Enhancing Investor Engagement with AI-Summarized Disclosures

T.J. WONG*
YANG YI√
GWEN YU†α
SHUBO ZHANG‡
TIANYU ZHANG#

April 2025

ABSTRACT

We conduct a field experiment where we provide investors with an AI-generated summary of annual reports during virtual conference calls. We find that providing investors with annual report summaries increases investor participation during the calls. Specifically, treatment firms with AI-generated summaries experience a 46% increase in the number of investor questions, relative to control firms. Moreover, the content of investors' questions directed toward the treatment firms is more aligned with the topics presented in the summaries; the treatment firms' questions are more likely to come from less experienced investors; and treatment firms' management provides longer and more detailed answers. The findings suggest that AI-generated summaries can lead to greater investor engagement by providing investors with focal points for more vibrant conversations with management.

*University of Southern California
√ Southwestern University of Finance and Economics
†University of Michigan
‡Shanghai JiaoTong University, China
Shenzhen Finance Institute, The Chinese University of Hong Kong, Shenzhen
α Corresponding author: 701 Tappan Street R3350, Ann Arbor MI 48109. gwyu@umich.edu.

JEL classification: G00, M40, M41.

Keywords: Retail investors, Generative AI, LLMs, information processing costs, field experiment, virtual conference calls.

Acknowledgements: We are grateful for helpful feedback from workshop participants at BYU, George Washington University, Harvard Business School, Shanghai University of Finance and Economics, USC, the University of Michigan Centennial Conference, and the Hosmer Interdisciplinary Research Lunch. We thank Fei Jiang, Jinxing Hao, and Wenqin Yu for excellent research assistance. Yu acknowledges financial support from the Ross School of Business.

1. INTRODUCTION

As company disclosures become increasingly complex, investors are faced with the increased burden of processing a large amount of information. While some argue that information intermediaries can help reduce investors' processing costs, such resources are not accessible to everyone—and are often inaccessible to the less sophisticated investors who most need them. In this paper, we test whether providing investors with summaries of annual reports leads to increased investor participation, which in turn could help investors process annual reports' content (Blankespoor et al. 2020).

Providing summaries could increase investor participation by serving as a focal point for communications with management. Summaries could help investors extract key insights and attend to information they might overlook in the full report (Hirshleifer and Teoh 2023).[1] Investors' questions based on a summary may thus be more informed and lead to deeper discussions of more relevant topics. However, summaries also eliminate information contained in the full report, some of which may be important. If investors rely solely on summaries without considering the full report, they may receive a limited information set. Consequently, summaries may lead investors to ignore relevant information from the report.

A challenge in examining summaries in archival settings is that they are often produced by managers, who may introduce bias by selectively highlighting favorable insights (Cardinaels et al. 2019). A key feature of our study is that the summary reports are not produced

[1] The idea that summary reports will help reduce information-processing costs, especially for less sophisticated investors, has been the primary motive for the SEC's proposal of the "two-tiered" financial reporting system (SEC 1995)—an earlier regulation that considered requiring firms to disclose summary reports of the 10-Ks.

by firms. Instead, we use AI-generated summaries of firms' annual reports, created by computer algorithms that have no incentives to favor the firm. In an experimental setting, we provide these AI-generated summaries and examine whether they lead to more investor participation in the form of greater attendance and more questions asked.

We conduct our experiment in China using the annual report briefing meetings known as earnings communication conferences (ECCs). Starting in 2004, all listed companies on China's main stock exchange were directed to hold an ECC within 15 trading days of their annual report release. One advantage of the ECC setting is that the calls are almost exclusively online.[2] The virtual setting allows us to include a broader audience range, as participants can join from anywhere by simply clicking on the platform. Another important feature of the ECC setting is that any participant can ask questions using the open chat feature. This contrasts with earnings conference calls in the U.S., where participants cannot ask a question unless they are called on by management.[3] The setting allows us to collect the complete set of questions raised by all investors and the corresponding answers. The observability of the complete set of questions and the subsequent dialogue allows us to understand how reducing investors' information-processing costs affects communication between management and investors.

We focus on annual reports for several reasons. Prior studies find that investors face significant information overload from annual reports, which have become increasingly lengthy

[2] While the regulation does not stipulate the format of the conference calls (it only requires firms to host calls in a format easily accessible to investors), virtual calls have become widely popular since the onset of COVID-19. As of 2024, more than 90% of companies on the main exchanges host their ECCs using an online platform, according to statistics released by the China Listed Companies Association (CLCA).
[3] Studies find that management discretion can lead to selective participation and skew the questions towards more sophisticated investors, such as security analysts and institutional investors (Mayew 2008; Brown, Call, Clement, and Sharp 2015).

over time (Guay et al. 2016; Bonsall et al. 2017; Bernard et al. 2023).[4] The reports' bloated content may not always contain new information and imposes significant processing costs on investors (Dyer et al. 2017; Kim et al. 2024). In addition, prior studies find that generative AI technology can be helpful in reducing annual reports' complexity.[5]

Following this literature, we use generative AI to populate summaries of five key topics in each company's annual report, then present these summaries to investors in a bullet point format. We choose five as the number of topics based on studies showing that the cognitive overload of human working memory starts after processing 7(+/-2) units of information (Atkinson and Shiffrin 1968).

Our sample includes all firms that hosted virtual calls on the Quanjing platform (also known as P5W Net) in 2023. The largest conference call platform in China, Quanjing hosts conference calls for more than 47% (1,301 out of 2,771) of the firms listed on the Shenzhen Stock Exchange.[6] Our sample includes 1,105 firms that are listed on the main board of any of the three major stock exchanges—Shenzhen (SZSE), Beijing, and Shanghai—and that held an ECC on Quanjing's platform in 2023.

For treatment firms, we inserted, on each firm's 2023 ECC announcement page, an "Annual Report Highlights" icon that linked to the AI-generated annual report summary [see Figure 1]. The icon was visible during the ECC presentations, which, based on the timestamps of investor

[4] The average length of annual reports in China is approximately 200 pages, which is comparable to the average length for U.S. firms.

[5] Cardinaels et al. (2019) find that AI-generated summaries are superior to summaries by managers and help investors arrive at more conservative valuation estimates.

[6] Quanjing was owned by the Shenzhen Stock Exchange before People's Daily acquired it. It hosts the largest number of conference calls in China, primarily covering firms listed in Shenzhen. Other platforms, such as China Securities and SEE, are smaller but focus on firms listed on the Beijing and Shanghai exchanges.

clicks, was when most investors accessed the summary. When investors clicked on the link, a pop-up window with our five-point summary of the annual report populated on their screens [see Figure 2].[7] This was our baseline treatment. We added an additional treatment where each of the five points was classified by its sentiment: positive, negative, or neutral. The second treatment was designed to reduce investors' information-processing cost of interpreting the sentiment of the summarized topic.

We find that including summaries of annual reports leads to a significant increase in investor engagement. Treatment firms receive significantly more investor questions: 19.86, compared to 14.27 for the control firms. The summaries do not necessarily lead to more investors attending the calls but do seem to reduce the information-processing costs of the investors in attendance.

While the increase in the number of questions is promising, it does not necessarily indicate improved engagement, especially if the questions raised are irrelevant.[8] It is possible that our summaries make participants ask more off-topic questions because they want to differentiate themselves from others by broaching issues not already raised. The fact that all questions are publicly visible on the platform could make investors more susceptible to such an audience effect.[9]

[7] The average number of clicks on each link was 23 (range: 0 to 200), which was 10% of the average number of participants on the calls.

[8] Off-topic questions can disrupt the flow of the call or take up time that could otherwise be used on more meaningful questions.

[9] The "audience effect," one of the oldest effects studied in psychology, refers to the tendency of individuals to change their behavior in the presence of others (Triplett 1898; Zajonc 1965).

To help rule out an audience effect, we analyze the content of the investor questions and test whether presenting our summaries increases the likelihood that the questions align with the topics in the summary. We also generated summary points and topics for the control firms, but did not make them visible to the investors. In the treatment firms, we find greater alignment between the topics investors raise and the topics of the summary points. Relative to control firms, the questions asked of treatment firms are 7% more likely (50% vs. 57%) to align with topics from the summary. We also examine what types of topics investors are more likely to pick up from the summaries. We find that investors are more likely to focus on firm-specific topics such as financial information, risks, strategy, and payout policy. Also, investors are more likely to pick up on summary points with negative sentiment than on summary points with positive sentiment.

We next examine the characteristics of the investors whose questions align with the summary topics. Due to the anonymous nature of ECCs, we do not have the identities of the investors who post questions, so we instead consider their track record on the platform using their anonymized IDs.[10] Specifically, we measure their experience level based on how active they were on the platform in the prior year. We find that silent investors—those who did not ask any questions in 2022—are more likely to ask questions that are guided by the summaries. Vocal investors—those who asked questions in 2022—pose slightly more questions than before but are less likely than silent investors to ask questions guided by the summaries. Interestingly,

[10] When participants register on the Quanjing platform, they are asked to authenticate their identities by providing their citizen number or phone number. Due to security reasons and IRB requirements, we were provided with the registrants' anonymized registered IDs (if available) but not their phone numbers.

the experienced investors shift to topics not raised in the annual report summary. These findings suggest that the summaries are more likely to guide investors with less experience. The concurrent, albeit modest, increase in the number of questions asked by experienced investors indicates that more participation by the inexperienced does not necessarily displace participation by the experienced.

When we examine the properties of the management responses, we find that the treatment firms' answers are significantly longer, averaging 35 more words than the control firms'. The treatment firms' answers also more directly address the investors' questions and provide more specifics, with more numerical data and supporting evidence. We also test whether the responses differ between questions that are, and questions that are not, topically aligned with the summaries. We find a significant improvement in response quality for all questions, regardless of alignment. This improvement for both question types indicates an overall enhancement in the information content of the conference call.

Finally, we examine capital market responses to the ECCs and other market-wide effects. Our findings indicate that treatment firms exhibit higher trading volume than control firms. Additionally, we observe increased activity for treatment firms on other online investor platforms (e.g., EasyIR). We interpret these results as evidence of spillover effects beyond the call itself.

Our paper contributes to the following literature. First, we add to the literature on investors' information-processing costs. For researchers, it is challenging to identify events that represent an exogenous change in information-processing costs. Studies of regulations that are designed to ease the processing burden typically lack a plausible benchmark due to market-wide implementation (Blankespoor 2019; Goldstein et al. 2023). Other studies rely on indirect measures that capture changes in the opportunity cost of processing information (Hirshleifer et al. 2009; deHaan et al. 2017; Darendeli 2024).[11] We take a different approach. Using a field experiment, our study provides causal evidence by using AI technology to directly lower information-processing costs.

Second, our findings contribute to the research on individual investors and their increasing use of technology. The rise of information technology has helped individual investors to emerge as a collective that can meaningfully impact the capital market (Wong et al. 2024; Brochet et al. 2023). However, studies find mixed evidence on the efficacy of retail investors' participation (Gao and Huang 2020; Bian, Li, and Yan 2021). Traditionally, individual investors have been viewed as uninformed and behaviorally biased (Barber and Odean 2008). Thus, their increased participation has led to concerns about whether they possess the necessary skills to invest wisely or only add volatility to the market. In this paper, we show how generative AI can be used to inform individual investors by helping them process a complex corporate disclosure.

Third, our paper contributes to the literature on conference calls and interactions between management and investors. Some prior studies attempt to infer the participants' engagement levels from the participants' attributes (Mayew et al. 2020). Others use the conversation itself as the unit of analysis for gauging managers' and market participants' engagement levels

[11] For example, busy earnings days or other extraneous events such as weather have been used to get closer to exogeneity.

(Rennekamp et al. 2019). We build on this line of literature (Croom et al. 2023; Markov and Yezegel 2023; Choi et al. 2024) and show that the extent to which investors engage with topics relevant to the existing disclosures can affect their overall engagement level.

2. INSTITUTIONAL SETTING

2.1. Annual reports and the role of summary

Studies show that annual reports have become increasingly lengthy and complex, which can create cognitive overload for investors. Providing summary reports to investors may help them better extract key insights from annual reports and attend to information they might otherwise neglect due to processing costs (Hirshleifer and Teoh 2023). In hopes of helping investors (particularly less sophisticated ones) process their annual reports' content, firms have thus begun providing various forms of summaries along with their annual reports.

In the 1990s, the SEC considered requiring firms to disclose summary reports with their 10K releases.[12] However, the Commission eventually discarded this proposal after significant pushback from the capital market. Two concerns that were raised at the time about summary reports may still apply today. The first concern is that the information-reduction process may eliminate important content from the full report. If it does, and if less sophisticated investors fixate on the summary report, then providing them with a summary might limit their

[12] In 1995, SEC proposed a "two-tiered" financial reporting system (SEC 1995), which involved producing a reduced-form annual report by abbreviating (or recasting) the complete financial statements. Bushman et al. (1996) show that providing a summary report can lead to improved liquidity of firms, even when there is less private information gathering by those less likely to rely on the summary report.

information set.[13] The second, and perhaps more significant, concern is that a summary may provide a biased view of a firm due to managerial opportunism. If firms are free to choose the key insights from their own reports, managers may select ones that are favorable to the firm. Consistent with this, Cardinaels et al. (2019) show that manager-generated summaries contain a more opportunistic tone and content than summaries provided by computer algorithms.

A key feature of our study is that the summary reports are not produced by firms. Instead, we use computer algorithms aided by AI technology, which has no clear incentive to select information favorable to the firm. However, other aspects of AI technology may add noise to the summary (e.g., hallucination). We therefore test the quality of the AI-generated summaries and find that they are of higher quality than human-generated ones. Another appealing feature of the AI technology is that it is readily available to all investors at very little cost, making the study's finding easily implementable even for less sophisticated investors.[14]

2.2. Earning Communication Calls (ECCs)

Although the release of annual reports is one of the most anticipated disclosure events for Chinese firms (Bian et al. 2021), prior studies suggest that the reports often result in significant information overload for investors. The reports are sometimes bloated with minimally informative text, which may add to the information asymmetry among investors (Kim et al.

[13] Bushman et al. (1996) show that the two-tiered system can be detrimental to unsophisticated investors when the summary report eliminates important value-relevant information. However, the liquidity effect of providing a summary report will always be positive, even when the summary results in less private information-gathering by sophisticated investors.

[14] Following our experiment, based on requests from many firms, the platform started to provide the summary service to *all* firms that host ECCs on it.

2024). The problem is particularly acute for individual investors, as they may lack the skills to process public information and may not have access to private communication channels.[15]

To help all investors, the China Securities Regulatory Commission (CSRC) advises that all public firms host a conference call within 15 trading days of their annual report's release. While this is not an explicit regulatory requirement, the majority of listed firms do host ECCs within that timeframe. The high adoption rate reflects the quasi-mandatory nature of the soft law system in China (Cheng, Hail, and Yu 2022).[16,17] One trend since COVID-19 is that these conference calls are increasingly held online. According to the latest records from the CLCA, 90% of ECCs were conducted virtually in 2022.

In 2005, the Quanjing platform, ultimately controlled by the Peoples' Daily, became the first online platform to host virtual ECCs. The platform was set up in response to a Shenzhen Stock Exchange regulation, included in its 2004 "Guidelines for the Protection of Investors' Rights and Interests on the SME Board," requiring that companies listed on the SME board host ECCs after publishing their annual financial reports. At that time, all ECCs were organized by and held on the Quanjing platform. Since then, other conference call platforms, such as Value Online, have been established by the Beijing and Shanghai Stock Exchanges. However,

[15] Individual investors account for approximately 85% of the trading in China's stock exchanges (Wong et al. 2024).

[16] The latest statistics from the China Listed Companies Association (CLCA) indicate that 5,130 companies from the Beijing, Shanghai, and Shenzhen stock markets (96.10% of all listed firms across the three exchanges) held conference calls for their 2024 annual reports.

[17] Our conversation with numerous board secretaries confirms that while the rules are not explicit, most firms view conference calls as mandatory in practice.

Quainjing continues to be the largest, hosting about one-fourth of all conference calls in China in 2024.

The annual report conference calls are open to current and prospective investors of the hosting firms. Each participant must register on the conference call platform using their resident ID number or phone number. Given that financial analysts and institutional investors can engage with management through many direct communication methods (e.g., site visits and phone conversations), retail investors constitute the predominant demographic of conference call attendees (Bian et al. 2021). Official records from the China Securities Regulatory Commission (CSRC) underscore the significance of conference calls for retail investors.[18]

2.3. Mechanics of Earnings Communication Calls

China's annual report conference calls are similar to U.S. earnings calls in that investors can ask questions about firms' annual earnings performance. A growing portion of firms in China hold their conference calls virtually. Unlike in the U.S., where the format of virtual conference calls varies widely (e.g., hybrid or virtual-only), the ECCs' format is uniform, with all calls being virtual-only and no in-person option. A typical ECC includes a presentation followed by a Q&A. The presentation can be a pre-recorded promotional video or a PowerPoint slide presentation delivered in real time.

[18] The records indicate that more than 700,000 individual investors participated in ECCs in 2022.

The Q&A session of ECCs is conducted using a chat function. Participants can submit questions at any time during the Q&A (and sometimes before the meeting), and can designate who should respond (e.g., CEO, CFO). The management team is strongly encouraged to answer all questions.[19,20] Both the answer and the question become visible to the public at the time the firm posts its response.

Another unique feature of ECCs is that participants can freely submit questions. This feature is distinct from conference calls in the U.S., where participants must be called on by the manager before they can ask a question. U.S. conference calls are predominantly attended by financial analysts and institutional investors, whose names are made known to the firms during the calls. These participants often have incentives to maintain access to management and may therefore be reluctant to ask confrontational questions. In contrast, ECC participants are predominantly individual investors who have no incentives to maintain access, and they are protected by anonymity. Thus, it is possible that their questions challenge management more directly than the questions during U.S. conference calls.

3. SAMPLE AND EXPERIMENT DESIGN

3.1 Sample Selection

[19] Firms are allowed to withhold inappropriate questions (i.e., those involving foul language or personal attacks) and redundant questions.

[20] Consistent with Bian et al. (2021), we find that firms withhold, on average, 13% of all submitted questions, which suggests that they answer most of the investors' questions.

Our experiment was conducted on firms that hosted their 2023 annual report conference calls on the Quanjing platform. We start with an initial sample of 1,168 listed firms that hosted ECCs on their prior-year (2022) annual reports on the platform. A majority of our sample firms are listed on the SZSE, where Quanjing has its market dominance.[21]

During the 2023 conference call season, some firms from the initial sample discontinued their use of the Quanjing platform, while other firms hosted their first calls on the platform. We randomly assign the newly participating firms to the control or treatment group once their conference call dates are confirmed (typically seven days prior to the call). Our final analysis includes a sample of 1,105 firms, of which 815 (73.49%) are from the initial sample and 290 are new additions.

Prior to the experiment, we randomly assign firms to either the control group or one of two treatment groups. Specifically, 30% of the firms are allocated to the control group, while the remaining 70% are evenly split between the two treatment groups: *Summary* (35%) and *Summary & Sentiment Label* (35%). Details on these treatment conditions are in Section 3.2.1.

Table 1, Panel A presents the distribution of firms across the treatment and control groups. We have 1,105 sample firms, consisting of 762 treatment firms and 343 control firms. Panel B presents the covariate balance between the treatment and control groups. The financial data and analysts following data for 2023 are collected from CSMAR. We report the means for variables such as the log of total assets at year-end (*Size*), return on assets (*ROA*), a binary variable

[21] Our sample covers approximately 40% of all firms listed on the SZSE. In untabulated results, we perform a balance test comparing the Quanjing sample firms with the entire population of SZSE-listed firms. Firms using the Quanjing platform are slightly smaller and more profitable than the average SZSE-listed firm, but in other characteristics the two groups are largely comparable.

indicating whether the firm is state-controlled (*SOE*), the percentage of shares held by institutional investors (*Institutional Holdings*), the number of analysts covering the firm (*Analysts Following*), and earnings surprise (*Earnings Surprise*). The descriptive statistics show that compared to the U.S. conference call firm sample, our sample firms have less institutional holdings (37%) and fewer analysts following (5.34), on average. The balance tests show that the observable covariates are well-balanced across the treatment and control groups, with no significant differences in these characteristics.

3.2 Experimental Design

Our experiment was conducted from April 7 to May 31, 2024. Our intervention involved posting summaries of annual reports on the conference call platform. Quanjing provided us with the conference call schedules once the meeting dates with the firms had been confirmed. Since all listed firms are required to hold their annual conference calls within 15 days *following* the publication of their annual reports,[22] we had time to create the five summary points of each firm's annual report prior to the meeting.

We use Kimi AI, an OpenAI Generative Pre-trained Transformer (GPT) alternative in China, to identify the five key summary points from each firm's 2023 annual report.[23] GPT is widely and effectively utilized in areas that deal with text, including text summarization (Goyal et al. 2022; Achiam et al. 2023). Kimi AI is one of the most widely used AI chatbots in China,

[22] The prompt for all firms was the following: "Please summarize five key points around the fundamentals based only on this annual report, and elaborate on each point in detail."

[23] We also considered having fewer than or more than five summary points. We chose to use five in order to strike a balance between providing enough information and not overloading investors, as discussed in prior research (https://www.ncbi.nlm.nih.gov/pmc/articles/PMC10322198/#).

according to the latest report from the World Bank (Liu and Wang 2024). It is known for its ability to process long text and is equipped to manage Chinese documents efficiently and directly.[24]

Our experiment includes the following two treatment groups.

***Summary* (35%)**: For this group, a summary consisting only of the five key points was posted on each company's page in the annual conference call section.

***Summary & Sentiment Label* (35%)**: Summaries were also posted for this group, but each key point was accompanied by a sentiment label: positive, negative, or neutral.[25] We manually assigned the label to each summary point.[26] See Appendix A1 for an example summary for each treatment group.

We introduced the second treatment group—and included the sentiment labels for each of the five summary points—to reduce investors' uncertainty about whether each key point represented good or bad news for the firm. There was no significant difference in the timing of the summary postings on the Quanjing platform between the two treatment groups.

---

[24] In untabulated tests, we assess the quality of AI-generated summaries by comparing them to human-generated summaries. We randomly sample 90 reports—60 from the two treatment groups and 30 from the control group—representing approximately 8% of our sample. We recruit 12 accounting majors at Southwest University of Finance and Economics to generate five key-point summaries of the annual reports, instructing them to (1) maintain a length similar to the AI-generated summaries and (2) use terms from the annual report rather than their own wording. We have two students summarize each annual report to mitigate the effect of individual errors. We evaluate the quality of AI-generated summaries using BERTScore, an evaluation metric that leverages contextual embeddings from the BERT model. The results show that AI-generated text achieves higher scores than human summaries (0.65 vs. 0.61), with the difference being statistically significant.

[25] In the experiment, we use the term "Future Growth" as the label for positive sentiment, "Steady Development" for neutral sentiment, and "Potential Risks" for negative sentiment. We avoided directly using "positive," "neutral," and "negative" because the treatment firms might have objected to such direct labeling.

[26] This was carried out by three accounting students at Southwest University of Finance and Economics.

For the control group, which composes 30% of the sample, no intervention was applied. A summary was generated for each firm but was not posted on the platform.

3.2.1 Posting Summaries of Annual Reports on the Conference Call Platform

For the two treatment groups, we posted, on the page announcing each firm's annual conference call, a link titled "Key Points of the Annual Report." Investors could access the summary points by clicking the link, which was available from the time when the upcoming conference call was announced on the platform (typically seven days before the call) until the conclusion of the call.

We began the experiment on April 7, 2024,[27] the first day after a public holiday for a Chinese festival, with two conference calls initiated on the Quanjing platform: Shandong Chenming Paper (SZSE code: 000488) from the treatment group and Suzhou Sushi Testing Group (SZSE code: 300416) from the control group. The experiment concluded on May 31, 2024, the day by which all listed firms were expected to have held their annual conference call.

3.2.2 Outcomes

We collect all questions and answers exchanged on the platform between the hosting firms and participants, along with information about the participants' identities and engagement levels during the meetings.

We consider three sets of outcomes. First, we assess overall participation in the conference calls by measuring (1) the number of participants, and (2) the total number of questions raised

[27] This was not the first day of the conference call season. Approximately 90 listed firms hosted conference calls over a three-month period prior to this date.

during the Q&As. These metrics were provided directly to us from Quanjing's records. Quanjing recorded every participant question (20,031 in total), including the small portion withheld by management (2,614). For our main analysis, we focus on the total number of questions, regardless of whether they were withheld. In additional analyses, we look deeper into the nature of the questions not addressed by management.

Next, we analyze the content of investors' inquiries to determine whether posting the summaries on the conference call webpage leads to more investor questions that align with the topics in the summary. We examine alignment based on the likelihood that the question topic (e.g., disclosure) matches the topic of any of the key points in the summary. Lastly, we assess whether posting the summaries improves the overall quality of the interactions, measured based on the quality of firms' responses.

4. EMPIRICAL TESTS AND FINDINGS

4.1 Descriptive Statistics of the Annual Report Summary

For each company, we generate a summary that identifies five key points derived from the 2023 annual report. We train Kimi AI to classify each of these points into one of 15 predefined topics: 1. Financial Information; 2. Production Management; 3. Product Market; 4. Supply Chain; 5. Innovation; 6. Risks; 7. Government Policy; 8. ESG; 9. Financing; 10. Strategy; 11. Payout; 12. Business Cooperation; 13. Investors' Relationship; 14. Capital Market; 15. Others.

The 15 categories are populated using the following process. We begin by manually reviewing 300 randomly selected key points, consolidating them into 15 distinct topics, and assigning appropriate labels. We present Kimi with 250 of these labelled observations (setting aside the remaining 50 for out-of-sample testing) and direct it to categorize each topic into one of the 15 topics, adhering to the classification logic we had established.[28] We repeat this process using 300 investor inquiries to classify the topic of each investor question.

Table 2, Panel A presents the distributions of topics of the key points in our sample. Overall, we find that the topic distribution is comparable between the control group and the treatment group. *Financial information* has the highest likelihood of being a key point in the annual reports, with a prevalence of 26%. This is consistent with our expectation that financial details are critical in annual reports because annual reports are designed to communicate firm performance. Next in significance are *Innovation* and *Risks*, which account for approximately 17% and 14%, respectively, of the key points across our treatment and control groups. These topics provide investors with insights into a company's environment, prospective growth opportunities, and potential risks.

We also manually assign a sentiment label (positive, neutral, or negative) to each summary point.[29] Table 2, Panel B presents the distribution of the three sentiment types for the key points and shows that it is similar in the control and treatment groups. Approximately 70% of the key points in the annual reports convey positive information; only about 10% convey negative

[28] Our detailed prompt is: "We have a total of 15 categories as follows, and based on these categories, we have provided a set of annotated samples for you to read first. To which category does XXX most directly belong?" The out-of-sample accuracy of the classification is 92%.

[29] This was carried out by three accounting students at Southwest University of Finance and Economics.

information. This distribution reflects firms' tendency, in annual reports, to present a favorable outlook and downplay the disclosure of adverse information.

Overall, we find no significant difference between the control and treatment groups in sentiment distribution, further confirming that the randomization process produces balanced samples.

4.2 Empirical Tests

4.2.1 Investor Participation

To examine whether posting a five-point summary of the annual report encourages investors to participate more actively during calls, we estimate the following firm-level regression equation:

$$Questions_i \text{ or } Participants_i = \alpha + \beta_1 T_i + \sum \beta_n Controls_i + \text{FE} + \varepsilon_i \qquad (1)$$

where the outcome variable, $Questions_i$, represents the total number of questions submitted by participants through the online platform during the conference call for firm i.[30] $Participants_i$ serves as another metric of investor engagement: it is the total headcount of individuals who joined the conference call for firm i. T_i represents our treatment group assignment, with *Treat* indicating that firm *i* was assigned to *either* treatment group and *Summary* and *Summary & Sentiment Label* indicating the specific assignment. $Controls_i$ includes the following control

[30] The exchange strongly recommends that listed companies actively respond to investors' questions during the conference call, allocate enough time for Q&A, and ensure a high response rate and quality of replies. All questions are published and addressed by the management team, except those that are abusive, involve personal attacks, or are deemed redundant. We were able to obtain all investor-submitted questions, including ones not made public by the firm. Our findings remain qualitatively the same when we exclude all the questions that were withheld and not answered by the firms.

variables measured in 2023: *Size*, the log of total assets at year-end; *MB*, the total market value of equity divided by book value of equity at year-end; *ROA*, net income divided by ending total assets; *SOE*, an indicator that equals one if the firm's ultimate shareholder is the government, zero otherwise; *Institutional Holdings*, the percentage of shares controlled by institutional investors; *Analysts Following*, the log of one plus the number of analysts following the firm; and *Earnings Surprise*, the difference between actual and mean of analyst forecast EPS, divided by the closing price of the last trading day before the annual report date. We use a Poisson regression model for all estimations. We also control for industry, province, and day fixed effects. Standard errors are clustered by industry.

Table 3, Panel A tabulates results of the univariate tests. We find that offering an annual-report summary during the conference call increases the engagement of retail investors. Specifically, the treatment firms' conference calls attract a significantly higher number of investor questions (z-stat = 3.54) than the control firms' calls. The number of participants, measured by headcount, is also higher, although the difference is statistically insignificant (z-stat = 0.73).

The regression results are reported in Table 3, Panel B. As shown in column (1), the coefficient on $Treat_i$ is positive and significant at the 1% level. On average, the inclusion of an annual report summary leads to a 46.65% increase in the number of questions posed by investors, as indicated by the comparison to the control group (calculated as $e^{0.383} - 1$). The results in column (3) reveal a similar pattern, with the summaries resulting in a 9.41% increase in conference call attendance (calculated as $e^{0.090} - 1$), compared to the control group. In column

(2), we further examine how the effect varies with the different treatment methods. We observe a significant increase in the number of questions across both treatment groups, and the groups' respective increases are not statistically different (Chi^2-stat=0.05). In column (4), we observe no significant increase in the number of conference call participants when a summary alone is posted, but a significant increase when the summary is accompanied by sentiment labeling. However, once again, the difference between the two treatment groups is not statistically significant (Chi^2-stat=0.32). The findings suggest that providing a summary of the annual report is associated with increased engagement and more questions by participants. Rather than increase the number of investors in attendance, the summary appears to mainly impact investors who were already involved in the disclosure event.[31]

In the next subsection, we link the investor questions to our summary by examining the type of questions asked. The increase in questions, while promising, does not necessarily mean that the topics in the summary are being broached. If investors feel the need to raise issues not known to others, then the presence of a summary could lead them to ask questions unrelated to the topics we identified. We therefore examine the content of the investor questions to see whether they pick up topics from the summaries.

4.3 Topical Alignment of Investors' Questions

[31] Following the taxonomy of Blankespoor et al. (2020), the findings are consistent with the summaries having a greater impact on reducing acquisition and integration costs. Their impact on reducing awareness cost is limited, because if investors had become aware that this disclosure existed, they would have attended it.

We analyze the content of investors' inquiries, then test whether posting summaries on conference call webpages increases the topical alignment between the investors' questions and the key points in the summaries.

Before proceeding with the regression results, we present a univariate comparison assessing the summaries' impact on the content of investors' inquiries. Table 4, Panel A reports the variation in the degree to which the topics of the investors' questions align with the topics of the key points in our summaries.

To identify the topics in the summary, we trained Kimi AI to classify each key point into one of the 15 topics we described in Section 4.1. We then had Kimi AI classify the investors' questions the same way. *Alignment* is a dummy variable that equals one if the topic of an investor's question matches any of the five topics referenced in the summary, zero otherwise. The 50% alignment rate observed in the control group establishes a baseline for the prevalence of questions that are relevant to the annual report's key points. Our analysis reveals a notable 7% increase in alignment (from 50% to 57%, with a z-statistic of 8.70) following the introduction of the summaries. The findings suggest that the summaries help direct investor attention to the key topics being presented.

To further examine this alignment effect, we estimate the following question-level regression equation:

$$\text{Logit}\ (Alignment_{i,j}) = \alpha + \beta_1\, T_i + \sum \beta_n\, Controls_i + \text{FE} + \varepsilon_{i,j} \qquad (2)$$

where the outcome variable is $Alignment_{i,j}$ as defined earlier. T_i represents our treatment group assignment for firm i: *Treat*, *Summary*, or *Summary & Sentiment Label*. To ensure the robustness of our analysis, we include an array of control variables as specified in equation (1), and we account for fixed effects at the industry, province, and day levels. Standard errors are clustered by industry.

Table 4, Panel B presents the regression results. The estimated coefficient in column (1) indicates that, on average, providing a summary to investors is positively correlated with the likelihood of topical alignment. Specifically, column (1) reveals that providing a summary results in a 28.0% (calculated as $e^{0.247} - 1$) greater likelihood that a question's topic aligns with one of the five topics in the annual report summary (t-stat= 3.78). Column (2) further indicates that the effect is significant in both treatment groups but is strongest when the summary is accompanied by sentiment labeling (Chi^2- stat = 2.80). In summary, the findings suggest that providing a summary of the annual report focuses investors' questions more on the key topics conveyed in the annual report, which is consistent with our hypothesis that summaries can influence investors' focus during conference calls. However, our intervention could have drawbacks, such as prompting investors to ask repetitive questions or reducing the overall quality of the questions and answers. Sections 4.4 and 4.5 explore these concerns in detail.

4.3.1 Conditional on Topic and Sentiment

Next, we explore which topics and sentiments in the summaries have a stronger effect on the topical alignment between investors' questions and the summaries. Our unit of analysis continues to be at the *question* level, but, for these analyses, we only retain the questions that align with a topic from the summaries. We estimate Equation (2), first using an indicator for each of the 15 topics as the dependent variable (Table 2, Panel A), then using an indicator for each of the sentiments as the dependent variable (Table 2, Panel B). Specifically, the topic indicator variable is set to one if the topic of the aligned key point matches that specific topic, zero otherwise. Similarly, the sentiment indicator variable is set to one if the sentiment of the aligned key point matches that sentiment, zero otherwise. Thus, we use these regressions to test which, of the 15 topics (or three sentiments) in the summaries' key points, investors are more likely to ask about, relative to the control sample.

The regression results on the summary's topics are displayed in Table 5. In Panel A, we present only the topics that yield significant outcomes. The coefficient on *Treat* is positively significant when the summary topic concerns financial information, risks, strategy, or payout, indicating that investors who receive the summaries are more inclined to pose questions that align with these four topics. This is consistent with findings (Choi et al. 2024) that retail investors on U.S. conference calls exhibit greater interest in firm-specific issues such as dividend policies, stakeholders, and technology, while analysts concentrate on macroeconomic topics. Descriptive analysis from Table 2 shows that the distribution of topics is well-balanced between the treatment and control groups, which implies that our regression results are unlikely to be driven by underlying differences in the groups' respective annual report topics.

In Table 5, Panel B, we focus on the different sentiments within the summaries. We observe a positively significant coefficient on *Treat* when the outcome indicator represents non-positive key points, suggesting that investors tend to probe more when the sentiment is non-positive. One possible reason for this is that investors may perceive summaries with positive sentiment as less credible due to managers' strong incentives to disclose good news. Additionally, unlike financial analysts, who must reveal their identities when they interview managers, the online participants enjoy anonymity in this setting, which may give them more freedom to ask negative questions.

4.3.2 Conditional on Investor Type

To examine whether the treatment effects on participation and question content differ across various investor groups, we further partition our sample based on specific investor characteristics. Participants can register on the Quanjing platform using either their citizen ID card or phone number. Quanjing assigns a registered ID to those that authenticate using a citizen ID card, but uses the phone number to identify those who use their phone number for authentication. Due to security reasons and requirements from the IRB, we were provided with participants' registered IDs (anonymized) but not their phone numbers. In our analysis of investor identity, we thus use 6,213 investors with registered IDs. These investors posed 13,726 questions, which represents 68.52% of the total sample of 20,031 questions. The following analysis is based on participants' anonymized IDs.

We first categorize investors based on whether they asked a question during any Quanjing platform conference call in 2022. We predict that silent investors—those that did not ask any questions in 2022—are less experienced and may therefore benefit more from the annual report summaries, leading them to ask more questions going forward. In contrast, we expect that vocal investors—those that asked one or more questions in 2022—are more experienced, so the summaries will have a weaker effect on them.

The cross-sectional regression results are presented in Table 6, Panel A. Column (1) shows that silent investors do tend to ask more questions after being provided a summary of the annual report in 2023. In column (3), which focuses on vocal investors, we note, in the treatment groups, a positive, albeit less significant (no statistical difference, Chi^2=1.27), increase in the number of questions asked during conference calls. This suggests that the annual report summaries also stimulate questioning by vocal investors, but to a lesser extent.

Next, we explore whether the annual report summaries' influence on the content of participants' questions differs by investor type. Table 6, Panel B presents the outcomes of this analysis. In column (1), the coefficient on *Treat* (Coeff. = 0.162) is positively significant, indicating that silent investors in the treatment group are more inclined to ask questions on topics that align with the topics in the summaries (relative to silent investors in the control group). More intriguingly, in column (3), the coefficient on *Treat* (Coeff. = -0.183) is significantly negative. This suggests that vocal investors in the treatment group tend to ask fewer questions that align with the topics in the summaries, and instead focus more on other topics.

We also categorize investors based on the number of conference calls they attended in 2023. Those attending five or fewer conference calls are likely to be less active and sophisticated than investors with broader participation. Table 6, Panel C shows the varying impacts of our treatment on these two groups. The significantly positive coefficient in column

(1) demonstrates that providing the annual report summaries prompts less active investors to pose questions that are more aligned with the summary's key points, which suggests that the summaries are especially valuable to these investors. In contrast, the insignificant coefficient in column (3) implies that more active investors, who are presumably more sophisticated, do not derive the same level of benefit from our annual report summaries.

The findings suggest that the summaries guide less experienced investors to ask more questions. While increased participation by these investors would seem likely to displace questions by experienced investors, we find that it does not. Rather, the increased participation of the inexperienced appears to make the experienced investors consider—and ask about—topics that go beyond the key points of the annual report. Next, we examine how both aligned and non-aligned questions impact the overall quality of the Q&A dialogue.

4.4 The Quality of Conference Calls

In this section, we investigate how providing annual report summaries affects the quality of managers' responses to investor questions. While the summaries encourage greater investor engagement, their impact on the overall quality of the call remains ambiguous because firms can avoid clearly answering the questions. To assess the implications of our intervention on the quality of the interaction, we focus on two aspects: the quality of the firms' responses to investors' questions, and the spillover effect (stock market movements and activities on online investor platforms) around the conference calls.

To evaluate the quality of firms' responses, we use two metrics: *Length*, which measures the response length by word count; and *Informative*, an AI-based measure of the comprehensiveness and quality of a firm's response.[32] We assess the response quality based on a scale from one to five using the following criteria: directness in addressing the investors' questions; provision of detailed information, including numerical data and supporting evidence; and the firm's attitude in responding. We define *Informative* as equal to one if the response receives a rating above the median score of three, zero otherwise.[33]

Table 7, Panel A presents the results of the univariate tests. We find that the treatment firms' answers are significantly longer, by 35 words on average (z-stat = 6.00), than the control firms'. Moreover, the treatment firms' answers are more informative (*Informative* = 1) (0.44 versus 0.41, z-stat = 3.75). The regression analyses confirm these findings. We replace the outcome variable with the quality metrics *Length* and *Informative* and re-estimate Equation (2). For *Length*, we use a Poisson regression model, and for *Informative*, we use a logit model. The regression results are reported in Table 7, Panel B. In column (1), the coefficient on *Treat* is positive and significant at the 1% level, indicating that providing annual report summaries leads to longer answers from firms. Column (3) shows a similar finding using the *Informative* measure.

We also test whether the response quality differs between questions whose topics align and questions whose topics do not align with the topics in the summaries. Table 7, Panel C, columns (1) and (2) show that when the summaries are available to investors, all questions, regardless of alignment, experience a significant improvement in response quality. Although

[32] We train Kimi AI to evaluate the quality of the firm's responses using a training sample of 500 randomly selected responses. The model's out-of-sample accuracy is 93% when verified against a hold-out sample based on manual coding.

the effect is slightly less pronounced for non-aligned questions, the difference is not statistically significant (Chi^2 =0.059). This could be partially explained by our earlier findings in Table 6: when the summaries are presented, previously silent investors pose questions related to the annual report, while vocal investors ask about other topics. The improvement in response quality for both types of questions indicates an overall enhancement in quality of the conference call. Columns (3) and (4) further substantiate this finding.

In Table 7, Panel D, we further analyze the impact of question alignment on the overall quality of the conference call interactions. We divide the sample into two groups based on the percentage of aligned questions. Our hypothesis is that conference calls that are more strongly influenced by the summaries (i.e., those with more aligned questions) will experience a greater improvement in response quality. The results show that the coefficient is significantly positive (t-stat = 5.00) for conference calls with a higher percentage (above the median) of aligned questions (Column 1) and not significant (t-stat = 0.61) for conference calls with a lower percentage of aligned questions (Column 2). The difference is statistically significant (Diff. = 0.174, Chi^2 = 20.62). These results support our earlier conclusion that the more the summaries engage and guide investors, the more the overall quality of conference call interactions is enhanced.

Finally, we examine market-wide effects to assess whether providing the summaries helps conference calls generate information for the market. We first focus on stock market movement, reflected by two key variables: *Turnover*, the cumulative turnover ratio (the number of shares traded divided by the number of shares outstanding) from one day before to five days after the conference call; and *Abs_CAR*, the absolute value of cumulated abnormal returns over the same period. Abnormal returns are calculated using the market model of raw returns minus market returns. Table 8, Panel A shows the estimated results. For the treatment group, we observe a significant increase in *Turnover* (Coeff. = 0.021, t-statistic = 1.91). Although the *Abs_CAR* increases, the change is not statistically significant (Coeff. = 0.004, t-statistic = 0.74).

Next, we examine the post-conference call activity of retail investors. If our summaries effectively engage retail investors, the investors should be more inclined to communicate with firms not only during the conference call but also in the days that follow. To capture this momentum, we monitor the interactions between investors and firms on online platforms. We collect all questions that investors raised about the sample firm around the conference call date, then count the number of questions raised from 0 to 30 days (and up to 90 days) post conference call. Table 8, Panel B presents the estimated results. For the treatment group, we observe a significant increase in posts following the calls (coefficient = 0.061, t-statistic = 3.34). However, this effect seems temporary and dissipates after the 90-day window. Overall, we conclude that providing summaries leads to a temporary increase in investors' incentive to communicate with firms. This finding supports our earlier finding that the summaries improve the overall quality of firms' interactions with investors during conference calls.

4.5 Potential Costs

In previous sections, we presented evidence supporting the idea that AI-generated annual report summaries can reduce information-processing costs for retail investors. Consequently, investors become more inclined to focus on annual reports and proactively engage with companies regarding the reports' key points. However, providing such guidance to investors may have a downside: it may lead investors to over-rely on the summaries, resulting in repeat inquiries about the same topics, which in turn could reduce the quality of the discussions and impose costs on firms.

We test for this downside by examining how frequently questions are dropped by the firm. In our setting, firms can withhold questions submitted by the participants. Any registered participant can submit questions to the firms' management during the meetings, but the submitted questions are not immediately posted online; they first undergo a real-time review by the firm. The firm does not have to publish questions that are abusive, involve personal attacks, or are deemed redundant, nor does management have to address them. We can observe the withheld questions because Quanjing, the host platform, provided us records of all questions, including ones not displayed. This allows us to test whether treatment firms become more aggressive in filtering out redundant questions, which would mitigate a potential cost associated with posting the summaries.

Before proceeding with the regression results, we first offer a univariate comparison. Table 9, Panel A reports the differences in the percentage of questions withheld by firms. Our findings reveal that firms in the treatment groups withheld 7.39% more questions than firms in the control group, which is significantly different (z-statistic=13.41). Empirically, we replicated the estimation from Equation (2) after replacing the outcome variable with *Withhold*, a binary indicator that equals one if question j for firm i was raised but not posted online, zero otherwise. Table 9, Panel B shows a positive and significant coefficient on *Treat* in column (1) (coefficient = 0.547, t-statistic = 2.02), suggesting that more questions are being withheld by treatment firms. Furthermore, Table 9, Panel C suggests that the treatment firms were more likely to withhold questions aligned with the key points' topics. These findings remain consistent when we split the treatment group into the two sub-treatment groups in column (2) in both Panel B and Panel C.

Next, we focus on the types of questions that firms withheld, and examine whether they differ between the treatment and control groups. Our conversations with Quanjing officials and managers of listed firms indicate that redundancy is a primary reason for withholding questions. To validate this explanation, we introduce a new variable, *Redundant*, to measure the redundancy of each question. We calculate the Jaccard similarity index for each question in relation to all other questions within the same firm and on the same topic, and we use the average of these similarity scores as the measure for *Redundant*.

We then conduct a regression analysis using two dependent variables: *Redundant*, a continuous measure of the similarity mean score defined earlier; and *Redundant Indicator*, an indicator that equals one if the similarity mean score is above 0.25 (which is in the top 10%), zero otherwise. We use *Redundant Indicator* to proxy for repetitive questions raised by investors. We also use *Withhold* as the key independent variable and follow the other specifications outlined in Equation (2).

Table 9, Panel D presents the results of these OLS regressions. The positively significant coefficient on *Withhold* in Column (1) (Coeff. = 0.021, t-statistic = 8.53) and Column (2) (Coeff. = 0.820, t-statistic = 10.35) suggests that the questions withheld by firms tend to be more redundant, which corroborates our observations from conversations with Quanjing and the firms. This finding also indicates that firms' screening of questions could help mitigate any redundancy caused by our intervention.

Additionally, we introduce the interaction term *Withhold* * *Treat* to evaluate whether the provision of summaries influences the level of redundancy in withheld questions. The coefficients on this interaction term (Coeff. = -0.011, t-statistic = -0.58) in column (3) and (Coeff. = -0.177, t-statistic = -0.48) column (4) are not significant, suggesting that the questions withheld by treatment firms are no more redundant than those withheld by control firms. This finding alleviates the concern that investors in the treatment group may over-rely on the summaries and thus ask more redundant questions.

We also find that the coefficient on *Treat* is marginally positive (t-statistic = 1.77) when we use the continuous dependent variable *Redundant* in column (3), indicating that firms in the treatment group answer more questions that are either more focused on the summary's topics or more redundant. When we use *Redundant Indicator* as the dependent variable in column (4), the coefficient on *Treat* is negative and statistically insignificant. The insignificant result for *Redundant Indicator* in column (4) suggests that the higher mean similarity score captured by *Redundant* in column (3) was of low magnitude. A low magnitude indicates that the participants' questions focus on the summary's topics rather than merely repeating its points.

4.6 Robustness Check

In this section, we conduct a series of robustness tests to validate our main conclusions. First, we perform a DID analysis, leveraging two years of conference call records. This analysis focuses on firms that hosted annual conference calls on the Quanjing platform for both their 2022 and 2023 annual reports. This subset consists of 815 firms, including 246 control firms and 569 treatment firms. Table 10, Panel A presents the univariate test results. They show that control group firms experience a decrease in the number of questions raised, from 18.80 to 14.35, while treatment firms see a slight increase, from 18.56 to 19.91. According to the Quanjing platform representatives we conversed with, the unfavorable capital market environment in 2024 reduced investors' enthusiasm for asking questions during conference calls, which could explain the drop in the control group. We also observe a rise in the number of participants for both control and treatment firms, but the increase is more pronounced in the treatment group.

We then repeat the regression analysis of Equation (1), modifying the independent variable *Treat* to *Treat* * *Post*, where *Post* is a binary variable that equals one for conference calls held in 2023 and zero for those held in 2022. We also include firm fixed effects and day fixed effects, replacing the industry, province, and day fixed effects. This approach is feasible because we possess two years of data for each firm, allowing us to control for firm-specific characteristics. Table 10, Panel B displays the regression results. In column (1), the coefficient on *Treat* * *Post* is positive and significant at the 1% level (Coeff. = 0.344, t-statistic = 6.06). Similarly, column (3) shows that conference call attendance significantly increases for treatment firms compared to control firms (Coeff. = 0.135, t-statistic = 3.16).

5. CONCLUSION

We use a field experiment to examine whether providing investors with AI-generated annual report summaries during virtual conference calls reduces their information-processing costs and enhances firm-investor interaction. Our experiment results show that providing the summaries significantly increases the number of questions raised, and that the content of these questions aligns more closely with the topics covered in the summaries. This topical alignment effect is stronger for firm-specific topics, such as financial information, risks, strategy, and payouts, and is more pronounced among less experienced investors. Additionally, we find that providing the summaries also increases the number of questions asked by experienced investors, though they tend to focus more on topics not covered by the summaries.

Our evidence further reveals that the summaries not only enhance investor engagement but also improve firms' responses to investor questions. We find that firms provide longer and more detailed responses when investors are given the summaries. This effect is observed for both topically aligned and non-aligned questions, suggesting that the quality of exchanges improves even for topics not covered in the summaries. The notion that providing the summaries enhances firm-investor interactions during conference calls is corroborated by a significantly greater trading volume of the firms' shares from one day before to five days after the calls, and a significantly greater number of investor questions directed to the firms on the stock exchange's online platform for up to 30 days after the calls.

Overall, these results suggest that AI-generated summaries can significantly reduce investors' information-processing costs, increase investor engagement, and lead to more informative responses from firms during conference calls.

References

Achiam, J., Adler, S., Agarwal, S., Ahmad, L., Akkaya, I., Aleman, F. L., ... & McGrew, B. (2023). Gpt-4 technical report. *arXiv preprint arXiv:2303.08774*.

Atkinson R. C., Shiffrin R. M. 1968. Human memory: a proposed system and its control processes, *The psychology of learning and motivation: Advances in research and theory 2nd ed*. eds. 89–195.

Barber, B.M., Odean, T., 2008. All that glitters: The effect of attention and news on the buying behavior of individual and institutional investors. *Review of Financial Studies* 21, 785-818

Bernard, D., Blankespoor, E., de Kok, T., & Toynbee, S. (2023). A Modular Measure of Business Complexity. *Available at SSRN 4480309*.

Bian, S., Li, F., Yan, Z. (2021). Do Retail Investors Matter?. Available at SSRN 3861763.

Blankespoor, E. (2019). The impact of information processing costs on firm disclosure choice: Evidence from the XBRL mandate. *Journal of Accounting Research*, *57*(4), 919-967.

Blankespoor, E., deHaan, E., & Marinovic, I. (2020). Disclosure processing costs, investors' information choice, and equity market outcomes: A review. *Journal of Accounting and Economics*, *70*(2-3), 101344.

Bonsall IV, S. B., Leone, A. J., Miller, B. P., & Rennekamp, K. (2017). A plain English measure of financial reporting readability. *Journal of Accounting and Economics*, *63*(2-3), 329-357.

Brochet, F., Chychyla, R., & Ferri, F. (2023). Virtual Shareholder Meetings. *Management Science* 70(9):5896-5930.

Brown, L. D., Call, A. C., Clement, M. B., & Sharp, N. Y. (2015). Inside the "Black Box" of Sell-Side Financial Analysts. Journal of Accounting Research, 53(1), 1–47.

Bushman, R. M., Gigler, F., & Indjejikian, R. J. (1996). A model of two-tiered financial reporting. *Journal of Accounting Research*, 34, 51-74

Cardinaels, E., Hollander, S., & White, B. J. (2019). Automatic summarization of earnings releases: attributes and effects on investors' judgments. Review of Accounting Studies, 24, 860-890.

Cheng, Q., L. Hail, and G., Yu, (2022), The past, present, and future of China-related accounting research, *Journal of Accounting and Economics* 74(2–3).

Choi, J. K., Huang, A. H., Qiu, L., & Zheng, Y. (2024). Investor Relations for the Rest of Us: Engaging with Retail Investors. *Available at SSRN*.

Croom, J., Grant, S. M., & Seto, S. C. (2023). Q&A interactions: Giving investors a voice and managers' withholding of information. *Available at SSRN 4508135*.

Darendeli, A. (2024). How do retail investors respond to summary disclosure? Evidence from mutual fund factsheets. *Review of Accounting Studies*, 1-45.

deHaan, E., Madsen, J. and Piotroski, J.D., 2017. Do Weather‐Induced Moods Affect the Processing of Earnings News? Journal of Accounting Research, 55(3), pp.509-550.

Dyer, T., Lang, M., & Stice-Lawrence, L. (2017). The evolution of 10-K textual disclosure: Evidence from Latent Dirichlet Allocation. *Journal of Accounting and Economics*, *64*(2-3), 221-245.

Gao, M., & Huang, J. (2020). Informing the market: The effect of modern information technologies on information production. *The Review of Financial Studies*, *33*(4), 1367-1411.

Goldstein, I., Yang, S., & Zuo, L. (2023). The real effects of modern information technologies: Evidence from the EDGAR implementation. *Journal of Accounting Research*, *61*(5), 1699-1733.

Goyal, T., Li, J. J., & Durrett, G. (2022). News summarization and evaluation in the era of gpt-3. *arXiv preprint arXiv:2209.12356*.

Guay, W., Samuels, D., & Taylor, D. (2016). Guiding through the fog: Financial statement complexity and voluntary disclosure. *Journal of Accounting and Economics*, *62*(2-3), 234-269.

Hirshleifer, D., Lim, S. S., & Teoh, S. H., 2009. Driven to distraction: Extraneous events and underreaction to earnings news. The Journal of Finance, 64(5), 2289-2325.

Hirshleifer, D. & Teoh, S.H., Limited attention, information disclosure, and financial reporting, *Journal of Accounting and Economics*, 36(1–3), 2003.

Kim, A., Muhn, M., & Nikolaev, V. (2024). Bloated Disclosures: Can ChatGPT Help Investors Process Financial Information?. arXiv preprint arXiv:2306.10224.

Liu, Y., & Wang, H. (2024). Who on Earth Is Using Generative AI?.

Markov, S., & Yezegel, A. (2023). Giving Retail Investors a Say in Disclosure. *Available at SSRN 4836378*.

Mayew, W. J. (2008). Evidence of Management Discrimination among Analysts during Earnings Conference Calls. Journal of Accounting Research, 46(3), 627–659

Mayew, W. J., Sethuraman, M., & Venkatachalam, M. (2020). Individual Analysts' Stock Recommendations, Earnings Forecasts, and the Informativeness of Conference Call Question and Answer Sessions. *The Accounting Review* 95(6):311-337.

Rennekamp, K.M., Sethuraman, M., Steenhoven, B.A., 2019. Engagement in earnings conference calls: A multi-method examination. Journal of Accounting and Economics, forthcoming.

Securities and Exchange Commission. "Use of Abbreviated Financial Statements in Document Delivered to Investors Pursuant to the Securities Act of 1933 and Securities Act of 1934." Rule Proposal, 1995.

Triplett, N. (1898). The dynamogenic factors in pacemaking and competition. *The American journal of psychology*, *9*(4), 507-533.

Wong, T. J., Yu, G., Zhang, S., & Zhang, T. (2024). Calling for transparency: Evidence from a field experiment. *Journal of Accounting and Economics*, *77*(1), 101604.

Zajonc, R. B. (1965). Social Facilitation: A solution is suggested for an old unresolved social psychological problem. *Science*, *149*(3681), 269-274.

Table 1: Pre-experiment Randomization

Before the experiment, we randomly assigned 30% of firms to the control group and 35% of firms to each of the two treatment groups—*Summary* and *Summary & Sentiment Label*—according to the list of firms that had used the Quanjing platform for their 2022 annual report conference calls. We randomly assigned firms that were new participants on the Quanjing platform in 2023 to one of the three groups once their conference call dates were confirmed with the platform. Panel A presents this sample selection process. Panel B presents the covariate balance between the treatment and control groups. We report the means for variables such as the log of total assets at year-end (*Size*), return on assets (*ROA*), a binary variable indicating whether the firm is state-controlled (*SOE*), the percentage of shares held by institutional investors (*Institutional Holdings*), the number of analysts covering the firm (*Analysts Following*), and the difference between actual and mean analyst forecast EPS, divided by the closing price on the last trading day before the annual report date (*Earnings Surprise*). We present the average number of each characteristic with T-statistics in parentheses for testing the difference. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Panel A: Final Sample

Groups	Listed firms	Percentage (%)	Preassigned	New
Control	343	31.04	246	97
Treatment1: Summary only	373	33.76	277	96
Treatment 2: Summary & Sentiment	389	35.20	292	97
In Total	1,105	100.00	815	290

Panel B: Balance Test of the Final Sample

	Control	Summary	Summary & Sentiment Label
	(1)	(2)	(3)
Size	22.21	22.05	22.04
		(1.44)	(1.53)
ROA	0.03	0.03	0.03
		(-0.98)	(-0.38)
SOE	0.13	0.15	0.10
		(-0.41)	(-1.06)
Institutional holdings	0.37	0.36	0.34
		(0.27)	(1.16)
Analysts following	5.34	4.34	4.54
		(1.60)	(1.27)
Earnings surprise	-0.02	-0.02	-0.02
		(-1.11)	(-0.60)
# of firms	343	373	389

Table 2: Description of the Annual Report Summary by Topic and Sentiment

Panel A: Distribution of Topics

Panel A presents the distribution of topics of the annual report summary for the conference call separately for the treatment and control samples. The only difference for control firms is that we did not post the summary publicly to investors. We instructed Kimi to classify each of the key points into one of 15 predefined topics: Financial Information, Production Management, Product Markets, Supply Chain, Innovation, Risks, Government Policy, ESG, Financing, Strategy, Payout, Business Cooperation, Investors' Relationship, Capital Market, and Others.

	CONTROL		TREATMENT	
Topics	Frequency	Percent (%)	Frequency	Percent (%)
1. Financial Information	447	26.06	1,001	26.27
2.Production Management	70	4.08	149	3.91
3.Product Markets	301	17.55	362	9.50
4.Supply Chain	28	1.63	45	1.18
5.Innovation	314	18.31	665	17.45
6.Risks	257	14.99	530	13.91
7. Government Policy	8	0.47	19	0.50
8.ESG	88	5.13	294	7.72
9.Financing	59	3.44	196	5.14
10.Strategy	83	4.84	354	9.29
11.Payout	52	3.03	164	4.30
12. Business Cooperation	4	0.23	12	0.31
13. Investors' Relationship	3	0.17	9	0.24
14. Capital Market	1	0.06	3	0.08
15. Others	0	0.00	7	0.18
Total	1,715	100.00	3,810	100.00

Panel B: Distribution of Sentiment

Panel B presents the distributions of sentiments of all key points separately for the treatment and control samples. We also manually assigned a sentiment label (positive, neutral, or negative) to each of the summary points.

	CONTROL		TREATMENT	
Topics	Frequency	Percent (%)	Frequency	Percent (%)
Negative	192	11.19	413	10.84
Neutral	308	17.96	723	18.98
Positive	1,215	70.85	2,674	70.18
Total	1,715	100.00	3,810	100.00

Table 3: Summary and Overall Participation

Panel A: Univariate Test

This table presents the difference between treatment and control firms in investors' level of participation during the conference call. We measure investors' level of participation using two variables: *Questions*, the number of questions submitted by participants through the online platform; and *Participants*, the total headcount of individuals who joined the conference call. *T*-statistics of difference tests are reported in parentheses. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

Outcomes	Control	Treatment	Difference (T-C)
N	343	762	
Questions	14.27	19.86	5.59***
			(3.54)
Participants	197.19	209.97	12.78
			(0.73)

Panel B: Regression

This table reports the Poisson regression results from estimating the following model using firm-level data:

$$\text{Poisson (Questions}_i \text{ or Participants}_i) = \alpha + \beta_1 T_i + \sum \beta_n \text{Controls}_i + \text{FE} + \varepsilon_i$$

In these estimations, the outcome variable *Questions$_i$* is measured by counting the number of questions submitted by participants through the online platform during the conference call for firm *i*. *Participants$_i$* is measured as the total headcount of individuals who joined the conference call for firm *i*. *T_i* represents the randomly assigned treatment groups of firm *i*, including *Treat*, *Summary*, and *Summary & Sentiment Label*. *T_i* represents our treatment group assignment: *Treat* indicates that firm *i* was assigned to *either* treatment group, while *Summary* and *Summary & Sentiment Label* indicate the specific treatment group to which firm *i* was assigned. *Controls$_i$* includes the following control variables measured in 2023: *Size*, the log of total assets at year-end; *MB*, the total market value of equity divided by book value of equity at year-end; *ROA*, net income divided by ending total assets; *SOE*, an indicator equal to one if the firm's ultimate shareholder is the government, zero otherwise; *Institutional Holdings*, the percentage of shares controlled by institutional investors; *Analysts Following*, the log of one plus the number of analysts following the firm; and *Earnings Surprise*, the difference between actual and mean of analyst forecast EPS, divided by the closing price of the last trading day before the annual report date. Industry FEs, province FEs, and day FEs are included in all columns. Standard errors are clustered by industry. ***, **, and * denote significance at the 1%, 5%, and 10% levels, respectively.

Dependent Variable:	Questions		Participants	
	(1)	(2)	(3)	(4)
Treat	0.383***		0.090**	
	(3.97)		(2.04)	
Summary		0.378***		0.065
		(4.41)		(0.83)
Summary & Sentiment Label		0.388***		0.115***
		(3.53)		(2.64)
Size	0.206*	0.206*	0.282***	0.281***

	(1.90)	(1.91)	(4.37)	(4.30)
MB	0.067***	0.067***	0.078***	0.078***
	(3.05)	(3.04)	(3.66)	(3.64)
ROA	0.293	0.291	-0.335	-0.341
	(0.39)	(0.39)	(-0.39)	(-0.40)
SOE	0.015	0.015	0.003	0.005
	(0.14)	(0.14)	(0.08)	(0.11)
Institutional holdings	-0.047	-0.046	-0.048	-0.044
	(-0.31)	(-0.31)	(-0.37)	(-0.34)
Analysts following	0.055	0.055	0.057	0.057
	(1.10)	(1.10)	(1.23)	(1.23)
Earnings surprise	1.698	1.701	0.613	0.620
	(1.12)	(1.11)	(0.62)	(0.62)
H0: T1-T2		-0.010		-0.050
		(0.05)		(0.32)
Industry FE, Province FE and Day FE	Yes	Yes	Yes	Yes
# of Observations	1,105	1,105	1,105	1,105
R-squared	0.20	0.20	0.48	0.48

Table 4: Topical Alignment of Investors' Questions

Panel A: Univariate Test

This table presents the properties of investors' questions across treatment and control firms. *Alignment* is a binary indicator that equals one if the topic of an investor's question matches any of the five key points' topics, zero otherwise. The annual report summary was made public for the treatment group but not for the control group. All analyses are conducted at the question level. *Z*-statistics of proportion tests are reported in parentheses. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

Control		Treatment		Difference (T-C)
N	*Alignment*	N	*Alignment*	
4,892	0.50	15,139	0.57	0.07**
				(8.70)

Panel B: Regression

This table reports the regression results from estimating the following model using question level data:

$$\text{Logit } (Alignment) = \alpha + \beta_1 T_i + \sum \beta_n Controls_i + \text{FE} + \varepsilon_i$$

In the estimation, the dependent variable *Alignment* is a b dummy variable that equals one if the topic of an investor's question matches any of the five key points' topics, zero otherwise. T_i represents our treatment group assignment for firm *i*: *Treat*, *Summary*, or *Summary & Sentiment Label*. The logit regression model is applied to all columns. We include the same array of control variables throughout the paper. Industry FEs, province FEs, and day FEs are included in all columns. Standard errors are clustered by industry. ***, **, and * denote significance at the 1%, 5%, and 10% levels, respectively.

Dependent Variable:	*Alignment*	
	(1)	(2)
Treat	0.247***	
	(3.78)	
Summary		0.201***
		(3.05)
Summary & Sentiment Label		0.294***
		(3.89)
H0: T1-T2		-0.093*
		(2.80)
Controls	Yes	Yes
Industry FE, Province FE and Day FE	Yes	Yes
# of Observations	20,031	20,031
R-squared	0.02	0.02

Table 5: Conditional on Topics and Sentiments

Panel A: Topics and Topical Alignment

This table reports the regression results from estimating the following model using question-level data. Only questions that were aligned with a topic in the summaries were used for the analyses (*Alignment* = 1):

$$\text{Logit}\ (\text{Financial}_{i,j}\ /\ \text{Risk}_{i,j}/\ \text{Strategy}_{i,j}/\ \text{Payout}_{i,j}) = \alpha + \beta_1\ T_i + \sum \beta_n\ \text{Controls}_{i,j} + \text{FE} + \varepsilon_{i,\,j}$$

In these estimations, the outcome variable Financial/ Risk/ Strategy/ Payout is an indicator that equals one if the topic of a certain key point is about Financial information/ Risk/ Strategy/ Payout, zero otherwise. T_i represents the randomly assigned treatment group of firm i: *Treatment*, *Summary*, or *Summary & Sentiment Label*. The logit regression model is applied to all columns. Industry FEs, province FEs, and day FEs are included in all columns. Standard errors are clustered by industry. ***, **, and * denote significance at the 1%, 5%, and 10% levels, respectively.

Dependent Variable:	Financial		Risk		Strategy		Payout	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Treat	0.111**		0.510***		0.758***		0.935***	
	(2.29)		(4.00)		(4.45)		(3.59)	
Summary		0.135***		0.486***		0.821***		1.145***
		(2.58)		(3.05)		(4.64)		(4.71)
Summary & Sentiment Label								
		0.086		0.536***		0.684***		0.747**
		(1.42)		(4.67)		(4.01)		(2.18)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Industry FE, Province FE and Day FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
# of Observations	11,078	11,078	11,078	11,078	11,078	11,078	11,078	11,078
R-squared	0.03	0.03	0.05	0.05	0.08	0.08	0.18	0.18

Panel B: Sentiments and Topical Alignment

This table reports the regression results from estimating the following model using question-level data and the converging sample (questions with *Alignment* = 1):

$$\text{Logit}\ (\text{Non-Positive}) = \alpha + \beta_1\ T_i + \sum \beta_n\ \text{Controls}_{i,j} + \text{FE} + \varepsilon_{i,\,j}$$

In this estimation, the outcome variable Non-Positive is an indicator that equals one if the sentiment of a certain key point is non-positive (neutral and negative), zero otherwise. T_i represents the randomly assigned treatment group of firm i: *Treat*, *Summary*, or *Summary & Sentiment Label*. The logit regression model is applied to all columns. Industry FEs, province FEs, and day FEs are included in all columns. Standard errors are clustered by industry. ***, **, and * denote significance at the 1%, 5%, and 10% levels, respectively.

Dependent Variable:	Non-Positive	
	(1)	(2)
Treat	0.204***	
	(3.13)	
Summary		0.203
		(1.54)
Summary & Sentiment Label		0.204**
		(2.26)
Controls	Yes	Yes
Industry FE, Province FE and Day FE	Yes	Yes
# of Observations	11,078	11,078
R-squared	0.18	0.18

Table 6: Questions by Investor Type

Panel A: Number of Questions Raised by Silent versus Vocal Investors

This table reports the regression results from estimating the following model using question-level data:

$$\text{Poisson } (Questions_i) = \alpha + \beta_1 T_i + \sum \beta_n Controls_i + \text{FE} + \varepsilon_i$$

In these estimations, the outcome variable $Questions_i$ is measured by counting the number of questions submitted by participants through the online platform during the conference call for firm i. T_i represents the randomly assigned treatment group of firm i: *Treat*, *Summary*, or *Summary & Sentiment Label*. In columns (1) and (2), we use the number of questions raised by silent investors (i.e., investors who did not ask any questions in the 2022 (the prior year) conference calls), while in columns (3) and (4) we use the number of questions raised by vocal investors (i.e., investors who asked questions in the 2022 conference calls). The Poisson regression model is applied to all columns. Industry FEs, province FEs, and day FEs are included in all columns. Standard errors are clustered by industry. ***, **, and * denote significance at the 1%, 5%, and 10% levels, respectively.

Dependent Variable:	Questions		Questions	
	(1)	(2)	(3)	(4)
	Silent Investors		Vocal Investors	
Treat	0.249**		0.143*	
	(2.35)		(1.95)	
Summary		0.253**		0.156***
		(2.55)		(2.73)
Summary & Sentiment Label		0.246**		0.131
		(2.12)		(1.06)
Controls	Yes	Yes	Yes	Yes
Industry FE, Province FE and Day FE	Yes	Yes	Yes	Yes
# of Observations	1,091	1,091	1,091	1,091
R-squared	0.16	0.16	0.15	0.15

Panel B: Alignment Conditional on Silent versus Vocal Investors

This table reports the regression results from estimating the following model using question-level data, using the full sample:

$$\text{Logit } (Alignment_{i,j}) = \alpha + \beta_1 T_i + \sum \beta_n Controls_i + \text{FE} + \varepsilon_i$$

In the estimation, the dependent variable $Alignment_{i,j}$ is a binary indicator that equals one if the topic of an investor's question j matches any of the five key points' topics from firm i's annual report, zero otherwise. T_i represents our treatment group assignment for firm i: *Treat*, *Summary*, or *Summary & Sentiment Label*. In columns (1) and (2), we use the questions raised by silent investors (i.e., investors who did not ask any questions in the 2023 (the prior year) conference calls), while in columns (3) and (4) we use the questions raised by vocal investors (i.e., investors who asked questions in the 2023 conference calls). The logit regression model is applied to all columns. Industry FEs, province FEs, and day FEs are included in all columns. Standard errors are clustered by industry. ***, **, and * denote significance at the 1%, 5%, and 10% levels, respectively.

Dependent Variable:	Alignment		Alignment	
	(1)	(2)	(3)	(4)
	Silent Investors		Vocal Investors	
Treat	0.162**		-0.183***	
	(2.54)		(-3.00)	
Summary		0.119*		-0.195***
		(1.87)		(-3.71)
Summary & Sentiment Label		0.205***		-0.170*
		(2.72)		(-1.83)
Controls	Yes	Yes	Yes	Yes
Industry FE, Province FE and Day FE	Yes	Yes	Yes	Yes
# of Observations	11,369	11,369	2,367	2,367
R-squared	0.18	0.18	0.12	0.12

Panel C: Alignment Conditional on More versus Less Active Investors

This table reports the regression results from estimating the following model using question-level data and the full sample:

$$\text{Logit } (Alignment_{i,j}) = \alpha + \beta_1 \, T_i + \textstyle\sum \beta_n \, Controls_i + \text{FE} + \varepsilon_i$$

In the estimation, the dependent variable $Alignment_{i,j}$ is a binary indicator that equals one if the topic of an investor's question j matches any of the five key points' topics from firm i's annual report, zero otherwise. T_i represents our treatment group assignment for firm i: *Treat*, *Summary*, or *Summary & Sentiment Label*. In columns (1) and (2) are investors who ask questions in less than six firms' 2023 conference calls; in columns (3) and (4) are investors who ask questions in more than five firms' 2023 conference calls. The logit regression model is applied to all columns. Industry FEs, province FEs, and day FEs are included in all columns. Standard errors are clustered by industry. ***, **, and * denote significance at the 1%, 5%, and 10% levels, respectively.

Dependent Variable:	Alignment		Alignment	
	(1)	(2)	(3)	(4)
	Less active Investors		More active Investors	
Treat	0.107**		0.053	
	(2.19)		(0.64)	
Summary		0.088		-0.028
		(1.40)		(-0.36)
Summary & Sentiment Label		0.128**		0.126
		(2.15)		(1.16)
Controls	Yes	Yes	Yes	Yes
Industry FE, Province FE and Day FE	Yes	Yes	Yes	Yes
# of Observations	7,841	7,841	5,093	5,093
R-squared	0.02	0.02	0.04	0.04

Table 7: The Quality of Firms' Responses

Panel A: Univariate Test

This table presents the difference in firms' response quality between treatment and control firms. In this estimation, the dependent variable *Length* is the total number of words contained in each response. *Informative* is an indicator variable that equals one if an answer receives an AI-generated rating above the median score of three, zero otherwise. All analyses are conducted at the question level. *Z*-statistics of proportion tests are reported in parentheses. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

Length	Control	Treatment	Difference (T-C)
	313.64	348.22	34.58***
			(6.00)
Informative			
	0.41	0.44	0.03***
			(3.75)

Panel B: Regression

This table reports the regression results from estimating the following model using question-level data:

$$\text{Poisson}(\text{Length}_{i,t}) / \text{Logit}(\text{Informative}_{i,t}) == \alpha + \beta_1 \text{ Ti} + \sum \beta_n \text{Controls}_{i,t} + \text{FE} + \varepsilon_{i,t}$$

In this estimation, the dependent variable *Length* is the total number of words contained in each response. *Informative* is an indicator variable that equals one if an answer receives an AI-generated rating above the median score of three, zero otherwise. *Ti* represents the firm's randomly assigned treatment group: *Treat*, *Summary*, or *Summary & Sentiment Label*. The Poisson model is reported in columns 1 and 2, and the Logit model is reported in columns 3 and 4. Control variables, industry FEs, province FEs, and day FEs are included in all columns. Standard errors are clustered by industry. ***, **, and * denote significance at the 1%, 5%, and 10% levels, respectively.

Dependent Variable:	Length		Informative	
	(1)	(2)	(3)	(4)
Treat	0.127***		0.146***	
	(5.87)		(2.94)	
Summary		0.135***		0.176**
		(4.77)		(2.04)
Summary & Sentiment Label		0.118***		0.114*
		(4.50)		(1.91)
Controls	Yes	Yes	Yes	Yes
Industry FE, Province FE and Day FE	Yes	Yes	Yes	Yes
Date Fixed Effects	Yes	Yes	Yes	Yes
# of Observations	17,417	17,417	17,417	17,417
R-squared	0.05	0.05	0.05	0.05

Panel C: Conditional on Alignment

This table reports the regression results from estimating the following model using question-level data, conditional on whether a question topic aligns with a topic in the summary of the annual report:

$$\text{Poisson}(\text{Length}_{i,t}) / \text{Logit}(\text{Informative}_{i,t}) == \alpha + \beta_1 \text{ Ti} + \sum \beta_n \text{Controls}_{i,t} + \text{FE} + \varepsilon_{i,t}$$

In this estimation, the dependent variable *Length* is the total number of words in each response. *Informative* is an indicator variable that equals one if an answer receives an AI-generated rating above the median score of three, zero otherwise. *Ti* represents the firm's randomly assigned treatment group: *Treat*, *Summary*, or *Summary & Sentiment Label*. The Poisson model is reported in columns 1 and 2, and the logit model is reported in columns 3 and 4. Control variables, industry FEs, province FEs, and day FEs are included in all columns. Standard errors are clustered by industry. ***, **, and * denote significance at the 1%, 5%, and 10% levels, respectively.

Dependent Variable:	Length		Informative	
	(1)	(2)	(3)	(4)
	Aligned	Not-aligned	Aligned	Not-aligned
Treat	0.144***	0.085***	0.162***	0.113*
	(5.83)	(2.69)	(2.97)	(1.90)
H1: A-NA	0.059		0.051	
	(2.47)		(0.34)	
Controls	Yes	Yes	Yes	Yes
Industry FE, Province FE and Day FE	Yes	Yes	Yes	Yes
# of Observations	9,705	7,712	9,705	7,712
R-squared	0.07	0.08	0.03	0.03

Panel D: Conditional on the Percentage of Alignment at the Firm Level

This table reports the regression results from estimating the following model using firm-level data, conditional on the percentage of aligned questions in each conference call:

$$\text{Poisson}(\text{Length}_{i,t}) / \text{Logit}(\text{Informative}_{i,t}) == \alpha + \beta_1 \text{ Ti} + \sum \beta_n \text{Controls}_{i,t} + \text{FE} + \varepsilon_{i,t}$$

In this estimation, the dependent variable *Length* is the total number of words in each response. *Informative* is an indicator variable that equals one if an answer receives an AI-generated rating above the median score of three, zero otherwise. *Ti* represents the firm's randomly assigned treatment group: *Treat*, *Summary*, or *Summary & Sentiment Label*. The Poisson model is reported in columns 1 and 2, and the logit model is reported in columns 3 and 4. Control variables, industry FEs, province FEs, and day FEs are included in all columns. Standard errors are clustered by industry. ***, **, and * denote significance at the 1%, 5%, and 10% levels, respectively.

Dependent Variable:	Length		Informative	
	(1)	(2)	(3)	(4)
	High percentage	Low percentage	High percentage	Low percentage
Treat	0.202***	0.027	0.260***	0.036
	(5.00)	(0.61)	(4.32)	(0.61)
H1: A-NA	0.175****		0.224***	
	(20.62)		(7.18)	
Controls	Yes	Yes	Yes	Yes

Industry FE, Province FE and Day FE	Yes	Yes	Yes	Yes
# of Observations	8,915	8,502	8,915	8,502
R-squared	0.09	0.10	0.04	0.04

Table 8: Market-wide effects

Panel A: Capital Market

This table reports the Poisson regression results from estimating the following model using firm-level data:

$$Abs\_CAR[-1, 5] \,/\, Turnover[-1, 5] = \alpha + \beta_1 \, Treat_t + \sum \beta_n \, Controls_{i,t} + FE + \varepsilon_{i,t}$$

In this estimation, the dependent variable *Turnover [-1, 5]* is the cumulative turnover ratio, which equals the ratio of the number of shares traded during the windows to the number of shares outstanding; and *Abs_CAR[-1, 5]* is the absolute value of cumulated abnormal returns in the window [-1, 5], where abnormal returns are calculated as raw returns less the market returns on the same day. *Ti* represents the firm's randomly assigned treatment group. Control variables, industry FEs, province FEs and day FEs are included in all columns. Standard errors are clustered by industry. ***, **, and * denote significance at the 1%, 5%, and 10% levels, respectively.

Dependent Variable:	Turnover [-1, 5]		Abs_CAR[-1, 5]	
	(1)	(2)	(3)	(4)
Treat	0.021*		0.004	
	(1.91)		(0.74)	
Summary		0.045**		0.011**
		(2.67)		(2.60)
Summary & Sentiment Label		-0.003		-0.003
		(-0.20)		(-0.39)
Controls	Yes	Yes	Yes	Yes
Industry FE, Province FE and Day FE	Yes	Yes	Yes	Yes
# of Observations	1,105	1,105	1,105	1,105
R-squared	0.01	0.24	0.14	0.04

Panel B: Investor Interaction Platform

This table reports the Poisson regression results from estimating the following model using firm-level data:

$$Ratio\ (Posts\ [0,30] /\ Posts\ [-90, -1]) = \alpha + \beta_1 \, Treat_t + \sum \beta_n \, Controls_{i,t} + FE + \varepsilon_{i,t}$$

In this estimation, the dependent variable, Posts [0,30]/ Posts [-90, -1], is defined as the ratio of posts from investors on the investor interaction platforms (EasyIR for the Shenzhen Stock Exchange and ehudong for the Shanghai Stock Exchange) within the period [0,30] to posts during the period [-90, -1]. We use this variable to measure the change in investor activity. The event day is the date of the conference call. *Ti* represents the firm's randomly assigned treatment group. Control variables, industry FEs, province FEs, and day FEs are included in all columns. Standard errors are clustered by industry. ***, **, and * denote significance at the 1%, 5%, and 10% levels, respectively.

Dependent Variable:	Posts [0,30]/ Posts [-90, -1]		Posts [0,90]/ Posts [-90, -1]	
	(1)	(2)	(3)	(4)
Treat	0.061***		0.090	
	(3.34)		(0.93)	
Summary		0.072***		0.126

		(5.75)		(1.20)
Summary & Sentiment Label		0.049		0.056
		(1.48)		(0.60)
Controls	Yes	Yes	Yes	Yes
Industry FE, Province FE and Day FE	Yes	Yes	Yes	Yes
# of Observations	942	942	942	942
R-squared	0.131	0.131	0.105	0.106

Table 9: Potential Costs: An Analysis of Questions Withheld by Firms

Panel A: Univariate Test

This table presents the difference in the percentage of questions raised by investors but not posted online between treatment and control firms. *Withhold* is an indicator variable that equals one for questions raised but not posted online, zero otherwise. All analyses are conducted at the question level. *Z*-statistics of proportion tests are reported in parentheses. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

Withhold	Control (4,892)	Treatment (15,139)	Difference (T-C)
	7.46%	14.85%	7.39%***
			(13.41)

Panel B: Regression

This table reports the regression results from estimating the following model using question-level data:

$$\text{Logit}(Withhold_{i,j}) == \alpha + \beta_1 \, Ti + \sum \beta_n \, Controls_{i,j} + FE + \varepsilon_{i,j}$$

In this estimation, the dependent variable *Withhold* is an indicator that equals one for questions raised but not posted online, zero otherwise. *Ti* represents the firm's randomly assigned treatment group: *Treat*, *Summary*, or *Summary & Sentiment Label*. The logit regression model is applied to all columns. Industry FEs, province FEs, and day FEs are included in all columns. Standard errors are clustered by industry. ***, **, and * denote significance at the 1%, 5%, and 10% levels, respectively.

Dependent Variable:	Withhold	
	(1)	(2)
Treat	0.547**	
	(2.02)	
Summary		0.653**
		(2.12)
Summary & Sentiment Label		0.445*
		(1.70)
Controls	Yes	Yes
Industry FE, Province FE and Day FE	Yes	Yes
# of Observations	20,031	20,031
R-squared	0.21	0.21

Panel C: Conditional on Alignment Questions

This table reports the regression results from estimating the following model using question-level data:

$$\text{Logit}(Withhold_{,t}) == \alpha + \beta_1 \, Ti + \beta_2 \, Ti * Alignment + \beta_3 \, Alignment + \sum \beta_n \, Controls_{i,j} + FE + \varepsilon_{i,j}$$

In this estimation, the dependent variable *Withhold* is an indicator that equals one for questions raised but not posted online, zero otherwise. *Alignment$_{i,j}$* is a binary indicator that equals one if the topic of an investor's question *j* matches any of the five key points' topics from firm *i*'s annual report, zero otherwise. *Ti* represents the firm's randomly assigned treatment group: *Treat*, *Summary*, or *Summary & Sentiment Label*. The logit regression model is applied to all columns. Control variables, industry FEs, province FEs, and day FEs are included in all columns. Standard errors are clustered by industry. ***, **, and * denote significance at the 1%, 5%, and 10% levels, respectively.

Dependent Variable:	Withhold	
	(1)	(2)
Treat * Alignment	0.286***	
	(2.81)	
Summary * Alignment		0.343***
		(2.74)
Summary & Sentiment Label * Alignment		0.241**
		(2.23)
Treat	0.427	
	(1.64)	
Summary		0.501
		(1.60)
Summary & Sentiment Label		0.351
		(1.45)
Alignment	-0.424***	-0.423***
	(-4.49)	(-4.41)
Controls	Yes	Yes
Industry FE, Province FE and Day FE	Yes	Yes
# of Observations	20,031	20,031
R-squared	0.21	0.21

Panel D: Redundancy in Withheld Questions

This table reports the regression results from estimating the following model using interaction-level data:

$$\text{Redundant or Redundant Indicator} == \alpha + \beta_1 \text{ Ti} + \beta_2 \text{ Ti} * \text{Withhold} + \beta_3 \text{ Withhold} + \sum \beta_n \text{Controls}_{i,j} + \text{FE} + \varepsilon_{i,j}$$

In this estimation, the dependent variable *Redundant* is the mean of the Jaccard similarity of each question with other questions for the same firm and same topic. *Redundant Indicator* is an indicator variable that equals one if the similarity mean score is above 0.25 (which is the top 10%), zero otherwise. *Withhold* is an indicator variable that equals one for questions raised but not posted online, zero otherwise. *Ti* represents the firm's randomly assigned treatment group: *Treat*, *Summary*, or *Summary & Sentiment Label*. Control variables, industry FEs, province FEs, and day FEs are included in all columns. Standard errors are clustered by industry. ***, **, and * denote significance at the 1%, 5%, and 10% levels, respectively.

Dependent Variable:	Redundant		Redundant Indicator	
	(1)	(2)	(3)	(4)
Withhold	0.021***	0.030*	0.820***	0.977***
	(8.53)	(2.04)	(10.35)	(3.49)
Withhold * Treat		-0.011		-0.177
		(-0.58)		(-0.48)
Treat		0.004*		-0.100

		(1.77)		(-0.95)
Controls	Yes	Yes	Yes	Yes
Industry FE, Province FE and Day FE	Yes	Yes	Yes	Yes
# of Observations	20,031	20,031	20,031	20,031
R-squared	0.041	0.042	0.077	0.077

Table 10: Alternative Specifications

Panel A: Univariate Test

This table presents the difference in the level of investor participation between treatment and control firms. We only keep sample firms that hosted both their 2022 and 2023 annual conference calls on the Quanjing platform, so that we can compare the *change* in investors' participation between the treatment and control firms. We measure investors' participation using *Questions*, which is the number of questions submitted by participants through the online platform during the conference call; and *Participants*, which is the total headcount of individuals who joined the conference call. *Z*-statistics of proportion tests are reported in parentheses. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

Questions	2023	2024	Difference
Control (246+246)	18.80	14.35	-4.45**
			(-2.57)
Treatment (569+569)	18.56	19.91	1.35
			(0.85)
Participants	2023	2024	Difference
Control (246+246)	200.85	204.95	4.11
			(0.18)
Treatment (569+569)	190.53	210.15	19.62
			(1.25)

Panel B: Regression

This table reports the Poisson regression results from estimating the following model using firm-level data:

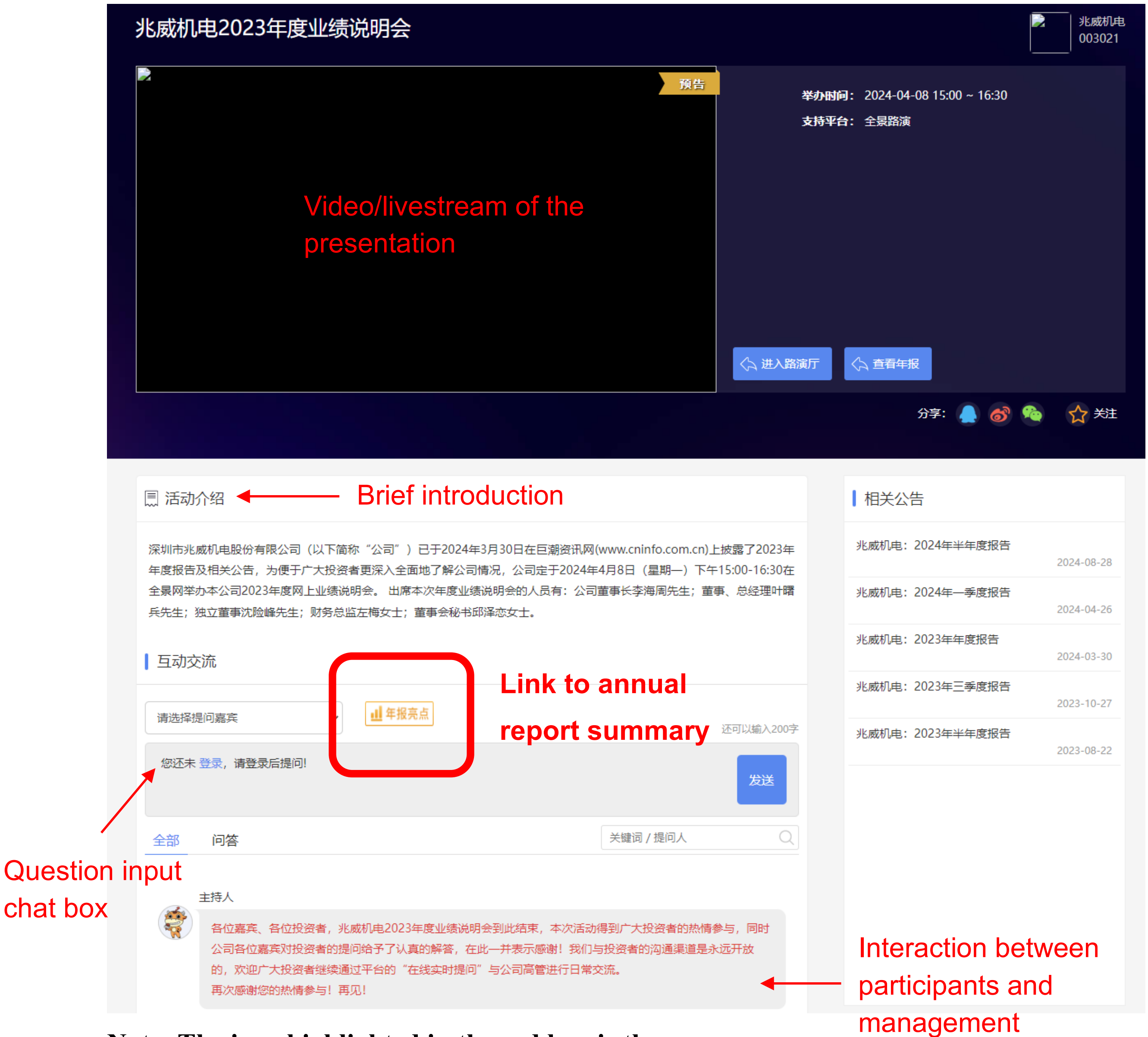
$$\text{Poisson}(\text{Questions}_{i,t} \text{ or Participants}_{i,t}) = \alpha + \beta_1 T_i * \text{POST} + \sum \beta_n \text{Controls}_{i,t} + \text{FE} + \varepsilon_{i,t}$$

We only keep sample firms that hosted both their 2022 and 2023 annual conference calls on the Quanjing platform, so that we can compare the *change* in investors' participation between treatment and control firms. In this estimation, the dependent variable $Questions_i$ is the number of questions submitted by participants through the online platform during the conference call for firm *i*. $Participants_i$ is the total headcount of individuals who joined the conference call for firm *i*. T_i represents the firm's randomly assigned treatment group: *Treat*, *Summary*, or *Summary & Sentiment Label*. *POST* is an indicator variable that equals one for the 2023 annual conference call and zero for the 2022 annual conference call. Control variables, firm FEs, and day FEs are included in all columns. Standard errors are clustered by industry. ***, **, and * denote significance at the 1%, 5%, and 10% levels, respectively.

Dependent Variable:	Questions		Participants	
	(1)	(2)	(3)	(4)
Treat * POST	0.344***		0.135***	
	(6.06)		(3.16)	
Summary * POST		0.353***		0.104**
		(5.68)		(2.07)
Summary & Sentiment Label * POST		0.336***		0.164***
		(3.62)		(4.18)

Controls	Yes	Yes	Yes	Yes
Firm FE and Day FE	Yes	Yes	Yes	Yes
# of Observations	1,630	1,630	1,630	1,630
R-squared	0.68	0.68	0.88	0.88

Figure 1: Example of a Conference Call Page



Note: The icon highlighted in the red box is the

Figure 2: An Example of an AI-generated Summary

a. Treatment: Summary Only

b. Treatment: Summary & Sentiment

Appendix A: Sample Summaries (Translated)

Panel A: A Sample Summary

1. Company Overview and Main Business

Shandong Chenming Paper Group Co., Ltd. is a limited liability company located in Shouguang City, Shandong Province, China. Originally known as Shouguang County Paper Mill, it has undergone several reorganizations and shareholding system reforms to become a large-scale company primarily engaged in the production of machine-made paper and paperboard. The company's main business activities include the production and sales of machine-made paper, electricity and thermal power, building materials, and papermaking chemical products. Additionally, the company is involved in forestry cultivation, seedling breeding, and processing and sales of timber and building materials, as well as hotel services, equipment financing leases, and other fields.

2. Financial Status

According to the 2023 annual report, the company is facing certain financial pressures. During the reporting period, the company achieved a revenue of 26,608,570,228.20 yuan, a decrease of 16.86% compared to the previous year; the net profit attributable to shareholders of the listed company was -1,281,289,649.82 yuan, indicating a loss. The company's total assets were 79,847,052,953.58 yuan, with total liabilities of 58,389,197,107.95 yuan and a debt ratio of 73.6%, showing that the enterprise has a certain financial pressure.

3. Research and Innovation

The company places great emphasis on technology research and development as well as innovation. During the reporting period, it applied for 33 patents and obtained 42 authorized patents, reflecting continuous investment in technological innovation and achievements. Through continuous technological improvements and product innovation, the company aims to enhance product quality and market competitiveness, laying the foundation for long-term development.

4. Market and Sales

The company's products are mainly sold in the Chinese mainland and overseas markets, with the machine-made paper business being the main source of the company's revenue. The company strives to expand the market and improve sales efficiency by establishing a sales network, implementing sales strategies, and optimizing customer service. However, during the reporting period, due to downstream demand and raw material price fluctuations, the company's sales revenue declined.

5. Environment and Social Responsibility

The company actively fulfills its environmental protection responsibilities, adhering to the concept of “green development and environment first,” and invests heavily in the construction of environmental protection facilities, such as alkali recovery systems and water reuse systems, to reduce environmental pollution in the production process. At the same time, the company also focuses on social responsibility, giving back to society through various public welfare activities, such as contributing to the "Good Quality Shandong" brand, enhancing the brand image.

Panel B: A Sample of Summary & Sentiment

1. Company's Potential Risks [Negative]

Financial Status and Profitability: Zhejiang Boyuan Electrical Co., Ltd., reported a revenue of 311,609,137.40 yuan for the year 2023, marking an 11.86% decrease from the previous year's 353,531,847.22 yuan. The net profit attributable to shareholders of the listed company was 32,990,152.49 yuan, a significant 52.40% decrease from 69,302,747.72 yuan in 2022. The basic earnings per share decreased from 1.07 yuan in 2022 to 0.41 yuan, a decline of 61.68%. However, the net cash flow from operating activities improved to a positive 25,956,925.88 yuan, a substantial increase of 197.37% from the previous year.

2. Company's Steady Development [Neutral]

Dividend Plan and Capital Situation: The company's board of directors has approved a profit distribution plan, which proposes to distribute a cash dividend of 0.86 yuan per 10 shares (before tax) to all shareholders based on the total share capital excluding repurchased shares at the time of the next profit distribution plan implementation. The company has a total share capital of 80,000,000 shares and a registered capital of 80,000,000.00 yuan as of the end of the reporting period.

3. Company's Future Growth [Positive]

Main Business and Market Layout: The company's main business focuses on the research and development, production, and sales of electrical insulating materials and other polymer composites. Its products include insulating resins, slot wedges and laminated products, fiber products, mica products, and binding products, which are used in various fields such as wind power generation, rail transit, industrial motors, household appliances, new energy vehicles, and hydroelectric power. The company has a stable market presence in China and is actively expanding into international markets.

4. Company's Future Growth [Positive]

R&D Investment and Technological Innovation: The company places a high priority on R&D investment, with a 2023 R&D expense of 25,108,088.78 yuan, a 5.71% increase from 23,751,502.28 yuan in 2022. It holds 101 invention patents and 25 utility model patents, has participated in the drafting of multiple national, industry, and group standards, and has undertaken key national and provincial scientific research projects.

5. Company's Steady Development [Neutral]

Risk Factors and Response Measures: The company faces risks such as high customer concentration, safety production risks, potential uncollectible accounts receivable, and raw material price volatility. To mitigate these risks, the company plans to adopt measures such as diversifying market layouts, strengthening safety production management, optimizing accounts receivable management, and procurement strategies to reduce the impact of potential risks.

Declaration of generative AI and AI-assisted technologies in the writing process.

During the preparation of this work the author(s) used KimiAI in order to translate Chinese versions of the sample summaries to English. After using this tool/service, the author(s) reviewed and edited the content as needed and take(s) full responsibility for the content of the published article.