Does Artificial Intelligence Reduce Investor Disagreement? Evidence from AI-Human Interactions Data*

Shiyang Huang
Faculty of Business and Economics, University of Hong Kong

Zhenghan Shao
School of Economics and Management, Tsinghua University

Yalin Song
School of Economics and Management, Tsinghua University

Tianyu Wang
School of Economics and Management, Tsinghua University

Jian Xue
School of Economics and Management, Tsinghua University

First Draft: January 2026
This Draft: May 2026

*We thank Danqing Mei (discussant), Da Xu for helpful comments. Please address correspondence to huangsy@hku.hk, shaozh24@mails.tsinghua.edu.cn, songyl@sem.tsinghua.edu.cn, wangty6@sem.tsinghua.edu.cn, xuejian@sem.tsinghua.edu.cn.

Does Artificial Intelligence Reduce Investor Disagreement? Evidence from AI-Human Interactions Data

Abstract

We use a proprietary granular user–chatbot (GenAI) interaction data to examine how users search for stock-related information and how human-AI interactions shape investors' beliefs. We first document that GenAI-generated answers shape users' beliefs, and users' beliefs become more polarized over the course of the AI-user interaction. This increased dispersion is driven primarily by between-topic dispersion, while within-topic dispersion weakly declines. We explore underlying mechanisms and confirm that the increased dispersion is not due to noise embedded in GenAI-generated answers, instead, GenAI-generated answers contain information and can predict future stock returns, which can explain the decline in within-topic dispersion. Meanwhile, users exhibit stickiness in topic searching, and this stickiness is not information-driven and contributes to the increased between-topic dispersion. We further examine the asset pricing implications in broad market and find that, following ChatGPT's deployment, both analyst forecasts and investor beliefs exhibit greater dispersion after earnings announcements.

Keywords: Generative AI, Information acquisition, Investor Belief, Dispersion, Disagreement
JEL Classification: D83; G11; G14

1 Introduction

Investors increasingly rely on AI tools (e.g., large language models, prediction algorithms, robo-analysts) in their daily investment activities. Survey evidence from a global sample of 11,000 retail investors highlights the rapid diffusion of generative AI in financial decision-making: approximately 50% of respondents report an intention to use tools such as ChatGPT or Gemini for portfolio selection, while 13% have already incorporated these technologies into their active investment workflows.[1]

Given the crucial role that AI plays in information processing and broad capital market, the deployment of ChatGPT in November 2022 has sparked a significant body of academic research. Cheng et al. (2025), Chang et al. (2023), and Even-Tov et al. (2025) find that AI tools help retail investors narrow the sophistication gap by improving information processing and supporting information-driven trading. Bertomeu et al. (2025) find that AI tools also support analysts' information production, and that restrictions on these tools reduce market liquidity and forecast accuracy. Sheng et al. (2026) document that institutional adoption of AI helps alpha generation. While prior studies find strong impacts of AI on investor trading and positive implications for capital market efficiency, it is still unclear whether and how AI usage would affect investors' belief formation and belief dispersion.

To fill this gap, we utilize a proprietary granular dataset of user–platform interactions from one of China's largest GAI service providers. This dataset includes detailed records of user inputs and corresponding AI responses, and we can identify stock-related queries at the investor level. Compared with recent studies on the effects of ChatGPT on financial markets relying on exogenous shocks (e.g., ChatGPT outages or regulatory bans), our dataset has advantages in granularity and completeness. These features make the data particularly well suited for our research questions, as its granularity and completeness allow us to precisely track investors' information acquisition and belief updating throughout the human-AI interaction process, which in turn affects the stock market.

[1] See: Reuters "'ChatGPT, what stocks should I buy?' AI fuels boom in robo-advisory market."

More importantly, we can observe multiple rounds of human–AI interaction each individual single user. By tracking changes in the same user's beliefs across multiple AI interactions, we can isolate the effect of AI on individual belief formation and on dispersion across users, and we can identify various drivers of belief dispersion, including heterogeneous information, AI sophistication, and users' interests in different topics.

AI tools could have ambiguous effects on investor belief and belief dispersion. On the one hand, AI processes earnings calls, filings, news, and alternative data at scale, providing investors with more accurate information about firm fundamentals. As a result, investors may form more homogeneous beliefs. On the other hand, heterogeneous access to AI tools, differences in prompts, data inputs, or interpretation skills can increase disagreement. For instance, investors may be interested in different topics, and they can dig deeper into their interests with the help of AI tools, amplifying the dispersion among investors. As another example, the false information embedded in AI-generated answers ("AI hallucination") could introduce noise to investors' belief formation, contributing to the dispersion. Wang (2025) shows that LLMs frequently hallucinate accounting facts—both by deviating from existing financials and by fabricating nonexistent items, implying that AI can shift costs from information processing to information verification. Consequently, it remains an open question whether AI tools enhance informational efficiency and whether they amplify or attenuate dispersion in investors' beliefs about firms.[2]

To conduct our study, we first apply natural language processing techniques to construct a measure of users' tone toward firms by extracting sentiment from user-generated prompts. Specifically, we quantify sentiment using the finance-domain sentiment model (FinBERT) and construct a continuous tone index from the predicted sentiment shares (Huang et al., 2023). This state-of-the-art large language model adapts to the finance domain, incorporates finance knowledge, and is better able to capture summarize con-

[2]Recent surveys characterize LLM hallucinations as a systematic factuality deficit, where models generate fluent yet unsupported or inconsistent statements (Zhang et al., 2025; Huang et al., 2025). This "distorted information" extends beyond blatant fabrication to subtle shifts that undermine the perceived reliability of a claim (Sun et al., 2024).

textual information in financial texts. Based on this tone measure, we can calculate the dispersion in beliefs across human users.

Our analysis yields several key findings. First, we document a pronounced increase in the dispersion of human users' tone as conversations progress from the initial round to the second (or final) round of AI–human interactions. On average, tone dispersion in second-round user prompts increases by 11% relative to the initial round. We confirm this pattern in regression analyses even after controlling for firm and time fixed effects. Specifically, we find that the absolute value of user tone increases by 25% from the first to the second round, suggesting that users' beliefs become more polarized over the course of the interaction. These results provide empirical evidence consistent with recent experimental and theoretical works (Glickman and Sharot, 2025; Acemoglu et al., 2025), showing that sentiment dispersion among humans increases with greater human–AI interaction.

Second, we explore the underlying sources of the increase in dispersions. To this end, we decompose overall dispersion into within-topic dispersion, which captures heterogeneity in tone among users discussing the same topic (e.g., price movements, firm fundamentals, or macroeconomic conditions)—and between-topic dispersion, which reflects differences in tone across topics emphasized by different users. We find that the increase in overall dispersion is primarily driven by a substantial rise in between-topic dispersion. Specifically, between-topic dispersion increases by 1.07 (approximately 14%) in the second round, relative to an initial-round mean of 7.47 in the initial round. By contrast, within-topic dispersion declines slightly, suggesting that users discussing the same topic become marginally more aligned in tone. Taken together, these findings imply that AI–human interactions cause investors to diverge more in what aspects of the firm they focus on, rather than in how they evaluate the same aspect.

Third, we find that users actively adjust the tone of their subsequent prompts to align with the tone embedded in AI-generated responses. Specifically, a one-standard-deviation increase in the tone difference between the AI response and the user's initial prompt is associated with an increase in user tone change equal to approximately 200% of its sample

mean, highlighting the role of AI in shaping human belief updating. This result remains robust after controlling for firm characteristics as well as a rich set of features of both user prompts and AI responses, including text length and readability.

We push further and attempt to understand how AI shapes human beliefs. AI may affect belief dispersion through two channels: by transmitting informative signals or by introducing noise into investors' beliefs. Although the noise channel cannot be completely ruled out, our evidence is more consistent with an information-based mechanism and Bayesian belief updating. Specifically, we find that the average tone measure derived from AI responses significantly predicts future stock returns, whereas tone measures derived from human prompts do not. This contrast suggests that AI answers convey incremental information beyond users' prior beliefs, rather than merely amplifying existing sentiment. Moreover, conditional on a given topic, tone in AI-generated content converges across responses, indicating greater signal precision and going against hallucination-driven noise as the primary driver of the increased belief dispersion. These results are generally consistent with the findings in prior studies that AI tools help investors process information and contribute to the price discovery (Cheng et al., 2025; Chang et al., 2023; Bertomeu et al., 2025).

So far, we have established two important facts: (1) users' belief dispersion increases after interacting with AI, and this increase is driven mainly by between-topic dispersion; (2) AI answers contain useful information. At first glance, these two facts appear contradictory, as information-based theories typically predict that more informative signals can help reduce dispersion. Thus, it is still unclear what drives the belief dispersion, particularly the between-topic component. Motivated by behavioral attention theories (see the survey by Gabaix, 2019), we exploit the granularity of our granular data and examine users' topic search behaviors throughout the human-AI interaction process. We find a striking persistence in users' topic choices across conversation rounds. Among more than ten distinct topics, a topic raised by a given investor in an earlier round is ten times more likely to be queried again in the next round than alternative topics. One potential explanation for this stickiness is that continuing within the same topic yields additional

information gains. However, we find this is not the case in the data. For instance, comparing the return predictability of AI answers in the last and second-to-last rounds reveals no incremental predictive power. Consequently, the persistence in user prompts is unlikely to be driven by rational information acquisition and instead appears more consistent with behavioral factors shaping investors' information search. Specifically, such topic search persistence is consistent with the theory in Nimark and Sundaresan (2019), who show that economic agents optimally select confirmatory information sources and may reduce learning about alternative topics as confidence rises, leading to persistent belief polarization even among fully rational investors. Taking the aforementioned results together, we confirm that the increased dispersion (particularly the between-topic dispersion) is driven by users' persistence in topic research.

Finally, we extend our study and examine whether investors become more or less dispersed in the post-ChatGPT period when processing firm-level information. To do so, we focus on the U.S. market, which is an important out-of-sample validation. To mitigate confounding factors, we zoom in on a salient information event—earnings announcements—during which investors and information intermediaries, such as financial analysts, rely on AI tools to process newly released public information. Investor belief dispersion is therefore likely to change sharply around these events. Employing a difference-in-differences design, we find that following the deployment of ChatGPT, both analyst forecast dispersion and investor belief dispersion increase more in the post-announcement window relative to the pre-announcement period. At the same time, trading volume and stock return volatility rise significantly. Further decomposing dispersion into difference-of-opinion (DO) and information asymmetry (IA) components by the machine learning methods in Liu et al. (2025), we show that the increase in dispersion is primarily driven by higher DO, while information asymmetry is alleviated. These findings suggest that AI adoption amplifies heterogeneity in interpretation rather than exacerbating information frictions. This evidence is also consistent with findings in Cheng et al. (2025), who document that ChatGPT outages are associated with reduced trading activity and that GenAI-assisted trading improves informational efficiency. Overall, the evidence from the

introduction of ChatGPT is broadly consistent with our earlier findings using granular user–AI interaction data.

In summary, our analysis shows that the use of GenAI tools increases dispersion in investors' beliefs. After receiving informative responses from AI, users obtain more precise signals regarding the specific topics they initially focus on. At the same time, investors exhibit strong persistence in topic search and are unlikely to switch to alternative topics. As a result, although each investor's signal becomes more precise, differences in the signals emphasized across investors lead to higher overall dispersion. Put differently, conditional on a common information set, investors focus on distinct dimensions of that information, generating greater differences of opinion. Within a given topic, however, AI delivers more precise information, thereby reducing information asymmetry. This mechanism reconciles rising belief dispersion (associated with higher trading volume) with improved informational efficiency documented in prior studies (Cheng et al., 2025).

2 Related Literature

Our paper is related to two main strands of the literature. The first examines how investor disagreement arises from heterogeneous information, priors, and frictions in information processing, and how such disagreement translates into trading and prices. The second strand examines how Artificial Intelligence (AI) and automation reshape investors' information environment by changing the costs and technologies of information acquisition and processing.

2.1 Formation of Disagreement

Extensive theoretical research studies the origins of investor disagreement. Classic models show that even with identical public news, investors may hold heterogeneous priors, which can generate trading volume and price dispersion (e.g., Harris and Raviv (1993); Bamber et al. (1997)). Disagreement is further amplified by heterogeneous interpretation, where agents map the same signal into different valuations because they rely on

different mental models to interpret the same public signal (Kandel and Pearson, 1995). Kondor (2012) show that public signals may reduce disagreement about fundamentals while increasing disagreement about prices, because traders use public information to infer others' private signals, amplifying higher-order belief dispersion and trading activity. Beyond higher-order beliefs, disagreement can also be driven by endogenous information acquisition. Social networks and echo chambers can further exacerbate belief dispersion, as investors self-segregate into like-minded communities that reinforce existing views and limit exposure to opposing information (Cookson et al., 2023).

A core mechanism is information frictions in disclosure processing. Even when information is formally public, investors may still hold incomplete information sets because acquiring and interpreting disclosures is costly (Grossman and Stiglitz, 1980), and limited processing capacity forces them to allocate scarce attention and effort unevenly across available signals (Sims, 2003; Blankespoor et al., 2020). Recent empirical work using social-media investor data finds that disagreement reflects both differences in investors' information sets and differences in interpretation; however, information heterogeneity is more strongly associated with trading volume (Cookson and Niessner, 2020). Antweiler and Frank (2004) find that message-board tone predicts volatility, and that disagreement across posts is linked to higher trading volume.

Our study complements this literature by examining how new technologies – specifically generative AI – shape investors' beliefs and affect belief dispersion. We investigate whether AI tools move investors toward a common information set or instead interact with existing frictions in ways that preserve or even exacerbate belief heterogeneity. Distinct from prior studies, we document that while GenAI delivers more precise information, it simultaneously widens between-topic dispersion due to the idiosyncratic nature of user queries.

2.2 AI, Investor Behavior and Financial Market

Emerging research studies how AI reshapes investor information acquisition and processing, and how these changes translate into trading behavior and market efficiency (Cao et al., 2025). Existing studies show that LLMs possess strong capabilities in processing complex economic and financial information, including extracting structured signals from unstructured corporate disclosures (Shaffer and Wang, 2024; Lu et al., 2025), identifying economically meaningful patterns and causal relationships across diverse textual sources such as analyst reports, earnings calls, and employee reviews (Lv, 2024; Li et al., 2026), and interpreting policy-related information and stance from central bank communications more accurately than traditional natural language processing methods (Hansen et al., 2024).

A central theme is that generative AI lowers the cost of extracting signals from complex disclosures and can therefore reduce information frictions. For example, Chang et al. (2023) show that after the diffusion of ChatGPT, retail trading becomes more aligned with AI-extracted earnings-call sentiment, especially for firms with high processing costs, and that improved information diffusion reduces information asymmetry. Related evidence at the portfolio choice stage suggests that algorithmic tools can attenuate behavioral frictions: investors adopting robo-advisors exhibit better diversification and weaker tendencies toward the disposition effect and return chasing (e.g., D'Acunto et al., 2019; Rossi and Utkus, 2024).

Consistent with these mechanisms, recent surveys, quasi-natural experiments, and empirical evidence indicate that AI is already embedded in capital market workflows. Using survey and interview evidence, Christ et al. (2025) document that sell-side equity analysts have begun adopting AI tools in their research processes. Cheng et al. (2025) exploit exogenous ChatGPT outages and find that trading activity declines when AI becomes unavailable, with stronger effects around corporate news and among firms held by more transient institutions, implying that AI facilitates information-driven trading. Bertomeu et al. (2025) leverage the unexpected ChatGPT ban in Italy and document

deteriorations in analyst information production and forecast performance, together with weaker market information efficiency, consistent with AI functioning as a general purpose information processing technology.

At the same time, AI can also generate forces that may amplify disagreement or distort price discovery. Reinforcement-learning trading agents may endogenously learn collusive-like strategies that impair liquidity and price efficiency, raising concerns that algorithmic interaction can produce correlated yet distortive trading behavior (Dou et al., 2025). Winder et al. (2025) show that large language models exhibit similar echo chambers as human investors, reinforcing existing investment biases inherent in their training data. Glickman and Sharot (2025) use experiment to discover that human–AI interactions can create a feedback loop in which AI amplifies users' pre-existing biases and users internalize the amplified signals, causing small distortions to snowball over time. Moreover, AI delivery is increasingly personalized by tailoring interpretations and recommendations to user-specific inputs, personalization can interact with prior beliefs and search patterns to fragment information diets and intensify polarization (Acemoglu et al., 2025).

While AI holds the potential to improve information efficiency and reconcile divergent beliefs, these effects are strictly constrained by the model's factual reliability, which remains a primary impediment in current LLM architectures. Recent surveys characterize LLM hallucinations as a systematic factuality deficit, where models generate fluent yet unsupported or inconsistent statements (Zhang et al., 2025; Huang et al., 2025). This "distorted information" extends beyond blatant fabrication to subtle shifts that undermine the perceived reliability of a claim (Sun et al., 2024). The stakes are particularly high in expert domains like finance, where precision in numerical data and time-sensitive institutional facts is paramount; here, minor inaccuracies can lead to catastrophic decision-making. Wang (2025) shows that LLMs frequently hallucinate accounting facts—both by deviating from existing financials and by fabricating nonexistent items, implying that AI can shift costs from processing to verification.

Our analysis sits at the intersection of these perspectives by examining whether, in

practice, chatbot usage is associated with a net reduction in investor disagreement or whether AI instead reinforces segmentation in beliefs across investors. Leveraging granular dialogue data, we provide direct evidence that investors update beliefs in response to AI answers, and mechanism tests point to an information channel, as AI responses reliably predict subsequent returns, although the AI response may contain noise.

3 Data and Variable Construction

3.1 Data Sources

Our primary dataset consists of anonymized AI-human conversation logs collected from one of the largest Chinese generative AI platforms during the period from September to December 2024. Each observation corresponds to a conversation round, i.e., one user prompt and the GenAI's response within an ongoing conversation. The raw logs contain (i) the prompt and answer texts, (ii) conversation identifiers that allow us to reconstruct multi-round interactions, and (iii) system metadata (e.g., model and tool types, and status flags). To connect conversational behavior to capital market settings, we match conversations to publicly listed firms using a firm identifier field available in the platform logs.

We begin with raw conversation logs and apply a rigorous sequence of cleaning and screening filters to construct a sample that facilitates the study of how users acquire firm-specific information. Our sample integrates two complementary datasets provided by the platform. Our main sample is the multi-round conversation data that containing multiple rounds human prompt–AI response interactions. This multi-round conversation dataset captures complete, multi-round dialogues referencing listed firms, allowing us to trace the intra-conversation dynamics of how human changes the prompts and adjusts according to AI answers. The second dataset is the single-round conversation dataset that recording conversation with only one round firm-related prompt–response, primarily supporting large-scale cross-sectional analyses and robustness checks.[3] Across both

[3]We compare the tone in single-round conversations with the tone in the first round of multi-round

sources, we implement a consistent filtering process to ensure the sample captures **capital market oriented information acquisition**, we exclude several types of observations: (i) interactions flagged with generation errors (e.g., `status = "error"`); (ii) multimodal interactions involving images, which preclude reliable text-based sentiment measurement; (iii) text-processing tasks—such as summarization, translation, and polishing—as these focus on language editing rather than information seeking and may mechanically bias sentiment scores; and (iv) conversations unrelated to capital markets, even if a firm name is mentioned. Specifically, we remove requests concerning product usage, customer support (e.g., account setup or troubleshooting), and general corporate inquiries (e.g., recruitment, office locations, or internal operations). This filtering process ensures the final sample isolates investment-related inquiries from consumer-oriented or administrative interactions. Appendix B provides anonymized, representative examples of conversations from our sample.

3.2 Basic Facts

3.2.1 Conversation Structure

The dataset is structured such that each observation represents a single query-response pair. Multiple observations may originate from the same conversation, allowing for the analysis of multi-turn dynamics. As illustrated in Figure 1, the distribution of conversation depth is heavily skewed toward short-horizon interactions. Two-round exchanges (one user query and one AI response) constitute the vast majority of the sample, accounting for 74.4% of all observations. This is followed by three-round conversations at 16.6%. Beyond this point, the density diminishes monotonically; the remaining probability mass is distributed across longer threads, which represent a marginal fraction of the total volume. This concentration suggests that most user intents are either addressed or abandoned within the first two turns. Complementing the conversation-level data, Figure 2 illustrates the distribution of unique users per firm. The pattern is strongly

conversations. As shown in Appendix Table A2, we find no statistically significant differences in either human or AI tone between the two settings.

right-skewed and mirrors the concentration observed in conversation depth. Most firms (76.5%) are associated with 0–10 unique users. The share declines sharply as the user count increases: 12.1% of firms involve 11–20 users and 4.5% fall in the 21–30 range. The distribution also features a long tail, with 5.5% of firms discussed by 41 or more unique users, indicating that user coverage varies substantially across firms.

3.2.2 Topic Composition

We classify each user query into economically meaningful topic categories using a hybrid procedure that combines rule-based seeding with supervised machine learning. First, we develop a topic taxonomy and a set of topic-specific seed keywords and pattern rules to generate high-recall candidate labels. Second, we draw a stratified sample of queries and manually verify and refine the preliminary labels, producing a clean labeled set for training. Third, we fine-tune a pretrained BERT model for multi-class sequence classification on this human-labeled data, and use the fine-tuned classifier to assign final topic labels to the full sample.[4] Figure 3 shows that user questions are concentrated in a small set of categories. Price Trends is the largest category (21.8%), followed by ESG and Governance Risk (20.4%). Together, these two categories account for 42.2% of all topic assignments. Several additional topics each represent nontrivial shares, including Trading Behavior (9.6%), Business Efficiency (8.7%), Macroeconomic Indicators (8.6%), Profitability (8.4%), and Balance Sheet (7.6%). The remaining categories appear less frequently, such as Financial Indicators (5.1%), Shareholder Payouts (3.2%), Trading Structure (3.0%), Market Sentiment (2.1%), and Valuation Metrics (1.5%). This distribution suggests that a large fraction of queries focus on market movements and risk assessment, while a sizable subset centers on fundamentals and macro conditions. One potential explanation is that users rely on conversational systems for rapid triage of firm-specific information. Topics such as price dynamics and ESG or governance risk are often

[4]The classification threshold is selected on the validation set to maximize macro F1. At the selected threshold of 0.45, the classifier achieves an accuracy of 96.5%, a macro F1 score of 95.5%, and a macro ROC-AUC of 98.9% on the validation set. The corresponding test-set performance is similar, with an accuracy of 96.6%, a macro F1 score of 95.9%, and a macro ROC-AUC of 99.2%. The comparable performance across the validation and test sets indicates stable out-of-sample classification performance.

linked to contemporaneous events and dispersed disclosures, which may generate demand for quick summarization and synthesis. In contrast, valuation and trading-structure questions may require more structured inputs and specialized terminology, which could limit their prevalence in natural-language queries. We treat these interpretations as descriptive and use them to motivate our subsequent analyses of content variation and dispersion across topic environments.

3.3 Tone and Dispersion Variables

3.3.1 Sentiment

We quantify tone using a finance-domain sentiment model and construct a continuous tone index from the predicted sentiment shares (Huang et al., 2023). This approach is well suited to our setting because FinBERT is designed for single-sentence classification, and the sentence is a natural unit for expressing sentiment. This makes the model particularly suitable for measuring the tone of short user prompts. For each text, we define

$$\text{Tone} = s_{\text{pos}} - s_{\text{neg}}. \tag{1}$$

where s_{pos} and s_{neg} denote the model-implied positive and negative sentiment shares, respectively. Higher values indicate a more positive tone, and we compute this measure consistently for both user queries and AI responses at the conversation–round level.

We also construct a set of *text-feature variables* to characterize both the user's prompt and the GAI's response in terms of content volume and linguistic accessibility. For each prompt and answer, we measure *Length* as the total number of words in the corresponding text. We also compute *Readability* to capture linguistic complexity, defined as the proportion of words that appear in a high-frequency dictionary; higher values indicate simpler, more readable text. Accordingly, we obtain $Length^{Human}$ and $Length^{GAI}$, as well as $Readability^{Human}$ and $Readability^{GAI}$, for every conversation round.

3.3.2 Disagreement

We measure belief disagreement in AI–human interactions by constructing tone dispersion at the conversation–round level and decomposing it into within-topic and between-topic components. For each conversation round r, we first compute the population variance of tone across all observations within the round. We then apply a variance decomposition based on the law of total variance, which expresses total dispersion as the sum of two components. We formulate the decomposition in variance terms, that is, as sums of squared deviations, because variance is additively decomposable and therefore provides a natural way to separate overall disagreement into within-topic and between-topic components.

The within-topic dispersion, $Dispersion_{i,t,r}^{Within}$, captures disagreement among observations discussing the same information topic within a given round. It is computed as the weighted average of topic-specific variances, where the weights reflect the relative frequency of each topic in that round. In contrast, the between-topic dispersion, $Dispersion_{i,t,r}^{Between}$, measures heterogeneity across information topics and is defined as the weighted variance of topic-level mean tones within the same round. Although the decomposition is derived in squared units, in the regression analysis we use the corresponding standard deviation measure rather than the variance itself. This transformation preserves the underlying dispersion ranking while expressing the variable in the same units as tone, which makes the magnitude of the dependent variable and the estimated coefficients more directly interpretable. In addition, standard deviation is less mechanically sensitive to extreme values than squared dispersion and is more comparable across specifications when discussing economic magnitude.

$$TotalDispersion_{i,t,r} = \underbrace{\sum_{k \in \mathcal{K}_r} \omega_{kr} \operatorname{Var}(\text{Tone} \mid k, r)}_{Dispersion_{i,t,r}^{Within}} + \underbrace{\sum_{k \in \mathcal{K}_r} \omega_{kr} \left(\mu_{kr} - \mu_r\right)^2}_{Dispersion_{i,t,r}^{Between}}, \tag{2}$$

where $\omega_{kr} = n_{kr}/N_r$ is the share of observations belonging to topic k in round r, μ_{kr} denotes the topic-level mean tone, and μ_r is the round-level mean tone. For empirical implementation, we then take the square root of each variance-based dispersion com-

ponent so that the regression outcome is a standard deviation measure rather than a variance measure. In the subsequent variable construction, we multiply the dispersion measure by 100 in regression analyses.

3.4 U.S. Data

Our U.S. equity sample consists of common stocks listed on the NYSE, AMEX, and NASDAQ, excluding firms in the financial and utilities sectors. To mitigate the influence of small and illiquid stocks, we further exclude stocks with prices below $5 per share. Stock level return and trading data are obtained from the Center for Research in Security Prices (CRSP), while accounting information is drawn from Compustat. The main sample spans January 2021 through December 2024. For the earnings-announcement analysis, we require complete data availability, with at least two months of observations both before and after each announcement window.

3.5 Summary Statistics

Table 2 reports summary statistics for the variables used in the empirical analysis. Panel A summarizes the core outcomes in the dispersion panel, including total disagreement and its decomposition into within-topic and across-topic components, together with a noise proxy. Panel A also presents text-feature measures constructed from both user prompts and GAI responses. Specifically, we quantify content volume using length and linguistic accessibility using readability, and we further report the cross-sectional dispersion of these text features across prompts and across answers. In addition, Panel A includes round-level tone measures for users and the GAI, along with prompt- and answer-level text characteristics at the conversation-round level. Panel B reports firm–date–conversation variables that capture conversation-level dynamics, including tone changes across rounds, the initial tone gap between the GAI and the user, and conversation-level text features of the GAI response. Panel C reports standard firm–date controls commonly used in asset-pricing and market-microstructure settings, together with firm-date news sentiment controls con-

structed from print and web news. Panel D reports benchmark variables from the U.S. equity market that characterize the broader information environment and market conditions used in the analysis. Our main investor-disagreement measure, $Dispersion_{i,t,r}$, has a mean of 10.392 at the firm–date–round level, with substantial cross-sectional and time-series variation (SD = 17.778). Decomposing Dispersion further, the $Dispersion_{i,t,r}^{Between}$ (mean = 8.228) is markedly larger than the $Dispersion_{i,t,r}^{Within}$ (mean = 3.174), indicating that overall disagreement is driven primarily by across-topic differences rather than the within-topic divergences.

4 Baseline Results

This section examines how investor disagreement evolves with continued interactions with AI and explores the underlying mechanisms driving these changes. Section 4.1 documents whether user disagreement increases or decreases across conversation rounds, using both univariate comparisons and panel regressions, and further decomposes overall dispersion into within-topic and between-topic components to identify the source of the change. Section 4.2 then studies the role of AI in belief updating by examining whether and how users adjust their expressed tones in response to the tone embedded in AI-generated answers during the interaction. Section 4.3 investigates whether the observed changes in disagreement reflect information transmission or noise introduction, by testing the return predictability of AI tones relative to human tones and by examining how information precision in AI answers evolves across interaction rounds.

4.1 Disagreement Increases or Decreases?

Table 3 reports univariate difference-in-means tests examining how user tone dispersion evolves across query rounds in conversations. The dependent variable, $Dispersion_{i,t,r}$, is measured at the firm–date–round level. The results show that user dispersion increases as conversations progress. Average dispersion rises from 9.581 in Round 1 to 10.688 in Round 2, implying a statistically significant increase of 1.107 (s.e. = 0.395). A similar

pattern holds when comparing the first and the last query rounds. Dispersion increases from 9.581 in Round 1 to 10.905 in the last round, corresponding to an increase of 1.324 (s.e. = 0.396), which is also statistically significant at the 1% level. These univariate results suggest that repeated interactions with AI are associated with higher user tone dispersion across rounds.

To mitigate potential endogeneity concerns, we next estimate panel regressions at the firm–date–round level. Our baseline specification is given by

$$Dispersion_{i,t,r} = \alpha + \beta D_{i,t,r} + \gamma X_{i,t,r} + \mu_i + \lambda_t + \varepsilon_{i,t,r}, \tag{3}$$

where $Dispersion_{i,t,r}$ measures user disagreement for firm i on date t at query round r, constructed identically to Table 3. The indicator variable $D_{i,t,r}$ captures the progression of AI-human interactions across query rounds. $D_{i,t,r}$ is defined as either $D^{Round2}_{i,t,r}$, which equals one for second-round queries and zero for first-round queries, or $D^{Final}_{i,t,r}$, which equals one for last-round queries and zero for first-round queries within the same conversation.

The vector $X_{i,t,r}$ includes, among others, controls for prompt-level characteristics such as the standard deviation of prompt length ($LengthSD^{Human}_{i,t,r}$) and prompt readability ($ReadabilitySD^{Human}_{i,t,r}$), as well as lagged firm characteristics, including stock return reversal, momentum, firm size, turnover, and the book-to-market ratio. Firm fixed effects μ_i absorb time-invariant heterogeneity across firms, while date fixed effects λ_t control for common shocks affecting all firms on the same day.

If later interactions are associated with greater divergence in user interpretations, we expect a positive association between the query-round indicators and overall dispersion. The coefficients on $D^{Round2}_{i,t,r}$ and $D^{Final}_{i,t,r}$ are expected to be positive, indicating higher user disagreement in later rounds relative to the first round.

Table 4 Columns (1)–(4) shows the panel regression results which are based on the regression specification in Equation 3. Across specifications, the results consistently indi-

cate that user disagreement increases with continued AI usage. In Columns (1) and (2), which are estimated on a firm–date–round level panel constructed from observations in the first and second query rounds, dispersion in the second round is significantly higher than in the first round, with an estimated increase of approximately 1.10. Columns (3) and (4), which analogously rely on a firm–date–round level panel constructed from observations in the first and the last query rounds within the same conversation, show that dispersion in the last round is also higher relative to the first round, with the estimated differences ranging from 1.30 to 1.32. These effects remain statistically significant after controlling for firm and date fixed effects as well as prompt-level characteristics.

Table 4 Columns (5)–(8) present results from a firm–date–conversation–round level specification that focuses on belief polarization during AI-human interactions. In these regressions, the dependent variable is the absolute value of human tone, $|Tone^{Human}_{i,t,c,r}|$, which serves as a proxy for belief polarization by capturing the extremeness of a user's expressed tone. This variable is measured at the prompt level, which corresponds to the firm–date–conversation–round level, as each conversation round belongs to a single firm and occurs within one day. Columns (6) and (8) further include user fixed effects. This specification helps absorb time-invariant user-specific factors, such as persistent language style, baseline communication habits, and the AI's memory with the same user. By controlling for the user fixed effects, the estimates are less likely to be driven by systematic differences across users and instead more cleanly capture within-user changes in expressed polarization across query rounds.

We estimate panel regressions analogous to the Equation 3, where the key explanatory variable indicates the progression of query rounds. $D^{Round2}_{i,t,c,r}$ equals zero for observations from the first round and one for observations from the second round; accordingly, Columns (5) and (6) are estimated using observations from the first and second rounds. Similarly, $D^{Final}_{i,t,c,r}$ equals zero for observations from the first round and one for observations from the last round within the same conversation. The regressions control for prompt-level characteristics, including prompt length and readability, as well as lagged firm characteristics such as return reversal, momentum, firm size, turnover, and the book-

to-market ratio. Firm and date fixed effects are included.

Complementary evidence from Columns (5)–(8) indicates that belief polarization intensifies as interactions progress. Both the second-round and last-round indicators enter positively and significantly, suggesting that repeated interactions with AI are associated with users expressing more extreme beliefs over the course of the conversation. Our empirical results complement experimental and theoretical models (Glickman and Sharot, 2025; Acemoglu et al., 2025) by showing that sentiment dispersion among humans increases with greater human–AI interaction.

4.1.1 Between-Topic versus Within-Topic

To further distinguish the sources of disagreement, we decompose overall dispersion into within-topic and between-topic components and estimate the following specification:

$$Dispersion^{k}_{i,t,r} = \alpha^{k} + \beta^{k} D^{Round2}_{i,t,r} + \gamma^{k} X_{i,t,r} + \mu_i + \lambda_t + \varepsilon^{k}_{i,t,r}, \tag{4}$$

where $k \in \{\text{Within}, \text{Between}\}$. $Dispersion^{Within}_{i,t,r}$ captures disagreement among users discussing the same topic, whereas $Dispersion^{Between}_{i,t,r}$ reflects disagreement across different topics within the same firm–date–round. This decomposition allows us to assess whether changes in disagreement across query rounds arise from increased divergence within topics or from dispersion across topic-level interpretations. In all specifications, standard errors are two-way clustered by firm and date. When overall dispersion is decomposed into within-topic and between-topic components, we expect the association to differ across dimensions. If continued interactions primarily broaden the range of topics or frames through which users engage with firm-related information, the positive association should be concentrated in between-topic dispersion. Accordingly, we expect a positive coefficient on the round indicator when $Dispersion^{Between}_{i,t,r}$ is the dependent variable, while the association with $Dispersion^{Within}_{i,t,r}$ is expected to be weak or statistically insignificant.

Table 5 decomposes user disagreement into within-topic and between-topic components to examine how disagreement evolves across query rounds. Columns (1) and (2) report results for $Dispersion_{i,t,r}^{Within}$, capturing disagreement among users discussing the same topic, while Columns (3) and (4) focus on $Dispersion_{i,t,r}^{Between}$, reflecting disagreement across different topics within the same firm–date–round.

The results reveal a clear contrast between within-topic and between-topic disagreement. The coefficient on the second-round indicator is small and statistically insignificant in Columns (1) and (2), suggesting that repeated interactions with AI do not significantly increase disagreement among users within the same topic. In contrast, Columns (3) and (4) show a large and statistically significant increase in between-topic dispersion in the second round. The estimated coefficients are economically meaningful and remain robust after controlling for firm and date fixed effects. These findings indicate that the overall increase in user disagreement documented earlier is primarily driven by divergence across topics rather than heightened disagreement within a given topic. This pattern suggests that later queries are associated with a greater dispersion across topics, indicating that users increasingly focus on different aspects of the same firm. As a result, interpretations of the firm's fundamentals become more dispersed across users as interactions with AI progress, while users discussing the same topic do not become systematically more diverged. It aligns with the notion that enhanced transparency can shift learning toward others' beliefs, amplifying disagreement even as fundamental information becomes easier to acquire (Banerjee et al., 2018).

4.2 The Role of GenAI

The preceding analyses show that user disagreement increases across query rounds, with the expansion driven primarily by dispersion across topics rather than within-topic disagreement. These aggregate patterns naturally raise the question of how much of this divergence reflects the influence of AI responses on individual users' belief updating. We therefore turn to a more granular analysis that examines the role of AI at firm–date–

conversation–round level, focusing on how exposure to the tone embedded in AI answers is associated with subsequent changes in human tone.

To quantify how AI answers are associated with subsequent updates in human tone within the same conversation, we estimate the following panel regression:

$$\Delta Tone^{Human}_{i,t,c,r=2} = \alpha + \beta\Big(Gap^{GAI-Human}_{i,t,c,r=1}\Big) + \Gamma' Z_{i,t,c,r=1} + \mu_i + \lambda_t + \varepsilon_{i,t,c}, \tag{5}$$

where the dependent variable is the change in human tone from Round 1 to Round 2,

$$\Delta Tone^{Human}_{i,t,c,r=2} = Tone^{Human}_{i,t,c,r=2} - Tone^{Human}_{i,t,c,r=1}, \tag{6}$$

measured at the firm–date–conversation level. The key independent variable is the AI-human tone difference in Round 1,

$$Gap^{GAI-Human}_{i,t,c,r=1} = Tone^{GAI}_{i,t,c,r=1} - Tone^{Human}_{i,t,c,r=1}. \tag{7}$$

The vector $Z_{i,t,c,r=1}$ includes query and response characteristics (e.g., $Length^{Human}$, $Readability^{Human}$, $Length^{GAI}$, and $Readability^{GAI}$) and lagged firm characteristics (e.g., $Reversal$, $Momentum$, $LnSize$, $Turn$, and BM). We incrementally add date fixed effects λ_t and firm fixed effects μ_i across columns. Standard errors are two-way clustered by firm and date. The coefficient β captures the association between exposure to the tone embedded in the Round 1 AI answer and subsequent changes in the user's tone. A positive β implies directional alignment: when the AI answer is more positive than the user's initial tone ($Gap^{GAI-Human}_{r=1} > 0$), the user's tone becomes more positive in Round 2 on average, and when the AI answer is more negative than the user's initial tone, the user's tone becomes more negative.

Table 6 provides direct evidence on how users adjust their expressed tone in response to the tone embedded in AI answers within the same conversation. Across all specifications, the coefficient on the AI-human tone difference is positive and highly statistically significant. This result indicates that users systematically adjust their tone in the direction

of the tone embedded in the AI response. The association remains stable in magnitude after controlling for prompt and response characteristics, lagged firm fundamentals, and fixed effects, suggesting that the observed tone updating is not driven by differences in query complexity, readability, or underlying firm conditions. These findings highlight the role of AI as an active reference point in the belief-updating process. Rather than merely reflecting user inputs, AI answers are associated with subsequent adjustments in human tone, implying that users partially align their expressed views with the tone conveyed by the model during continued interaction.

4.3 Mechanism Analysis: Information or Noise?

AI can lower information-processing costs and help users extract signals more quickly (e.g., Cheng et al., 2025). However, improved fluency and accessibility do not guarantee informativeness and accuracy. Prior work shows that LLMs can produce plausible but incorrect content (Huang et al., 2025), ranging from misstatements of reported figures to outright invention of nonexistent items (Wang, 2025). And AI-generated outputs may induce unwarranted confidence and lead users to overreact to low-value or uninformative cues (Messeri and Crockett, 2024; Croom, 2025). Consistent with this concern, we therefore ask whether the tone embedded in AI outputs contains return-relevant information that predicts future stock returns, or instead primarily reflects noise with limited predictive content for market outcomes.

4.3.1 Return Predictability of AI and Human Tones

To address this question, we examine whether the tone embedded in AI answers contains predictive information for subsequent stock returns. Specifically, we regress future stock returns over horizons of one to five days on the average tone of AI answers and, for comparison, the average tone expressed by users in the preceding period. All specifications include firm and date fixed effects to absorb time-invariant firm characteristics and common market-wide shocks, as well as standard return predictors and firm-date news

sentiment controls constructed from print and web news. We estimate the following panel regression:

$$Ret_{i,t,t+h} = \alpha + \beta_1 AveTone_{i,t-1}^{Human} + \beta_2 AveTone_{i,t-1}^{GAI} + \Gamma' X_{i,t-1} + \mu_i + \lambda_t + \varepsilon_{i,t}, \quad (8)$$

where $Ret_{i,t,t+h}$ denotes the cumulative stock return for firm i from day t to $t+h$. Table 7 reports the estimation results. Across all specifications, the coefficient on AI sentiment ($AveTone_{i,t-1}^{GAI}$) is positive and highly statistically significant. Notably, the magnitude of these coefficients increases monotonically with the return horizon, suggesting that the sentiment embedded in AI responses possesses persistent predictive power for subsequent stock returns. In contrast, the coefficients on human sentiment ($AveTone_{i,t-1}^{Human}$) are economically small and statistically insignificant, indicating that users' expressed sentiment lacks comparable forecasting ability. These findings imply that AI-generated tone is not merely noise; rather, it provides incremental information beyond contemporaneous user sentiment. Importantly, the return predictability of AI-generated tone remains significant after controlling for firm-date news sentiment from both print and web news, suggesting that AI answers capture value-relevant information beyond contemporaneous public news sentiment. Furthermore, the lack of predictability in the original user prompts suggests that AI does not simply restate investor priors. Instead, AI appears to integrate disparate signals to produce value-relevant insights that are subsequently incorporated into equity prices.

4.3.2 AI Answer and Information Precision

Earlier sections document that investor belief dispersion increases with continued GenAI interaction. One concern is that the increase in dispersion may be driven by hallucination-induced noise in AI-generated content. To distinguish between noise amplification and information supply, we examine whether AI-generated tone converges or diverges conditional on a given topic.

The empirical design closely parallels that in Table 5. Here, instead of investor be-

liefs, we focus on information noise in AI-generated answers, measured as within-topic dispersion in tone across the first and second interaction rounds and denoted by $Noise$. Specifically, we construct $Noise_{i,t,r}$ as the within-topic dispersion of AI-answer tone:

$$Noise_{i,t,r} = 100 \times \sqrt{\sum_{k \in \mathcal{K}_{i,t,r}} \omega_{i,t,k,r} \operatorname{Var}\left(Tone^{GAI}_{i,t,c,r} \mid k\right)}, \tag{9}$$

where $\mathcal{K}_{i,t,r}$ denotes the set of topics discussed for firm i on date t in round r, $\omega_{i,t,k,r} = n_{i,t,k,r}/N_{i,t,r}$ is the share of AI answers associated with topic k, and $Tone^{GAI}_{i,t,c,r}$ is the tone of the AI-generated answer in conversation c. Table 8 reports the estimation results. The dependent variable is $Noise_{i,t,r}$, measured at firm–date–round level. The key explanatory variable is $D^{Round2}_{i,t,r}$, which equals one for second-round queries and zero for first-round queries. Control variables include dispersion in answer length ($LengthSD^{GAI}_{i,t,r}$) and readability ($ReadabilitySD^{GAI}_{i,t,r}$) to account for mechanical heterogeneity in AI responses, as well as lagged firm characteristics.

Coefficients on $D^{Round2}_{i,t,r}$ are negative and highly statistically significant. The estimates indicate that generated answer noise declines by approximately 20% from the first to the second interaction round. This pattern holds when adding answer-level dispersion controls, lagged firm characteristics, and date and firm fixed effects.

The decline in $Noise_{i,t,r}$ provides direct evidence against hallucination-induced noise as the primary driver of the observed increase in investor belief dispersion. If hallucinations were dominant, AI-generated content would be expected to become more dispersed rather than more convergent across responses. Instead, conditional on a given topic, tone in AI-generated answers converges across interaction rounds, consistent with increasing signal precision.

5 Disagreement, Belief Polarization, and Stickiness

While we document the evolution of AI usage and tone dispersion, it remains unclear whether these dynamics accumulate into belief polarization. This section explores how

heterogeneity in the information content of AI-generated answers and users' persistent information search behavior jointly contribute to belief polarization. Section 5.1 examines whether the return-relevant informativeness of AI answers varies across conversation rounds. Section 5.2 then documents strong topic stickiness in user–AI interactions, showing that users are highly persistent in the topics they query across conversation rounds. Finally, Section 5.3 evaluates whether this persistence reflects rational information acquisition or behavioral search behavior by testing whether additional conversation rounds deliver incremental return-relevant information beyond earlier responses.

5.1 Heterogeneity of Return Predictions in Different Rounds

We next examine whether the return-relevant information content of AI answers varies across conversation rounds. This question is motivated by our earlier evidence that, conditional on a given topic, information noise in AI answers decline from round 1 to round 2 (Table 8) during the conversation, suggesting increasing signal precision in the first two rounds. If later-round answers are indeed more precise along the topic dimension on which a user focuses, then the tone embedded in AI responses should exhibit stronger return predictability in later rounds than in the first round. More broadly, potential heterogeneity in answer informativeness across rounds is central to understanding how AI shapes human's belief updating and subsequent investor behavior during multi-turn interactions.

Table 9 examines whether the informativeness of AI answers is stronger in later rounds (round 2 and the last round) relative to the first round. The specification follows Table 7 and keeps the same setting. The key difference across the three panels in Table 9 lies in the construction of the main explanatory variables. Specifically, the regressors are constructed using the average tone in (i) round 1, (ii) round 2, and (iii) the last round within a conversation, respectively.

Table 9 shows clear heterogeneity in informativeness across rounds. When average AI tone is measured using round 1 answers, the coefficients are small and statistically

insignificant, indicating little return predictability in the initial interaction. In contrast, AI tone measured from round 2 or the last round answers strongly predicts future stock returns across horizons. Later-round AI answers are substantially more informative than first-round answers. This pattern is consistent with the evidence in Table 8, which shows that information noise in AI-generated tone declines from round 1 to round 2. This result indicates that the heterogeneity in return predictability across conversation rounds is not driven by variation in human prompts, but instead reflects differences in the informational content embedded in AI-generated answers.

5.2 Topic Stickiness

We next document a striking feature of AI-human interactions: users exhibit strong persistence in what they ask about across conversation rounds. Conditional on discussing a given topic in the previous round, users are much more likely to continue querying the same topic in the next round than to switch to alternative topics.

Figure 4 visualizes topic persistence using heatmaps of topic transition probabilities across query rounds. In each panel, the vertical axis corresponds to the topic raised in the first round, and the horizontal axis corresponds to the topic queried in the destination round. Each cell reports the probability that a given topic appears in the destination round conditional on that topic appearing in the first round. Darker colors indicate higher probabilities. Panel A reports transitions from the first round to the second round, while Panel B reports transitions from the first round to the last round within a conversation. In both panels, the darkest regions of the heatmap are concentrated along the main diagonal, indicating strong within-topic persistence: conditional on a topic being discussed in the first round, the same topic is highly likely to reappear in subsequent rounds. Figure 5 extends this analysis to transitions across adjacent rounds throughout the entire conversation. The heatmap again exhibits pronounced diagonal dominance across round pairs.

This pronounced topic stickiness has important implications for belief dispersion.

When users repeatedly engage with the same topic, their interaction with AI concentrates information acquisition along a narrow, topic-specific dimension. As a result, AI can deliver precise information within topic which the user focuses on. However, because different users persistently focus on different topics, the set of signals emphasized across investors remains heterogeneous. Consequently, even as information precision improves within topics, differences in topic emphasis across users can translate into greater dispersion in beliefs at the aggregate level.

We next examine whether GAI-generated information shapes the evolution of users' topic choices during a conversation. This analysis speaks to the role of GAI in directing users' attention: if GAI answers resolve users' current information needs, users may be less likely to continue discussing the same topic in the next round. Alternatively, if GAI answers provide additional evidence that invites further inquiry, they may increase users' topic stickiness.

We construct a round-level indicator, denoted as $TopicStickiness_{i,t,c,r}$, to capture whether users continue discussing topics from the previous round.[5] We estimate the following specification:

$$\begin{aligned} TopicStickiness_{i,t,c,r} = \beta_1 InfoSource^{GAI}_{i,t,c,r-1} + \beta_2 HardInfo_{i,t,c,r-1} + \Gamma' Controls_{i,t,c,r-1} \\ + \alpha_i + \delta_t + \varepsilon_{i,t,c,r}. \end{aligned} \tag{10}$$

where $InfoSource^{GAI}_{i,t,c,r-1}$ captures the extent to which the GAI answer in the previous round provides explicit information sources. We use three alternative measures: an indicator for whether the answer provides any source, the proportion of source-providing sentences, and the number of sources provided. $HardInfo_{i,t,c,r-1}$ is an indicator equal to one if the previous-round topic is classified as hard-information intensive. $Controls$ include previous-round GAI answer length, previous-round GAI answer readability, and lagged firm characteristics. α_i and δ_t denote firm and date fixed effects, respectively.

[5] $TopicStickiness_{i,t,c,r}$ equals one if the topic set in the current round shares at least one topic with the topic set in the previous round within the same conversation, and zero otherwise. Thus, this measure captures whether the user continues at least one topic from the previous round.

Standard errors are two-way clustered by firm and date.

Table 10 reports the results. Across specifications, information sourcing by GAI is negatively associated with users' topic stickiness. In column (2), after including firm and date fixed effects, the coefficient on $Source^{GAI}$ is negative and statistically significant. This result indicates that when the previous GAI answer provides explicit sources, users are less likely to continue discussing overlapping topics in the next round. Column (3) shows a similar pattern when information sourcing is measured by the proportion of source-providing sentences. Column (4) restricts the sample to rounds in which the GAI answer provides at least one source and examines variation in the number of sources. The coefficient on $SourceCount^{GAI}$ remains negative and statistically significant, suggesting that, conditional on providing sources, a larger number of sources is also associated with weaker topic stickiness.

The results suggest that sourced GAI answers do not mechanically lock users into their existing topics of attention. Instead, information sourcing appears to facilitate topic transitions. One interpretation is that explicit sources help users resolve their information demand in the current topic, thereby reducing the need to continue asking about the same issue. Consistent with this interpretation, GAI answer length is negatively associated with topic stickiness in columns (2)–(4), and the association is statistically significant. This suggests that more detailed answers may provide sufficient information for users to complete one line of inquiry and move to other topics. GAI answer readability is also generally negatively associated with topic stickiness, especially in columns (1)–(3), suggesting that easier-to-process responses may reduce users' need to remain on the same topic.

The coefficient on $HardInfo_{i,t,c,r-1}$ is negative and significant across specifications. This pattern suggests that users are less likely to continue with overlapping topics after hard-information-intensive discussions. Because hard-information topics are more likely to rely on objective facts, figures, or verifiable evidence, they may be more readily resolved once relevant information is provided.

Taken together, the evidence indicates that GAI answer characteristics systematically shape users' subsequent topic choices. Rather than increasing topic stickiness, sourced, detailed, and readable GAI answers are associated with weaker persistence in users' discussion topics. This finding is consistent with the view that GAI functions not only as a responder to user queries, but also as an information organizer that helps users close one topic and redirect attention to other issues.

5.3 Rational Information Acquisition versus Behavioral Persistence

This strong topic stickiness naturally raises the question of whether continued interaction reflects rational information acquisition, namely whether users remain within a topic because successive interactions generate incremental information gains, or whether it instead reflects behavioral forces that shape investors' information search. In this section, we provide direct evidence on this question by examining whether AI-generated answers differ in their return-relevant information content across later conversation rounds, excluding the first round.

Table 11 examines whether the informativeness of second-round AI answers differs depending on whether a conversation ends after two rounds or continues for more than two rounds. We compare the return predictability of second-round AI answer mean tone in conversations with exactly two rounds ($R = 2$) to that in conversations with three or more rounds ($R \geq 3$). We construct two alternative measures of second-round AI tone: one based on conversations that end after two rounds ($AveTone^{AI,r=2} \mid R = 2$) and the other based on conversations that proceed to three or more rounds ($AveTone^{AI,r=2} \mid R \geq 3$). Across all return horizons, the coefficients on $AveTone^{AI,r=2}$ are positive and statistically significant under both constructions, indicating that second-round AI answers contain return-relevant information regardless whether the conversation continues. Coefficient difference tests also fail to reject equality between the two estimates at conventional significance levels across all horizons. This evidence suggests that the decision to continue the conversation beyond the second round is not associated with information content

already embedded in the second-round AI response.

Table 12 further investigates whether additional rounds generate incremental information within multi-round conversations by comparing the return predictability of AI tone in the second-last and last rounds of the same conversation. Both $AveTone^{AI,r=-2}$ and $AveTone^{AI,r=-1}$ significantly predict future stock returns in all horizons, with quantitatively similar coefficient magnitudes. Coefficient difference tests consistently fail to reject equality between the two estimates. These results indicate that additional rounds do not deliver systematically stronger return-relevant information than the immediately preceding round.

These results show little evidence of incremental information gains across non-initial conversation rounds. While AI-generated answers beyond the first round are informative about future returns, their informational content does not increase monotonically with continued interaction. This pattern suggests that users' persistent engagement with the same topic is unlikely to be driven by rational information acquisition aimed at extracting progressively richer signals. Instead, the absence of incremental informativeness across later rounds points to behavioral persistence in information search, whereby users continue querying within a familiar topic even after the primary informational content has already been revealed.

6 ChatGPT, Information Processing and Disagreement

In this section, we conduct one additional analysis to explore the potential effect of AI on investors' belief dispersions. As a prominent example of generative AI models, ChatGPT was launched in November 2022 and has since made advanced GenAI tools widely accessible. Several studies exploit exogenous shocks to ChatGPT usage to establish a causal link between GenAI tools and capital market trading activity and market efficiency. Cheng et al. (2025) document that a substantial number of investors use ChatGPT in ways that influence their trading decisions and market outcomes. Bertomeu et al. (2025) examine how the emergence of generative artificial intelligence reshapes the information environ-

ment in capital markets. Both studies find that GenAI plays a positive role in improving the market information environment and overall market efficiency.

Earnings announcements serve as one of the most important channels in which companies disclose material information. Public information would reduce uncertainty about fundamentals, however, investors still differ in how they process and interpret the same announcement. They may emphasize different components of the release and draw different implications for fundamentals, leading to heterogeneous belief updating and dispersion in expectations following the announcement.

We therefore examine whether the release of ChatGPT changes how investors process earnings-announcement information, as reflected in the post-announcement response of investor dispersion and related market outcomes. This analysis provides an external, market-level check on our micro evidence that GenAI can reduce within-topic frictions and information asymmetry while increasing overall disagreement.

We study a set of variables that capture disagreement and information asymmetry around earnings announcements. Our primary disagreement proxy is analyst forecast dispersion ($AnalystDispersion$), which directly measures cross-analyst heterogeneity in expectations and has been widely used as a benchmark for belief disagreement. We also examine market-based measures that reflect disagreement-driven trading and heterogeneous belief updating. Higher disagreement generally implies a larger fraction of investors taking offsetting positions, which tends to increase trading activity; we therefore include turnover ($Turn$) as a quantity-based proxy for disagreement. In addition, heterogeneous interpretations of earnings news can translate into greater dispersion in price impacts and trading demands, leading to higher idiosyncratic volatility ($IVOL$). We include bid–ask spreads ($Spread$) as a proxy for market-wide information asymmetry. We also use two machine-based measures from Liu et al. (2025): difference of opinion (DO) and information asymmetry (IA), which are constructed from textual signals and designed to capture disagreement and asymmetric information in the information environment. Following Liu et al. (2025), we use the orthogonalized versions of these measures, where DO and IA are

residualized with respect to each other to isolate their distinct components.

To focus on the earnings-announcement response, we construct event-window changes by comparing the post-announcement window to a pre-announcement benchmark. For each firm i and earnings-announcement month t, we define

$$\Delta Y_{i,t} = \overline{Y}_{i,post} - \overline{Y}_{i,pre}, \tag{11}$$

where Y denotes one of the six outcomes described above ($AnalystDispersion$, $Turn$, $IVOL$, $Spread$, DO, and IA). $\overline{Y}_{i,post}$ denotes the average level of Y over the announcement month and the subsequent two months, and $\overline{Y}_{i,pre}$ denotes the average over the two months preceding the announcement month. This construction isolates how the results move after the earnings information is released, relative to the firm's own recent baseline.

We then estimate the following specification:

$$\Delta Y_{i,t} = \alpha + \beta\, PostGPT_t + \Gamma' X_{i,t-1} + \mu_i + \varepsilon_{i,t}, \tag{12}$$

where $PostGPT_t$ equals one for December 2022 and later and zero otherwise, $X_{i,t-1}$ includes standard firm controls, and μ_i denotes firm fixed effects. Standard errors are clustered by firm and time and the sample spans from 2021 to 2024.

We focus on the earnings-announcement window for two reasons. First, earnings releases deliver concentrated, firm-specific information at a well-defined time, which reduces ambiguity in timing relative to diffuse news flows. Second, defining outcomes as within-firm pre–post changes helps mitigate concerns that would arise in a simple event-study design. Table 13 reports the main results. We find that, following the release of ChatGPT, earnings announcements are associated with larger changes in several disagreement-related outcomes. Analyst forecast dispersion, turnover, idiosyncratic volatility, and the machine-based difference-of-opinion measure all show significantly larger post-announcement increases relative to the pre-announcement period in the post-ChatGPT sample. This pattern indicates that earnings announcements induce

larger post–pre increases in overall investor disagreement once GenAI tools become widely available. In contrast, measures associated with information asymmetry move in the opposite direction. The post–announcement change in information asymmetry declines, and bid–ask spreads do not increase following earnings announcements in the post-ChatGPT period.

Market-level evidence is consistent with our micro findings. After ChatGPT's release, earnings announcements induce larger post–pre increases in overall investor disagreement. At the same time, information-asymmetry measures do not increase. This combination mirrors our micro evidence that AI usage raises total dispersion while reducing information frictions and information asymmetry.

7 Conclusion

This paper examines whether generative AI tools (e.g., ChatGPT-like chatbots) reduce or amplify investor belief dispersion. Using a novel, granular dataset of multi-round user–AI conversations from one of China's largest GAI service providers, linked to firm-level and market outcomes, we provide direct evidence on belief updating and the evolution of disagreement.

Our main finding is that AI usage *increases* dispersion in investors' expressed beliefs over the course of interaction. As conversations move from the initial to later rounds, user tone becomes more heterogeneous and more extreme, consistent with polarization rather than random noise. Decomposing dispersion further reveals that this increase is driven primarily by a rise in *between-topic* dispersion: investors diverge in *which* aspects of the firm they focus on (e.g., price trends, fundamentals, macro conditions), while *within-topic* dispersion slightly declines, suggesting that investors discussing the same topic become marginally more aligned.

We also document strong belief adjustment toward AI outputs. Users shift their subsequent tone in the direction of the tone embedded in AI answers, indicating that AI-generated content shapes belief updating. Mechanism tests point more strongly to an

information-based channel than to hallucination-driven noise. Specifically, tone extracted from AI answers predicts subsequent stock returns, whereas tone from user prompts does not, implying that AI responses convey incremental, value-relevant information beyond investors' priors. Conditional on topic, AI tone is relatively consistent across responses, consistent with improved signal precision rather than fabrication as the force behind rising dispersion.

We further find a striking persistence in users' topic choices across rounds: investors are substantially more likely to continue querying the same topic than to switch to alternative signals, even though later-round AI answers do not exhibit incremental return predictability relative to earlier rounds. This stickiness can explain the increased between-topic dispersion.

We also study the asset pricing implications. We zoom in on earnings announcements and use a difference-in-differences design around the deployment of ChatGPT in the U.S. market. We find that after the deployment of ChatGPT, analyst forecast dispersion and investor belief dispersion rise more in the post-announcement window, while decompositions show that the increase is primarily attributable to higher differences of opinion rather than greater information asymmetry.

Overall, our evidence reconciles two seemingly competing views of AI in capital markets. AI can improve information efficiency by delivering more precise signals *within* a given information dimension, yet it can increase belief dispersion by inducing heterogeneous attention allocation *across* dimensions. These findings highlight that improvements in information processing need not imply belief convergence, and they underscore the importance of understanding how AI shapes investors' attention, search persistence, and the resulting structure of disagreement in financial markets.

References

Acemoglu, D., Ozdaglar, A., and Siderius, J. (2025). Ai and social media: A political economy perspective. Technical report, National Bureau of Economic Research.

Antweiler, W. and Frank, M. Z. (2004). Is all that talk just noise? the information content of internet stock message boards. *The Journal of Finance*, 59(3):1259–1294.

Bamber, L. S., Barron, O. E., and Stober, T. L. (1997). Trading volume and different aspects of disagreement coincident with earnings announcements. *The Accounting Review*, pages 575–597.

Banerjee, S., Davis, J., and Gondhi, N. (2018). When transparency improves, must prices reflect fundamentals better? *The Review of Financial Studies*, 31(6):2377–2414.

Bertomeu, J., Lin, Y., Liu, Y., and Ni, Z. (2025). The impact of generative ai on information processing: Evidence from the ban of chatgpt in italy. *Journal of Accounting and Economics*, page 101782.

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

Cao, S., Chen, W., Ma, G., and Srinivasan, S. (2025). Exploring generative ai in accounting: Information production and consumption. *Available at SSRN 5677522*.

Chang, A., Dong, X., Martin, X., and Zhou, C. (2023). Ai democratization, return predictability, and trading inequality. *Available at SSRN 4543999*.

Cheng, Q., Lin, P., and Zhao, Y. (2025). Does generative ai facilitate investor trading? early evidence from chatgpt outages. *Journal of Accounting and Economics*, page 101821.

Christ, M. H., Kim, M., and Yip, M. A. (2025). Word on the street: Field evidence of artificial intelligence use in sell-side equity research. *Available at SSRN 5071836*.

Cookson, J. A., Engelberg, J. E., and Mullins, W. (2023). Echo chambers. *The Review of Financial Studies*, 36(2):450–500.

Cookson, J. A. and Niessner, M. (2020). Why don't we agree? evidence from a social network of investors. *The Journal of Finance*, 75(1):173–228.

Croom, J. (2025). Interactivity and illusions of ability: How using generative ai affects investor judgments. *Journal of Accounting Research*.

Dou, W. W., Goldstein, I., and Ji, Y. (2025). Ai-powered trading, algorithmic collusion, and price efficiency. *Jacobs Levy Equity Management Center for Quantitative Financial Research Paper, The Wharton School Research Paper*.

D'Acunto, F., Prabhala, N., and Rossi, A. G. (2019). The promises and pitfalls of robo-advising. *The Review of Financial Studies*, 32(5):1983–2020.

Even-Tov, O., Lourie, B., Nekrasov, A., et al. (2025). The effect of ai on retail investor behavior: Evidence from account-level trading data. *Available at SSRN 5778246*.

Gabaix, X. (2019). Behavioral inattention. *NBER Working Paper*.

Glickman, M. and Sharot, T. (2025). How human–ai feedback loops alter human perceptual, emotional and social judgements. *Nature Human Behaviour*, 9(2):345–359.

Grossman, S. J. and Stiglitz, J. E. (1980). On the impossibility of informationally efficient markets. *The American Economic Review*, 70(3):393–408.

Hansen, A. L., Horton, J. J., Kazinnik, S., Puzzello, D., and Zarifhonarvar, A. (2024). Simulating the survey of professional forecasters. *Available at SSRN*.

Harris, M. and Raviv, A. (1993). Differences of opinion make a horse race. *The Review of Financial Studies*, 6(3):473–506.

Huang, A. H., Wang, H., and Yang, Y. (2023). Finbert: A large language model for extracting information from financial text. *Contemporary Accounting Research*, 40(2):806–841.

Huang, L., Yu, W., Ma, W., Zhong, W., Feng, Z., Wang, H., Chen, Q., Peng, W., Feng, X., Qin, B., et al. (2025). A survey on hallucination in large language models: Principles, taxonomy, challenges, and open questions. *ACM Transactions on Information Systems*, 43(2):1–55.

Kandel, E. and Pearson, N. D. (1995). Differential interpretation of public signals and trade in speculative markets. *Journal of Political Economy*, 103(4):831–872.

Kondor, P. (2012). The more we know about the fundamental, the less we agree on the price. *The Review of Economic Studies*, 79(3):1175–1207.

Li, K., Mai, F., Shen, R., Yang, C., and Zhang, T. (2026). Dissecting corporate culture using generative ai. *The Review of financial studies*, 39(1):253–296.

Liu, Y., Guo, K., and Wang, T. (2025). Different opinion or information asymmetry: Machine-based measure and consequences. *Working Paper*.

Lu, S., Riedl, E., Xu, S., and Serafeim, G. (2025). Climate solutions, transition risk, and stock returns. *FEB-RN Research Paper*, pages 25–024.

Lv, L. (2024). The value of information from sell-side analysts. *arXiv preprint arXiv:2411.13813*.

Messeri, L. and Crockett, M. J. (2024). Artificial intelligence and illusions of understanding in scientific research. *Nature*, 627(8002):49–58.

Nimark, K. P. and Sundaresan, S. (2019). Inattention and belief polarization. *Journal of Economic Theory*, 180:203–228.

Rossi, A. G. and Utkus, S. (2024). The diversification and welfare effects of robo-advising. *Journal of Financial Economics*, 157:103869.

Shaffer, M. and Wang, C. C. (2024). Scaling core earnings measurement with large language models. *Available at SSRN*.

Sheng, J., Sun, Z., Yang, B., and Zhang, A. L. (2026). Generative ai and asset management. *The Review of Financial Studies*, forthcoming.

Sims, C. A. (2003). Implications of rational inattention. *Journal of Monetary Economics*, 50(3):665–690.

Sun, Y., Sheng, D., Zhou, Z., and Wu, Y. (2024). Ai hallucination: towards a comprehensive classification of distorted information in artificial intelligence-generated content. *Humanities and Social Sciences Communications*, 11(1):1–14.

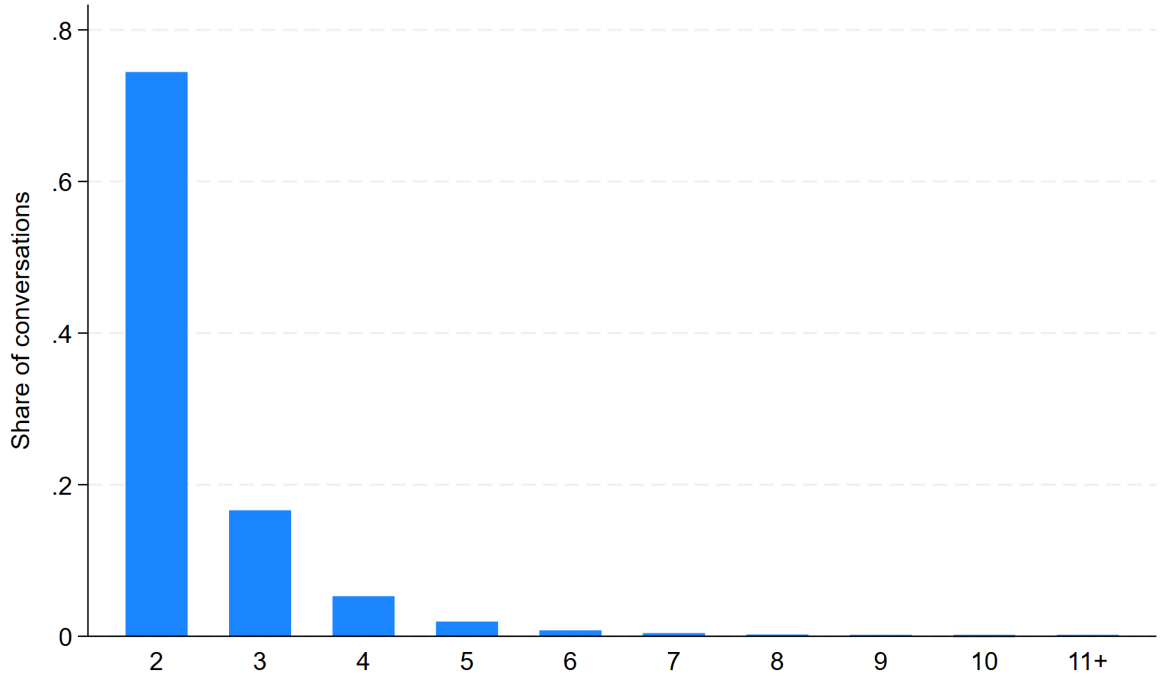
Wang, A. D. (2025). F(r)iction in machines: Accounting hallucinations of large language models. *Available at SSRN 5629978*.

Winder, P., Hildebrand, C., and Hartmann, J. (2025). Biased echoes: Large language models reinforce investment biases and increase portfolio risks of private investors. *PLOS ONE*, 20(6):1–24.

Zhang, Y., Li, Y., Cui, L., Cai, D., Liu, L., Fu, T., Huang, X., Zhao, E., Zhang, Y., Chen, Y., et al. (2025). Siren's song in the ai ocean: A survey on hallucination in large language models. *Computational Linguistics*, pages 1–46.

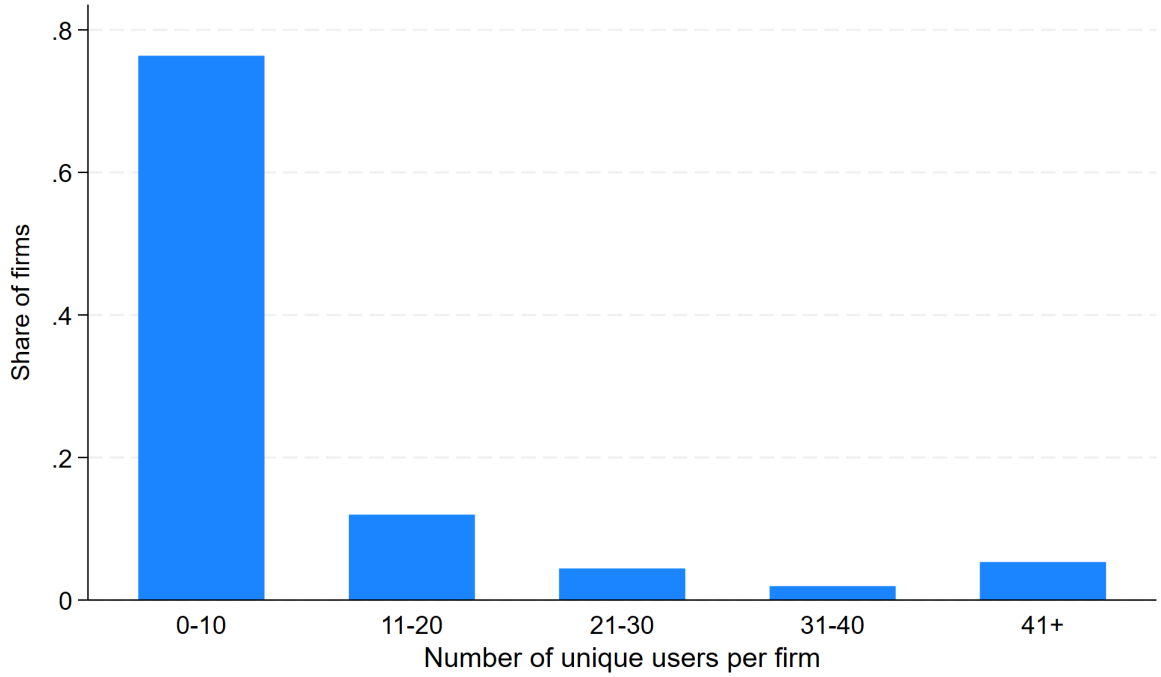
Tables and Figures

Figure 1: Distribution of Total Rounds



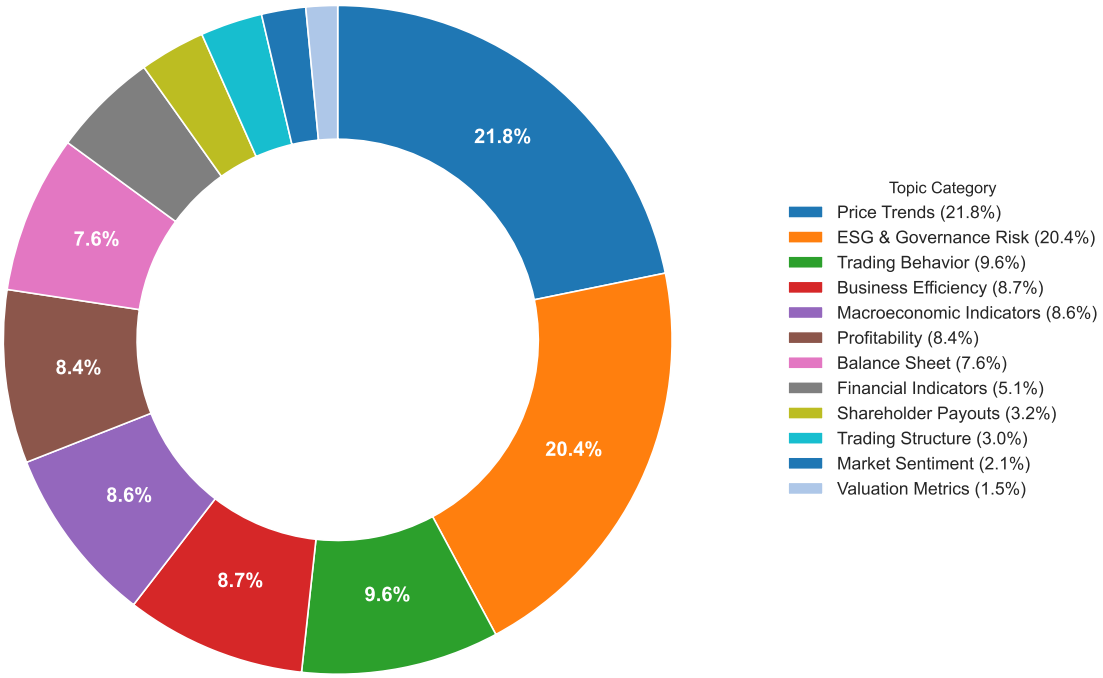
Notes: This figure plots the distribution of the total number of rounds in multi-turn conversations. The x-axis reports the conversation length (total rounds), and the y-axis shows the share of observations in each length bin. The "11+" bin aggregates conversations with 11 or more rounds.

Figure 2: Distribution of Unique Users per Firm



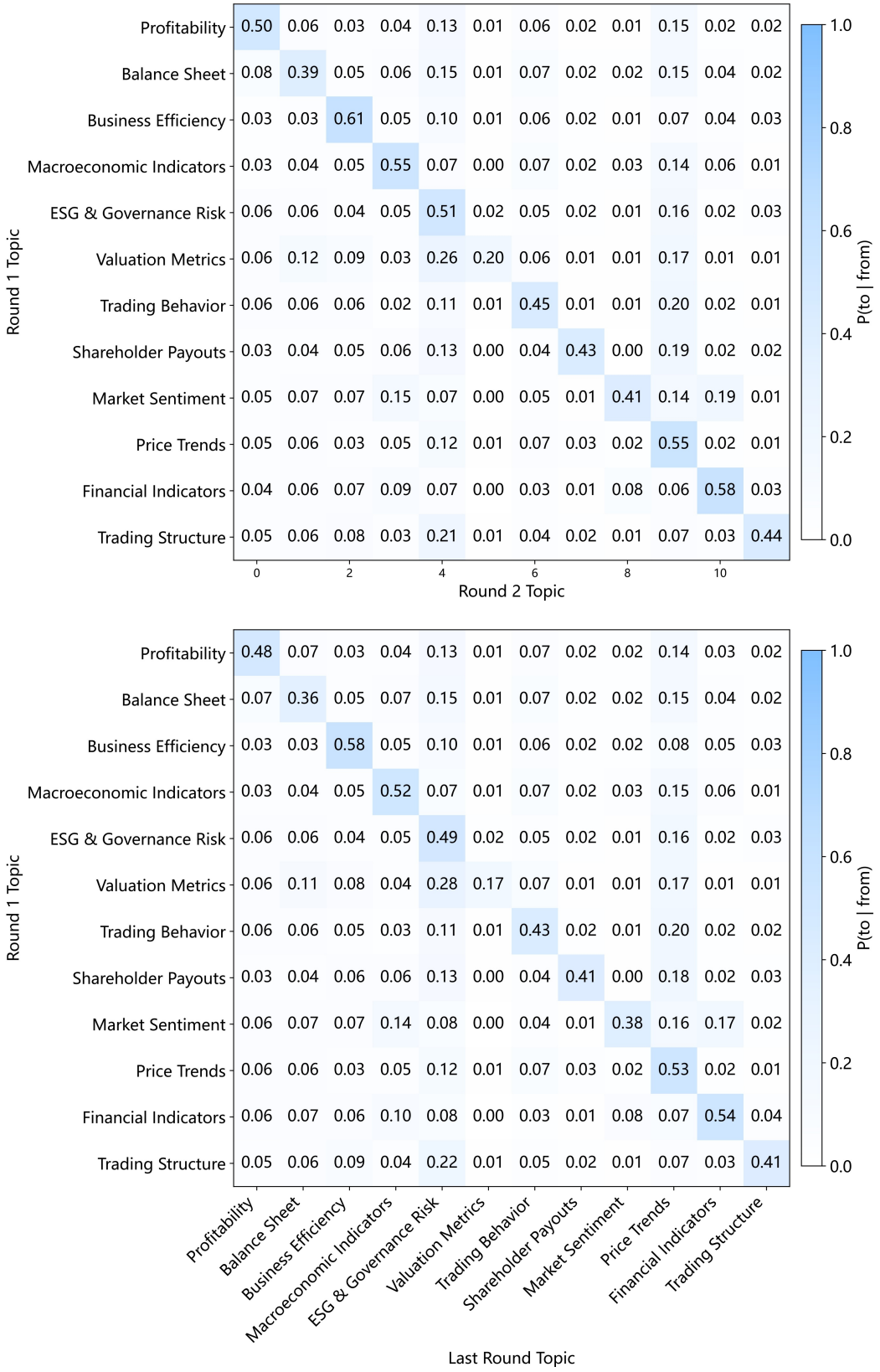
Notes: This figure shows the distribution of the number of unique users per firm in the conversation sample. The x-axis groups firms by the count of distinct users who interacted with the AI about that firm (0–10, 11–20, 21–30, 31–40, and 41+). The y-axis reports the share of firms in each bin; "41+" pools firms with 41 or more unique users.

Figure 3: Topic Composition of User Queries



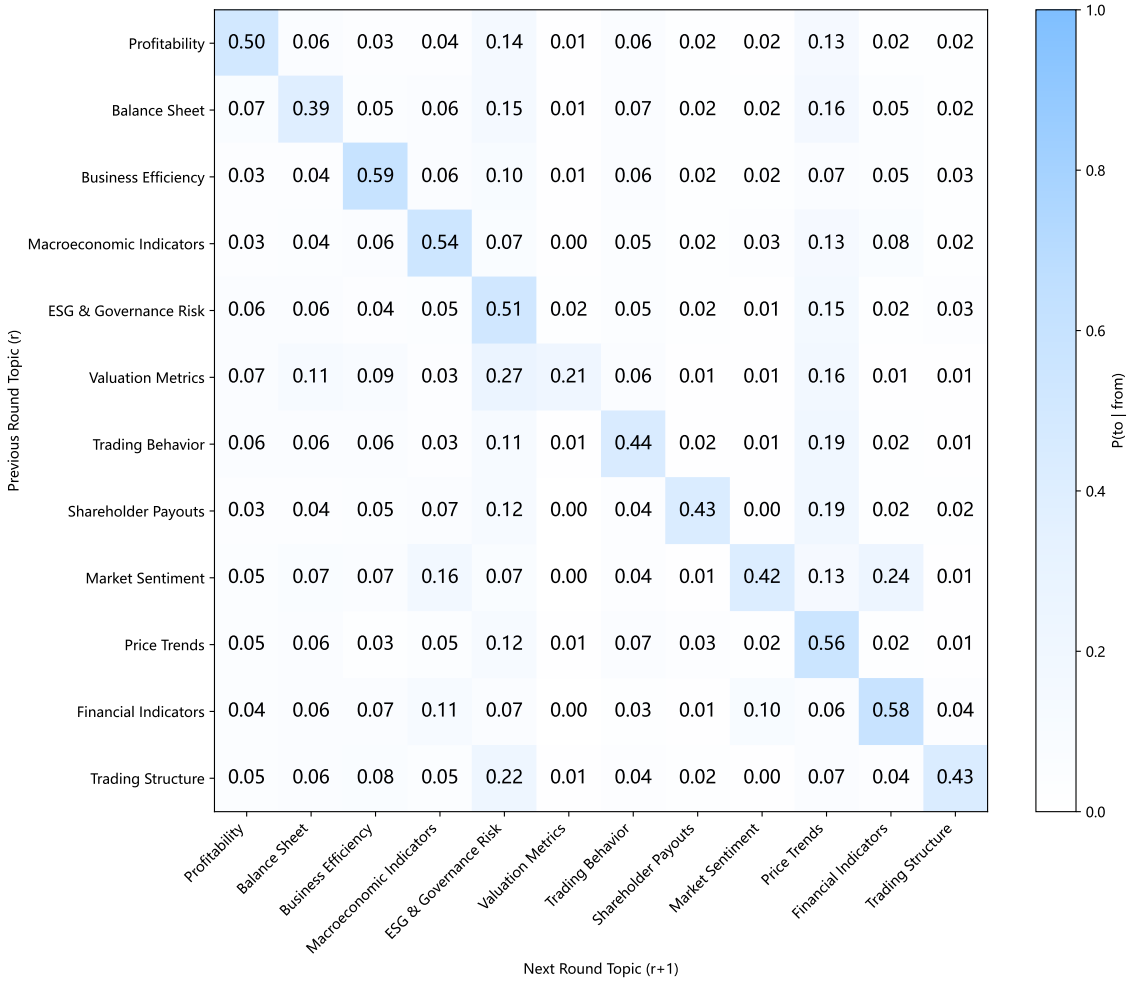
Notes: This figure presents the distribution of conversation topics in the final sample. If a single user query involves multiple topics, the observation is counted multiple times, once for each associated topic. Topic categories are constructed to capture distinct dimensions of financial discussions, such as price dynamics, firm fundamentals, and macroeconomic conditions. Representative examples for each topic category are provided in the Appendix C to illustrate the conceptual distinctions across classifications.

Figure 4: Topic Transition: First Round to Second/Last Round



Notes: This figure reports topic transition probabilities across query rounds within the same conversation. Panel A shows transitions from the first-round query to the second-round query, and Panel B shows transitions from the first-round query to the last-round query. Each cell reports the probability that topic y appears in the destination round conditional on topic x appearing in the origin round. Topics are assigned in a multi-label setting, so row and column sums are not constrained to equal one.

Figure 5: Topic Transition Matrix: Adjacent Rounds



Notes: This figure reports transition probabilities of query topics across adjacent query rounds within the same conversation. Each entry represents the probability that a query topic (y) appears in next round conditional on query topic (x) appearing in current round. Because query topics are assigned in a multi-label setting, row and column sums are not constrained to sum to one.

Table 1: Sample Selection

This table summarizes the sample selection procedure starting from the capital-market-oriented information acquisition conversation sample.

Selection Step	# Conversation–round	# Firm–date–round
Original conversations focus on financial markets	89,997	62,495
Exclude conversations spanning multiple calendar days	88,251	61,350
Exclude conversations involving multiple stocks	88,251	61,350
Exclude users with multiple conversations on the same firm–day	84,028	59,723
Retain the first and second rounds, and append the final round observations	103,165	70,802
Require at least two observations within each firm–date–round	45,873	13,510
Exclude observations with missing control variables	44,046	12,966
Retain firm–dates with observations in both the first/second rounds and the final round	41,558	11,847

Table 2: **Summary Statistics**

This table reports descriptive statistics for the variables used in the empirical analysis. Panel A presents the main variables from the investor dispersion panel measured at the firm—date-–round and firm-–date—conversation-–round levels. Panel B reports key variables measured at the firm–date–conversation level. Panel C reports variables measured at the firm—date level. Panel D reports major variables from the U.S. equity market. The sample period for Panels A-–C spans September 1, 2024 through December 31, 2024, while the sample period for Panel D spans December 2022 through November 2024.

	N	Mean	SD	P10	P90
Panel A: Dispersion Panel Variables					
$Dispersion_{i,t,r}$	11,847	10.392	17.778	0.015	46.272
$Dispersion^{Within}_{i,t,r}$	11,847	3.174	9.873	0.000	7.477
$Dispersion^{Between}_{i,t,r}$	11,847	8.228	15.789	0.000	41.825
$Noise_{i,t,r}$	11,847	9.361	17.751	0.000	40.026
$LengthSD^{Human}_{i,t,r}$	11,847	0.012	0.025	0.001	0.032
$ReadabilitySD^{Human}_{i,t,r}$	11,847	0.114	0.080	0.021	0.228
$LengthSD^{GAI}_{i,t,r}$	11,847	0.060	0.037	0.013	0.109
$ReadabilitySD^{GAI}_{i,t,r}$	11,847	0.063	0.046	0.012	0.120
$\lvert Tone^{Human}_{i,t,c,r}\rvert$	91,374	9.238	26.759	0.018	24.621
$Length^{Human}_{i,t,c,r}$	91,374	0.015	0.035	0.004	0.022
$Readability^{Human}_{i,t,c,r}$	91,374	0.621	0.192	0.357	0.833
Panel B: Firm-Date-Conversation Variables					
$\Delta Tone^{Human}_{i,t,c,r=2}$	30,458	1.524	31.570	-0.784	4.740
$Gap^{GAI-Human}_{i,t,c,r=1}$	30,458	29.825	59.957	-13.297	99.738
$Length^{GAI}_{i,t,c,r=1}$	30,458	0.182	0.092	0.057	0.297
$Readability^{GAI}_{i,t,c,r=1}$	30,458	0.679	0.100	0.563	0.789
Panel C: Firm-Date Variables					
$AveTone^{GAI}$	24448	0.310	0.458	-0.120	0.995
$AveTone^{Human}$	24448	0.023	0.198	-0.009	0.159
$Reversal_{i,t}$	24,448	0.155	0.270	-0.075	0.461
$Momentum_{i,t}$	24,448	0.009	0.366	-0.372	0.421
$LnSize_{i,t}$	24,448	17.043	1.541	15.218	19.228
$Turn_{i,t}$	24,448	0.042	0.048	0.005	0.106
$BM_{i,t}$	24,448	0.691	0.821	0.134	1.500
$PrintNews_{i,t}$	24,448	0.116	0.314	0.000	1.000
$WebNews_{i,t}$	24,448	0.157	0.462	0.000	1.000
Panel D: U.S. Equity Market Variables					
$\Delta AnalystDispersion_{i,t}$	13,398	−1.290	23.368	-9.609	6.095
$\Delta Turn_{i,t}$	13,398	0.868	7.249	-5.736	8.125
$\Delta IVOL_{i,t}$	13,398	0.165	0.651	-0.535	0.922
$\Delta Spread_{i,t}$	13,398	0.696	23.752	-27.920	28.651
$\Delta DO_{i,t}$	13,398	−1.951	18.281	-25.105	19.230
$\Delta IA_{i,t}$	13,398	0.458	4.141	-4.281	5.334
$PostGPT_{i,t}$	13,398	0.553	0.497	0.000	1.000
$Reversal_{i,t}$	13,398	0.005	0.110	-0.125	0.142
$Momentum_{i,t}$	13,398	0.231	0.567	-0.329	0.888
$LnSize_{i,t}$	13,398	8.356	1.669	6.307	10.729
$Turn_{i,t}$	13,398	18.283	14.320	6.651	34.031
$BM_{i,t}$	13,398	0.450	0.386	0.082	0.952

Table 3: Differences in User Dispersion across Query Rounds

This table reports univariate difference-in-means tests for investor tone disagreement across query rounds in AI conversations. The dependent variable is *Dispersion*, measured at the firm–date–round level. For each panel, the columns report the average dispersion for the each round and the difference between two rounds. The standard error of the difference in means from a two-sample t-test is reported in parentheses. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

Dispersion	Round 1	Round 2	R_2 - R_1
	9.581	10.688	1.107***
			(0.395)
	Round 1	Last Round	R_{Last} - R_1
	9.581	10.905	1.324***
			(0.396)

Table 4: GAI Usage and User Disagreement

This table reports panel regression results on how user disagreement evolves with the use of generative AI across query rounds using the Equation 3. The dependent variable in Columns (1)–(4) is $Dispersion$, measured at the firm–date–round level, and the dependent variable in Columns (5)–(8) is $|Tone^{Human}|$, measured at the firm–date–conversation–round level. Columns (1)–(2) and (5)–(6) compare the second round to the first round using D^{Round2}, which equals one for second-round queries and zero for first-round queries. Columns (3)–(4) and (7)–(8) compare the last round to the first round using D^{Final}, which equals one for last-round queries and zero for first-round queries. Control variables include the standard deviation of prompt length ($Length^{Human}$), the standard deviation of prompt readability ($Readability^{Human}$), and lagged firm characteristics $Reversal$, $Momentum$, $LnSize$, $Turn$, and BM, where $Reversal$ is defined as the stock return over the past one month, $Momentum$ is defined as the cumulative stock return from the second to the twelfth month prior, $LnSize$ is the logarithm of tradable market capitalization, $Turn$ is the average daily turnover rate over the past one month, and BM is the book-to-market ratio. Standard errors are two-way clustered by firm and date and are reported in parentheses. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

	$Dispersion_{i,t,r}$					$\lvert Tone^{Human}_{i,t,c,r}\rvert$			
	(1)	(2)	(3)	(4)		(5)	(6)	(7)	(8)
$D^{Round2}_{i,t,r}$	1.107***	1.100***			$D^{Round2}_{i,t,c,r}$	1.959***	1.723***		
	(0.350)	(0.357)				(0.234)	(0.297)		
$D^{Final}_{i,t,r}$			1.324***	1.302***	$D^{Final}_{i,t,c,r}$			2.102***	1.817***
			(0.375)	(0.382)				(0.240)	(0.307)
$LengthSD^{Human}_{i,t,r}$		114.747***		114.453***	$Length^{Human}_{i,t,c,r}$		148.407***		161.390***
		(13.360)		(13.151)			(16.747)		(14.827)
$ReadabilitySD^{Human}_{i,t,r}$		-2.432		-1.620	$Readability^{Human}_{i,t,c,r}$		7.072***		7.195***
		(2.450)		(2.199)			(1.199)		(1.250)
$Reversal_{i,t-1}$		0.300		1.178	$Reversal_{i,t-1}$		1.777		2.570
		(2.631)		(2.878)			(3.625)		(3.758)
$Momentum_{i,t-1}$		0.346		1.038	$Momentum_{i,t-1}$		1.927		2.544
		(2.654)		(2.843)			(3.271)		(3.195)
$LnSize_{i,t-1}$		-0.201		-0.631	$LnSize_{i,t-1}$		3.745		3.803
		(3.875)		(4.207)			(5.761)		(5.768)
$Turn_{i,t-1}$		-7.128		-9.812	$Turn_{i,t-1}$		-14.625		-23.792
		(18.907)		(20.381)			(21.811)		(23.736)
$BM_{i,t-1}$		0.417		1.016	$BM_{i,t-1}$		8.795*		4.805
		(0.615)		(0.703)			(5.159)		(4.079)
Constant	9.581***	11.611	9.581***	18.084	Constant	7.884***	-75.662	7.884***	-71.700
	(0.458)	(70.639)	(0.458)	(76.445)		(0.251)	(102.649)	(0.251)	(101.909)
Firm FE	No	Yes	No	Yes	Firm FE	No	Yes	No	Yes
Date FE	No	Yes	No	Yes	Date FE	No	Yes	No	Yes
User FE	No	No	No	No	User FE	No	Yes	No	Yes
# of Obs	7,898	7,898	7,898	7,898	# of Obs	60,916	60,916	60,916	60,916
R^2	0.001	0.189	0.001	0.185	R^2	0.001	0.622	0.002	0.614

Table 5: Within-Topic versus between-topic Dispersion

This table reports panel regression results on how within-topic and between-topic user disagreement evolve with the use of generative AI between the first and second rounds, as specified in Equation 4. The dependent variable in Columns (1)–(2) is $Dispersion_{i,t,r}^{Within}$, and the dependent variable in Columns (3)–(4) is $Dispersion_{i,t,r}^{Cross}$, both measured at the firm–date–round level. The key independent variable is D^{Round2}, which equals one for second-round queries and zero for first-round queries within the same conversation. Control variables include the standard deviation of prompt length ($LengthSD^{Human}$) and the standard deviation of prompt readability ($ReadabilitySD^{Human}$), and lagged firm characteristics $Reversal$, $Momentum$, $LnSize$, $Turn$, and BM. Columns (2) and (4) include both firm and date fixed effects, as indicated at the bottom of the table. Standard errors are two-way clustered by firm and date and are reported in parentheses. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

	$Dispersion_{i,t,r}^{Within}$		$Dispersion_{i,t,r}^{Between}$	
	(1)	(2)	(3)	(4)
$D_{i,t,r}^{Round2}$	-0.113	-0.049	1.068***	1.038***
	(0.226)	(0.228)	(0.304)	(0.306)
$LengthSD_{i,t,r}^{Human}$		32.490***		99.174***
		(5.995)		(13.203)
$ReadabilitySD_{i,t,r}^{Human}$		2.919*		-3.378
		(1.696)		(2.453)
$Reversal_{i,t-1}$		-0.471		0.810
		(1.720)		(2.014)
$Momentum_{i,t-1}$		0.250		0.163
		(1.595)		(2.059)
$LnSize_{i,t-1}$		2.084		-1.629
		(2.412)		(3.212)
$Turn_{i,t-1}$		3.388		-8.959
		(7.715)		(16.926)
$BM_{i,t-1}$		0.516		0.150
		(0.345)		(0.509)
Constant	3.207***	-37.006	7.474***	36.616
	(0.424)	(43.826)	(0.321)	(58.727)
Firm FE	No	Yes	No	Yes
Date FE	No	Yes	No	Yes
# of Obs	7,898	7,898	7,898	7,898
R^2	0.000	0.172	0.001	0.173

Table 6: GAI Answer and Human Tone Updating

This table reports panel regression results on how human tone updates from round 1 to round 2 within the same conversation in response to tone embedded in AI's answer, as specified in Equation 5. The dependent variable is $\Delta Tone^{Human}_{i,t,c,r=2}$. The key independent variable is the AI-human tone difference in round 1, $Gap^{GAI-Human}_{i,t,c,r=1}$. Columns (2)–(4) additionally control for query and response characteristics, including $Length^{Human}$, $Readability^{Human}$, $Length^{GAI}$, and $Readability^{GAI}$, as well as lagged firm characteristics $Reversal$, $Momentum$, $LnSize$, $Turn$, and BM. Column (3) includes date fixed effects and Column (4) includes both firm and date fixed effects, as indicated at the bottom of the table. Standard errors are two-way clustered by firm and date and are reported in parentheses. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

	$\Delta Tone^{Human}_{i,t,c,r=2}$			
	(1)	(2)	(3)	(4)
$Gap^{GAI-Human}_{i,t,c,r=1}$	0.059***	0.061***	0.062***	0.065***
	(0.005)	(0.005)	(0.005)	(0.005)
$Length^{Human}_{i,t,c,r=1}$		5.892	2.503	0.550
		(6.933)	(6.169)	(6.135)
$Readability^{Human}_{i,t,c,r=1}$		2.173**	1.971**	1.457
		(0.888)	(0.891)	(0.959)
$Length^{GAI}_{i,t,c,r=1}$		-0.877	-2.567	-3.286
		(3.896)	(2.908)	(2.652)
$Readability^{GAI}_{i,t,c,r=1}$		-2.236	-2.128	-0.883
		(1.941)	(1.833)	(1.786)
$Reversal_{i,t-1}$		0.412	0.829	1.258
		(0.832)	(0.875)	(1.550)
$Momentum_{i,t-1}$		0.524	0.318	1.007
		(0.751)	(0.738)	(1.575)
$LnSize_{i,t-1}$		-0.142	-0.164	1.462
		(0.153)	(0.144)	(2.307)
$Turn_{i,t-1}$		14.697**	15.538**	-0.671
		(6.355)	(6.338)	(11.268)
$BM_{i,t-1}$		-0.016	-0.001	0.439
		(0.035)	(0.024)	(0.281)
Constant	-0.249	1.831	2.455	-26.732
	(0.325)	(2.952)	(2.846)	(40.315)
Firm FE	No	No	No	Yes
Date FE	No	No	Yes	Yes
# of Obs	30,458	30,458	30,458	29,791
R^2	0.013	0.014	0.019	0.104

Table 7: AI Answer and Future Stock Returns

This table reports panel regression results on the relation between lagged query tone and future stock returns. The dependent variables are $Ret_{i,t,t+1}$, $Ret_{i,t,t+2}$, $Ret_{i,t,t+3}$, and $Ret_{i,t,t+5}$ in Columns (1)-(4), respectively. The key independent variables are $AveTone^{Human}_{i,t-1}$ and $AveTone^{GAI}_{i,t-1}$. Control variables include $Reversal$, $Momentum$, $LnSize$, $Turn$, BM, $PrintNews$, and $WebNews$. All specifications include firm fixed effects and date fixed effects, as indicated at the bottom of the table. Standard errors are two-way clustered by firm and date and are reported in parentheses. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

	$Ret_{i,t,t+1}$	$Ret_{i,t,t+2}$	$Ret_{i,t,t+3}$	$Ret_{i,t,t+5}$
	(1)	(2)	(3)	(4)
$AveTone^{Human}_{i,t-1}$	0.174	0.072	-0.136	-0.131
	(0.143)	(0.234)	(0.313)	(0.408)
$AveTone^{GAI}_{i,t-1}$	0.209***	0.328***	0.495***	0.604***
	(0.062)	(0.102)	(0.145)	(0.190)
$Reversal_{i,t-1}$	1.419*	2.988**	4.777***	8.075***
	(0.757)	(1.266)	(1.686)	(2.178)
$Momentum_{i,t-1}$	-0.625	-0.851	-1.381	-1.670
	(0.441)	(0.838)	(1.196)	(1.684)
$LnSize_{i,t-1}$	-7.285***	-14.850***	-23.231***	-36.101***
	(0.978)	(1.776)	(2.621)	(3.507)
$Turn_{i,t-1}$	-5.905	-9.548	-11.362	-13.558
	(3.801)	(6.307)	(9.063)	(12.038)
$BM_{i,t-1}$	0.865**	1.450*	1.717	2.700
	(0.409)	(0.797)	(1.204)	(1.739)
$PrintNews_{i,t-1}$	-0.112	-0.290**	-0.386**	-0.491*
	(0.085)	(0.129)	(0.186)	(0.277)
$WebNews_{i,t-1}$	0.060	-0.036	-0.073	-0.152
	(0.067)	(0.114)	(0.157)	(0.207)
Firm FE	Yes	Yes	Yes	Yes
Date FE	Yes	Yes	Yes	Yes
# of Obs	24,448	24,448	24,448	24,448
R^2	0.456	0.509	0.548	0.599

Table 8: AI Answer and Information Precision

This table reports panel regression results on how information noise in AI answer evolves. The dependent variable is $Noise$, measured as the within-topic tone dispersion in AI answer. The key independent variable is D^{Round2}, which equals one for the second-round query and zero for the first-round query. Control variables include answer-level dispersion measures and lagged firm characteristics $Reversal$, $Momentum$, $LnSize$, $Turn$, and BM. Columns (3) include date fixed effects, and Columns (4) include both firm and date fixed effects, as indicated at the bottom of the table. Standard errors are two-way clustered by firm and date and are reported in parentheses. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

	$Noise_{i,t,r}$			
	(1)	(2)	(3)	(4)
$D^{Round2}_{i,t,r}$	-2.217***	-2.087***	-2.090***	-2.143***
	(0.414)	(0.399)	(0.400)	(0.407)
$LengthSD^{GAI}_{i,t,r}$		42.280***	40.656***	23.575***
		(10.488)	(10.619)	(4.644)
$ReadabilitySD^{GAI}_{i,t,r}$		32.440***	32.407***	19.296***
		(8.638)	(8.786)	(5.318)
$Reversal_{i,t-1}$		3.315***	3.581***	4.876*
		(1.083)	(1.105)	(2.899)
$Momentum_{i,t-1}$		0.731	1.135	3.414
		(0.891)	(0.954)	(2.908)
$LnSize_{i,t-1}$		1.093***	1.150***	2.050
		(0.388)	(0.388)	(4.455)
$Turn_{i,t-1}$		6.533	5.881	-14.348
		(8.312)	(8.824)	(15.187)
$BM_{i,t-1}$		0.000	-0.002	0.358
		(0.057)	(0.057)	(0.367)
Constant	10.855***	-14.973**	-15.987**	-31.157
	(0.937)	(7.323)	(7.319)	(81.088)
Firm FE	No	No	No	Yes
Date FE	No	No	Yes	Yes
# of Obs	7,898	7,898	7,898	7,898
R^2	0.004	0.039	0.058	0.203

Table 9: Return Predictability Across Conversation Rounds

This table reports panel regression results on the relation between average tone and future stock returns across different query rounds. The dependent variables are $Ret_{i,t,t+1}$, $Ret_{i,t,t+2}$, $Ret_{i,t,t+3}$, and $Ret_{i,t,t+5}$ in Columns (1)-(4) within each panel, respectively. Columns (1)-(4), (5)-(8), and (9)-(12) use *AveTone* measured in round 1, round 2, and the last round in conversations, respectively. The key independent variables are $AveTone_{i,t}^{GAI}$ and $AveTone_{i,t}^{Human}$. Control variables include *Reversal*, *Momentum*, *LnSize*, *Turn*, *BM*, *PrintNews*, and *WebNews*. All specifications include firm fixed effects and date fixed effects, as indicated at the bottom of the table. Standard errors are two-way clustered by firm and date and are reported in parentheses. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

	AveTone (Round 1)				*AveTone* (Round 2)				*AveTone* (Last round)			
	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)
	$Ret_{i,t,t+1}$	$Ret_{i,t,t+2}$	$Ret_{i,t,t+3}$	$Ret_{i,t,t+5}$	$Ret_{i,t,t+1}$	$Ret_{i,t,t+2}$	$Ret_{i,t,t+3}$	$Ret_{i,t,t+5}$	$Ret_{i,t,t+1}$	$Ret_{i,t,t+2}$	$Ret_{i,t,t+3}$	$Ret_{i,t,t+5}$
$AveTone_{i,t}^{GAI}$	0.080	0.117	0.172	0.220	0.183***	0.261***	0.398***	0.481***	0.156***	0.207**	0.328**	0.449**
	(0.050)	(0.076)	(0.108)	(0.143)	(0.056)	(0.093)	(0.134)	(0.177)	(0.059)	(0.102)	(0.144)	(0.186)
$AveTone_{i,t}^{Human}$	0.169	0.105	0.021	0.168	-0.002	-0.109	-0.276	-0.102	0.039	0.035	-0.085	0.027
	(0.119)	(0.182)	(0.252)	(0.353)	(0.116)	(0.190)	(0.265)	(0.336)	(0.124)	(0.207)	(0.290)	(0.351)
$Reversal_{i,t-1}$	1.144	2.596**	4.355**	7.601***	1.134	2.583**	4.336**	7.574***	1.138	2.588**	4.344**	7.580***
	(0.775)	(1.293)	(1.715)	(2.152)	(0.774)	(1.293)	(1.714)	(2.148)	(0.775)	(1.293)	(1.715)	(2.150)
$Momentum_{i,t-1}$	-0.762	-1.058	-1.686	-2.064	-0.778	-1.079	-1.716	-2.107	-0.775	-1.077	-1.714	-2.104
	(0.478)	(0.904)	(1.279)	(1.727)	(0.477)	(0.903)	(1.278)	(1.725)	(0.477)	(0.904)	(1.279)	(1.726)
$LnSize_{i,t-1}$	-7.299***	-15.261***	-23.939***	-36.967***	-7.301***	-15.261***	-23.936***	-36.963***	-7.299***	-15.257***	-23.929***	-36.956***
	(1.016)	(1.831)	(2.689)	(3.508)	(1.016)	(1.830)	(2.688)	(3.505)	(1.017)	(1.832)	(2.690)	(3.506)
$Turn_{i,t-1}$	-6.108	-9.396	-11.090	-14.703	-6.084	-9.360	-11.030	-14.625	-6.128	-9.417	-11.118	-14.732
	(3.852)	(6.455)	(9.319)	(12.427)	(3.845)	(6.449)	(9.310)	(12.419)	(3.844)	(6.449)	(9.309)	(12.418)
$BM_{i,t-1}$	0.767*	1.226	1.330	2.411	0.765*	1.220	1.320	2.403	0.774*	1.234	1.340	2.428
	(0.399)	(0.762)	(1.153)	(1.739)	(0.398)	(0.760)	(1.152)	(1.738)	(0.399)	(0.761)	(1.153)	(1.736)
$PrintNews_{i,t-1}$	-0.088	-0.251*	-0.388*	-0.535*	-0.089	-0.254*	-0.392*	-0.542*	-0.088	-0.254*	-0.393*	-0.542*
	(0.090)	(0.145)	(0.213)	(0.310)	(0.091)	(0.145)	(0.213)	(0.310)	(0.091)	(0.145)	(0.213)	(0.310)
$WebNews_{i,t-1}$	0.086	0.016	0.021	-0.086	0.087	0.016	0.021	-0.086	0.087	0.017	0.022	-0.084
	(0.074)	(0.126)	(0.175)	(0.230)	(0.074)	(0.126)	(0.174)	(0.230)	(0.074)	(0.126)	(0.174)	(0.230)
Firm FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Date FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
# of Obs	21,198	21,198	21,198	21,198	21,198	21,198	21,198	21,198	21,198	21,198	21,198	21,198
R^2	0.455	0.510	0.548	0.597	0.455	0.510	0.548	0.598	0.455	0.510	0.548	0.598

Table 10: GAI Answer Characteristics and Human Topic Stickiness

This table reports panel regression results on how GAI answer characteristics are associated with human users' topic stickiness across consecutive rounds within the same conversation, following Equation (10). The dependent variable is $TopicStickiness_{i,t,c,r}$, an indicator equal to one if the topic set in the current round shares at least one topic with the topic set in the previous round, and zero otherwise. $Source^{GAI}_{i,t,c,r-1}$ is an indicator equal to one if the GAI answer in the previous round provides explicit information sources. $PropSource^{GAI}_{i,t,c,r-1}$ is the proportion of source-providing sentences in the previous-round GAI answer. $SourceCount^{GAI}_{i,t,c,r-1}$ is the number of sources provided in the previous-round GAI answer. Column (4) restricts the sample to observations in which the previous-round GAI answer provides at least one source and examines variation in the number of sources conditional on source provision. $HardInfo_{i,t,c,r-1}$ is an indicator equal to one if the previous-round topic is classified as hard-information intensive. Control variables include previous-round GAI answer length, previous-round GAI answer readability, and lagged firm characteristics, including $Reversal$, $Momentum$, $LnSize$, $Turnover$, and BM. Columns (2)–(4) include firm and date fixed effects. Standard errors are two-way clustered by firm and date and are reported in parentheses. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

	$TopicStickiness_{i,t,c,r}$			
	(1)	(2)	(3)	(4)
$Source^{GAI}_{i,t,c,r-1}$	-0.135***	-0.100***		
	(0.014)	(0.012)		
$PropSource^{GAI}_{i,t,c,r-1}$			-0.123***	
			(0.026)	
$SourceCount^{GAI}_{i,t,c,r-1}$				-0.006*
				(0.004)
$HardInfo_{i,t,c,r-1}$	-0.067***	-0.073***	-0.075***	-0.070***
	(0.016)	(0.015)	(0.014)	(0.017)
$Length^{GAI}_{i,t,c,r-1}$	-0.071	-0.137**	-0.230***	-0.177*
	(0.082)	(0.066)	(0.067)	(0.093)
$Readability^{GAI}_{i,t,c,r-1}$	-0.181**	-0.180***	-0.157**	-0.098
	(0.070)	(0.067)	(0.070)	(0.087)
$Reversal_{i,t-1}$	-0.140***	-0.104*	-0.104*	-0.086
	(0.031)	(0.062)	(0.061)	(0.069)
$Momentum_{i,t-1}$	0.026	0.044	0.044	0.034
	(0.024)	(0.057)	(0.057)	(0.063)
$LnSize_{i,t-1}$	0.023***	0.025	0.027	-0.004
	(0.006)	(0.091)	(0.090)	(0.108)
$Turnover_{i,t-1}$	-0.563**	-0.021	-0.023	0.079
	(0.226)	(0.299)	(0.298)	(0.351)
$BM_{i,t-1}$	0.005***	-0.099***	-0.099***	-0.148***
	(0.001)	(0.036)	(0.037)	(0.043)
Constant	-0.777***	-0.652	-0.734	-0.207
	(0.110)	(1.620)	(1.607)	(1.927)
Firm FE	No	Yes	Yes	Yes
Date FE	No	Yes	Yes	Yes
# of Obs	44,230	44,230	44,230	34,354
Pseudo R^2	0.00418	0.0247	0.0244	0.0262

Table 11: Return Predictability Comparison by Conversation Rounds: Two or More Rounds

This table reports panel regression results on the predictive power of second-round AI tone for future stock returns using the two-rounds or multi-rounds (>2) conversation subsamples. The dependent variables are $Ret_{i,t,t+1}$, $Ret_{i,t,t+2}$, $Ret_{i,t,t+3}$, and $Ret_{i,t,t+5}$. The key independent variables are $AveTone^{AI,r=2}|R = 2$ and $AveTone^{AI,r=2}|R \geq 3$, constructed from second-round AI tone in conversations with total rounds $R = 2$ and $R \geq 3$, respectively; Columns (1), (3), (5), and (7) report regressions using $AveTone^{AI,r=2}|R = 2$, and Columns (2), (4), (6), and (8) report regressions using $AveTone^{AI,r=2}|R \geq 3$. Control variables include $Reversal$, $Momentum$, $LnSize$, $Turn$, BM, $PrintNews$, and $WebNews$. All specifications include date fixed effects, as indicated at the bottom of the table. Standard errors are two-way clustered by firm and date and are reported in parentheses. Coefficients difference test reports a χ^2 value (p-value) for equality of the coefficients on $AveTone^{AI,r=2}|R = 2$ and $AveTone^{AI,r=2}|R \geq 3$ within each return horizon. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

	$Ret_{i,t,t+1}$		$Ret_{i,t,t+2}$		$Ret_{i,t,t+3}$		$Ret_{i,t,t+5}$	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
$AveTone^{GAI,r=2}_{i,t}|R = 2$	0.164***		0.201**		0.260**		0.374**	
	(0.051)		(0.085)		(0.119)		(0.160)	
$AveTone^{GAI,r=2}_{i,t}|R \geq 3$		0.280***		0.456***		0.599***		0.651***
		(0.079)		(0.131)		(0.182)		(0.244)
χ^2 (Coef. diff.)	1.386		2.315		2.043		0.798	
P-value	0.239		0.128		0.153		0.372	
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Date FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
# of Obs	17,235	6,758	17,235	6,758	17,235	6,758	17,235	6,758
R^2	0.345	0.349	0.392	0.386	0.426	0.419	0.463	0.470

Table 12: Return Predictability: Last vs. Second-Last Round

This table reports panel regression results on the predictive power of late-round AI tone for future stock returns using the multi-round conversation sample. The dependent variables are $Ret_{i,t,t+1}$, $Ret_{i,t,t+2}$, $Ret_{i,t,t+3}$, and $Ret_{i,t,t+5}$. The key independent variables are $AveTone^{AI,r=-2}$ and $AveTone^{AI,r=-1}$, constructed from the second-last round and the last round of the same conversation, respectively; Columns (1), (3), (5), and (7) report specifications using $AveTone^{AI,r=-2}$, and Columns (2), (4), (6), and (8) report specifications using $AveTone^{AI,r=-1}$. Control variables include $Reversal$, $Momentum$, $LnSize$, $Turn$, BM, $PrintNews$, and $WebNews$. All specifications include date fixed effects, as indicated at the bottom of the table. Standard errors are two-way clustered by firm and date and are reported in parentheses. Coefficients difference test reports a χ^2 value (p-value) for equality of the coefficients on $AveTone^{AI,r=-2}$ and $AveTone^{AI,r=-1}$ within each return horizon. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

	$Ret_{i,t,t+1}$		$Ret_{i,t,t+2}$		$Ret_{i,t,t+3}$		$Ret_{i,t,t+5}$	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
$AveTone_{i,t}^{GAI,r=-2}$	0.171***		0.250***		0.327***		0.407***	
	(0.044)		(0.073)		(0.102)		(0.138)	
$AveTone_{i,t}^{GAI,r=-1}$		0.175***		0.218***		0.287***		0.406***
		(0.045)		(0.075)		(0.104)		(0.142)
χ^2 (Coef. diff.)	0.005		0.125		0.095		0.000	
P-value	0.946		0.723		0.758		0.996	
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Date FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
# of Obs	21,852	21,852	21,852	21,852	21,852	21,852	21,852	21,852
R^2	0.348	0.348	0.393	0.393	0.427	0.427	0.468	0.468

Table 13: ChatGPT Release and U.S. Equity Market Outcomes

This table reports panel regression results on the effects of ChatGPT release on investor disagreement and market outcomes in the U.S. equity market from 2021 to 2024. The dependent variables are changes in disagreement and stock characteristics measures , denoted by ΔY, including analyst forecast dispersion, stock turnover, idiosyncratic volatility, bid-ask spread, machine-based different in opioions (DO) and information asymmetry (IA) as in Liu et al. (2025). ΔY is constructed as the average value of Y over the earnings-announcement month and the subsequent two months minus the average value of Y over the two months preceding the earnings-announcement month. The key independent variable is $PostGPT$, which equals zero for November 2022 and earlier and equals one for December 2022 and later. Control variables include $Reversal$, $Momentum$, $LnSize$, $Turn$, and BM. All specifications include firm fixed effects. Standard errors are two-way clustered by firm and date and are reported in parentheses. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

	(1) $\Delta AnalystDispersion_{i,t}$	(2) $\Delta Turn_{i,t}$	(3) $\Delta IVOL_{i,t}$	(4) $\Delta Spread_{i,t}$	(5) $\Delta DO_{i,t}$	(6) $\Delta IA_{i,t}$
$PostGPT_{i,t}$	2.180**	1.303***	0.123**	-1.038	10.154**	-1.109*
	(0.997)	(0.445)	(0.057)	(3.453)	(3.782)	(0.577)
$Reversal_{i,t-1}$	2.919	-2.623*	-0.712***	-15.267*	3.839	-6.155***
	(6.483)	(1.488)	(0.146)	(8.460)	(11.029)	(2.062)
$Momentum_{i,t-1}$	-3.143***	-0.685*	-0.021	-1.508	-4.496**	0.693**
	(0.938)	(0.358)	(0.040)	(2.944)	(2.154)	(0.319)
$LnSize_{i,t-1}$	0.756	-0.463	0.014	7.137**	0.941	-0.423
	(1.457)	(0.533)	(0.037)	(3.401)	(2.344)	(0.551)
$Turn_{i,t-1}$	0.012	-0.356***	-0.023***	-0.273***	0.201***	-0.057***
	(0.041)	(0.018)	(0.002)	(0.065)	(0.049)	(0.011)
$BM_{i,t-1}$	-2.325	-0.789	0.110**	9.137**	2.598	0.351
	(2.213)	(0.686)	(0.052)	(3.717)	(2.351)	(0.469)
Firm FE	Yes	Yes	Yes	Yes	Yes	Yes
# of Obs	13,398	13,398	13,398	13,398	13,398	13,398
R^2	0.114	0.293	0.243	0.119	0.212	0.170

Appendix A Variable Definitions

Table A1: Variable Definitions

This table defines the main variables used in the empirical analyses.

Variable	Construction	Description
Panel A: Tone and Text Variables		
$Tone$	$Tone = s_{pos} - s_{neg}$, where s_{pos} and s_{neg} are the FinBERT-implied positive and negative sentiment shares of the corresponding text, following Huang et al. (2023). Superscripts indicate the text source, with $Human$ for investor prompts and GAI for AI responses.	Text tone.
$\lvert Tone^{Human} \rvert$	Absolute value of $Tone^{Human}$.	Belief polarization.
$\Delta Tone^{Human}$	Second-round investor prompt tone minus first-round investor prompt tone within the same conversation.	Human tone update.
$Gap^{GAI-Human}$	First-round AI response tone minus first-round investor prompt tone within the same conversation.	Initial AI-human tone gap.
$Length$	Number of words in the corresponding text, computed separately for investor prompts and AI responses. Superscripts indicate the text source when needed. The variable is expressed in thousands of words in the empirical tables.	Text length.
$Readability$	Share of words in the corresponding text that appear in a high-frequency dictionary, computed separately for investor prompts and AI responses. Superscripts indicate the text source when needed.	Text readability.
Panel B: Disagreement and Dispersion Measures		
$Dispersion$	Population standard deviation of investor prompt tone across all prompts in the same firm-date-round cell.	Total disagreement.
$Dispersion^{Within}$	Square root of the topic-share-weighted average within-topic population variance of investor prompt tone.	Within-topic disagreement.
$Dispersion^{Between}$	Square root of the topic-share-weighted variance of topic-level mean investor prompt tone around the firm-date-round mean.	Between-topic disagreement.
$Noise$	Within-topic dispersion of AI response tone within a firm-date-round cell.	AI answer noise.
$LengthSD$	Population standard deviation of text length within a firm–date–round cell. Superscripts indicate the text source, with $Human$ for investor prompts and GAI for AI responses.	Dispersion of text length.

Table A1 – Continued

Variable	Construction	Description
$ReadabilitySD$	Population standard deviation of text readability within a firm–date–round cell. Superscripts indicate the text source, with $Human$ for investor prompts and GAI for AI responses.	Dispersion of text readability.
Panel C: Stock-Day Tone Aggregates		
$AveTone$	Arithmetic mean of tone across all texts about the same stock on the same day. Superscripts indicate the text source, with $Human$ for investor prompts and GAI for AI responses.	Average tone.
$AveTone^{r}$	Firm-date mean tone computed using texts from a specified conversation round r. Superscripts may additionally indicate the text source when needed.	Round-specific average tone.
$AveTone^{GAI,r=2} \mid R = 2$	Firm-date mean AI response tone computed from second-round responses in conversations with exactly two rounds.	Second-round AI tone in two-round conversations.
$AveTone^{GAI,r=2} \mid R \geq 3$	Firm-date mean AI response tone computed from second-round responses in conversations with three or more rounds.	Second-round AI tone in longer conversations.
Panel D: Topic Stickiness and AI Answer Characteristics		
$TopicStickiness$	Indicator equal to one if the current round shares at least one topic with the previous round within the same conversation.	Continuation of prior topic.
$Source^{GAI}$	Indicator equal to one if the previous-round GAI response provides explicit information sources.	Source provision.
$PropSource^{GAI}$	Proportion of sentences in the previous-round GAI response that provide explicit information sources.	Intensity of source provision.
$SourceCount^{GAI}$	Number of information sources provided in the previous-round GAI response.	Number of provided sources.
$HardInfo$	Indicator equal to one if the previous-round topic is classified as hard-information intensive.	Hard-information topic.
Panel E: Return Variables and Stock Characteristics		
Ret	Stock return over the corresponding future horizon after the conversation date.	Future stock return.
$Reversal$	Cumulative stock return over the prior 21 trading days, using adjusted closing prices.	Short-term reversal.
$Momentum$	Cumulative return from trading day $t-252$ to trading day $t-21$, using adjusted closing prices.	Momentum.
$LnSize$	Natural logarithm of market capitalization.	Firm size.

Table A1 – Continued

Variable	Construction	Description
$Turn$	Average daily turnover over the prior 21 trading days, where daily turnover equals traded shares divided by float shares.	Past-month turnover.
BM	Book equity from the 2024 semiannual consolidated financial statement divided by market capitalization.	Book-to-market ratio.
$PrintNews$	$(N^{Print,+} - N^{Print,-})/(N^{Print,+} + N^{Print,-})$, where $N^{Print,+}$ and $N^{Print,-}$ are positive and negative firm-related newspaper financial news counts. The variable is set to zero if there is no non-neutral newspaper financial news on that day.	Newspaper news sentiment.
$WebNews$	$(N^{Web,+} - N^{Web,-})/(N^{Web,+} + N^{Web,-})$, where $N^{Web,+}$ and $N^{Web,-}$ are positive and negative firm-related online financial news counts. The variable is set to zero if there is no non-neutral online financial news on that day.	Online news sentiment.
Panel F: Market-Level Event-Window Variables		
$PostGPT$	Indicator equal to one for December 2022 and later, and zero otherwise.	Post-ChatGPT period.
$\Delta AnalystDispersion$	Dispersion of analysts' earnings forecasts across analysts following the same firm.	Analyst forecast dispersion.
$\Delta Turn$	Stock turnover ratio, measured as trading volume divided by shares outstanding.	Trading activity.
$\Delta IVOL$	Idiosyncratic volatility estimated from residual stock returns.	Idiosyncratic volatility.
$\Delta Spread$	Relative bid-ask spread constructed from daily trading data.	Market liquidity.
ΔDO	Difference-of-opinion component of market friction disagreement (MFD), estimated using the machine-learning decomposition method in Liu et al. (2025).	Difference of opinion.
ΔIA	Information-asymmetry component of market friction disagreement (MFD), estimated using the machine-learning decomposition method in Liu et al. (2025).	Information asymmetry.

Appendix B Examples of AI-human conversations

This section presents translated and de-identified examples of AI-human conversations used for illustration purposes.

Example 1

User: Global sales figures for **Company A** in September are out. The data suggest year-over-year growth.

AI: **Company A** reported strong global sales performance in September 2023. According to the latest figures, the firm sold approximately 121,600 vehicles during the month, representing a year-over-year increase of about 29.9%. This growth suggests solid momentum in the firm's global operations and an expanding market presence.

Over the first nine months of 2023, **Company A**'s cumulative vehicle sales reached roughly 864,000 units, a year-over-year increase of about 7.7%. New-energy vehicle sales were a key driver, totaling around 170,000 units over the same period and rising by approximately 75.8%, indicating strong demand in emerging segments.

The firm's major brands also posted growth to varying degrees. For example, year-to-date sales for key brands were reported at approximately 495,900 units (Brand H), 33,000 units (Brand W), 78,000 units (Brand O), 105,000 units (Brand T), and 153,000 units (Pickup segment). In September, Brand H sold about 73,766 units, up roughly 37% year-over-year, reflecting solid competitiveness in its core market.

International expansion was another highlight: overseas sales for the first nine months were reported at around 212,000 units, increasing by approximately 89.5% year-over-year, which is consistent with rapid growth in global markets.

Overall, these figures point to steady global expansion and particularly strong traction in new-energy vehicles.

User: **Company A** does not pursue aggressive market expansion, which seems more consistent with a sustainable growth strategy.

AI: This strategy reflects a long-term orientation toward sustainable development. Avoiding overly aggressive expansion can help firms better manage risk, maintain product quality, preserve financial stability, and adapt more effectively to changing market conditions. Such an approach may also support organizational learning and long-run competitiveness.

Example 2

User: The government is reportedly deploying a large fund to support the stock market. **Company B** surged to a new high. Will it keep rising?

AI: Whether **Company B** continues to rise is uncertain, because a new-high rally can reflect both improved fundamentals and short-term market forces.

First, it is important to distinguish market expectations from intrinsic value. If the stock has already priced in very optimistic assumptions about future growth, subsequent upside may require earnings or cash-flow outcomes that exceed those already-high expectations. In that case, even strong operational performance may not translate into the same pace of price appreciation.

Second, company-specific fundamentals still matter. If **Company B** has durable competitive advantages, strong execution, and credible long-run growth drivers, these factors can support the valuation over time. However, the stock's sensitivity to delivery of targets tends to increase when the valuation is elevated.

Third, policy and macro conditions can affect the short-run path. Reports of broad market-support measures may boost risk appetite and liquidity, which can lift prices. But such effects are often time-varying and can be offset by macro uncertainty, sector competition, or shifts in investor sentiment.

Overall, the probability of continued gains depends on (i) whether fundamentals keep improving, (ii) whether expectations move even higher, and (iii) how policy, liquidity, and risk sentiment evolve. This is not investment advice; a prudent approach is to examine valuation relative to realistic scenarios, track key operating indicators, and consider downside risk as well as upside potential.

Example 3

User: **Company C** experienced a toxic wastewater leakage incident.

AI: Public accounts describe a major environmental incident involving **Company C** on July 3, 2010. The event was linked to a failure in a wastewater containment system at one of the firm's mining-related facilities, after which contaminated acidic effluent reportedly entered a drainage pathway and ultimately flowed into a nearby river system. The leakage was described as substantial in volume and was associated with downstream water-quality deterioration and widespread losses in aquaculture operations.

Investigations described in public reports attributed the incident to both engineering and management failures. In addition to the physical containment failure, the event was reportedly aggravated by improper site practices and inadequate construction, production, and environmental management controls. These deficiencies were characterized as systematic rather than accidental.

The consequences were described as severe for local communities and the surrounding environment. Reports noted significant impacts on fisheries and related livelihoods, and also indicated that parts of a nearby urban area faced temporary disruption to water supply services.

Regulatory and legal outcomes reported publicly included a large monetary penalty, accountability measures for responsible personnel, and requirements to suspend certain operations and undertake remediation and rectification. The incident has often been cited as an example of governance, compliance, and environmental-risk management breakdown in heavy-industry operations.

Example 4

User: Accounts receivable may expose **Company D** to overdue, doubtful, or uncollectible balances, which could weaken earnings quality and increase operating risk.

AI: Yes. Elevated accounts receivable can create credit-risk exposure for **Company D**, including the possibility that some balances become overdue, doubtful, or ultimately uncollectible. These outcomes can reduce the reliability of reported earnings because revenue may be recognized before cash is received, while later write-downs or provisions can reverse prior income.

From an operating perspective, problematic receivables also tighten cash flow and may increase financing needs and related costs. If collections are delayed or default risk rises, the firm may face higher working-capital pressure and greater uncertainty in day-to-day operations.

To mitigate these risks, **Company D** typically needs disciplined receivables management, such as customer credit screening, clear payment terms, active collection and follow-up procedures, and timely loss provisioning. These steps help contain bad-debt risk and support more sustainable earnings quality.

Appendix C Topic Classification

Topic Definitions and Keyword Examples

This table provides detailed definitions and representative keyword examples for each topic category used in the analysis.

Topic Name	**Description**	**Keyword Examples**
Price Trends	Short-horizon stock price direction and technical price movements.	price trend; momentum; price movement
Trading Behavior	Trading activity and market participation surrounding stock prices.	stock price; trading; turnover; volume
Trading Structure	Market microstructure and trading mechanisms.	order structure; trading mechanism; execution
Financial Indicators	Quantitative indicators used for trading or financial analysis.	financial indicators; trading indicators
Valuation Metrics	Relative or absolute valuation based on prices and accounting ratios.	valuation metrics; P/E; valuation ratios
Profitability	Firm earnings performance and operating cash generation.	profitability; earnings; cash flow
Balance Sheet	Financial position reflected in assets, liabilities, and capital structure.	balance sheet; assets; liabilities; capital structure
Business Efficiency	Operational efficiency and resource utilization.	efficiency; return; operating efficiency
Macroeconomic Indicators	Aggregate economic conditions affecting firms and markets.	macroeconomy; economic indicators
ESG & Governance Risk	Environmental, social, governance, and regulatory risks.	ESG; governance; regulatory risk
Shareholder Payouts	Distribution of firm value to shareholders.	dividends; share repurchases
Market Sentiment	Investor sentiment and market mood not directly tied to fundamentals.	market sentiment; investor sentiment

Appendix D Online Tables

Table A2: Tone Differences between Single- and Multi-Round Conversations

This table reports univariate comparisons of tone at the conversation level between single-round and multi-round conversations. The sample is restricted to the single-round conversations (Single) and the first round of multi-round conversations (Multi). Differences are computed as Single − Multi. t-statistics from two-sample t tests are reported in parentheses.

	Mean		
	Single	Multi	Single - Multi
Human tone (Prompt)	0.017	0.012	0.004
			(1.01)
AI tone (Response)	0.325	0.311	0.014
			(1.38)