Social Finance in the Age of AI: Evidence from Machine-Generated Content

Lin Peng Qiguang Wang Weichen Zhang Xiaoyan Zhang*

July 2025

Abstract

This paper examines how AI adoptions in financial discourse in online social media improve or harm market efficiency. Analyzing AI content on Seeking Alpha and Reddit's r/WallStreetBets from 2022-2024, we document stark differences in how AI is utilized and its market impact. On curated Seeking Alpha, authors use AI to overcome information frictions; AI content enhances price discovery and improves market quality. On unmoderated r/WallStreetBets, users turn to AI during high retail buying periods; AI contents stimulate more user discussions, amplify noise trading, and predict lottery-like returns. These findings demonstrate that platform governance and user motivations, not the technology itself, determines whether AI serves market efficiency or fuels speculation.

JEL Codes: G12, G14, G15

Keywords: Generative AI, Large Language Models, Financial Discourse, Retail Investors, Information Frictions, Market Microstructure, Social Media

*Lin Peng: Zicklin School of Business, Baruch College/CUNY (corresponding author, lin.peng@baruch.cuny.edu). Qiguang Wang: Hong Kong Baptist University, Hong Kong, China, email: qiguangw@hkbu.edu.hk. Weichen Zhang, Tsinghua University, Beijing, China, email: zhangwch.20@pbcsf.tsinghua.edu.cn. Xiaoyan Zhang, Tsinghua University, Beijing, China, email: zhangxiaoyan@pbcsf.tsinghua.edu.cn. We thank GPTZero for providing API access and technical support to detect AI usage.

Social Finance in the Age of AI: Evidence from Machine-Generated Content

July 2025

Abstract

This paper examines how AI adoptions in financial discourse in online social media improve or harm market efficiency. Analyzing AI content on Seeking Alpha and Reddit's r/WallStreetBets from 2022-2024, we document stark differences in how AI is utilized and its market impact. On curated Seeking Alpha, authors use AI to overcome information frictions; AI content enhances price discovery and improves market quality. On unmoderated r/WallStreetBets, users turn to AI during high retail buying periods; AI contents stimulate more user discussions, amplify noise trading, and predict lottery-like returns. These findings demonstrate that platform governance and user motivations, not the technology itself, determines whether AI serves market efficiency or fuels speculation.

JEL classification: G11, G14, G15.

Keywords: Generative AI, Large Language Models, Financial Discourse, Retail Investors, Information Frictions, Market Microstructure, Social Media

1 Introduction

Information is crucial of capital markets, and the cost of producing and processing information is a first-order determinant of market efficiency.[1] The recent advent of powerful generative article intelligence (AI), particularly Large Language Models (LLM) like ChatGPT, has suddenly compressed those costs, reshaping how investors produce, acquire, and process information (Blankespoor, Croom, and Grant 2024). At first glance, AI appears to lift the classic information-efficiency constraint.[2]

Yet regulators are unsettled. SEC Chair Gary Gensler warned that reliance on a handful of base models could create "monocultures" that breed systemic risk (Securities and Exchange Commission 2024b). The Financial Stability Board echoed this concern, arguing that "widespread use of AI models … could lead to increased correlation in trading, amplify market stress, exacerbate liquidity crunches, and increase asset-price vulnerabilities"(Financial Stability Board 2024). The U.S. Department of the Treasury likewise issued an alert on AI-assisted financial crimes and market manipulation (Financial Crimes Enforcement Network 2024). The emerging debate therefore turns on a simple but unresolved question: Does the adoption of AI in financial information market improve market efficiency, or does it drown markets in correlated noise and manipulation?

Answering that question is empirically difficult for two reasons. First, the net market effect is likely context-dependent. The same LLM tool that accelerates due-diligence for a stock analyst (Bertomeu et al. 2025) can, in the hands of fraudsters of a pump-and-dump scheme, fuel speculative herding (Securities and Exchange Commission 2024a). How can we disentangle the information-driven adoption from manipulative adoption? Second, AI adoption is still invisible in most disclosures; detecting it requires specialized text analysis.

To isolate the different ex-ante motives, we focus on information intermediaries—writers

[1] See Hayek (1945) and Fama (1970) for the critical role of information in financial markets, and Grossman and Stiglitz (1980), Diamond and Verrecchia (1981) and Kim and Verrecchia (1994) for the effect of information acquisition and processing cost on market efficiency.

[2] See Brynjolfsson, Li, and Raymond (2025), Noy and Zhang (2023), and Bertomeu et al. (2025) for evidence on AI reducing cost and improving output quality include.

and users who produce or disseminate public equity commentary—and exploit variation across online platforms with markedly different governance regimes and user characteristics. To this end, we collect over 130 thousands articles and messages from June 2022 to December 2024 on two influential social media platform: Seeking Alpha (SA), a curated portal that screens every submission, and Reddit's r/WallStreetBets (WSB), an almost unmoderated message board popular with retail traders. To overcome the second hurdle, we using GPTZero, a state-of-the-art commercial detector that outperforms all other models in our testing.[3]

Applying GPTZero to measure the probability of AI-genearated text, we uncover three stylized patterns. First, AI is not adopted at random: authors reach for it when information is costly to gather or when a rhetorical punch is most lucrative. Second, AI-generated commentary dampens open debate. AI contents draw fewer comments on SA while attracting more discussion on WSB, but user comments are less divergent, indicating that LLMs homogenise, rather than diversify opinions. Third, the market consequence hinges on platform governance and user base sophistication. On a curated venue like SA, AI functions as an information supplement and is associated with greater market efficiency. On an unmoderated forum like WSB, the same technology behaves as a noise amplifier, fuelling volume without improving liquidity and increasing the probability of lottery-like extreme return events. These patterns show that AI's impact is conditional, not intrinsic: whether it promotes efficiency or speculation depends on the rules of the marketplace in which it appears.

We document that the average AI probability rises from virtually zero in mid-2022 to roughly one-in-ten on SA towards the end of 2023 before trailing down and 7% on WSB by end-2024. These number are an order of magnitude larger than those documented in corporate filings (see Blankespoor, Croom, and Grant 2024). In other words, social-media pundits have adopted AI earlier and more aggressively than have corporate filers.

Consistent with strategic adoption for cost-reduction or productivity gains, we find SA authors are 22% more likely to use AI when covering a firm for the first time in the past 6 months. AI usage also increases during periods with sparse news flow. A one-standard-deviation de-

[3] Appendix Table A2 provides a detailed comparison of GPTZero against other mainstream detection methods found in the literature (e.g., Wu et al. 2025)

crease in news coverage is associated with a 0.32 percentage point higher AI probability, while a one-standard-deviation decrease in analyst following corresponds to a 0.39 percentage point increase. These effects are economically meaningful given SA's average AI probability of 10.6%. All of These are scenarios of high marginal cost of human research, where the benefits of AI is greatest.

However, these information-friction effects are far smaller or entirely disappear on WSB. Instead, we find evidence supporting rhetorical-engineering motives. AI probability is 46% to 55% higher relative to the sample average after high retail buying periods. Together, these patterns validate our premise that AI adoption is driven both by information-production frictions and by opportunities to shape audience sentiment.

How do users react to the AI content? We document a sizable, though modest, homogenizing effect of AI across the two platforms. A fully AI-generated content attracts 6%–7% less comments in the next ten days for SA but draw 7%–9% more comments on WSB. The conversation that does emerge is also more uniform: the cross-comment sentiment dispersion falls by roughly 10% relative to its mean, indicating fewer dissenting viewpoints. The platform differences suggest distinct mechanisms. On SA, AI may raise content quality through productivity gains, leading to clearer contents that generate consensus. On WSB, AI appears to streamline the crafting of punchy narratives that herd sentiment rather than enrich debate.

The market impact of AI content diverges sharply by platform. Moving to a stock-day level analysis, we first examine return-sentiment sensitivity. On SA, AI presence amplifies how markets respond to article sentiment on the same day. The return-sentiment sensitivity nearly quadruples on AI days. SA's typical 10-14 hour review ensures authors cannot react to contemporaneous returns. As a result, any return-sentiment sensitivity is more likely to reflects genuine information transmission. And we are more comfortable to interpret this enhanced sensitivity as AI helping synthesize and clarify information, thereby accelerating price discovery. We also find a similar pattern on WSB. However, the instant publication on WSB makes interpretation ambiguous: the correlation may simply reflect AI crafting narratives that chase intraday price movements rather than inform them. Notably, this heightened sensitivity appears only on day zero. We find no predictive power for future returns on either platform, confirming that AI primarily accelerates the incorporation of existing information rather than generating new alpha.

We observe more divergence when examining trading activities. On SA, AI presence is associated with improved market quality across multiple dimensions. Over the five days following publication, stocks with AI content experience a 2.0% standard deviation decrease in abnormal volume, a 3.0% decrease in volatility, and a 1.9% decrease in bid-ask spreads. These effects begin immediately—spreads already tighten by 1.3% on the publication day itself, suggesting market makers quickly recognize AI content as reducing information asymmetry. The pattern indicates AI helps traders reach consensus more efficiently, reducing the need for price discovery through trading.

On WSB, we observe the polar opposite. AI contents trigger immediate trading actitives: a 9.4% standard deviation surge in abnormal volume and 10.2% increase in absolute returns on day zero. These effects persist and intensify over the next week, with volume remaining elevated by 10.5% and volatility by 11.4%. Crucially, this frenetic activity provides no liquidity benefit—spreads remain unchanged or widen depending on regression specifications. This combination of higher volume and volatility without improved liquidity is consistent with of noise trading.

While these trading patterns on WSB align with noise trading, they could theoretically arise from informed trading as well. Information shocks can also generate increased volume and volatility as prices adjust to new fundamentals and widen bid-ask spread. To provide more direct evidence of the detrimental role of AI contents on WSB, we extreme return events. We find that AI content strongly predicts lottery-like outcomes that are inconsistent with informed trading. The odds of MAX events (daily return reaching 21-day highest) increases by 26.9% on AI publication days and remains elevated by 17.2% over the next five days. For lottery events—MAX events with 21-day maximum return also ranking in the top decile market-wide—the effects are even more dramatic: AI presence increases the same-day odds by 57.0% and next-five-day odds by 31.9%. On SA, AI shows no association with either MAX or lottery events.

Taken together, these results show that platform governance is crucial in determining the market implications of AI content. SA's editorial review process, which screens all submissions for quality regardless of how they were created, ensures that any AI-assisted content that makes it through serves an informational role. WSB's unmoderated publication allows AI to become a tool for sentiment amplification, generating noise trading that persists for days and culminates in lottery-like payoffs that attract retail speculators.

Our paper makes several contributions to the rapidly evolving literature on AI in financial markets. First, we extend the study of AI's financial impact by focusing on its growing presence within online social media platforms. While prior research has begun to explore AI in formal corporate disclosures (Blankespoor, Croom, and Grant 2024) and professional analyst reports (Bertomeu et al. 2025), our work provides one of the first large-scale, comparative empirical analyses of AI's footprint in the more dynamic and often less regulated sphere of public equity commentary on retail-facing forums.

Second, we provide evidence on the micro-foundations of AI adoption and its immediate consequences for user engagement. We empirically demonstrate *why* and *when* content creators on these platforms turn to AI on each platform. These platform-specific motivations add nuance to existing research on productivity effects of AI (e.g., Brynjolfsson, Li, and Raymond 2025; Noy and Zhang 2023). Moreover, we show that AI-enabled content consistently leads to a homogenization of discourse. The evidence aligns with broader findings on LLMs' tendency to reduce textual diversity (Anderson, Shah, and Kreminski 2024; Padmakumar and He 2024) and the potential for audiences to discount AI-generated text (Cong et al. 2024; Plate, Voshaar, and Zimmermann 2025).

Last, and critically, our study is among the first to empirically establish and quantify the decisive moderating role of platform governance in shaping the market impact of AI. While theory suggests that lower information costs should enhance market efficiency (e.g., Grossman and Stiglitz 1980; Diamond and Verrecchia 1981), and some studies find AI enhancing informational efficiency in curated settings (Kim, Muhn, and Nikolaev 2025; Bertomeu et al.

2025), concerns about AI-induced noise, herding, and manipulation persist (Dou, Goldstein, and Ji 2025). Our cross-platform analysis directly addresses this tension. The association between AI and lottery-like events on WSB but not SA provides the cleanest evidence that platform governance determines whether AI serves price discovery or market manipulation.

In this regard, our work is contemporaneous with and complements Bradshaw et al. (2025), who also examine AI use on Seeking Alpha. While their study provides valuable insights into AI adoption by information intermediaries on that specific platform, our paper offers several distinct contributions. Foremost, we conduct a cross-platform analysis comparing SA with WSB. This comparative approach allows us to isolate the impact of platform governance and user sophistication on AI's role, documenting heterogeneous adoption motivations and divergent effects on market efficiency.

2 Hypothesis Development

In a market with information frictions, any mechanism, including AI assistance, that (i) lowers author's production cost or (ii) alters the way readers process the content can reshape the both the discussion patterns and price and trading dynamics. Building along this line of thought, we structure our hypotheses around three core questions: (1) Why do content creators turn to AI? (2) How do users engage with and interpret AI-generated financial content? (3) How does the presence of AI-generated content on different platforms influence market outcomes? Throughout, we emphasize that the same AI shock can manifest differently across platforms that differ in editorial governance and user sophistication and trading motive.

2.1 Adoption of AI in content creation

We first investigate the motivation of AI adoption in content creation. Recent studies show that AI significantly lowers the cost and time required to produce textual content for customer-support agents and professional writers, with this benefit being more pronounced when humans lack prior knowledge or experience (e.g., Brynjolfsson, Li, and Raymond 2025; Noy and

Zhang 2023). In the context of financial content creation, we posit that AI's ability to lower information production costs becomes most appealing when unassisted human research would be particularly expensive. This occurs when authors cover unfamiliar firms or industries, or when public news is scarce and producing novel insights independently is costly. In these scenarios, AI flattens the learning curve and enables publication that might otherwise be uneconomical. We therefore expect AI adoption to be greater in first-time coverage settings and during periods of low firm-specific news flow.

Beyond cost savings, AI also allows authors to engineer stylistic and emotional attributes that persuade and resonate with specific audiences (Matz et al. 2024; Cong et al. 2024). LLMs make it effortless to incorporate emotional language, dramatic analogies, and meme-worthy phrases. This rhetorical engineering is particularly valuable on lightly moderated platforms where contents face no pre-publication review and users' trading responds more to compelling narratives than to fundamental analysis (Shleifer and Summers 1990; Barber et al. 2022; Hirshleifer, Peng, and Wang 2025; Bali et al. 2025). We thus hypothesize that AI adoption increases when sentiment coordination offers the greatest payoff—after sharp price run-ups, during periods of high volatility, or for potential "meme" stocks. In other words, author may use AI in a strategic attempt to induce herding rather than to convey new information. Based on these considerations, we propose our first hypothesis:

Hypothesis 1a *The adoption of AI for creating financial content is driven by authors' efforts to overcome information production costs and information frictions, particularly when addressing unfamiliar topics or operating in information-scarce environments.*

Hypothesis 1b *Authors employ AI to craft stylistic and emotional cues that resonate with platform-specific audiences, with adoption most pronounced when sentiment coordination is valuable.*

2.2 Engagement and Interpretation of AI-Generated Content

Recent evidence indicates that LLM assistance tends to homogenizes text, reducing lexical and semantic diversity. Controlled studies show higher cross-writer similarity when AI is adopted for creative ideation (Anderson, Shah, and Kreminski 2024), essay writing (Padmakumar and

He 2024), and Q&A tasks (Reviriego et al. 2024), and that AI autocomplete pushes authors of diverse cultural backgrounds toward a common, Western-centric style (Agarwal, Naaman, and Vashistha 2025). Finance research echoes this pattern: ChatGPT-edited loan applications converge on polished templates (Cong et al. 2024), and practitioners fear AI will compress investment commentary into "essentially a single view" (Bradshaw et al. 2025).

Such textual homogenization could lead to content that offers fewer novel angles or controversial points that typically stimulate vibrant discussion and debate (Berger and Milkman 2012; Vosoughi, Roy, and Aral 2018). As such, AI content could lead to reduced engagement and lower disagreement.

Alternatively, as audiences become adept at identifying AI-generated content, they may discount it as less credible or authentic, leading to less engagement. Evidence is already emerging: lenders on a large peer-to-peer platform rely less on loan narratives displaying ChatGPT-like homogeneity once that pattern becomes detectable (Cong et al. 2024), and investors react negatively when MD&A disclosures exhibit markers of generative AI usage (Plate, Voshaar, and Zimmermann 2025). Combined, we posit that AI-generated financial commentary will elicit less user interaction and debate.

Hypothesis 2a *AI-generated financial content leads to lower levels of user engagement (e.g., fewer comments) and reduced disagreement (e.g., lower sentiment dispersion in comments) compared to human-written content.*

However, this perspective may not fully capture the dynamics of sentiment-driven communities, where engagement is often driven by social reinforcement rather than pure information seeking. Foundational work in psychology identifies a powerful "confirmation bias," the tendency for individuals to seek out and favor information that confirms their pre-existing beliefs (Nickerson 1998). This bias gives rise to selective exposure, where individuals actively choose to consume media that aligns with their views, often to manage their mood and reinforce their self-concept (Knobloch-Westerwick 2014).

The most relevant evidence comes from Cookson, Engelberg, and Mullins (2023), who show that retail investors on financial social media form "echo chambers," disproportionately following and interacting with like-minded peers. Engagement within these chambers is not just about processing information but about achieving social validation. Content that reinforces the dominant group narrative is psychologically rewarding and thus stimulates higher levels of interaction. Given that AI can be used to effortlessly create and scale precisely the kind of punchy, persuasive, and highly-aligned narratives that thrive in such echo chambers, it may act as a catalyst for engagement on certain platforms. This leads to a competing, platform-specific hypothesis:

Hypothesis 2b *AI-generated content designed to reinforce prevailing community narratives will lead to higher user engagement. By facilitating the creation of confirmatory content, AI acts as a tool for social validation and strengthens the echo chamber effect, thereby increasing comment volume.*

2.3 Market Impact of AI-Generated Content

The effect of AI-generated content on stock markets is unlikely to be uniform. We argue that market impact is critically moderated by platform governance structures and the characteristics of their user bases, which lead to divergent outcomes even when article-level reactions are similar.

On platforms like Seeking Alpha, which employ editorial screening and have policies against undisclosed AI use (Bradshaw et al. 2025), any AI-assisted content that passes through is likely to be of higher quality or at least align with the platform's standards. Users on such platforms tend to be more sophisticated and information-driven (Cookson et al. 2024), and likely perceive AI assistance as lowering information processing costs or providing clear summaries. Theoretically, lowering information production or acquisition cost improves price informativeness and market efficiency (Grossman and Stiglitz 1980; Diamond and Verrecchia 1981; Kim and Verrecchia 1994). Empirically, greater public information supply in more easily-accessible format accelerate price discovery (Chang et al. 2022; Luo et al. 2023), and AI summaries of lengthy financial reports greatly reduce information processing cost and lead to tighter bid-ask spread (Kim, Muhn, and Nikolaev 2025).

Moreover, AI can directly enhance informativeness of financial content. Bertomeu et al. (2025) provide compelling evidence that AI usage by stock analysts significantly improves forecast accuracy, which in turn enhances informational efficiency and market quality. This finding is particularly relevant to our setting, as contributors on platforms like Seeking Alpha share key characteristics with professional analysts—both prioritize fundamental analysis and aim to provide valuable insights. Taken together, we hypothesize that on a curated platform where such content predominates, AI assistance leads to lower speculative trading, reduced return volatility, tighter bid-ask spreads, and greater price informativeness.

Hypothesis 3a *On platforms with strong editorial oversight and a user base inclined towards fundamental analysis (e.g., Seeking Alpha), AI-generated content improved market efficiency.*

However, on largely unmoderated platforms like WSB, the ease of generating and disseminating AI content can be exploited to create and amplify narratives, regardless of their fundamental accuracy. AI assisted content could facilitate a shallow consensus around a particular sentiment or meme, rather than a deep, shared understanding of fundamentals. This can lead to increased trading volume, particularly retail churn (Hirshleifer, Peng, and Wang 2025) or herding (Barber et al. 2022), driven by amplified sentiment rather than new information. We also expect no improvement, or even a deterioration, in market liquidity because the herding by unsophisticated investors can create inventory risk and harm market liquidity (Eaton et al. 2022).

Hypothesis 3b *On platforms with minimal editorial oversight and a user base more susceptible to sentiment and noise trading, AI-generated content will act as a noise amplifier, potentially degrading market quality for affected stocks.*

3 Data

3.1 Social Media Financial Contents

Our primary data is the textual contents collected from two distinct online platforms from June 2022 to December 2024. This time frame is selected to encompass the period following the launch of ChatGPT in November 2022 and to include the most recent data available for the study. From Seeking Alpha, a platform recognized for its crowd-sourced equity research contributed by semi-professional analysts, we gather all articles published within the "Analysis" section. Consistent with the platform's content categorization, items identified as earnings call transcripts or corporate presentation slides were excluded to only use original analytical work. To study the stock market implications, we follow the literature to only select (1) single-ticker articles (Chen et al. 2014), or (2) articles with a valid primary ticker, which identifies the main focus of the article (Campbell, DeAngelis, and Moon 2019).

From Reddit's r/WallStreetBets forum, a popular venue for retail investor discussion known for its distinct culture and shorter-form content, we collect all submissions (original posts) and subsequent comments over the same period. We deliberately pool submissions and comments in our analysis for several reasons that reflect the fundamental nature of financial discourse on Reddit. First, unlike traditional publishing platforms where content follows a clear hierarchy, Reddit operates through fluid, conversational threads where comments frequently contain substantive analysis that rivals or exceeds the original post. Users routinely post detailed analysis or market insights in comments, blurring the distinction between primary and secondary content. Second, this pooling captures how information actually propagates on WSB—through rapid exchanges rather than standalone posts. Restricting analysis to submissions alone would miss the majority of the platform's financial discourse and misrepresent how retail investors actually communicate. Third, the decision to use AI assistance likely reflects similar motivations whether crafting a submission or comment, particularly given our focus on rhetorical engineering versus information production.

To ensure the relevance and substance of the WSB data, we apply several filters. First, we retain only single-ticker messages explicitly mentioning one valid stock ticker, excluding common abbreviations or terms that incidentally match tickers but denote different meanings. Second, we restrict the WSB sample to messages exceeding 50 words in length. This threshold serves a dual purpose: it filters out very brief, potentially uninformative remarks, and critically, it ensures sufficient text length for reliable processing by our chosen AI detection tool, GPTZero, which requires a minimum input of 250 characters (approximately 40-45 words). The 50-word minimum provides a conservative buffer above this technical requirement.

Our filtering yields 57,581 SA articles and 64,848 WSB messages (including both submissions and comments) for analysis. Notably, there are only 6,487 WSB submissions, accounting for 10% of the total WSB messages. Thus, there are substantive financial discussions on WSB comment section than the original posts. In Section 7, we demonstrate that our results remain robust to alternative sample definitions: first, when analyzing submissions only, and second, when including all WSB messages regardless of length (assigning zero AI probability to those below 50 words). Both robustness checks confirm that our findings reflect fundamental platform differences rather than these methodological choices.

Each filtered SA article and qualifying WSB messages is processed using the GPTZero. GPTZero is a leading commercial AI content detection tool that employs a seven-component model to analyze text at sentence, paragraph, and document levels.[4] The algorithm classifies text into three categories AI-ONLY, HUMAN-ONLY, or MIXED, where MIXED refers to human with AI assistance, and provides corresponding probability scores that sum to one. GPTZero classifies an article into the category with the highest probability. Hereby, we refer to AI contents as those classified as either AI-ONLY or MIXED.

We calculate the fraction of AI contents in a month on each platform and present the time trend in Figure 1. Following ChatGPT's introduction, we observe a sharp increase in AI content on SA, with the fraction peaking at 10% towards the end 2023. However, an abrupt decline begins thereafter, with the fraction falling to 3% at the end of 2024. The drop in the AI usage on SA coincides with an editorial prohibition on AI in article submissions. Bradshaw et al. (2025) provide an in-depth discussion of this ban and study its implications.

In contrast, the trajectory on WSB is much more gradual. AI contents account for less than 2% of monthly submissions through the first half of 2023, then climb steadily to roughly 7% by December 2024. The slower increase is possibly due to the nature of the WSB content,

[4]In the online appendix, we provide detail comparison of GPTZero against other AI detectors using a corpus of human-written and AI-generated financial contents.

which tends to be much shorter, slang-laden, and meme-centric. In these format, human spontaneity is already cheap and AI's benefit is not immediate.

3.2 Variables Construction

We collect financial market data from a number of sources. We obtain daily stock returns, trading volume, and pricing information from the Center for Research in Security Prices (CRSP), firm-level accounting from Compustat. Information regarding quarterly institutional ownership is from the Thomson Reuters Institutional Holdings (13F) database. Analysts coverage is from Institutional Brokers' Estimate System (IBES). To construct measures pertinent to market microstructure, we use data from the Trade and Quote (TAQ) database. We construct variables at two levels: the individual article/message level and the aggregated stock-day level (summed over individual articles/messages for the same stock on the same day). Thereby, we focus on the key AI variables and defer the detailed definitions of all other variables to Appendix Table B..3.

At the article/message level, the primary variable is the continuous AI probability score:

$$AIProb_{ijt} = \Pr(\text{AI-ONLY}_{ijt}) + \Pr(\text{MIXED}_{ijt}). \tag{1}$$

$AIProb_{i,j,t}$, based on the content of article j about firm i published on day t, is the sum of probability scores for the AI-ONLY and MIXED categories. To illustrate, if an article j has an AI-ONLY probability of 0.3, MIXED probability of 0.1, and a HUMAN-ONLY probability of 0.6, its $\text{AIProb}_{i,j,t}$ would be 0.4.

We calculate a number of article characteristics, including the word count (length), sentiment, textual complexity, Fog Index, count of numbers (quantitative), and count of images (graphical). Post-publication engagement is measured by the number of user comments received on a content from its publication day through 10 days post-publication, and user reactions are measured by user disagreement, calculated the standard deviation of sentiment scores from user comments on the content during the same 10-day period, using the respective platform-specific sentiment dictionaries.

Another key methodological consideration is matching articles to trading days. We assign articles to trading days as follows: articles published before 4:00 PM on a trading day are matched to that day, while articles published after 4:00 PM or on non-trading days are matched to the next trading day.

At the stock-day level, our primary measure of AI exposure for firm i on trading day t is an indicator variable

$$AIDay_{it} = \mathbb{1}\{\#\,\text{AI article}_{it} > 0\}. \tag{2}$$

Thus, AIDay flags days with at least one AI-ONLY or MIXED articles. This daily AI presence indicator serves as a key independent variable in our market-level analyses. The corresponding dependent variables include a number of market outcomes measured subsequent to the content's publication, typically starting on day $t + 1$. These include cumulative abnormal returns (CAR), abnormal trading volume (AVOL), effective bid-ask spreads (Spread), and realized volatility (Vol).

We present the summary statistics in Table 1, with panel A for the article-level sample and panel B for the stock-day sample. Our article-level sample consists of 57,729 SA articles and 65,148 WSB submissions over the period from December 2022 to December 2024. Due to the cross-platform nature of our study, we focus on the differences in these variables across SA and WSB. The last column in panel A shows that all mean differences across the two samples are statistically significant at the 1% level. Textual characteristics reveal fundamental differences between platforms. SA articles have nearly three times higher AI probability (10.6% vs 3.6%), are substantially longer (1,454 vs 114 words), more positive in sentiment, more complex, and contain more visual elements. These patterns align with SA's role as a platform for in-depth, analytical content produced by semi-professional authors, contrasting with WSB's focus on shorter, narrative-driven contents.

Engagement patterns also differ dramatically across platforms. SA articles attract significantly more comments (11.3 vs 0.3) and generate higher disagreement levels, pointing to more intensive debate. The lower engagement and disagreement on WSB aligns with theories of financial social media "echo chambers," where users interact with like-minded peers, leading to social reinforcement rather than substantive debate. Firm characteristics show that SA covers smaller companies ($146B vs $344B) with higher book-to-market ratios (0.46 vs 0.38) and greater institutional ownership (71% vs 65%), while WSB focuses on larger, more visible firms with greater analyst coverage (5.1 vs 4.0 analysts). These differences, while statistically significant, are economically modest. Author activity patterns are stark. SA authors are far more prolific (49.5 vs 4.7 articles over six months) and generate substantially more user comments per article. This reinforces the view of SA contributors as dedicated, semi-professional analysts compared to WSB's more casual user base.

Panel B of Table 1 provides a stock-day level summary statistics. The SA article covers significantly more stocks compared to WSB (3,193 vs 1,351), resulting in a larger stock-day sample size. AI days occur in only 0.24% of SA stock-days and 0.18% of WSB stock-days. On average, SA generates 3.59 articles per 100 stock-days compared to WSB's 9.43 articles per 100 stock-days, indicating more concentrated posting activity on WSB when content is published. The results on firm characteristics confirms that SA coverage gravitate toward smaller firms with higher book-to-market ratios and less analyst coverage, whereas WSB focuses on larger, more visible companies. A key distinction at this level is market liquidity. Stocks discussed on SA exhibit significantly wider bid-ask spreads on average over next five days ($\text{Spread}_{[t+1,t+5]}$: 0.41% vs. 0.29%).

Taken together, the summary statistics in Table 1 establish SA and WSB as distinct ecosystems with different informational and liquidity environments, laying the groundwork for our analysis of how this context shapes the market impact of AI.

4 AI Adoption

In this section, we empirically investigate the motivations behind AI adoption in financial content creation, as outlined in Hypotheses 1a and 1b. We examine how article characteristics, author familiarity with the subject matter, and the prevailing information environment influence the likelihood of AI usage.

4.1 Article Characteristics and AI Content

To understand how the fundamental characteristics of an article relate to the likelihood of AI assistance, we estimate the following regression:

$$AIProb_{ijt} = \gamma' \boldsymbol{X}_{ijt} + \epsilon_{ijt} \tag{3}$$

where AIProb_j is the AI probability score from GPTZero for content j on stock i at time t. $\mathbf{X}_j$ include various textual features such as length, sentiment, complexity, Fog Index, image count, and number count. For the WSB sample, we include a dummy variable to indicate whether the content is a submission or a comment. All non-dummy independent variables are standardized to zero mean and unit variance. As a result, the coefficients can be interpreted as the percentage points change in AI probability for a one-standard-deviation change in a independent variable. Standard errors are two-way clustered by stock and day.

The results, presented in Table 2, reveal different patterns for SA and WSB. On SA (columns 1), articles that are shorter, with more positive sentiment, and have higher complexity and Fog Index scores tends to have a higher AI probability. Conversely, a higher proportion of images or numbers is associated with a lower AI probability. Economically, a one standard deviation increase in article sentiment score, complexity, and Fog index are associated with a 2.94, 1.87, and 1.77 percentage point (pp) increase in AI probability, respectively. In column 2, we further add author fixed effects and theses effects remain robust except for article length, which becomes insignificant. These findings suggest that on SA, AI might be used to generate more complicated and positively-toned narratives, but less so for content rich in specific visual or numerical data.

On WSB (columns 3–4), AI content is associated with longer messages, more positive sentiment, higher complexity, a higher Fog Index, and, unlike SA, a higher proportion of numbers. The strong positive association with word count on WSB could reflect a self section issue: users intending to produce longer, more elaborate messages may be more inclined to leverage AI to assist with content generation. A submission on WSB is associated with a 193% higher AI probability score than a comment, relative to the sample mean. This finding suggests that users specifically rely on AI tools for higher-effort tasks like drafting a substantive, standalone submission.

These results provide initial support for our first hypothesis. For example, the associations with complexity and positive sentiment are consistent with rhetoric-engineering motives in Hypothesis 1b. Alternatively, authors are more likely to resort to AI when dealing with more challenging task, so the evidence also aligns with cost/productivity driven motives in Hypothesis 1a. However, given the contemporaneous nature of these regressions, we acknowledge that these characteristics might either motivate the use of AI or be a byproduct of its application.

Table 2 also reveals a dramatic increase in explanatory power when author fixed effects are included. The R^2 jumps from 10.4% to 59.9% on SA and 57.4% on WSB, suggesting that AI adoption is primarily driven by unobserved, time-invariant author characteristics.

4.2 Information Production Cost and AI Usage

To test whether authors turn to AI when facing information frictions due to unfamiliarity (Hypothesis 1a), we examine the relationship between AI probability and proxies for an author's inexperience in covering a specific firm or industry. We estimate:

$$AIProb_{ijt} = \beta_1 \cdot First\ Time\ Firm_{ijt} + \gamma' \boldsymbol{X}_{ijt} + \epsilon_{ijt}, \tag{4}$$

where $First\ Time\ Firm_{ijt}$ is an indicator variable that equals one if the author/user covers the stock i for the first time in the past six months.[5] $\boldsymbol{X}$ include stock size, book-to-market ratio and institutional ownership. We also include the author's past six-month activities, such as the number of articles, the number of AI articles, and the per-article average number of comments

[5] We test the effect of author unfamiliarity by using first-time industry coverage and find similar results. See Appendix Table A3.

received during the first 20 days after publication. In addition to stock and day fixed effects, we also add author fixed effects to control for time-invariant heterogeneity.Standard errors are three-way clustered by stock, day, and author.

The results in Table 3 strongly support Hypothesis 1a. On Seeking Alpha (column 1), first-time firm coverage is associated with a 2.32 pp higher AI probability ($t = 9.42$). This effect is economically significant relative to an mean of 10.65 pp on SA: when venturing into new domains, authors are 22% more likely to use AI. Given that AI adoption exhibit strong author heterogeneity (e.g., Table 2), we include author fixed effects in column 2 and find that the effect of first-time firm coverage remains positive and robust, albeit with a smaller magnitude.

In contrast, on WSB (columns 3–4), there is no statistically significant increase in AI adoption when contributors cover firm for the first time. This divergence from the SA findings suggests that AI adoption on WSB, even when contributors encounter unfamiliar topics, is not primarily driven by an effort to overcome information production costs or information frictions, as proposed in Hypothesis 1a.

These findings are consistent with the notion that AI flattens the learning curve, especially in SA, and reduces the marginal cost of research, making it a valuable tool when authors lack deep prior knowledge. This aligns with studies like Brynjolfsson, Li, and Raymond (2025) and Noy and Zhang (2023), which find that AI's productivity benefits are often more pronounced for tasks where individuals have less experience.

We further test Hypothesis 1a by examining whether AI adoption increases during periods of firm-specific information scarcity. We measure information availability using the daily count of high-relevance news items from RavenPack and proxy ex-ante information supply by the number of analysts following the stocks. We then estimate the following specification from Table 4:

$$AIProb_{ijt} = \beta_1 News_{it} + \beta_2 Analyst_{it} + \gamma' \boldsymbol{X}_{ijt} + \epsilon_{ijt}, \tag{5}$$

where $News_{it}$ is the logarithm of one plus the number of Dow Jones news articles with a relevance score above 90 concerning firm i in the last three days (sourced from Ravenpack), and $Analysts$ is the logarithm of one plus the number of analysts following the stock i in the most recent quarter end.

The results in Table 4 provide additional support for Hypothesis 1a. In the Seeking Alpha sample (columns 1-2), we find statistically significant negative associations between news coverage and AI probability, and between analyst following and AI probability. Specifically, a one-standard-deviation increase in news coverage is associated with a 0.32 to 0.21 percentage point decrease in AI probability, while a one-standard-deviation increase in analyst following corresponds to a 0.73 to 0.33 percentage point decrease in AI probability. These findings suggest that SA contributors are more likely to utilize AI when external information sources are limited, consistent with AI serving as a substitute for costly information acquisition when the marginal cost of original content production is high.

In contrast, for content on WSB (columns 3-4), we find no statistically significant relationship between news coverage and AI probability. Interestingly, analyst following exhibits a positive association with AI probability on WSB (though statistically significant only in column 3), which is opposite to the negative relationship observed in the SA sample.This divergence suggests that information scarcity is a less critical driver of AI adoption for WSB users, whose content generation might be more influenced by ongoing narratives or sentiment (Hypothesis 1b). The SA finding resonates with the idea that AI can help "fill the void" when traditional information sources are quiet.

4.3 Rhetorical Engineering and AI Adoption

Beyond cost-saving, we test Hypothesis 1b, which posits that authors employ AI for "rhetorical engineering"—strategically crafting stylistic and emotional content to persuade audiences, particularly when sentiment coordination is most valuable. This motive is expected to be most pronounced on lightly moderated platforms like WSB, where compelling narratives can drive trading behavior more than fundamental analysis.

To test this directly, we examine whether AI adoption correlates with surges in retail investor attention. We estimate the following regression.

$$AIProb_{ijt} = \beta_1 \cdot \textit{Top 10\% OIB}_{it} + \gamma' \boldsymbol{X}_{ijt} + \epsilon_{ijt}, \tag{6}$$

where *Top 10% OIB*$_{it}$ is an indicator that identifies if a stock's retail order imbalance over the preceding three days ranked in the top 10% cross-sectionally. We classify retail orders using the algorithm in Boehmer et al. (2021) and calculate the retail order imbalance (OIB) as the number of shares bought minus the number of shares sold by retail investors, scaled by the sum of the two. A spike in retail OIB serves as a clear proxy for a stock becoming a focus of intense retail sentiment and trading activity.

The results, presented in Table 5, provide strong, platform-dependent support for our hypothesis. For messages on WSB (columns 3–4), the coefficient on high retail order imbalance is positive and highly significant, both statistically and economically. A surge in retail buying predicts a 1.67 to 1.97 percentage point increase in a content's AI probability, an effect corresponding to a 46% to 55% rise relative to the platform's sample average. This indicates that WSB users are substantially more likely to use AI when a stock is already experiencing a wave of retail-driven activity.

In stark contrast, this effect is entirely absent on Seeking Alpha. As shown in columns 1 and 2, the coefficient for high retail order imbalance is statistically insignificant for SA articles. This divergence provides compelling evidence that the drivers of AI adoption are fundamentally different across platforms. While the strategic use of AI for rhetorical engineering is a key factor on a sentiment-driven forum like WSB, it does not motivate content creators on the more fundamentals-oriented SA platform.

5 User Engagement

5.1 AI Content and Subsequent User Comments

To assess the impact of AI-generated content on user engagement, we use the cumulative number of comments a content receives. We estimate the following regression model:

$$\log(1 + Comments_{ij[t,t+10]}) = \beta_1 \cdot AIProb_{ijt} + \gamma' \boldsymbol{X}_{ijt} + \epsilon_{ijt}, \tag{7}$$

where $Comments_{ijt[t,t+10}$ is the cumulative number of comments received for content j (on stock i by user u) between its publication on day t and day $t+10$. We select the cutoff day of $t+10$ because 85% of comments are received within 10 days after publication. Controls include the article characteristics used in Table 2, firm characteristics, and author past activities performance. Standard errors are two-way clustered by stock and day.

The results, presented in Table 6, show platform-specific associations between the AI probability of initial content and the volume of subsequent comments. On Seeking Alpha (columns 1–2), the findings suggest a negative relationship between AI probability and user engagement. An AI-generated article (AIProb = 1) has 6% fewer comments compared to a human-generated article (AIProb = 0). Conversely, on WSB (columns 3–4), the results consistently show an opposite and statistically significant trend: a higher AI probability in the initial content is associated with an increase in subsequent comments. A fully AI-generated message is associated with 7% to 9% more comments than a fully human-generated message.

These findings on Seeking Alpha lend support to Hypothesis 2a. The reduction in comments could stem from several factors outlined in our hypothesis development: AI-generated content, due to its potential homogenization (Anderson, Shah, and Kreminski 2024; Padmakumar and He 2024) or perceived completeness, may offer fewer novel angles or controversial points that typically stimulate discussion (Berger and Milkman 2012). Alternatively, if users are becoming adept at identifying and discounting AI content as less authentic or credible (Cong et al. 2024; Plate, Voshaar, and Zimmermann 2025), this could also lead to reduced interaction.

The opposite result on WSB provides strong support for Hypothesis 2b. The significant increase in comment volume for AI-generated messages on this platform aligns with the theory of echo chambers and social reinforcement (Cookson, Engelberg, and Mullins 2023; Nickerson 1998). Rather than presenting complex, debatable information, AI-generated content on WSB can be optimized to produce persuasive, meme-friendly narratives that confirm the community's prevailing sentiment. The higher comment volume is therefore consistent with AI being used as a tool to amplify sentiment and strengthen community consensus, as predicted.

5.2 AI Content and Subsequent User Disagreement

Next, we examine whether AI-generated content is associated with lower disagreement or sentiment dispersion among users. We measure disagreement as the standard deviation of sentiment scores in the comments received following the initial content. The regression specification, as detailed in Table 7, is:

$$Disagreement_{ij[t,t+10]} = \beta_1 \cdot AIProb_{ijt} + \gamma' \boldsymbol{X}_{ijt} + \epsilon_{ijt}, \tag{8}$$

where $\text{Disagreement}_{ij[t,t+10]}$ is the sentiment dispersion in comments for content j on stock i over the window $[t, t+10]$, standardized to zero mean and unit variance. As before, AI Prob_{ijt} is the AI probability of the initial article.

The results in Table 7 provide support for the hypothesis that AI content reduces user disagreement. On SA (column 1–2), the coefficient on AI-related probability is consistently negative and statistically significant. In the specification with day, stock, and author fixed effects, the coefficient is -0.004. Economically, an AI-generated article is associated with 0.004 standard-deviation lower in disagreement, which is a decrease of 5% relative to the mean of disagreement. On WSB (column 3–4), we observe a similar disagreement-reducing effect, although it is not statistically significant.

These findings suggest that AI-generated content leads to more homogeneous expressed opinions. This aligns with the idea that LLM-produced text might be less nuanced or present fewer contentious points, thereby reducing the scope for divergent interpretations (Bradshaw et al. 2025). The convergence of opinion could also be a consequence of the reduced engagement observed earlier; if fewer people comment, and those who do are perhaps less inclined to express strong or outlying views in response to AI-like text, then measured disagreement would naturally fall. This evidence supports the notion that AI content, whether through its intrinsic properties or user perception, tends to narrow the spectrum of subsequent discussion.

6 Market Impact

6.1 AI Content and Price Reactions

To understand the market impact of AI-generated content, we first examine AI contents' potential informational role. If AI helps by synthesizing existing information, then its presence should accelerate how quickly the concurrent information gets incorporated into stock prices.

To test this, we use a stock-day level analysis. We expect that on days when AI-generated articles about a stock are published, the market will react more quickly to the sentiment of those articles. We run the following regression:

$$\begin{aligned}\mathrm{AR}_{it}/\mathrm{CAR}_{i[t+1,t+5]} =& \beta_1 \mathit{AIDay}_{it} \times \mathit{Sentiment}_{it} + \beta_2 \mathit{AIDay}_{it} + \beta_3 \mathit{Sentiment}_{it} + \\ & \beta_4 \mathit{Attention}_{it} + \gamma' \boldsymbol{X}_{it} + \epsilon_{it},,\end{aligned} \tag{9}$$

where AR_{it} is the DGTW-adjusted abnormal return on day t (Daniel et al. 1997) and $\mathrm{CAR}_{i[t+1,t+5]}$ is the cumulative abnormal return from $t+1$ to $t+5$. The main variable of interest is the interaction between AIDay_{it}, an indicator for at least one AI articles for stock i on day t, and $\mathit{Sentiment}_{it}$, the average sentiment of all content for that stock on that day.[6] We control for

[6] Following Cookson et al. (2024), we set the daily average sentiment to a neutral score of zero if no content is published about the stock.

the total number of articles ($Attention_{it}$) as a proxy for social media attention (e.g., Cookson et al. 2024) and a vector of stock controls, $\boldsymbol{X}_{it}$, including size, book-to-market ratio, institutional ownership, analyst coverage, an earnings announcement dummy, and past one-month abnormal return and volatility.

Table 8 shows the results. On Seeking Alpha (column 1) with the same-day abnormal return as the dependent variable, the interaction term is positive and significant. This result shows the market reacts more strongly to article sentiment when AI-generated content is present. Specifically, the return-sentiment sensitivity on AI days is 0.162 ($0.120 + 0.045$), nearly four times the size on non-AI days (0.045). This supports Hypothesis 3a: on a curated platform, AI-generated content helps investors process information faster, accelerating price discovery.

This effect is also present on WSB, with even stronger economic magnitude (column 3). The combined model in column (3), which runs a direct "horse race" between the platforms, confirms this pattern. The interaction for both SA and WSB remains significant. At first glance, this seems to contradict our hypothesis that AI content would enhance informational efficiency only on curated platforms.

However, a critical difference in publication timing resolves this puzzle. On SA, articles undergo editorial review that typically takes 10–24 hours.[7] This delay ensures authors cannot react to "contemporaneous" returns in our test. Articles submitted during trading hours undergo review and publish after market close, appearing on the next trading day. Articles submitted after hours (4pm to 9:30am) could be published within the same same trading day (a 24-hour window from close to close) but miss entirely that trading day's price movements. In both cases, authors write without knowing the publication day's price movements. The positive interaction we observe therefore represents genuine information transmission —the market responding to AI-enhanced analysis of prior information.

On WSB, instant publication eliminates this temporal separation. Authors can observe

[7] Most articles are reviewed or published within 10–14 hours on weekdays. For submission on weekends or by new authors, the review process can take several days. See SA submission guidelines `https://about.seekingalpha.com/article-submission-guidelines` and Q&A `https://help.seekingalpha.com/contributors/how-long-does-it-take-to-publish-my-article`.

intraday price movements and immediately post AI-generated narratives that explain or amplify them. Given our earlier evidence that WSB users strategically adopt AI during high retail trading periods (Table 5), the positive interaction likely reflects AI being used to chase prices rather than inform them.

The lack of predictive power for future returns on both platforms (columns 2, 4, and 6) confirms that AI content primarily accelerates the incorporation of existing information rather than generating alpha. But only on SA's curated environment can we confidently interpret this as information flowing from content to prices. On WSB, the same statistical pattern may simply capture AI's role in crafting real-time narratives that follow market movements.

Another concern for our results is that there is strategic timing of AI articles. Authors could use AI precisely when contemporaneous human articles or messages are more informative, so that at the stock-day level, so the AI day dummy coincides with human articles with stronger return predictability, thereby driving the result. However, such strategic timing is infeasible on SA with publication delays. To directly address this concern, we run the same regression at the article-level sample, replacing the AI day dummy with article-level AIProb and interacting it with article-level sentiment. The results, tabulated in Appendix Table **??**, shows that the return-sentiment sensitivity increases with AI probability on SA and the result is significant at 10%. In contrast, this sensitivity decreases with AI probability among WSB articles, though the effect is insignificant.

This finding reinforces our central theme: platform governance shapes AI's market impact. Even identical regression coefficients can have opposite economic interpretations depending on the institutional context that governs content creation and publication.

6.2 AI Content and Trading Behavior

We next examine how the presence of AI-generated content influences trading behavior and market liquidity. We estimate the following regression specification:

$$\text{Trading Metric}_{it^+} = \beta_1 AIDay_{it} + \beta_2 Attention_{it} + \beta_3 Sentiment_{it} + \gamma' \boldsymbol{X}_{it} + \epsilon_{it}, \qquad (10)$$

The key independent variable is $AIDay_{i,t}$, an indicator for firm-days with at least one AI-generated article. We examine market quality both on the publication day (Panel A) and over the subsequent five days (Panel B): abnormal trading volume, return volatility (absolute returns on day t or standard deviation over $[t+1, t+5]$), and the bid-ask spread. All dependent variables are standardized, so the coefficients represent changes in standard deviations.

The findings reveal strikingly different patterns across platforms that evolve over time. On Seeking Alpha (Table 9, columns 1-3), AI content on day t shows no significant impact on volume or absolute returns, but already reduces bid-ask spreads by 0.013 standard deviations. This immediate liquidity improvement suggests market makers recognize AI content as reducing information asymmetry even before the full price discovery process unfolds.

Over the next five trading days (panel B), the effects become more pronounced. AI presence predicts a 0.020 standard deviation decrease in abnormal volume, a 0.030 decrease in volatility, and a 0.019 decrease in bid-ask spreads. These results paint a coherent picture: AI-synthesized information on SA reduces disagreement, leading to less speculative trading, lower volatility, and tighter spreads. The pattern aligns with our return results. AI accelerates consensus formation around fundamental values.

On WSB, we observe the opposite dynamics. On the publication day (columns 4-6), AI messages are associated with a 0.094 standard deviation surge in abnormal volume and a 0.102 increase in absolute returns, though bid-ask spreads remain unchanged. This immediate spike in trading activity and volatility is consistent with the notion that AI content could trigger rapid sentiment-driven trading. The effects persist and intensify over the following week (Panel B). AI presence predicts a 0.105 standard deviation increase in abnormal volume and a 0.114 increase in volatility. Crucially, this heightened activity provides no liquidity benefit—spreads remain statistically unchanged in the WSB-only specification. In the combined model (column 9), we even see evidence of deteriorating liquidity, with spreads widening by 0.021 standard deviations.

This pattern—higher volume and volatility without improved liquidity—is the hallmark of noise trading. The temporal persistence from day t through day $t+5$ suggests AI content on

WSB not only react to price movements but potentially sustains momentum-driven trading over multiple days.

The divergence across platforms provides compelling evidence that AI's market impact depends critically on institutional context. On SA, where editorial review ensures content quality, AI enhances market efficiency both immediately (through reduced spreads) and over time (through lower volume and volatility). On WSB, where anyone can instantly post AI-generated narratives, the same technology amplifies noise trading that begins immediately and persists for days.

6.3 AI Content and Extreme Return Events

While our trading behavior results show clear platform differences, the patterns on WSB—higher volume, volatility, and spreads—could theoretically arise from either noise trading or informed trading. Information shocks can also generate increased volume and volatility as prices adjust to information, and market makers may widen spreads when facing better-informed traders. To distinguish between these explanations, we examine whether AI content predicts extreme positive return events that are inconsistent with information-based trading.

Following Bali, Cakici, and Whitelaw (2011) and Bali et al. (2025), we identify two types of events that capture lottery-like payoffs. A MAX event occurs when a stock's daily return is the highest over a trailing 21-trading-days window. A lottery event represents an even more extreme outcome: it is a MAX event and the MAXRET falls into the top decile of its daily cross-sectional distribution, where MAXRET is the 21-day highest return. These events reflect the lottery-like characteristics that attract sentiment-driven retail traders but are unlikely to result from fundamental information arrival, which typically generates more symmetric return distributions. We estimate logistic regressions predicting the probability of these events:

$$\text{MAX}_{it^+}/\text{Lottery}_{it^+} = \beta_1 AIDay_{it} + \beta_2 Attention_{it} + \beta_3 Sentiment_{it} + \gamma' \boldsymbol{X}_{it} + \epsilon_{it}, \qquad (11)$$

where the dependent variables are indicators for whether stock i experiences a MAX or lottery event either on day t or during the subsequent five-day window $[t+1, t+5]$.

Tables 10 and 11 present striking results. On WSB, AI content predicts both types of extreme events. For MAX events (Table 10), AI presence is associated with an increase in odds by 26.9% in the same day ($e^{0.238} - 1$, column 3) and 17.2% ($e^{0.159} - 1$, column 4) in the next five days. The lottery event results are even more dramatic (Table 11): AI content is associated with an increase in odds by 57.0% ($e^{0.451} - 1$) and 31.9% ($e^{0.277} - 1$) for the same day and the next five days using the combined sample.

On SA, we find no such relationship. The AI day coefficients are statistically insignificant and economically negligible for both MAX and lottery events across all specifications. These findings lend further support to Hypothesis 3b. Informed trading, even with private information about positive developments, would not systematically produce the extreme right-tail events we observe. The fact that AI content predicts lottery-like returns suggests that AI on WSB facilitates sentiment-driven rather than information-driven trading.

The platform divergence reinforces our central thesis. On WSB, AI appears alongside or helps create the extreme positive returns that trigger lottery-like trading, with effects persisting for days. On SA, where editorial review ensures content quality, AI shows no association with these speculative events. This provides the cleanest evidence yet that platform governance determines whether AI serves market efficiency or market exuberance.

7 Robustness

This section provides further validation for our core empirical findings by addressing potential methodological concerns. We show that our results for WSB hold when restricting the sample to original posts (submissions); our findings are not sensitive to the initial filtering of very short messages, which cannot be reliably tested by GPTZero.

7.1 WSB Submissions Only

We address the potential concern that our WSB findings are driven by pooling high-effort original posts (submissions) with subsequent comments. Considering the potential different characteristics of posts and reactive comments, we restrict the sample to only WSB submissions and replicate our core analysis from Table 3 to 7. Results are presented in Appendix Table A4.

Despite the dramatically reduced sample size, our key findings remain consistent. Proxies for information cost, such as covering a firm for the first time (First Time Firm) or number of news articles (News), remain statistically insignificant determinants of AI probability. Conversely, the rhetorical-engineering motive remains significantly positive. The probability of using AI in crafting WSB submissions increases by 13.408–17.534 percentage points after extreme retail buying. To put it into perspective, the average AI probability of WSB submissions is 12.54 percentage points.

The user engagement patterns also persist, with AI content associated with reduced comment volume and higher disagreement. These consistent results in a sample one-tenth the size provide strong validation that our findings reflect fundamental platform differences rather than our content aggregation choices.

7.2 Including Short Messages on WSB

Our main analysis excludes WSB messages shorter than 50 words to ensure sufficient text length for reliable AI detection by GPTZero. To verify that this filtering does not systematically bias our conclusions, we replicate the article-level analyses using all WSB messages, assigning a zero AI probability to those below the 50-word threshold. Importantly, we include an indicator variable for messages exceeding 50 words in all specifications. It ensures that our zero AI probability assignment for short messages does not create spurious correlations with variables that predict message length. Moreover, it allows us to test whether the motivations for AI adoption differ conditional on the decision to write substantive content.

Appendix Table A5 reports the results of this test. The findings on the determinants of AI adoption and user engagement remain robust. The proxies for information-production cost remain insignificant. Furthermore, the proxy for the rhetorical-engineering motive (Top 10% OIB) remains positive and marginally significant, confirming that this motive is not an artifact of our sample filtering.

8 Conclusion

This paper provides a comprehensive empirical analysis of how generative AI is reshaping financial discourse and market dynamics across distinct online platforms. By comparing the curated Seeking Alpha with the unmoderated Reddit's r/WallStreetBets, we document how platform governance fundamentally determines whether AI enhances market efficiency or amplifies market noise.

We find that AI adoption surged following ChatGPT's launch, but with distinct motivations across platforms. On SA, authors primarily use AI to overcome information production costs, particularly when covering unfamiliar firms or during information-scarce periods. On WSB, AI adoption aligns more with rhetorical engineering—users deploy AI strategically during high retail trading periods to craft compelling narratives. These different adoption patterns foreshadow the divergent market impacts we document.

AI-generated content consistently reduces user engagement diversity on both platforms, leading to fewer comments on SA and lower sentiment dispersion overall. This homogenization of discourse suggests AI may be creating echo chambers rather than fostering the debate essential to price discovery. Yet the market consequences of this convergence depend entirely on platform context.

Our market impact analysis reveals a fundamental asymmetry. On SA, AI content enhances market quality through multiple channels. The editorial review process creates temporal separation between content creation and market outcomes, allowing us to identify genuine information transmission. AI presence accelerates price discovery, reduces trading volume and volatility, and tightens bid-ask spreads. Crucially, AI shows no association with extreme return events, confirming its informational rather than speculative role.

On WSB, identical statistical patterns have opposite economic interpretations. Instant publication eliminates the temporal separation needed to establish causality, and the apparent return-sentiment sensitivity may simply reflect AI-generated narratives chasing price movements. AI presence predicts higher volume and volatility without improving liquidity, classic symptoms of noise trading. Most tellingly, AI content strongly predicts lottery-like return events, increasing their probability by up to 57%. This association with extreme positive returns - inconsistent with informed trading but characteristic of sentiment-driven speculation - provides decisive evidence that AI on WSB amplifies market exuberance rather than efficiency.

These findings contribute to several important debates. First, they demonstrate that generative AI's economic impact is not inherent to the technology but shaped by institutional context. The same tool that enhances price discovery on a curated platform can fuel speculation on an unmoderated one. Second, they highlight a critical tension in democratizing financial discourse: while AI lowers barriers to participation, without proper governance it may degrade rather than improve market quality. Third, they suggest that platform regulators cannot remain neutral since rules governing AI content directly determine whether it serves or subverts market efficiency.

References

Agarwal, D., M. Naaman, and A. Vashistha. 2025. AI suggestions homogenize writing toward western styles and diminish cultural nuances. In *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems*, CHI ’25, 1–21. ACM.

Anderson, B. R., J. H. Shah, and M. Kreminski. 2024. Homogenization effects of large language models on human creative ideation. In *Proceedings of the 16th Conference on Creativity & Cognition*, C&C ’24, 413–25. New York, NY, USA: Association for Computing Machinery.

Bali, T. G., N. Cakici, and R. Whitelaw. 2011. Maxing out: Stocks as lotteries and the cross-section of expected returns. *Journal of Financial Economics* 99:427–46.

Bali, T. G., D. Hirshleifer, L. Peng, Y. Tang, and Q. Wang. 2025. Social interactions and lottery stock mania. Working paper, Georgetown University.

Barber, B. M., X. Huang, T. Odean, and C. Schwarz. 2022. Attention-induced trading and returns: Evidence from robinhood users. *Journal of Finance* 77:3141–90.

Berger, J., and K. L. Milkman. 2012. What makes online content viral? *Journal of Marketing Research* 49:192–205.

Bertomeu, J., Y. Lin, Y. Liu, and Z. Ni. 2025. The impact of generative AI on information processing: Evidence from the ban of ChatGPT in Italy. *Journal of Accounting and Economics* 101782.

Blankespoor, E., J. Croom, and S. M. Grant. 2024. Generative AI and investor processing of financial information. Working paper, University of Washington.

Boehmer, E., C. M. Jones, X. Zhang, and X. Zhang. 2021. Tracking Retail Investor Activity. *The Journal of Finance* 76:2249–305.

Bradshaw, M., H. Ma, A. Yost, and Z. Zou. 2025. Generative AI use by capital-market information intermediaries: Evidence from seeking alpha. Working paper, Boston College.

Brynjolfsson, E., D. Li, and L. Raymond. 2025. Generative AI at work. *Quarterly Journal of Economics* 140:889–942.

Campbell, J. L., M. D. DeAngelis, and J. R. Moon. 2019. Skin in the game: Personal stock holdings and investors’ response to stock analysis on social media. *Review of Accounting Studies* 24:731–79. doi: 10.1007/s11142-019-09498-9.

Chang, Y.-C., P.-J. Hsiao, A. Ljungqvist, and K. Tseng. 2022. Testing disagreement models. *Journal of*

Finance 77:2239–85.

Chen, H., P. De, Y. J. Hu, and B.-H. Hwang. 2014. Wisdom of crowds: The value of stock opinions transmitted through social media. *Review of Financial Studies* 27:1367–403. doi: 10.1093/rfs/hhu001.

Cong, L. W., Y. Guo, X. Zhao, and W. Zhou. 2024. Writing quality and soft information in the GenAI age: Evidence from online credit markets. Working paper, Cornell University.

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

Cookson, J. A., R. Lu, W. Mullins, and M. Niessner. 2024. The social signal. *Journal of Financial Economics* 158:103870–.

Daniel, K., M. Grinblatt, S. Titman, and R. Wermers. 1997. Measuring mutual fund performance with characteristic-based benchmarks. *The Journal of Finance* 52:1035–58.

Diamond, D. W., and R. E. Verrecchia. 1981. Information aggregation in a noisy rational expectations economy. *Journal of Financial Economics* 9:221–35.

Dou, W. W., I. Goldstein, and Y. Ji. 2025. Ai-powered trading, algorithmic collusion, and price efficiency. Working paper, University of Pennsylvania.

Eaton, G. W., T. C. Green, B. S. Roseman, and Y. Wu. 2022. Retail trader sophistication and stock market quality: Evidence from brokerage outages. *Journal of Financial Economics* 146:502–28.

Fama, E. F. 1970. Efficient capital markets: A review of theory and empirical work. *Journal of Finance* 25:383–417.

Financial Crimes Enforcement Network. 2024. Fraud schemes involving deepfake media targeting financial institutions. FinCEN Alert FIN-2024-Alert004.

Financial Stability Board. 2024. The financial stability implications of artificial intelligence.

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

Hayek, F. A. 1945. The use of knowledge in society. *American Economic Review* 35:519–30.

Hirshleifer, D., L. Peng, and Q. Wang. 2025. News diffusion in social networks and stock market reactions. *Review of Financial Studies* 38:883–937.

Hu, D., C. M. Jones, S. Li, V. Zhang, and X. Zhang. 2025. The rise of reddit: How social media affects belief formation and price discovery. Working paper, Peking University.

Kim, A. G., M. Muhn, and V. V. Nikolaev. 2025. Bloated disclosures: Can chatgpt help investors process information? Working paper, University of Chicago Booth.

Kim, O., and R. E. Verrecchia. 1994. Market liquidity and volume around earnings announcements. *Journal of Accounting and Economics* 17:41–67.

Knobloch-Westerwick, S. 2014. *Choice and preference in media use: Advances in selective exposure theory and research.* New York, NY: Routledge.

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

———. 2024. Measuring firm complexity. *Journal of Financial and Quantitative Analysis* 59:2487–514.

Luo, X., T. Wang, L. Yang, X. Zhao, and Y. Zhang. 2023. Initial evidence on the market impact of the ixbrl adoption. *Accounting Horizons* 37:143–71.

Matz, S. C., J. D. Teeny, S. S. Vaid, H. Peters, G. M. Harari, and M. Cerf. 2024. The potential of generative AI for personalized persuasion at scale. *Scientific Reports* 14:4692–.

Mitchell, E., Y. Lee, A. Khazatsky, C. D. Manning, and C. Finn. 2023. DetectGPT: Zero-Shot Machine-Generated Text Detection using Probability Curvature. ArXiv:2301.11305 [cs].

Nickerson, R. S. 1998. Confirmation bias: A ubiquitous phenomenon in many guises. *Review of General Psychology* 2:175–220.

Noy, S., and W. Zhang. 2023. Experimental evidence on the productivity effects of generative artificial intelligence. *Science* 381:187–92.

Padmakumar, V. S., and T. He. 2024. Does writing with language models reduce content diversity? In *Proceedings of the 12th International Conference on Learning Representations (ICLR 2024).* Poster paper.

Plate, B. J., J. Voshaar, and J. Zimmermann. 2025. Investor reactions to generative AI usage in md&a disclosures. Working paper, University of Bremen.

Reviriego, P., R. Valenzuela, M.-T. Pardo, and A. González. 2024. Playing with words: Comparing the vocabulary and lexical diversity of chatgpt and humans. *Machine Learning with Applications* 18:100247–.

Securities and Exchange Commission. 2024a. Artificial intelligence (ai) and investment fraud: Investor alert.

———. 2024b. Office hours with gary gensler: Systemic risk in artificial intelligence. Transcript.

Shleifer, A., and L. H. Summers. 1990. The noise trader approach to finance. *Journal of Economic*

Perspectives 4:19–33.

Su, J., T. Y. Zhuo, D. Wang, and P. Nakov. 2023. DetectLLM: Leveraging Log Rank Information for Zero-Shot Detection of Machine-Generated Text. *arXiv* .

Vosoughi, S., D. Roy, and S. Aral. 2018. The spread of true and false news online. *Science* 359:1146–51.

Wu, J., S. Yang, R. Zhan, Y. Yuan, D. F. Wong, and L. S. Chao. 2025. A survey on LLM-generated text detection: Necessity, methods, and future directions. *Computational Linguistics* .

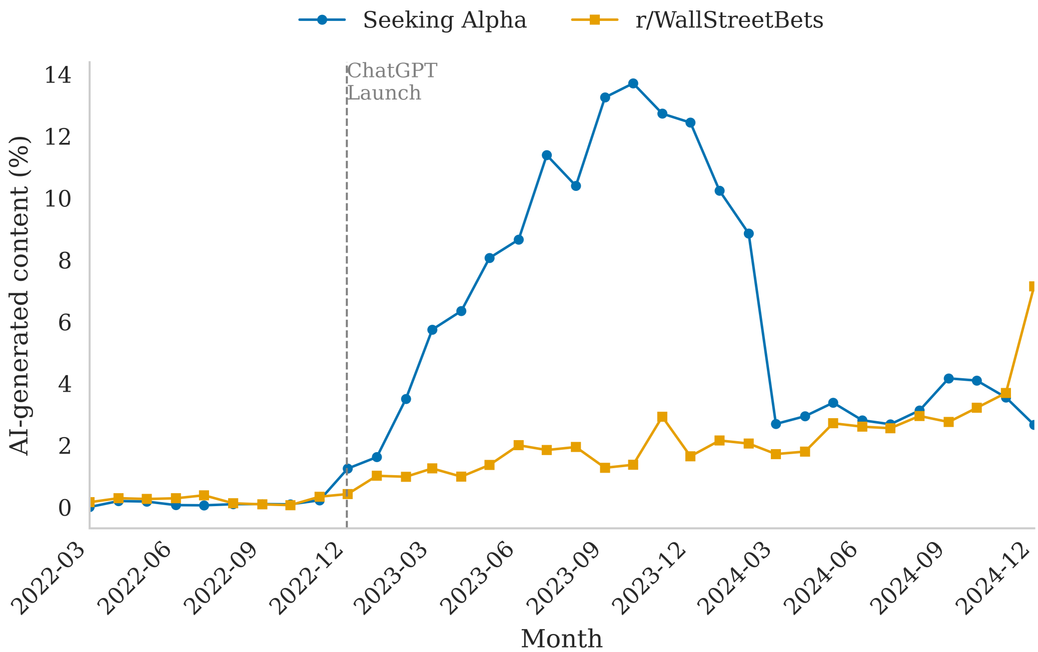

Figure 1: AI Proportion Time-Series

This figure illustrates the monthly trend of AI-related content on two online platforms: Seeking Alpha and Reddit r/WallStreetBets. It displays the monthly AI Proportion for each platform, covering the period from March 2022 to December 2024. The AI Proportion plotted is derived from GPTZero's classification of all articles from each platform within our sample. A vertical dashed line on November 30, 2022, indicates the launch date of ChatGPT.

Table 1: Summary statistics

This table shows the textual attributes of articles on Seeking Alpha and posts on Reddit's r/WallStreetBets, together with the characteristics of the firms they discuss, over the sample period (December 2022–December 2024). Panel A reports summary statistics for each platform, covering AI probability, article characteristics, firm-level fundamentals, and author past activities. Panel B provides summary statistics of variables used in the stock-day sample. The detailed definitions of all variables are in Appendix Table A1.

Panel A: Article-level sample

	SA						WSB						Diff
	N	Mean	SD	p25	p50	p75	N	Mean	SD	p25	p50	p75	SA-WSB
AI Prob (%)	57,581	10.65	21.65	0.91	2.75	6.28	64,848	3.60	14.36	0.52	0.68	1.57	7.05***
Length	57,581	1453.67	675.77	1022	1309	1705	64,848	114.18	124.75	61	79	119	1339.49***
Sentiment	57,581	0.03	0.01	0.02	0.02	0.03	64,848	0.02	0.03	0.00	0.02	0.04	0.01***
Complexity (%)	57,581	0.22	0.27	0.05	0.14	0.31	64,848	0.09	0.37	0.00	0.00	0.00	0.13***
Fog Index	57,581	14.57	2.38	12.96	14.39	15.88	64,848	10.66	6.93	7.59	9.48	12.00	3.91***
Quantitative (%)	57,581	3.80	2.24	2.38	3.41	4.73	64,848	2.67	3.52	0.00	1.69	3.85	1.13***
Graphical (%)	57,581	0.46	0.31	0.24	0.41	0.62	64,848	0.02	0.25	0.00	0.00	0.00	0.44***
Comments	57,581	11.32	20.00	1	4	13	64,848	0.34	2.25	0	0	0	10.98***
Disagreement	57,581	0.04	0.05	0.00	0.04	0.07	64,848	0.01	0.04	0.00	0.00	0.00	0.03***
Size ($B)	54,429	146.29	416.43	2.04	12.21	73.54	63,400	343.38	565.22	3.08	46.31	637.67	-197.09***
BM	54,429	0.46	0.55	0.12	0.30	0.62	63,400	0.38	0.45	0.05	0.20	0.54	0.08***
IO	54,429	0.71	0.25	0.60	0.75	0.86	63,400	0.65	0.29	0.58	0.66	0.79	0.06***
Analysts	54,429	3.97	3.00	1.47	3.31	5.89	63,400	5.11	3.42	2.21	4.78	7.36	-1.14***
Author Articles	57,581	49.54	38.98	20	42	70	64,848	4.71	11.66	0	0	7	44.83***
Author AI Articles	57,581	2.95	7.40	0	0	2	64,848	0.17	2.34	0	0	0	2.78***
Author Comments	57,581	11.11	9.17	4.59	9.93	16.00	64,848	0.12	0.46	0	0	0	11.00***

Panel B: Stock-level sample

	SA						WSB					
	N	Mean	SD	p25	p50	p75	N	Mean	SD	p25	p50	p75
AIDay (×100)	1,607,466	0.24	4.85	0	0	0	687,418	0.18	4.28	0	0	0
Articles (×100)	1,607,466	3.59	22.37	0	0	0	687,418	9.43	136.15	0	0	0
Sentiment (×100)	1,607,466	0.08	0.50	0	0	0	687,418	0.06	0.60	0	0	0
AR_t (%)	1,459,930	0.02	5.18	-1.25	-0.04	1.15	648,718	0.03	4.31	-1.17	-0.04	1.08
$CAR_{[t+1,t+5]}$ (%)	1,594,720	0.11	11.68	-3.46	-0.18	3.07	682,020	0.13	12.65	-3.31	-0.18	2.92
$AVOL_{[t+1,t+5]}$ (%)	1,591,465	2.87	60.09	-29.55	-2.29	28.54	680,946	3.02	56.58	-27.32	-2.10	26.51
$Volatility_{[t+1,t+5]}$ (%)	1,594,720	2.85	4.64	1.31	2.09	3.42	682,020	2.78	3.94	1.24	1.98	3.28
$Spread_{[t+1,t+5]}$ (%)	1,520,352	0.41	1.21	0.07	0.14	0.41	653,059	0.29	0.60	0.05	0.09	0.21
Size ($B)	1,484,486	14.11	91.72	0.30	1.35	5.32	653,497	26.67	135.58	0.76	3.01	13.88
BM	1,484,486	0.62	0.96	0.22	0.45	0.80	653,497	0.56	0.65	0.19	0.40	0.73
IO	1,484,486	0.68	0.35	0.50	0.77	0.90	653,497	0.71	0.35	0.60	0.80	0.90
Analysts	1,484,486	1.91	1.86	0.37	1.10	2.58	653,497	2.46	2.19	0.74	1.84	3.68
Earnings Day	1,484,486	0.01	0.12	0	0	0	653,497	0.02	0.12	0	0	0
$CAR_{[t-21,t-1]}$ (%)	1,462,355	1.60	23.58	-6.78	0.27	7.73	649,907	1.66	20.29	-6.49	0.41	7.54
$Volatility_{[t-21,t-1]}$ (%)	1,605,333	3.19	4.32	1.68	2.47	3.83	686,727	3.12	3.52	1.59	2.33	3.69

Table 2: Characterizing AI-Generated Articles: Key Textual Attributes

This table presents results of the following regression:

$$AIProb_{ijt} = \gamma' \boldsymbol{X}_{ijt} + \epsilon_{ijt}$$

where AI Prob$_{i,j,t}$ is AI-related probability score assigned by GPTZero to content j about firm i posted on day t. Article Characteristics$_{j,t}$ is the content j's characteristics (standardized to zero mean and unit variance). Columns 1–2 report results for articles published on Seeking Alpha; Columns 3–4 report analogous results for posts on Reddit's r/WallStreetBets. Standard errors are two-way clustered by stock and day, and the resultant t-statistics are shown in parentheses. *p <.1; **p <.05; ***p <.01.

	SA		WSB			
	(1)	(2)	(3)	(4)	(5)	(6)
Length	0.345***	-0.043	2.737***	1.479***	2.106***	1.286***
	(3.44)	(-0.32)	(16.88)	(10.13)	(13.70)	(8.49)
Sentiment	2.943***	0.774***	0.858***	0.220***	0.804***	0.216***
	(22.31)	(8.02)	(8.82)	(3.15)	(8.83)	(3.07)
Complexity	1.866***	0.410***	0.577***	0.260***	0.550***	0.255***
	(14.94)	(5.04)	(6.51)	(2.88)	(6.55)	(2.83)
Fog Index	1.771***	3.839***	0.397***	0.171**	0.424***	0.178**
	(13.29)	(20.35)	(6.97)	(2.35)	(7.66)	(2.44)
Graphical	-0.402***	-0.962***	0.054	-0.030	-0.382***	-0.164
	(-3.27)	(-7.35)	(0.63)	(-0.24)	(-4.10)	(-1.37)
Quantitative	-1.332***	-0.313***	0.278***	0.200*	0.139*	0.170
	(-9.78)	(-2.87)	(3.17)	(1.81)	(1.75)	(1.58)
Submissions					6.964***	2.834***
					(9.00)	(4.09)
Day FE	X	X	X	X	X	X
Stock FE	X	X	X	X	X	X
Author FE		X		X		X
Obs.	57,581	57,581	64,848	64,848	64,848	64,848
Adj. R^2 (%)	9.9	59.5	10.3	56.9	12.0	57.1

Table 3: AI Adoption: Marginal Cost of Information Production

This table presents results of the following regression:

$$AIProb_{ijt} = \beta_1 \cdot First\ Time\ Firm_{ijt} + \gamma' \boldsymbol{X}_{ijt} + \epsilon_{ijt},$$

where AI $\text{Prob}_{i,j,t}$ is AI-related probability score assigned by GPTZero to content j about firm i posted on day t. First Time $\text{Firm}_{i,j,t]}$ is a dummy variable that equals 1 if the author has not mentioned the firm i in the previous 6 months. $\mathbf{X}_{i,j,t}$ includes the firm characteristics and author history performance (standardized to zero mean and unit variance). Columns 1–3 report results for articles published on Seeking Alpha; Columns 4–6 report analogous results for posts on Reddit's r/WallStreetBets. Standard errors are two-way clustered by stock and day, and the resultant t-statistics are shown in parentheses. $^*p<.1$; $^{**}p<.05$; $^{***}p<.01$.

	SA		WSB	
	(1)	(2)	(3)	(4)
First Time Firm	2.322***	0.555***	0.007	-0.012
	(9.42)	(3.26)	(0.04)	(-0.08)
Size	-0.573	-1.781**	-1.398*	-0.527
	(-0.54)	(-2.20)	(-1.86)	(-0.54)
BM	0.919***	0.291	0.162	-0.320
	(2.84)	(0.86)	(1.00)	(-1.40)
IO	0.943*	0.452	0.363	-0.480
	(1.69)	(1.21)	(0.95)	(-1.19)
Author Article	-3.418***	-1.109***	-0.808***	-0.055
	(-20.60)	(-8.10)	(-8.50)	(-0.58)
Author AI Article	8.358***	2.497***	2.356***	-0.431***
	(36.41)	(19.15)	(6.99)	(-2.68)
Author Comments	0.490***	0.468***	-0.168*	0.061
	(3.32)	(3.74)	(-1.78)	(0.56)
Submissions			9.089***	4.364***
			(10.95)	(5.88)
Day FE	X	X	X	X
Stock FE	X	X	X	X
Author FE		X		X
Obs.	54,429	54,429	63,400	63,400
Adj. R^2 (%)	18.7	59.5	11.5	56.6

Table 4: AI Adoption: Information Scarcity

This table presents results of the following regression:

$$AIProb_{ijt} = \beta_1 News_{it} + \beta_2 Analyst_{it} + \gamma' \boldsymbol{X}_{ijt} + \epsilon_{ijt},$$

where AI $\text{Prob}_{i,j,t}$ is AI-related probability score assigned by GPTZero to content j about firm i posted on day t. News $\text{Intensiveness}_{i,t}$ is transformation $log(1+n)$ of the number of news articles concerning firm i in the past three days, sourced from RavenPack. $\mathbf{X}_{i,j,t}$ includes the firm characteristics and author history performance (standardized to zero mean and unit variance). Columns 1–3 report results for articles published on Seeking Alpha; Columns 4–6 report analogous results for posts on Reddit's r/WallStreetBets. Standard errors are two-way clustered by stock and day, and the resultant t-statistics are shown in parentheses. *p <.1; **p <.05; ***p <.01.

	SA		WSB	
	(1)	(2)	(3)	(4)
News	-0.319**	-0.206**	-0.050	-0.215
	(-2.15)	(-2.11)	(-0.37)	(-1.41)
Analysts	-0.729***	-0.332**	0.527**	0.162
	(-3.42)	(-1.98)	(2.42)	(0.79)
Size	-0.389	-1.673**	-1.727**	-0.520
	(-0.37)	(-2.06)	(-2.06)	(-0.50)
BM	0.920***	0.288	0.221	-0.326
	(2.81)	(0.84)	(1.34)	(-1.39)
IO	0.960*	0.454	0.429	-0.463
	(1.73)	(1.22)	(1.14)	(-1.18)
Author Article	-3.678***	-1.131***	-0.809***	-0.053
	(-22.16)	(-8.26)	(-9.29)	(-0.59)
Author AI Article	8.383***	2.497***	2.356***	-0.430***
	(36.57)	(19.16)	(6.99)	(-2.67)
Author Comments	0.479***	0.472***	-0.169*	0.061
	(3.22)	(3.76)	(-1.78)	(0.56)
Submissions			9.088***	4.363***
			(10.93)	(5.88)
Day FE	X	X	X	X
Stock FE	X	X	X	X
Author FE		X		X
Obs.	54,429	54,429	63,400	63,400
Adj. R^2 (%)	18.5	59.5	11.5	56.6

Table 5: AI Adoption: Rhetoric Engineering

This table presents results of the following regression:

$$AIProb_{ijt} = \beta_1 \cdot Top\ 10\%\ OIB_{it} + \gamma' \boldsymbol{X}_{ijt} + \epsilon_{ijt},$$

where AI Prob$_{i,j,t}$ is AI-related probability score assigned by GPTZero to content j about firm i posted on day t. *HighOIB* is a dummy variable that equals 1 if the retail order imbalance ranks in the preceeding three days ranks in the top 10% across-sectionally. $\mathbf{X}_{i,j,t}$ includes the firm characteristics and author history performance (standardized to zero mean and unit variance). Columns 1–2 report results for articles published on Seeking Alpha; Columns 3–4 report analogous results for posts on Reddit's r/WallStreetBets. Standard errors are two-way clustered by stock and day, and the resultant t-statistics are shown in parentheses. *p <.1; **p <.05; ***p <.01.

	SA		WSB	
	(1)	(2)	(3)	(4)
Top 10% OIB	-0.224	-0.451	1.673**	1.970***
	(-0.38)	(-1.08)	(2.52)	(3.31)
Size	-0.437	-1.663**	-1.586**	-0.766
	(-0.40)	(-2.04)	(-2.13)	(-0.73)
BM	0.959***	0.345	0.218	-0.233
	(2.88)	(0.99)	(1.51)	(-1.18)
IO	-0.065	-0.109	0.208***	-0.088
	(-0.18)	(-0.58)	(5.80)	(-1.14)
Author Articles	-3.719***	-1.143***	-0.885***	-0.035
	(-21.92)	(-8.07)	(-11.53)	(-0.41)
Author AI Articles	8.480***	2.515***	0.829***	-0.371***
	(36.76)	(18.83)	(6.20)	(-2.60)
Author Comments	0.419***	0.489***	0.144*	0.056
	(2.74)	(3.82)	(1.69)	(0.48)
adjusted_date	X	X	X	X
permno	X	X	X	X
author_id		X		X
Obs.	52,991	52,991	61,073	61,073
Adj. R^2 (%)	18.9	60.0	6.6	58.4

Table 6: User Engagement

This table presents results of the following regression

$$\log(1 + \mathit{Comments}_{ij[t,t+10]}) = \beta_1 \cdot \mathit{AIProb}_{ijt} + \gamma' \boldsymbol{X}_{ijt} + \epsilon_{ijt},$$

where $\log(1+\mathit{Comments}_{ij[t,t+10]})$ is the transformation $\log(1+n)$ of the number of comments following content j in the next ten days. AIProb_{ijt} is AI-related probability score assigned by GPTZero to content j about firm i posted on day t. $\mathbf{X}_{i,j,t}$ includes the article characteristics, firm characteristics, and author history performance (standardized to zero mean and unit variance). Columns 1–3 report results for articles published on Seeking Alpha; Columns 4–6 report analogous results for posts on Reddit's r/WallStreetBets. Standard errors are two-way clustered by stock and day, and the resultant t-statistics are shown in parentheses. *p <.1; **p <.05; ***p <.01.

	SA		WSB	
	(1)	(2)	(3)	(4)
AIProb	-0.075***	-0.068**	-0.088***	0.002
	(-3.16)	(-2.57)	(-5.80)	(0.06)
Length	0.111***	0.111***	0.017***	0.023***
	(13.42)	(11.39)	(6.68)	(6.92)
Sentiment	-0.047***	-0.082***	-0.004***	-0.003
	(-7.17)	(-11.72)	(-3.08)	(-1.35)
Complexity	-0.058***	-0.032***	0.000	-0.001
	(-7.08)	(-5.09)	(-0.29)	(-0.57)
Fog Index	-0.068***	-0.033***	-0.004***	-0.005
	(-14.07)	(-4.03)	(-2.77)	(-1.53)
Graphical	0.045***	-0.014**	0.003	0.002
	(7.83)	(-2.06)	(0.87)	(0.55)
Quantitative	-0.010*	-0.040***	-0.002*	0.002
	(-1.85)	(-5.96)	(-1.80)	(0.88)
Size	0.198*	0.211***	-0.020	0.000
	(1.84)	(2.66)	(-0.79)	(-0.00)
BM	-0.013	-0.001	0.009	0.003
	(-0.48)	(-0.06)	(1.22)	(0.21)
Analysts	0.009	0.005	0.004	0.004
	(0.76)	(0.54)	(1.02)	(0.58)
IO	-0.027	-0.056**	0.003	0.014
	(-0.86)	(-2.05)	(0.29)	(1.02)
Author Article	-0.104***	-0.015**	0.000	0.008**
	(-15.17)	(-2.04)	(-0.03)	(2.52)
Author AI Article	0.027***	0.006	0.004**	0.002
	(4.97)	(1.29)	(2.52)	(0.99)
Author Comments	0.196***	0.013**	0.007***	-0.013***
	(20.52)	(1.98)	(2.89)	(-4.67)
Submissions			0.572***	0.699***
			(20.11)	(17.56)
Day FE	X	X	X	X
Stock FE	X	X	X	X
Author FE		X		X
Obs.	54,429	54,429	63,400	63,400
Adj. R^2 (%)	51.4	64.1	19.6	27.1

Table 7: Disagreement

This table presents results of the following regression

$$Disagreement_{ij[t,t+10]} = \beta_1 \cdot AIProb_{ijt} + \gamma' \boldsymbol{X}_{ijt} + \epsilon_{ijt},$$

where Disagreement$_{ij[t,t+10]}$ is the standard deviation of the sentiment score of comments following content j in the next ten days. The dependent variable is standardized to zero mean and unit variance. AI Prob$_{ijt}$ is AI-related probability score assigned by GPTZero to content j about firm i posted on day t. $\mathbf{X}_{i,j,t}$ includes the article characteristics, firm characteristics, and author past activities (standardized to zero mean and unit variance). Columns 1–2 report results for articles published on Seeking Alpha; Columns 3–4 report analogous results for contents on Reddit's r/WallStreetBets. Standard errors are two-way clustered by stock and day, and the resultant t-statistics are shown in parentheses. *p <.1; **p <.05; ***p <.01.

	SA		WSB	
	(1)	(2)	(3)	(4)
AIProb	-0.004***	-0.004***	0.006***	0.000
	(-4.35)	(-3.42)	(4.25)	(0.08)
Length	0.003***	0.003***	0.001***	0.001***
	(12.11)	(7.10)	(3.65)	(3.87)
Sentiment	0.000	-0.001**	-0.001***	-0.001*
	(0.94)	(-1.98)	(-3.27)	(-1.92)
Complexity	-0.001***	0.000**	0.000	0.000
	(-4.54)	(-2.15)	(-0.29)	(-0.21)
Fog Index	-0.001***	0.001*	0.000	0.000
	(-6.32)	(1.72)	(0.59)	(0.00)
Graphical	0.002***	0.001**	0.000	0.000
	(9.68)	(2.45)	(0.02)	(1.13)
Quantitative	0.000	0.000	0.001***	0.001***
	(0.91)	(1.02)	(8.35)	(3.30)
Size	0.006*	0.007**	-0.001	0.002
	(1.83)	(2.52)	(-0.30)	(0.47)
BM	0.000	0.000	0.000	0.001
	(-0.51)	(-0.41)	(0.25)	(0.81)
Analysts	0.000	0.000	0.000	0.000
	(-0.65)	(-0.98)	(0.43)	(-0.06)
IO	0.000	0.000	0.000	-0.003*
	(0.34)	(-0.38)	(-0.31)	(-1.81)
Author Article	-0.002***	-0.001*	0.000	0.000
	(-8.49)	(-1.80)	(1.44)	(0.71)
Author AI Article	0.001***	0.001**	0.000	0.000
	(4.95)	(2.50)	(-1.15)	(0.02)
Author Comments	0.004***	0.000	0.000*	0.000
	(13.67)	(0.79)	(-1.78)	(-1.38)
Submissions			0.030***	0.019***
			(40.15)	(11.05)
Day FE	X	X	X	X
Stock FE	X	X	X	X
Author FE		X		X
Obs.	54,429	54,429	63,400	63,400
Adj. R^2 (%)	20.5	25.1	6.8	13.6

Table 8: AI Content: Return

This table presents results of the following regression:

$$\begin{aligned}\text{AR}_{it}/\text{CAR}_{i[t+1,t+5]} =&\beta_1 AIDay_{it} \times Sentiment_{it} + \beta_2 AIDay_{it} + \beta_3 Sentiment_{it} + \\ &\beta_4 Attention_{it} + \gamma' \boldsymbol{X}_{it} + \epsilon_{it},\end{aligned}$$

where $\text{AR}_{[t]}$, is the DGTW-adjusted abnormal return for firm i on day t and $\text{CAR}_{[t+1,t+5]}$, is the cumulative abnormal return over the subsequent five-day window. Columns 1 and 2 use the SA sample and WSB sample, respectively. Column 3 uses the union of the SA sample and WSB sample and presents a combined model that includes all variables from both platforms. Stock and day fixed effects are included. Standard errors are two-way clustered by stock and day. t-statistics are reported in parentheses. Significance levels are denoted as: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

	SA		WSB		Combined	
	AR_t (1)	$\text{CAR}_{[t+1,t+5]}$ (2)	AR_t (3)	$\text{CAR}_{[t+1,t+5]}$ (4)	AR_t (5)	$\text{CAR}_{[t+1,t+5]}$ (6)
SA AIDay × SA Sentiment	0.120***	0.024			0.115***	0.022
	(2.84)	(0.49)			(2.75)	(0.44)
SA AIDay	-0.824***	-0.395			-0.795***	-0.380
	(-2.72)	(-1.18)			(-2.62)	(-1.14)
SA Sentiment	0.045***	-0.025**			0.047***	-0.023**
	(5.54)	(-2.09)			(5.96)	(-1.97)
SA Attention	-0.024**	0.022			-0.028***	0.021
	(-2.39)	(1.61)			(-2.88)	(1.52)
WSB AIDay × WSB Sentiment			0.169**	-0.004	0.124**	-0.006
			(1.98)	(-0.07)	(2.02)	(-0.17)
WSB AIDay			-0.089	0.045	-0.192	0.151
			(-0.36)	(0.12)	(-0.78)	(0.39)
WSB Sentiment			0.029***	-0.003	0.019***	-0.001
			(3.86)	(-0.32)	(3.90)	(-0.15)
WSB Attention			0.052***	0.004	0.034***	0.005
			(3.59)	(0.22)	(3.62)	(0.38)
Size	-0.649***	-3.356***	-0.725***	-3.489***	-0.678***	-3.453***
	(-10.40)	(-13.80)	(-9.24)	(-11.32)	(-10.68)	(-14.00)
BM	-0.053**	-0.231**	-0.031	-0.163	-0.049*	-0.226**
	(-1.99)	(-2.36)	(-1.06)	(-1.39)	(-1.88)	(-2.33)
IO	-0.008	-0.040	-0.024**	-0.067	-0.003	-0.006
	(-0.69)	(-0.93)	(-2.04)	(-0.96)	(-0.24)	(-0.10)
Analysts	-0.028**	-0.117**	-0.008	-0.049	-0.029**	-0.120**
	(-1.97)	(-2.34)	(-0.72)	(-0.99)	(-2.00)	(-2.40)
Earnings Day	0.113	0.060	0.200*	0.071	0.091	0.043
	(1.55)	(0.56)	(1.80)	(0.63)	(1.23)	(0.41)
Abn. Ret. [t-21,t-1]	-0.071***	-0.338***	-0.069***	-0.216**	-0.081***	-0.374***
	(-5.49)	(-7.18)	(-3.35)	(-2.45)	(-5.20)	(-7.98)
Volatility [t-21,t-1]	0.004	0.044	0.014	0.069	0.000	0.046
	(0.19)	(0.57)	(0.43)	(0.65)	(0.01)	(0.62)
Day FE	X	X	X	X	X	X
Stock FE	X	X	X	X	X	X
Obs.	1,432,850	1,424,099	639,099	635,343	1,455,530	1,446,761
Adj. R^2 (%)	0.0	3.6	0.2	5.1	0.1	3.5

Table 9: AI Content: Trading Behavior

This table presents results of the following regression

$$\text{Trading Metric}_{it^+} = \beta_1 AIDay_{it} + \beta_2 Attention_{it} + \beta_3 Sentiment_{it} + \gamma' \boldsymbol{X}_{it} + \epsilon_{it},$$

The dependent variable, Trading Metric$_{it^+}$, measures trading activity for stock i. Panel A examines metrics on day t, while Panel B uses metrics averaged over the subsequent five-day window, $[t+1, t+5]$. The columns correspond to abnormal daily volume (1, 4, 7), a return-based measure (absolute return in Panel A and return volatility in Panel B; columns 2, 5, 8), and the average bid-ask spread (3, 6, 9). The dependent variable is standardized to zero mean and unit variance. $AIDay_{it}$ is an indicator variable of AI content of stock i on day t. $Sentiment_{it}$ is the sentiment score averaged across these articles. $Attention_{it}$ is the logarithm of one plus the number of articles of stock i on day t and All regressions include controls for firm characteristics, and include stock and day fixed effects. Standard errors are two-way clustered by stock and day. t-statistics are reported in parentheses. Significance levels are denoted as: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Panel A: Time Period $[t]$									
	SA			WSB			Combined		
	Volume (1)	Abs(Ret) (2)	Spread (3)	Volume (4)	Abs(Ret) (5)	Spread (6)	Volume (7)	Abs(Ret) (8)	Spread (9)
SA AIDay	0.014	0.020	-0.013***				0.017	0.024	-0.013***
	(1.03)	(1.01)	(-3.03)				(1.31)	(1.22)	(-3.32)
SA Sentiment	-0.005***	-0.012***	0.000				-0.003*	-0.009***	0.000
	(-3.11)	(-6.33)	(-0.55)				(-1.83)	(-5.13)	(-0.48)
SA Attention	0.033***	0.033***	-0.002***				0.029***	0.029***	-0.002***
	(16.97)	(13.40)	(-3.58)				(14.99)	(12.17)	(-3.70)
WSB AIDay				0.094**	0.102**	0.013	0.091**	0.118**	0.021**
				(2.54)	(2.24)	(1.32)	(2.49)	(2.43)	(2.42)
WSB Sentiment				0.006***	0.003	0.000	0.004***	0.002*	0.000
				(4.22)	(1.63)	(-0.57)	(4.28)	(1.93)	(-0.28)
WSB Attention				0.044***	0.045***	0.002**	0.026***	0.028***	0.001*
				(11.77)	(10.71)	(2.46)	(10.34)	(9.77)	(1.68)
Size	-0.100***	-0.119***	-0.339***	-0.060**	-0.123***	-0.256***	-0.105***	-0.122***	-0.329***
	(-3.82)	(-5.88)	(-11.06)	(-2.05)	(-4.69)	(-5.11)	(-4.12)	(-6.14)	(-11.02)
BM	0.005	-0.021**	0.018*	-0.008	-0.006	0.020	0.003	-0.019**	0.023**
	(0.70)	(-2.48)	(1.70)	(-0.73)	(-0.58)	(1.43)	(0.43)	(-2.33)	(2.31)
IO	-0.013**	-0.004	-0.015**	-0.005*	0.000	-0.012	-0.012*	-0.002	-0.016**
	(-1.99)	(-0.77)	(-2.00)	(-1.76)	(0.04)	(-1.19)	(-1.95)	(-0.43)	(-1.99)
Analysts	-0.034***	-0.003	-0.008**	-0.034***	-0.002	-0.009**	-0.034***	-0.002	-0.007**
	(-7.73)	(-0.88)	(-2.51)	(-5.57)	(-0.47)	(-2.41)	(-7.72)	(-0.78)	(-2.47)
Earnings Day	1.081***	1.442***	0.085***	1.329***	1.539***	0.079***	1.045***	1.396***	0.079***
	(58.11)	(50.65)	(16.80)	(54.09)	(42.40)	(16.70)	(57.24)	(49.28)	(15.96)
Abn. Ret. [t-21,t-1]	0.205***	0.027***	-0.024***	0.238***	0.022***	-0.031***	0.205***	0.028***	-0.025***
	(3.98)	(3.70)	(-2.99)	(24.13)	(3.64)	(-4.70)	(4.12)	(3.76)	(-3.01)
Volatility [t-21,t-1]	0.051	0.048	-0.005	0.169***	0.107***	-0.029***	0.059	0.053	-0.008
	(1.07)	(1.48)	(-1.50)	(4.69)	(5.42)	(-3.92)	(1.16)	(1.55)	(-1.45)
Day FE	X	X	X	X	X	X	X	X	X
Stock FE	X	X	X	X	X	X	X	X	X
Obs.	1,435,318	1,435,318	1,361,546	640,288	640,286	609,892	1,458,189	1,458,187	1,382,586
Adj. R^2 (%)	19.2	23.6	60.9	24.3	27.3	65.5	18.8	23.5	61.3

Panel B: Time Period $[t+1, t+5]$									
	SA			WSB			Combined		
	Volume (1)	Vol. (2)	Spread (3)	Volume (4)	Vol. (5)	Spread (6)	Volume (7)	Vol. (8)	Spread (9)
SA AIDay	-0.020*	-0.025*	-0.017***				-0.017	-0.022*	-0.017***
	(-1.73)	(-1.94)	(-4.97)				(-1.46)	(-1.76)	(-5.28)
SA Sentiment	-0.004***	-0.011***	0.000				-0.002	-0.009***	0.000
	(-2.75)	(-8.16)	(-0.47)				(-1.59)	(-7.39)	(-0.30)
SA Attention	0.022***	0.014***	-0.001***				0.018***	0.011***	-0.001***
	(12.04)	(7.98)	(-3.40)				(10.43)	(7.08)	(-3.69)
WSB AIDay				0.105***	0.109**	0.010	0.102***	0.104**	0.020***
				(3.20)	(2.36)	(1.26)	(3.07)	(2.27)	(2.62)
WSB Sentiment				0.006***	0.002	-0.001	0.004***	0.001	0.000
				(4.54)	(1.06)	(-1.57)	(4.46)	(1.09)	(-0.79)
WSB Attention				0.035***	0.031***	0.002***	0.022***	0.019***	0.001**
				(10.12)	(8.21)	(2.90)	(8.85)	(7.72)	(2.55)
Size	-0.183***	-0.197***	-0.375***	-0.157***	-0.214***	-0.271***	-0.189***	-0.198***	-0.362***
	(-5.80)	(-7.31)	(-10.86)	(-4.28)	(-5.94)	(-4.91)	(-6.21)	(-7.58)	(-10.80)
BM	-0.007	-0.043***	0.021*	-0.029**	-0.027**	0.023	-0.010	-0.039***	0.027**
	(-0.71)	(-4.13)	(1.75)	(-1.99)	(-2.00)	(1.44)	(-1.01)	(-3.95)	(2.32)
IO	-0.019**	-0.010	-0.017**	-0.008***	-0.003	-0.014	-0.018**	-0.007	-0.019*
	(-2.34)	(-1.38)	(-1.98)	(-3.03)	(-0.70)	(-1.15)	(-2.27)	(-0.91)	(-1.96)
Analysts	-0.039***	0.002	-0.009**	-0.037***	0.001	-0.011**	-0.039***	0.002	-0.009**
	(-7.83)	(0.33)	(-2.56)	(-5.10)	(0.16)	(-2.45)	(-7.80)	(0.39)	(-2.56)
Earnings Day	0.430***	0.128***	0.003	0.465***	0.108***	0.002	0.413***	0.123***	0.001
	(29.46)	(8.57)	(0.95)	(28.61)	(7.34)	(0.57)	(29.57)	(8.72)	(0.45)
Abn. Ret. [t-21,t-1]	0.241***	0.024***	-0.027***	0.271***	0.017**	-0.034***	0.238***	0.024***	-0.028***
	(4.03)	(4.11)	(-3.02)	(24.74)	(2.28)	(-4.82)	(4.18)	(3.95)	(-3.05)
Volatility [t-21,t-1]	0.032	0.038	-0.006	0.120***	0.084***	-0.031***	0.040	0.042	-0.009
	(0.84)	(1.47)	(-1.50)	(3.93)	(4.80)	(-3.58)	(0.96)	(1.55)	(-1.45)
Day FE	X	X	X	X	X	X	X	X	X
Stock FE	X	X	X	X	X	X	X	X	X
Obs.	1,424,099	1,424,099	1,358,845	635,343	635,343	608,811	1,446,761	1,446,761	1,379,841
Adj. R^2 (%)	17.5	34.8	78.1	20.6	39.0	80.1	17.2	34.7	78.3

Table 10: AI Content: MAX Events

This table presents results of the following logistic regression:

$$\text{MAX}_{it^+} = \beta_1 AIDay_{it} + \beta_2 Attention_{it} + \beta_3 Sentiment_{it} + \gamma' \boldsymbol{X}_{it} + \epsilon_{it},$$

The dependent variable, MAX_{it^+}, is a dummy equal to 1 if stock i experiences a MAX event either on day t or in the subsequent five-day window, $[t+1, t+5]$, and 0 otherwise. A MAX event is when the daily return is the highest over the prior 20 trading days. $AIDay_{it}$ is an indicator variable of AI content of stock i on day t. $Sentiment_{it}$ is the sentiment score averaged across these articles. $Attention_{it}$ is the logarithm of one plus the number of articles of stock i on day t and All regressions include controls for firm characteristics, and include stock and day fixed effects. Standard errors are two-way clustered by stock and day. t-statistics are reported in parentheses. Significance levels are denoted as: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

	SA		WSB		Combined	
	MAX_t (1)	$\text{MAX}_{[t+1,t+5]}$ (2)	MAX_t (3)	$\text{MAX}_{[t+1,t+5]}$ (4)	MAX_t (5)	$\text{MAX}_{[t+1,t+5]}$ (6)
SA AIDay	-0.019	0.023			-0.017	0.027
	(-0.21)	(0.49)			(-0.19)	(0.57)
SA Sentiment	0.009	-0.018***			0.014*	-0.015***
	(1.32)	(-4.56)			(1.94)	(-4.04)
SA Attention	0.038***	0.029***			0.031***	0.025***
	(5.31)	(6.93)			(4.29)	(6.21)
WSB AIDay			0.238**	0.159**	0.242**	0.158**
			(2.08)	(2.09)	(2.12)	(2.05)
WSB Sentiment			0.024***	0.010***	0.017***	0.007***
			(3.46)	(2.73)	(3.70)	(2.82)
WSB Attention			0.071***	0.046***	0.043***	0.028***
			(7.29)	(6.51)	(6.69)	(6.21)
Size	-0.410***	-0.480***	-0.463***	-0.538***	-0.417***	-0.476***
	(-5.79)	(-9.66)	(-5.31)	(-8.31)	(-5.81)	(-9.58)
BM	-0.063**	-0.098***	-0.063*	-0.100***	-0.060**	-0.094***
	(-2.01)	(-4.46)	(-1.82)	(-3.82)	(-2.00)	(-4.41)
IO	-0.004	-0.006	0.002	0.002	-0.002	-0.004
	(-0.48)	(-0.53)	(0.40)	(0.24)	(-0.24)	(-0.40)
Analysts	-0.020*	-0.006	-0.023*	-0.004	-0.020*	-0.006
	(-1.73)	(-0.57)	(-1.74)	(-0.34)	(-1.72)	(-0.55)
Earnings Day	2.066***	1.063***	2.235***	1.116***	2.040***	1.051***
	(30.22)	(37.57)	(28.38)	(34.83)	(29.98)	(37.74)
Abn. Ret. [t-21,t-1]	-1.030***	-0.925***	-0.889***	-0.802***	-1.033***	-0.928***
	(-27.78)	(-37.39)	(-20.70)	(-26.25)	(-28.00)	(-37.58)
Volatility [t-21,t-1]	-1.711***	-1.743***	-1.457***	-1.443***	-1.746***	-1.769***
	(-29.75)	(-39.12)	(-21.55)	(-24.00)	(-29.33)	(-37.77)
Day FE	X	X	X	X	X	X
Stock FE	X	X	X	X	X	X
Obs.	1,435,289	1,424,072	640,286	635,342	1,458,159	1,446,735

Table 11: AI Content: Lottery Events

This table presents results of the following logistic regression:

$$\text{Lottery}_{it^{+}} = \beta_1 AIDay_{it} + \beta_2 Attention_{it} + \beta_3 Sentiment_{it} + \gamma' \boldsymbol{X}_{it} + \epsilon_{it},$$

The dependent variable, $Lottery_{it^{+}}$, is a dummy equal to 1 if stock i experiences a lottery event either on day t or in the subsequent five-day window, $[t+1, t+5]$, and 0 otherwise. A lottery event is a MAX event where the corresponding 20-day maximum return also ranks in the top decile compared to all other stocks on that same day. $AIDay_{it}$ is an indicator variable of AI content of stock i on day t. $Sentiment_{it}$ is the sentiment score averaged across these articles. $Attention_{it}$ is the logarithm of one plus the number of articles of stock i on day t and All regressions include controls for firm characteristics, and include stock and day fixed effects. Standard errors are two-way clustered by stock and day. t-statistics are reported in parentheses. Significance levels are denoted as: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

	SA		WSB		Combined	
	Lottery_t (1)	$\text{Lottery}_{[t+1,t+5]}$ (2)	Lottery_t (3)	$\text{Lottery}_{[t+1,t+5]}$ (4)	Lottery_t (5)	$\text{Lottery}_{[t+1,t+5]}$ (6)
SA AIDay	0.029	-0.040			0.023	-0.030
	(0.14)	(-0.28)			(0.10)	(-0.21)
SA Sentiment	0.046**	-0.019			0.064***	-0.008
	(2.34)	(-1.38)			(3.22)	(-0.59)
SA Attention	0.187***	0.091***			0.154***	0.071***
	(9.34)	(6.63)			(7.34)	(5.22)
WSB AIDay			0.451*	0.277	0.526**	0.289*
			(1.82)	(1.64)	(2.05)	(1.71)
WSB Sentiment			0.047**	0.014	0.037***	0.012*
			(2.39)	(1.30)	(2.75)	(1.65)
WSB Attention			0.264***	0.146***	0.153***	0.090***
			(9.21)	(6.65)	(7.92)	(6.50)
Size	-0.729***	-0.767***	-0.816***	-0.744***	-0.784***	-0.785***
	(-8.18)	(-9.23)	(-6.02)	(-6.53)	(-8.72)	(-9.60)
BM	-0.068*	-0.137***	-0.013	-0.090*	-0.058	-0.125***
	(-1.75)	(-3.42)	(-0.24)	(-1.67)	(-1.53)	(-3.27)
IO	-0.111	-0.102	-0.109	-0.090	-0.085*	-0.085*
	(-1.62)	(-1.55)	(-1.17)	(-1.47)	(-1.81)	(-1.75)
Analysts	-0.053*	-0.038	-0.085*	-0.056	-0.058**	-0.042
	(-1.81)	(-1.23)	(-1.65)	(-1.13)	(-2.00)	(-1.39)
Earnings Day	3.197***	1.537***	3.441***	1.638***	3.108***	1.497***
	(34.84)	(29.16)	(27.50)	(24.81)	(33.47)	(28.79)
Abn. Ret. [t-21,t-1]	-0.151***	-0.131***	-0.144***	-0.124***	-0.173***	-0.147***
	(-7.34)	(-7.07)	(-5.41)	(-4.92)	(-8.89)	(-8.09)
Volatility [t-21,t-1]	0.029***	-0.028	0.039**	-0.012	0.032***	-0.027
	(4.09)	(-1.27)	(2.26)	(-0.49)	(4.43)	(-1.22)
Day FE	X	X	X	X	X	X
Stock FE	X	X	X	X	X	X
Obs.	687,841	680,696	253,316	274,574	707,467	700,185

Appendix

A. Figure and Variable List

B. AI Detection Method Comparison

B..1 Simulated Sample Construction

To evaluate the detectors, we require a corpus in which the true source of every document is known. We therefore build a test set that pairs real finance commentary with machine-generated text written to the same brief. We randomly select 1,000 long-form equity articles from Seeking Alpha and 1,000 discussion messages from r/WallStreetBets, all dated between December 2020 and November 2021. The window lies well before public LLM deployment.

We then generate AI text based on Wall Street Journal news. We start from collecting 2,854 articles that appeared in the Wall Street Journal “Markets," “Stocks," “U.S. Markets," and “Heard on the Street" section between December 2020 and November 2021. Each article is fed to three frontier models-Llama 3.3, GPT-4o, and Claude 3.5-under two fixed instructions. Prompt 1 asks the model to recast the story in Seeking Alpha's house style, while Prompt 2 requests a Reddit r/WallStreetBets “DD" post. The exact instructions are reproduced below. The exercise yields 17,124 synthetic documents (2,854 articles times 3 models times 2 styles), bringing the evaluation set to 19,124 observations, roughly 10 percent human and 90 percent AI.

Prompt 1 - Seeking Alpha style:

Your task is to write a short-form Seeking Alpha analysis.
Here is the article I want you to write based on:

`{article}`

Important Instructions – Carefully craft a single analysis that:
Use Seeking Alpha's voice: write as a seasoned, opinionated investor, offering a clear perspective.
Paraphrase, never copy sentences from the article provided.
Total length 1,000-1,500 characters, including spaces.
Output your analysis directly, without any preamble.

Prompt 2 - Reddit style:

Your task is to write a short-form r/wallstreetbets Due Diligence (DD).
Here is the article I want you to write based on:

`{article}`

Important Instructions – Carefully craft a single analysis that:
Use r/wallstreetbets DD tone, potentially using slang and emphasizing high-risk/high-reward scenarios.

Paraphrase, never copy sentences from the article provided.
Total length 1,000-1,500 characters, including spaces.
Output your analysis directly, without any preamble.

B..2 Detectors

We compare four families of detectors that capture the main strands of the literature (See, e.g., Wu et al. (2025)).

GPTZero is a proprietary ensemble of multiple lexical and probabilistic features calibrated to minimize false positives. It returns probabilities for "Human," "AI," and "MIXED", as documented in detail above. Perturbation curvature is represented by DetectGPT (Mitchell et al. 2023). The algorithm measures how log-likelihood falls when the text is lightly masked, exploiting the idea that AI prose sits at a local optimum of its own model. We implement the authors' public code and methodologies. Statistical signatures follow Su et al. (2023). We report the Normalised Perplexity Ratio (NPR) and five auxiliary scores-Log-Likelihood Ratio, entropy, token rank, raw log-likelihood, and log-rank-computed under a Llama 3.3 reference. Cut-offs are set at the 28.6-percentile of each score's distribution in the human subset, a value that equates type-I and type-II costs in a validation grid.

B..3 Detection Results

Appendix Table A2 summarises accuracy and F1 statistics averaged across the full benchmark. GPTZero identifies provenance with 98 percent accuracy and an F1 of 0.99. DetectGPT attains 81 percent accuracy ($F1 = 0.89$); NPR matches that level, and the remaining statistics hover between 76 and 80 percent. The Log-Likelihood Ratio is the strongest of the auxiliary scores, yet still trails GPTZero by roughly twenty percentage points. Because following empirical discussions rely on the accurate classification of each document, we adopt GPTZero as the detector for all subsequent analysis and treat its output as the measure of AI involvement in platform content.

Table A1: Description of Variables

Variable	Definition
Article level	
AIProb_{ijt}	The sum of probability scores of the content j for the AI-ONLY and MIXED categories, assigned by GPTZero.
Length_{ijt}	Natural logarithm of word count of content j. For clarity, the descriptive statistics in Table 1 are based on the original, untransformed word count.
Sentiment_{ijt}	Sentiment scores of the content j derived from platform-specific dictionaries (Loughran and McDonald (2011) for Seeking Alpha; Hu et al. (2025), modified for Reddit slang and emojis, for r/WallStreetBets). The sentiment score is calculated as follows: $$\text{Sentiment}_{ijt} = \frac{N_{\text{positive}} - N_{\text{negative}}}{N_{\text{positive}} + N_{\text{negative}}}$$ Where N_{positive} and N_{negative} are respectively the count of positive/negative words in the platform-specific content j. To reduce the impact of extreme values, the resulting ratio is winsorized at the 1st and 99th percentiles.
Complexity_{ijt}	Complexity scores derived from Loughran and McDonald (2024)'s dictionaries, calculated as follows: $$Complexity_{ijt} = \frac{\text{Words in Dictionary}}{textTotalWords}$$ Where #Words in Dictionary are the count of the specific complexity words identified in Loughran and McDonald (2024).
$Quantitative_{i,j,t}$	Total count of numbers mentioned divided by word count of a article, winsorized at the 1st and 99th percentiles.
$Graphical_{i,j,t}$	Total count of images divided by word count, winsorized at the 1st and 99th percentiles.
Fog Index_{ijt}	Fog Index is a readability test for English writing, calculated as follows: $$\text{Fog Index}_{ijt} = 0.4 \times \left(\frac{\text{Words}}{\text{Sentences}} + 100 \times \frac{\text{Complex Words}}{\text{Words}} \right),$$ where Words represents the total number of words in the content j; Sentences represents the total number of sentences in the content j; Complex Words represents the count of words with three or more syllables (excluding common suffixes like -es, -ed, -ing, proper nouns, and familiar jargon). To reduce the impact of extreme values, the resulting ratio is winsorized at the 1st and 99th percentiles.
$Comments_{i,j,[t,t+10]}$	Natural logarithm of number of user comments received on an article from its publication day (t) through 10 days post-publication ($t+10$).
$Disagreement_{i,j,[t,t+10]}$	Standard deviation of sentiment scores from user comments posted in response to content j during the period $[t, t+10]$. The sentiment scores for these comments are calculated using the dictionary from Loughran and McDonald (2011) for Seeking Alpha content, and a dictionary from Hu et al. (2025), modified for Reddit slang and emojis, for r/WallStreetBets content.
Author Articles	Natural logarithm of one plus the number of articles published by a given author in the past six months.
Author AI Articles	Natural logarithm of one plus the number of AI-ONLY or MIXED articles by a given author in the past six months.

Variable	Definition
Author Comments	Natural logarithm of one plus the average number of comments per article within 21 days, where the average is taken across all articles that a given author publishes in the past six months.
Stock-day level	
$AIDay_{i,t}$	An indicator that equals one if at least one article about firm i published on day t is classified as AI-ONLY or MIXED by GPTZero.
Sentiment$_{i,t}$	The average sentiment score of articles about firm i on day t. Value set to a neutral score of zero for days without any content of firm i.
Attention$_{it}$	Natural logarithm of one plus the number of articles on SA (or posts on WSB) of firm i on day t.
$Size$	Natural logarithm of the market capitalization for firm i.
BM	Natural logarithm of book-to-market ratio for firm i.
IO	Institutional ownership, calculated as the percentage of common shares outstanding of firm i held by institutional investors as of the last quarter ending before day t.
AR$_{it}$	Abnormal return adjusted for size, book-to-market, and momentum following Daniel et al. (1997) on day t.
CAR$_{i[t+1,t+5]}$	Cumulative abnormal returns adjusted for size, book-to-market, and momentum following Daniel et al. (1997) over the 3 trading days from $t+1$ to $t+5$.
AVOL$_{i[t+1,t+5]}$	Average abnormal volume for firm i over the 5 trading days from $t+1$ to $t+5$:
	$$\text{AVOL}_{i[t+1,t+5]} = \frac{1}{5}\sum_{s=1}^{5}\log(1+Volume_{i,t+s}) - \frac{1}{31}\sum_{s=-41}^{-11}\log(1+Volume_{i,t+s}),$$ where $Volume$ is the dollar volume, calculated as the numbers of shares traded multiplied by the price per share.
Volatility$_{i[t+1,t+5]}$	Standard deviation of the daily stock returns for firm i over the 5 trading days from $t+1$ to $t+5$.
$Spread_{i,t}$	Average daily volume-weighted effective bid-ask spread for firm i over the 5 trading days from $t+1$ to $t+5$.
CAR$_{i[t-21,t-1]}$	Cumulative abnormal returns adjusted for size, book, and momentum following Daniel et al. (1997) over the preceding 21 trading days from $t-21$ to $t-1$.
Volatility$_{i[t-21,t-1]}$	Standard deviation of the daily stock returns for firm i over the preceding 21 trading days from $t-21$ to $t-1$.

Table A2: AI Detection Method Comparison

This table presents the performance of various AI text detection methods, evaluated on our full benchmark dataset using overall accuracy and the F1 score. Accuracy measures the proportion of documents whose provenance (AI-generated or human-written) is correctly identified. The F1 score represents the harmonic mean of precision and recall, providing a balanced assessment of classification performance. The comparison includes GPTZero, DetectGPT, NPR, Log-Likelihood Ratio (LLR), Entropy, Rank, Log Likelihood, and Log Rank.

Detection Method	**Accuracy**	**F1 Score**
GPTZero	98%	99%
DetectGPT	81%	89%
NPR	80%	89%
LLR	78%	87%
Entropy	76%	86%
Rank	76%	86%
Log Likelihood	76%	86%
Log Rank	76%	86%

Table A3: AI Adoption: Marginal Cost of Information Production

This table presents results of the following regression:

$$AIProb_{ijt} = \beta_1 \cdot First\ Time\ Industry_{ijt} + \gamma' \boldsymbol{X}_{ijt} + \epsilon_{ijt},$$

where AI Prob$_{i,j,t}$ is AI-related probability score assigned by GPTZero to content j about firm i posted on day t. First Time Industry$_{i,j,t]}$ is a dummy variable that equals 1 if the author has not mentioned firm i's industry in the previous 6 months. $\mathbf{X}_{i,j,t}$ includes the firm characteristics and author history performance (standardized to zero mean and unit variance). Columns 1–3 report results for articles published on Seeking Alpha; Columns 4–6 report analogous results for posts on Reddit's r/WallStreetBets. Standard errors are two-way clustered by stock and day, and the resultant t-statistics are shown in parentheses. *p <.1; **p <.05; ***p <.01.

	SA		WSB	
	(1)	(2)	(3)	(4)
First Time Industry	2.034***	0.294	0.019	-0.079
	(7.04)	(1.39)	(0.11)	(-0.52)
Size	-0.644	-1.808**	-1.397*	-0.536
	(-0.61)	(-2.24)	(-1.86)	(-0.55)
BM	0.933***	0.293	0.161	-0.319
	(2.87)	(0.87)	(0.99)	(-1.40)
IO	0.933*	0.453	0.363	-0.480
	(1.69)	(1.22)	(0.95)	(-1.19)
Author Article	-3.449***	-1.120***	-0.806***	-0.060
	(-20.85)	(-8.15)	(-8.56)	(-0.63)
Author AI Article	8.371***	2.497***	2.356***	-0.431***
	(36.57)	(19.15)	(6.99)	(-2.68)
Author Comments	0.495***	0.472***	-0.168*	0.061
	(3.34)	(3.77)	(-1.78)	(0.55)
Submissions			9.088***	4.364***
			(10.95)	(5.88)
Day FE	X	X	X	X
Stock FE	X	X	X	X
Author FE		X		X
Obs.	54,429	54,429	63,400	63,400
Adj. R^2 (%)	18.6	59.5	11.5	56.6

Table A4: Robustness: WSB submissions-only sample

This table presents results of replicating the results from Table 3 to Table 7 for the r/WallStreetBets (WSB) submissions-only sample. Columns 1—6 examine the determinants of AI adoption. Columns 7—8 and 9—10 examine the impact of AI content on subsequent user comments and disagreement, respectively.
Standard errors are two-way clustered by stock and day, and the resultant *t*-statistics are shown in parentheses. **p* <.1; ***p* <.05; ****p* <.01.

	AIProb						Comments		Disagreement	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
First Time Permno	0.532	-1.112								
	(0.69)	(-0.73)								
News			-0.653	-3.484						
			(-0.80)	(-1.51)						
Top 10% OIB					13.408^{***}	17.534^{*}				
					(2.73)	(1.87)				
AIProb							-0.208^{***}	-0.107	0.213^{***}	-0.186
							(-5.01)	(-0.38)	(2.91)	(-0.52)
Day FE	X	X	X	X	X	X	X	X	X	X
Stock FE	X	X	X	X	X	X	X	X	X	X
Author FE		X		X		X		X		X
Obs.	6,487	6,487	6,487	6,487	6,127	6,127	6,487	6,487	6,487	6,487
Adj. R^2 (%)	10.4	23.1	10.5	23.8	11.5	25.9	5.3	-67.4	2.5	-71.2

Table A5: Robustness: Including Messages below 50 Words on WSB

This table replicates the results from Table 3 to Table 7 using all r/WallStreetBets (WSB) messages (i.e., including below-50-word submissions and comments). Columns 1—6 examine the determinants of AI adoption. Columns 7—8 and 9—10 examine the impact of AI content on subsequent user comments and disagreement, respectively. Standard errors are two-way clustered by stock and day, and the resultant *t*-statistics are shown in parentheses. **p* <.1; ***p* <.05; ****p* <.01.

	AIProb						Comments		Disagreement	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Word Count > 50	2.925***	2.675***	2.924***	2.677***	2.761***	2.572***	-0.489***	-0.510***	-0.066***	-0.028***
	(26.15)	(28.10)	(26.19)	(27.99)	(27.01)	(27.49)	(-45.99)	(-63.84)	(-9.26)	(-4.20)
First Time Permno	-0.008	0.017								
	(-0.56)	(1.53)								
News			-0.006	-0.005						
			(-0.56)	(-0.61)						
Top 10% OIB					0.159*	0.112*				
					(1.85)	(1.95)				
AIProb							-0.374***	-0.263***	0.152***	0.018
							(-5.39)	(-10.17)	(3.97)	(0.42)
Day FE	X	X	X	X	X	X	X	X	X	X
Stock FE	X	X	X	X	X	X	X	X	X	X
Author FE		X		X		X		X		X
Obs.	949,674	949,674	949,674	949,674	901,387	901,387	949,674	949,674	949,674	949,674
Adj. R^2 (%)	9.7	46.6	9.7	46.6	9.0	46.0	11.5	17.9	2.2	2.4