	132.04***	23.93**	144.87***	125.06***	19.81***
<b>Executives</b>	166.23***	134.78***	31.46*	148.77***	128.19***	20.58***
Lengthy	155.88***	139.25***	16.64***	156.78***	128.05***	28.72***
Informal	145.29***	141.72***	3.58	170.33***	128.99***	41.35***
Dial	156.20***	141.29***	14.91*	163.07***	130.21***	32.86***
Emotion	142.12***	142.15***	−0.03	151.90***	132.13***	19.76
Evasive	144.15***	141.89***	2.26	143.03***	132.88***	10.16
Unprep	283.55**	137.84***	145.71	160.21***	132.63***	27.58
Conflict	140.48***	142.18***	−3.85	146.71***	132.84***	12.98
<b>Analysts</b>	154.91***	139.56***	15.36***	158.38***	127.77***	30.61***
RepPart	157.92***	139.28***	18.64***	158.92***	128.27***	30.65***
RepQ	178.95***	140.94***	38.01***	217.46***	130.23***	87.22***
QVolume	165.62***	141.65***	23.96*	207.10***	131.77***	75.33*
Off-Top	133.15***	142.26***	−9.11	188.57***	131.47***	57.10*
LackCrit	170.93***	142.04***	28.89	227.56***	131.97***	95.59
FinQuery	188.24***	142.07***	46.17	231.88***	132.59***	99.94
<b>Content</b>	165.25***	131.75***	33.51**	154.12***	124.28***	29.84***
NonFin	157.06***	139.05***	18.01***	157.60***	127.64***	29.95***
ProdDisc	159.63***	138.72***	20.91***	158.75***	127.78***	30.97***
FinDisc	153.49***	139.96***	13.53**	150.42***	129.79***	20.62***
Strategy	154.36***	139.85***	14.51***	149.12***	129.84***	19.27***
Changes	177.49***	135.54***	41.96	161.22***	128.56***	32.67***
Extern	157.02***	140.07***	16.95***	154.40***	130.12***	24.29***
Macro	151.33***	140.68***	10.65	146.47***	130.66***	15.81**
Market	161.61***	139.80***	21.81***	162.01***	129.64***	32.37***
OpIssue	168.45***	139.20***	29.26***	154.04***	130.38***	23.66*
Legal	140.54***	142.28***	−1.74	132.90***	132.97***	−0.07
Surprise	175.89***	140.91***	34.98***	201.35***	130.45***	70.90**
<b>Technical</b>	144.61***	142.23***	2.38	131.76***	133.30***	−1.54
TechDiff	144.61***	142.23***	2.38	131.76***	133.30***	−1.54

firms and 29.81pp for value-weighted firms. The largest significant increases are driven by discussions of corporate changes, analyses of the firm's competitors or market, or its products and services. Equally- and value-weighted results tend to agree. Overall, the trading activity of firms with unusual communication is significantly elevated for 13 (equally-weighted; 17 when value-weighting) out of the 25 dimensions.

**Panel Regressions.** Our previous results show that earnings announcement returns are significantly more negative upon encountering unusual communication patterns in earnings calls. At the same time, market participants increase their trading activity, potentially resulting in disagreement about the conflicting information obtained. In panel regressions, we now rule out confounding effects. We control for the released earnings, analyst expectations thereof, and their forecast errors in panel regressions. We also control for differences in characteristics. We include market capitalization (Size), book-to-market ratio (B2M), investment behavior (Inv) and profitability (Prof), return momentum (Mom), 21-day MAX return following Bali et al. (2011), Amihud (2002) illiquidity (Illiq), and the stock’s nominal share price (PRC). Finally, we rule out that our results are exclusively driven by industry selection effects.

Table 9: Panel Regression: Unusual Communication and Announcements Returns

The table shows the results of regressing the cumulative earnings announcement return from  $t - 1$  to  $t + 1$  on the four broader categories of unusual firm communication, *Executives*, *Analysts*, *Contents* and *Technical Difficulties*. We furthermore control for differences in characteristics of the afflicted firms (“Chars”), standardized unexpected earnings, following the definition of Engelberg et al. (2018), analyst forecast dispersion, and the number of analysts covering a particular firm (“Disp. Cvg.”). We also control for the firm’s industry affiliation. Finally, we include calendar quarter fixed effects. Standard errors are clustered by calendar quarter and industry. \*\*\* (\*\*, \*) denotes statistical significance at the 1% (5%, 10%) level.

Dimension	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Unusual	−0.35**						
Executives		0.19	0.42	0.34	0.31	0.28	0.33
Analysts		−0.92***	−0.64***	−0.60**	−0.56**	−0.41**	−0.60**
Content			−0.61***	−0.58***	−0.58***	−0.66***	−0.57***
Technical				0.56**	0.55**	0.55*	0.55**
Chars					X		
Disp. Cvg.						X	
Industry FE							X
Time FE	X	X	X	X	X	X	X
$R^2$	0.07	0.22	0.32	0.40	1.08	0.52	0.39
N	9,414	9,414	9,414	9,414	9,414	8,400	9,414

Table 9 shows the results. Unusual communication of any form is associated with 35bps lower earnings announcement returns, mirroring the results of Table 7. We include both the *Executives* and the *Analysts* indicator in regression (2). Here, we find that only the impact of unusual communication by analysts remains negative and statistically significant. If we also include unusual *Contents* in regression (3), we find that both unusual communication by analysts is associated with significantly lower announcement returns of −64bps, as well as unusual contents (−61bps). In regression (4), we also include

the occurrence of technical difficulties, for which we find a positive return impact.

We control for differences in firm characteristics between firms with and without unusual communication in regression (5). The results from regression (4) remain largely unchanged: unusual communication by analysts and discussions of unusual contents both coincide with significantly lower earnings announcement returns. Controlling for standardized unexpected earnings, analyst forecast dispersion, and the number of analysts covering a firm produces similar results. We do find, however, that the coefficient for *Analysts* reduces from  $-64bps$  to  $-41bps$ . Finally, we include industry fixed effects using the classification based on SIC codes on Kenneth French's website. As in Table 5, we classify firms into 12 distinct industries. The inclusion of the fixed effects leaves the coefficients on the different categories of unusual communication unchanged.

## 5. Conclusion

We systematically identify various dimensions of unusual financial communication using earnings calls of S&P 500 companies and study how the stock market reacts to it. Traditionally, financial analysts in particular try to identify unusual patterns in earnings calls. However, this approach is inherently subjective due to the human factor and the resulting assessment may vary substantially between analysts and over time. This subjectivity can lead to potential biases in perception. In contrast, the use of ChatGPT provides us with a standardized and consistent approach to identifying unusual financial communication patterns, which minimizes the analysis's subjectivity. In addition, GPT-4-Turbo's ability to process large bodies of text allows for a more comprehensive and in-depth analysis of financial communication than the model's previous iterations.

Our research makes a significant theoretical contribution by expanding the understanding of the impact of financial communication on market dynamics. On the one hand, we uncover unusual communication patterns that have been elusive to prior research. On the other hand, our findings also provide a more detailed understanding of how unusual communication is perceived by market participants. We find a robust negative stock market reaction to various dimensions of unusual communication, including by the firm's executives, financial analysts, as well as the discussion's content. In addition, the study expands our understanding of the potential applications for modern machine learning tools in financial market research. Tools like ChatGPT are particularly useful in extract-

ing information from large and diverse textual datasets.

Our study also has important implications for practitioners. Financial analysts, asset managers, and other market participants can use our prompting approach for ChatGPT to identify unusual financial communication patterns, which allows them to make more informed decisions. We demonstrate how artificial intelligence can benefit the day-to-day applications of financial analysts.

## References

- Amihud, Y., 2002. Illiquidity and Stock Returns: Cross-Section and Time-Series Effects. *Journal of Financial Markets* 5, 31–56.
- Bai, J. J., Boyson, N. M., Cao, Y., Liu, M., Wan, C., 2023. Executives vs. Chatbots: Unmasking Insights through Human-AI Differences in Earnings Conference Q&A. Working Paper.
- Baker, S. R., Bloom, N., Davis, S. J., 2016. Measuring Economic Policy Uncertainty. *The Quarterly Journal of Economics* 131, 1593–1636.
- Baker, S. R., Bloom, N., Davis, S. J., Sammon, M. C., 2021. What Triggers Stock Market Jumps? Working Paper.
- Bali, T. G., Cakici, N., Whitelaw, R. F., 2011. Maxing Out: Stocks as Lotteries and the Cross-Section of Expected Returns. *Journal of Financial Economics* 99, 427–446.
- Banerjee, S., 2011. Learning from Prices and the Dispersion in Beliefs. *The Review of Financial Studies* 24, 3025–3068.
- Banerjee, S., Kremer, I., 2010. Disagreement and Learning: Dynamic Patterns of Trade. *The Journal of Finance* 65, 1269–1302.
- Breuer, W., Knetsch, A., et al., 2023. It’s not What you Say, But How you Say it – Charismatic Rhetoric in Earnings Conference Calls. Working Paper.
- Bybee, L., 2023. The Ghost in the Machine: Generating Beliefs with Large Language Models. Working Paper.
- Bybee, L., Kelly, B. T., Manela, A., Xiu, D., 2023a. Business News and Business Cycles. *The Journal of Finance* (Forthcoming).
- Bybee, L., Kelly, B. T., Su, Y., 2023b. Narrative Asset Pricing: Interpretable Systematic Risk Factors from News Text. *The Review of Financial Studies* 36, 4759–4787.
- Cao, S., Jiang, W., Yang, B., Zhang, A. L., 2023. How to Talk When a Machine Is Listening: Corporate Disclosure in the Age of AI. *The Review of Financial Studies* 36, 3603–3642.
- Chen, J., Tang, G., Zhou, G., Zhu, W., 2023. ChatGPT, Stock Market Predictability and Links to the Macroeconomy. Working Paper.
- de Kok, T., 2023. Generative LLMs and Textual Analysis in Accounting: (Chat) GPT as Research Assistant? Working Paper.

- Doherty, K., Marques, F., 2023. Citadel Negotiating Enterprise Wide ChatGPT License, Griffin Says. Bloomberg Available at: <https://www.bloomberg.com/news/articles/2023-03-07/griffin-says-trying-to-negotiate-enterprise-wide-chatgpt-license?embedded-checkout=true> (Accessed: December 8th, 2023).
- Dong, M. M., Stratopoulos, T. C., Wang, V. X., 2023. A Scoping Review of ChatGPT Research in Accounting and Finance. Working Paper.
- Engelberg, J., McLean, R. D., Pontiff, J., 2018. Anomalies and News. *The Journal of Finance* 73, 1971–2001.
- Ernst, R., Gilbert, T., Hrdlicka, C. M., 2019. More than 100% of the Equity Premium: How Much is Really Earned on Macroeconomic Announcement Days? Working Paper.
- Fama, E. F., French, K. R., 1993. Common Risk Factors in the Returns on Stocks and Bonds. *Journal of Financial Economics* 33, 3–56.
- Fama, E. F., MacBeth, J. D., 1973. Risk, Return, and Equilibrium: Empirical Tests. *Journal of Political Economy* 81, 607–636.
- Gentzkow, M., Kelly, B., Taddy, M., 2019. Text as Data. *Journal of Economic Literature* 57, 535–574.
- Goetzmann, W. N., Massa, M., 2005. Dispersion of Opinion and Stock Returns. *Journal of Financial Markets* 8, 324–349.
- Gorodnichenko, Y., Pham, T., Talavera, O., 2023. The Voice of Monetary Policy. *American Economic Review* 113, 548–584.
- Goyal, T., Li, J. J., Durrett, G., 2022. News Summarization and Evaluation in the Era of GPT-3. ArXiv preprint arXiv:2209.12356.
- Greenwood, R., Hanson, S. G., 2012. Share Issuance and Factor Timing. *The Journal of Finance* 67, 761–798.
- Hassan, T. A., Hollander, S., Van Lent, L., Schwedeler, M., Tahoun, A., 2020a. Firm-Level Exposure to Epidemic Diseases: Covid-19, SARS, and H1N1. Working Paper.
- Hassan, T. A., Hollander, S., Van Lent, L., Tahoun, A., 2019. Firm-Level Political Risk: Measurement and Effects. *The Quarterly Journal of Economics* 134, 2135–2202.
- Hassan, T. A., Hollander, S., Van Lent, L., Tahoun, A., 2020b. The Global Impact of Brexit Uncertainty. Working Paper.
- He, Z., Kelly, B., Manela, A., 2017. Intermediary Asset Pricing: New Evidence from Many Asset Classes. *Journal of Financial Economics* 126, 1–35.

- Hong, H., Stein, J. C., 2007. Disagreement and the Stock Market. *Journal of Economic perspectives* 21, 109–128.
- Hou, K., Moskowitz, T. J., 2005. Market Frictions, Price Delay, and the Cross-Section of Expected Returns. *The Review of Financial Studies* 18, 981–1020.
- Hu, K., 2023. ChatGPT Sets Record for Fastest-Growing User Base – Analyst Note. Reuters Available at: <https://www.reuters.com/technology/chatgpt-sets-record-fastest-growing-user-base-analyst-note-2023-02-01/> (Accessed: December 8th, 2023).
- Jensen, T. I., Kelly, B. T., Pedersen, L. H., 2023. Is There A Replication Crisis In Finance? *Journal of Finance* 78.
- Jha, M., Qian, J., Weber, M., Yang, B., 2023. ChatGPT and Corporate Policies. Working Paper.
- Katz, D. M., Bommarito, M. J., Gao, S., Arredondo, P., 2023. GPT-4 passes the bar exam. Working Paper.
- Kim, A., Muhn, M., Nikolaev, V., 2023a. From Transcripts to Insights: Uncovering Corporate Risks Using Generative AI. ArXiv preprint arXiv:2310.17721.
- Kim, A. G., Muhn, M., Nikolaev, V. V., 2023b. Bloated Disclosures: Can ChatGPT Help Investors Process Information? Working Paper.
- Kim, J. H., 2023. What if ChatGPT Were a Quant Asset Manager. *Finance Research Letters* 58, 104580.
- Lee, J., 2016. Can Investors Detect Managers' Lack of Spontaneity? Adherence to Predetermined Scripts During Earnings Conference Calls. *The Accounting Review* 91, 229–250.
- Li, K., Mai, F., Shen, R., Yan, X., 2021. Measuring Corporate Culture Using Machine Learning. *The Review of Financial Studies* 34, 3265–3315.
- Liang, L., 2003. Post-Earnings Announcement Drift and Market Participants' Information Processing Biases. *Review of Accounting Studies* 8, 321–345.
- Lopez-Lira, A., Tang, Y., 2023. Can ChatGPT Forecast Stock Price Movements? Return Predictability and Large Language Models. ArXiv preprint arXiv:2304.07619.
- Loughran, T., McDonald, B., 2011. When is a Liability not a Liability? Textual Analysis, Dictionaries, and 10-Ks. *The Journal of Finance* 66, 35–65.
- McLean, R. D., Pontiff, J., Reilly, C., 2020. Taking Sides on Return Predictability. Working Paper.

- Pelster, M., Val, J., 2023. Can ChatGPT Assist in Picking Stocks? *Finance Research Letters* p. 104786.
- Pontiff, J., Woodgate, A., 2008. Share Issuance and Cross-Sectional Returns. *The Journal of Finance* 63, 921–945.
- Savor, P., Wilson, M., 2013. How Much do Investors Care About Macroeconomic Risk? Evidence from Scheduled Economic Announcements. *Journal of Financial and Quantitative Analysis* 48, 343–375.
- Savor, P., Wilson, M., 2016. Earnings Announcements and Systematic Risk. *The Journal of Finance* 71, 83–138.
- Spataro, J., 2023. Introducing Microsoft 365 Copilot – your Copilot for Work. Microsoft Available at: <https://blogs.microsoft.com/blog/2023/03/16/introducing-microsoft-365-copilot-your-copilot-for-work/> (Accessed: December 8th, 2023).
- Welch, I., Goyal, A., 2008. A Comprehensive Look at the Empirical Performance of Equity Premium Prediction. *The Review of Financial Studies* 21, 1455–1508.
- Wu, S., Irsoy, O., Lu, S., Dabravolski, V., Dredze, M., Gehrmann, S., Kambadur, P., Rosenberg, D., Mann, G., 2023. BloombergGPT: A Large Language Model for Finance. Working Paper.
- Zhang, X. F., 2006. Information Uncertainty and Stock Returns. *Journal of Finance* 61, 105–137.
- Zhang, Y., Ni, A., Yu, T., Zhang, R., Zhu, C., Deb, B., Celikyilmaz, A., Awadallah, A. H., Radev, D., 2021. An Exploratory Study on Long Dialogue Summarization: What Works and What’s Next. ArXiv preprint arXiv:2109.04609.



# Appendix

## A. Cumulative Abnormal Earnings Announcement Returns

Table A.1: Unusual Communication and Abnormal Returns – FF4 (–1/ + 1)

The table shows the average cumulative abnormal earnings announcement return from  $t - 1$  to  $t + 1$  for firms with Unusual and firms with Usual communication along each of the 25 identified dimensions, as well as the five broader categories. We also show the difference between the two (Unusual – Usual). In the columns on the left, we equally-weight firms in the two portfolios, in the column on the right we value-weight firms by their market capitalization at the end of the previous month. We use the Fama and MacBeth (1973) factor model augmented with momentum and estimated over a 200-day period ending 30 days before event day  $t$  to generate abnormal returns. We require at least 100 valid return observations during this period. \*\*\* (\*\*, \*) denotes statistical significance at the 1% (5%, 10%) level.

Dimension	Equally-weighted			Value-weighted		
	Unusual	Usual	Diff	Unusual	Usual	Diff
<b>Unusual</b>	0.02	0.31**	–0.29**	–0.32	0.26	–0.57**
<b>Executives</b>	–0.12	0.29**	–0.41***	–0.48*	0.19	–0.67**
Lengthy	–0.38**	0.31***	–0.68***	–0.58*	0.17	–0.75**
Informal	0.31*	0.18	0.13	–0.45	0.08	–0.53
Dial	0.21	0.19*	0.03	–0.77	0.09	–0.86
Emotion	–0.13	0.20*	–0.33	–0.31	0.04	–0.34
Evasive	–0.33	0.21*	–0.54	–0.62	0.04	–0.65
Unprep	–0.24	0.20*	–0.44	–0.25	0.02	–0.27
Conflict	0.63	0.19*	0.45	0.51	0.02	0.49
<b>Analysts</b>	–0.40***	0.31***	–0.72***	–0.73**	0.19	–0.92**
RepPart	–0.51***	0.32***	–0.83***	–0.76**	0.19	–0.95**
RepQ	–1.54***	0.26**	–1.80***	–2.50**	0.14	–2.64**
QVolume	–0.20	0.20*	–0.40	–0.65	0.04	–0.68
Off-Top	0.30	0.18*	0.12	0.53	0.01	0.52
LackCrit	–1.17	0.21**	–1.38	–3.18	0.07	–3.26
FinQuery	–1.29	0.20**	–1.49	–2.56	0.03	–2.60
<b>Content</b>	–0.22**	0.38***	–0.60***	–0.73***	0.34**	–1.07***
NonFin	–0.26*	0.29**	–0.54***	–0.66**	0.19	–0.85**
ProdDisc	–0.31*	0.29**	–0.60***	–0.52	0.16	–0.68*
FinDisc	–0.36**	0.30***	–0.66***	–0.62*	0.16	–0.78**
Strategy	–0.30*	0.28**	–0.58***	–0.59	0.15	–0.74*
Changes	–0.15	0.25**	–0.40**	–0.54**	0.12	–0.65**
Extern	–0.44**	0.27**	–0.71***	–0.58	0.11	–0.69
Macro	–0.49***	0.29**	–0.78***	–0.45	0.09	–0.54
Market	–0.13	0.24**	–0.37*	–0.46	0.09	–0.55
OpIssue	–1.04***	0.34***	–1.38***	–1.19***	0.17	–1.36***
Legal	–0.32	0.22**	–0.54**	–0.64	0.06	–0.70
Surprise	–0.88**	0.23**	–1.11***	–1.98***	0.11	–2.09***
<b>Technical</b>	0.73***	0.13	0.59***	0.46	–0.06	0.52
TechDiff	0.73***	0.13	0.59***	0.46	–0.06	0.52

Table A.2: Unusual Communication and Abnormal Returns – FF4 (–1/ + 3)

The table shows the average cumulative abnormal earnings announcement return from  $t - 1$  to  $t + 3$  for firms with Unusual and firms with Usual communication along each of the 25 identified dimensions, as well as the five broader categories. We also show the difference between the two (Unusual – Usual). In the columns on the left, we equally-weight firms in the two portfolios, in the column on the right we value-weight firms by their market capitalization at the end of the previous month. We use the Fama and MacBeth (1973) factor model augmented with momentum and estimated over a 200-day period ending 30 days before event day  $t$  to generate abnormal returns. We require at least 100 valid return observations during this period. \*\*\* (\*\*, \*) denotes statistical significance at the 1% (5%, 10%) level.

Dimension	Equally-weighted			Value-weighted		
	Unusual	Usual	Diff	Unusual	Usual	Diff
<b>Unusual</b>	0.02	0.35**	–0.33***	–0.28	0.18	–0.46**
<b>Executives</b>	–0.09	0.30**	–0.40**	–0.39	0.14	–0.53
Lengthy	–0.38**	0.33**	–0.72***	–0.63	0.14	–0.77*
Informal	0.26	0.21	0.06	–0.44	0.05	–0.49
Dial	0.25	0.21	0.04	–0.60	0.04	–0.64
Emotion	–0.08	0.22*	–0.31	–0.18	0.00	–0.18
Evasive	–0.44	0.24*	–0.67	–0.34	–0.01	–0.34
Unprep	–0.05	0.22*	–0.26	0.28	–0.02	0.30
Conflict	0.38	0.21*	0.18	0.17	–0.01	0.18
<b>Analysts</b>	–0.38**	0.34**	–0.71***	–0.61	0.14	–0.75*
RepPart	–0.51***	0.35***	–0.86***	–0.78*	0.15	–0.93**
RepQ	–1.70***	0.29**	–1.99***	–2.73**	0.11	–2.85**
QVolume	–0.39	0.22*	–0.61	–0.98	0.01	–0.99
Off-Top	0.41	0.20*	0.21	0.84	–0.04	0.88
LackCrit	–1.45	0.23*	–1.68	–3.81	0.04	–3.85*
FinQuery	–1.14	0.22*	–1.36	–2.75	0.01	–2.73
<b>Content</b>	–0.26**	0.43***	–0.69***	–0.72***	0.30**	–1.02***
NonFin	–0.26	0.31**	–0.58***	–0.57	0.14	–0.71*
ProdDisc	–0.33	0.32**	–0.64***	–0.57	0.14	–0.71
FinDisc	–0.37*	0.32**	–0.69***	–0.68*	0.14	–0.82**
Strategy	–0.31*	0.31**	–0.62***	–0.58	0.12	–0.70
Changes	–0.21	0.29**	–0.50**	–0.63**	0.10	–0.72**
Extern	–0.41**	0.30**	–0.71***	–0.51	0.08	–0.60
Macro	–0.45**	0.32**	–0.77***	–0.51	0.08	–0.59
Market	–0.16	0.26**	–0.42*	–0.49	0.07	–0.56
OpIssue	–1.04***	0.36***	–1.41***	–1.05**	0.13	–1.18**
Legal	–0.29	0.24**	–0.52*	–0.36	0.02	–0.37
Surprise	–0.87**	0.26**	–1.13***	–1.99***	0.08	–2.07***
<b>Technical</b>	0.82***	0.15	0.67***	0.79**	–0.13	0.92**
TechDiff	0.82***	0.15	0.67***	0.79**	–0.13	0.92**

Table A.3: Unusual Communication and Abnormal Returns – FF4 (−1/ + 10)

The table shows the average cumulative abnormal earnings announcement return from  $t - 1$  to  $t + 10$  for firms with Unusual and firms with Usual communication along each of the 25 identified dimensions, as well as the five broader categories. We also show the difference between the two (Unusual – Usual). In the columns on the left, we equally-weight firms in the two portfolios, in the column on the right we value-weight firms by their market capitalization at the end of the previous month. We use the Fama and MacBeth (1973) factor model augmented with momentum and estimated over a 200-day period ending 30 days before event day  $t$  to generate abnormal returns. We require at least 100 valid return observations during this period. \*\*\* (\*\*, \*) denotes statistical significance at the 1% (5%, 10%) level.

Dimension	Equally-weighted			Value-weighted		
	Unusual	Usual	Diff	Unusual	Usual	Diff
<b>Unusual</b>	0.04	0.36	−0.32**	−0.26	−0.09	−0.17
<b>Executives</b>	−0.16	0.34*	−0.49**	−0.50	−0.03	−0.46
Lengthy	−0.46*	0.36*	−0.82***	−0.77*	−0.02	−0.76
Informal	0.12	0.24	−0.11	−0.40	−0.11	−0.28
Dial	−0.09	0.25	−0.34	−0.70	−0.09	−0.61
Emotion	−0.51	0.25	−0.76*	−0.27	−0.14	−0.13
Evasive	−0.59	0.26	−0.85	0.13	−0.16	0.29
Unprep	−0.15	0.23	−0.38	0.73	−0.17	0.91
Conflict	0.27	0.23	0.08	0.37	−0.15	0.54
<b>Analysts</b>	−0.45**	0.36*	−0.81***	−0.80**	−0.02	−0.78
RepPart	−0.54**	0.36*	−0.91***	−0.95**	−0.00	−0.95*
RepQ	−1.46***	0.30	−1.75***	−2.56**	−0.05	−2.51**
QVolume	−0.90	0.25	−1.15**	−1.19	−0.13	−1.06
Off-Top	0.08	0.22	−0.14	0.81	−0.17	0.99
LackCrit	−1.87	0.24	−2.11	−3.65	−0.12	−3.52
FinQuery	−1.34	0.24	−1.58	−2.52	−0.14	−2.33
<b>Content</b>	−0.29*	0.46**	−0.75***	−0.75***	0.11	−0.86***
NonFin	−0.35	0.35*	−0.70***	−0.70*	−0.01	−0.69
ProdDisc	−0.32	0.33*	−0.65**	−0.66	−0.03	−0.63
FinDisc	−0.41	0.33*	−0.74***	−0.69*	−0.04	−0.65
Strategy	−0.34	0.32*	−0.66***	−0.69	−0.05	−0.65
Changes	−0.24	0.31	−0.55**	−0.54*	−0.08	−0.46
Extern	−0.36	0.31	−0.67***	−0.32	−0.09	−0.23
Macro	−0.48*	0.33*	−0.81***	−0.55	−0.08	−0.47
Market	−0.17	0.28	−0.44*	−0.58	−0.09	−0.49
OpIssue	−0.97***	0.37*	−1.34***	−1.21**	−0.02	−1.19*
Legal	−0.11	0.25	−0.35	−0.31	−0.12	−0.19
Surprise	−1.18**	0.29	−1.46**	−2.63***	−0.05	−2.58***
<b>Technical</b>	0.85***	0.16	0.69***	0.80	−0.30*	1.10*
TechDiff	0.85***	0.16	0.69***	0.80	−0.30*	1.10*