AI, Opinion Ecosystems, and Finance*

David Hirshleifer Lin Peng Qiguang Wang
Weichen Zhang Xiaoyan Zhang

December 2025

Abstract

This paper studies how investors use generative AI in discussions across two major investing social media platforms with distinct governance and user bases: Seeking Alpha and Reddit's r/WallStreetBets. We document large cross-platform differences in adoption patterns and their associated market outcomes. On Seeking Alpha, AI adoption is more prevalent when information is scarce or when contributors cover unfamiliar stocks. AI article sentiment positively predicts returns and is associated with greater information discovery: more informative retail order flow, reduced user disagreement, and narrower bid-ask spread. On WallStreetBets, in contrast, AI adoption tends to rise following surges in retail buying and is linked to sentiment contagion. Adoption is also followed by higher abnormal trading volume, greater volatility, and more lottery-like return distributions. These results indicate that both the adoption of AI and its relationship with market outcomes depend upon the institutional and behavioral context in which the technology is deployed.

JEL Codes: G12, G14, G15

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

---

*David Hirshleifer, Marshall School of Business, University of Southern California. 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. Weichen Zhang, Singapore Management University, Singapore. Xiaoyan Zhang, Tsinghua University, Beijing, China. The paper was previously circulated under the title "*Social Finance in the Age of AI: A Tale of Two Platforms*." We are grateful to GPTZero for providing access to their AI detection API and Alex Cui and Jonathan White for technical assistance. We appreciate helpful comments and discussions from Jeremy Bertomeu, Xi Dong, Michael Farrell, Ioannis V. Floros, Hong Liu, Fangzhou Lu (discussant), Xiumin Martin, Rich Frankel, Bernard Yeung, Anastasia Zakolyukina, Wu Zhu, Dexin Zhou, seminar participants at Baruch College, the University of Wisconsin-Milwaukee, Washington University in St. Louis, and participants at the Third Annual Conference on Capital Market Research in the Era of AI (AICM) in Beijing, China and the Wolfe Research 9th Annual Global Quantitative and Macro Investment Conference in New York.

AI, Opinion Ecosystems, and Finance

December 2025

Abstract

This paper studies how investors use generative AI in discussions across two major investing social media platforms with distinct governance and user bases: Seeking Alpha and Reddit's r/WallStreetBets. We document large cross-platform differences in adoption patterns and their associated market outcomes. On Seeking Alpha, AI adoption is more prevalent when information is scarce or when contributors cover unfamiliar stocks. AI article sentiment positively predicts returns and is associated with greater information discovery: more informative retail order flow, reduced user disagreement, and narrower bid-ask spread. On WallStreetBets, in contrast, AI adoption tends to rise following surges in retail buying and is linked to sentiment contagion. Adoption is also followed by higher abnormal trading volume, greater volatility, and more lottery-like return distributions. These results indicate that both the adoption of AI and its relationship with market outcomes depend upon the institutional and behavioral context in which the technology is deployed.

JEL Codes: G12, G14, G15

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

1 Introduction

The cost of producing and processing information is a central determinant of the efficiency of capital markets.[1] The recent advent of powerful generative artificial intelligence (AI), particularly large language models (LLM) such as ChatGPT, has sharply reduced those costs, reshaping how investors produce, acquire, and process information. Consistent with this view, Cao et al. (2024) find that a machine-learning-trained AI analyst outperforms most human analysts in return forecasts and that combining humans with AI delivers the most accurate forecasts.[2]

However, recent work highlights that the same technological forces lowering information costs may also undermine information quality. In particular, Stiglitz and Ventura-Bolet (2025) warn that AI can distort the information ecosystem by making it cheaper to produce misinformation and harder for consumers to evaluate content quality. Absent adequate safeguards—such as accountability mechanisms and content moderation—they argue that information quality may deteriorate and polarization may intensify. Furthermore, AI could be used to induce mistaken beliefs or impulsive behavior through manipulation of mood and cognitive bias. In a similar spirit, Acemoglu, Ozdaglar, and Siderius (2025) raise concerns that the use of AI by social media platforms can cause misleading persuasion and exacerbate belief polarization.Financial regulators have expressed related concerns. For example, the Financial Stability Board warns that "GenAI could increase financial fraud and the ability of malicious actors to generate and spread disinformation in financial markets" (Financial Stability Board 2024).

Synthesizing these contrasting perspectives, and drawing particularly on the models of Stiglitz and Ventura-Bolet (2025) and Acemoglu, Ozdaglar, and Siderius (2025) as well as behavioral principles, we postulate that AI tools on digital platforms drastically reduce the cost

[1] See Hayek (1945) (on information and markets in general), Fama (1970), Grossman and Stiglitz (1980), Diamond and Verrecchia (1981) and Kim and Verrecchia (1994).

[2] Relatedly, Cheng, Lin, and Zhao (2025) provide evidence that ChatGPT outages reduce informed trading and that GenAI-assisted trading enhances long-run stock-price informativeness. Similarly, exploiting a ban in Italy, Bertomeu et al. (2025) find that reduced ChatGPT usage adversely affects information production by financial analysts.

of producing both meaningful and misleading content. The net impact of this technological shift on the opinion ecosystem of a platform is therefore ambiguous. On one hand, it may enhance the efficiency of information processing and transmission (*Information Enhancement*); on the other, it may degrade the ecosystem by enabling the spread of misinformation, triggering emotional reactions, and overwhelming users' ability to discern truth and to draw correct conclusions (*Information Distortion*). We hypothesize that the balance between these competing effects depends critically upon a platform's institutional and behavioral environment, including its governance, content moderation policies, accountability mechanisms, and the sophistication of its user base. Overall, the question is whether AI adoption enhances the production and dissemination of value-relevant information or instead introduces markets with useless or misleading content, and whether institutional setting affects the balance between them.

To evaluate this hypothesis, we empirically examine the adoption of AI in financial discourse, particularly in articles posted on financial social media platforms, and the association of AI use with investor beliefs and market outcomes. We consider three key issues: what circumstances prompt content creators to adopt GenAI, how users engage with and interpret AI-generated content, and how such content influences market prices and trading.

A key challenge for this enterprise is determining whether content has been generated by AI. We address this challenge by using a state-of-the-art and validated commercial AI detector, and by performing further out-of-sample validation tests. Furthermore, we seek to provide insight into whether the effects of AI adoption are context dependent, which is relevant for issues of platform design and policy. We therefore perform tests across across online social media platforms that differ in governance regimes and user sophistication.

Specifically, we collect articles and messages on U.S. publicly listed firms from June 2022 to December 2024, spanning the November 30, 2022 introduction of ChatGPT, across two influential social media platforms: SeekingAlpha (SA) and Reddit's r/WallStreetBets (WSB). We explore whether the very different institutional environments of these platforms are associated with important differences in the opinion ecosystems. Seeking Alpha is by subscription only and operates under strong editorial oversight: articles are reviewed prior to publication, and authors are financially compensated for producing original investment analysis. Its user base consists largely of semi-professionals with investment expertise.[3] In contrast, WallStreetBets is a decentralized Reddit community with no pre-publication review and far weaker oversight. Posts on WallStreetBets sometimes highlight screenshots of high-risk trades as well as memes GIFs. Its users are generally less sophisticated and more prone to behavioral biases and peer effects. Together, these two platforms provide a natural setting for assessing how institutional and user contexts shape the adoption and market impact of generative AI.

To measure the likelihood that a textual item is AI-generated, we use GPTZero, a state-of-the-art AI content detector. This tool is based on a multi-faceted detection approach. It analyzes linguistic complexity through metrics such as 'perplexity' and 'burstiness', employs a deep learning model for sentence-level classification, and cross-reference text against a vast corpus of internet and academic writing to limit false positives, particularly the common misclassification of highly predictable text, such as famous quotes or legal boilerplate, as AI-generated. In a study of leading commercial detectors, Jabarian and Imas (2025) find that GPTZero achieves high accuracy in detecting AI-generated general texts across models, genres, and even against tools that promise to artificially generate humanized versions of original AI-generated texts. Similarly, Blankespoor, deHaan, and Li (2025) perform validation analyses that confirm the ability of GPTZero to detect AI-generated content in financial reports.

To further validate the effectiveness of GPTZero for financial discourse on social media, we construct a custom test sample. This sample consists of (1) human-written articles from SA and WSB randomly drawn from the pre-ChatGPT period and (2) LLM-generated social media posts prompted with pre-ChatGPT *Wall Street Journal* news to mimick SA and WSB styles. On this dataset, GPTZero achieved an accuracy of 98% and a 0.99 F1 score,[4] outper-

[3]See, for example, Chen et al. (2014a), Chen and Hwang (2022), Farrell et al. (2022), Kogan, Kostyuk, and Solovyeva (2023), Dim (2024).

[4]The F1 score is a metric that balances a model's precision and recall. It ranges from 0 to 1, where 1 indicates perfect performance.

forming the other open-source detectors that we evaluated.

Applying GPTZero to our sample, we document that the average *AIProb*—the probability that an article or post is generated by AI—rises markedly on SA, from nearly zero in November 2022 to a peak of 13.7% on SA by late 2023, before declining due to a platform policy change that discourages AI-generated articles. This trend is consistent with the findings of Bradshaw et al. (2025). For WSB, *AIProb* also rises from nearly zero pre-ChatGPT to 7.1% on WSB by the end of 2024. In comparison, Blankespoor, deHaan, and Li (2025) finds that the greatest use of AI in corporate filings is in S1 filings (for new security issuances), with rates reaching between 2.5% and 4.5% in 2024. This suggests that social media commentators adopted AI more aggressively than did corporate filers.

We next examine how, when, and with what consequences market participants adopt AI across these platforms. There are four key findings. First, adoption motives differ across platforms. On SA, authors rely on AI when information is costly to gather; they are 22% more likely to use AI when covering a firm for the first time in six months, and usage is higher when analyst coverage or news availability is low. This evidence is consistent with authors turning to AI when marginal costs of conducting research and writing reports are high. In contrast, on WSB, adoption is unrelated to information availability. Instead, *AIProb* is 44–55% higher following periods of heavy retail buying. This suggests that WSB adoption is likely associated with noise trading and speculation rather than information production.

Second, we find that AI promotes consensus on SA but not WSB. Third, it promotes sentiment contagion on WSB but not SA. AI-generated content, to the extent that it provides value-relevant advice can resolve uncertainty and thereby foster investor consensus. Alternatively, persuasive AI-generated content may incite misperceptions and amplify sentiment contagion. We find that these distinct outcomes manifest on each platform. On SA, AI-generated articles are followed by a 10% reduction in the dispersion of comment sentiment over the next ten days, relative to the mean. In contrast, on WSB, AI content is not significantly associated with lower sentiment dispersion. Instead, it is linked to greater sentiment transmission: the sentiment of AI posts strongly predicts the sentiment of subsequent comments, a pattern absent on SA. These differences likely reflect platform-specific mechanisms. On SA, AI appears to be used to provide value-relevant fundamental information, reducing uncertainty and encouraging consensus. On WSB, AI appears to be used more for refining punchy narratives that promote sentiment contagion.

Fourth, the market outcomes associated with AI adoption diverge sharply across platforms. On SA, the sentiment of AI-generated articles positively predicts future stock returns, with an highly bullish sentiment corresponding to a positive one-week-ahead return of 35 basis points and no evidence of subsequent reversal. In contrast, sentiment of other articles on SA does not exhibit such predictive power. Moreover, SA-based AI content is associated with more informative retail order flows, as measured by the predictive power of retail net buys for future returns.

The evidence is consistent with the view that AI content on SA conveys value-relevant information that is not yet fully incorporated into prices at the publication date. As investors subsequently adopt this information, either by paying attention to the AI article itself or via related post-publication news, prices on average gradually adjust. Consequently, SA AI views predict future stock performance and coincides with more informative retail order flows. Additionally, stocks with AI-generated content experience a tightening of bid-ask spreads over the five days following publication. This evidence is consistent with fundamental relevant content helping resolve uncertainty and reducing the adverse selection cost faced by liquidity providers (see, e.g., Glosten and Milgrom 1985; Glosten 1989; Glosten and Harris 1988; Huang and Stoll 1997; Stoll 1989).

In contrast, on WSB, AI posts are on average followed by surges in abnormal trading volume and significantly higher volatility over the next five trading days. Moreover, this content strongly predicts destabilizing, lottery-like outcomes. The odds of a MAX event (a 21-day high) increase by 18% over the five days following publication, relative to the sample mean. The odds of a *lottery event*, defined as a MAX event that also ranks in the top decile cross-sectionally, rises by 44% over the following week. In contrast, AI content on SA is associated with a reduction rather than an increase in the frequency of such lottery-like events. Furthermore, AI content on WSB AI does not enhance the informativeness of retail order flow. These findings indicate that AI content on WSB is associated with speculative trading and followed by extreme price run-ups.

Together, this evidence is consistent with SA AI use being associated with better price discovery, perhaps because AI use enhances the production of value-relevant information; whereas on WSB AI use is linked to noise trading and predicts lottery-like outcomes. These findings underscore the importance of platform governance and community characteristics in shaping the market implications of AI adoption. These differences may derive from the editorial oversight of SA in the form of filtering and curating content. Other contributors to these differences may include differences in user sophistication, incentives (possibly nonpecuniary), or community norms.[5] More broadly, the findings indicate that the use of AI and its relationship with market outcomes depend on the institutional and behavioral environment. This conclusion is consistent with the theoretical models by Stiglitz and Ventura-Bolet (2025), which emphasize the role of the platform environment, and by Acemoglu, Ozdaglar, and Siderius (2025), which show how AI can induce belief polarization and manipulative persuasion.

Our paper makes several contributions to the rapidly evolving literature on AI in financial markets. First, we are among the first to empirically quantify the moderating role of platform governance and user characteristics in shaping the relationship between AI deployment and market outcomes. Theory predicts that lower information costs should enhance the informational efficiency of markets (e.g., Grossman and Stiglitz 1980; Diamond and Verrecchia 1981), and recent studies show evidence consistent with this prediction (Kim, Muhn, and Nikolaev 2025; Bertomeu et al. 2025). However, concerns have been raised that AI induces noise, herding, and collusion (Dou, Goldstein, and Ji 2025). Moreover, the persuasive capa-

[5]The dynamics of investment platforms that we study reflect a broader societal shift toward AI-curated information consumption. For instance, on September 25, 2025, OpenAI launched ChatGPT Pulse, a service that provides personalized daily briefings by analyzing private user data such as chat history and calendars (https://openai.com/index/introducing-chatgpt-pulse/, accessed September 28, 2025). By synthesizing far-flung information into authoritative summaries, such services may have large effects. Access to personal data together with the ability to shape narratives may position providers to shape what users see and believe.

bilities of AI can amplify behavioral biases such as attention-induced trading (Barber et al. 2022) and social transmission bias (Hirshleifer 2020b; Han, Hirshleifer, and Walden 2022) and echo chambers (Cookson, Engelberg, and Mullins 2023).

Our cross-platform analysis addresses this tension: the association between AI and lottery-like events on WSB, but not SA, highlights how institutional context determines the association of AI with better price discovery versus speculative trading. In this respect, our analysis complements Bradshaw et al. (2025), who examine AI deployment exclusively on the SA platform. Our tests contrast the SA platform with WSB, allowing us to draw comparative conclusions about AI adoption in relation to investor behavior and market outcomes.

Second, we provide evidence on the microfoundations of AI adoption and its behavioral consequences. We document why and when content creators turn to AI across platforms, showing that platform-specific motivations shape both the production and reception of AI-generated content. These findings add nuance to existing research on the productivity effects of AI (e.g., Noy and Zhang 2023; Chang et al. 2023; Cao et al. 2024; Bertomeu et al. 2025; Brynjolfsson, Li, and Raymond 2025). This paper also broadens the scope of the literature on AI analysis of textual communication in finance by studying retail-facing online platforms rather than disclosure channels such as corporate disclosures (Blankespoor, deHaan, and Li 2025) and professional analyst reports (Bertomeu et al. 2025). Our study provides one of the first large-scale comparative analyses of the footprint of AI in the more dynamic and loosely regulated domain of public equity commentary. We further show that this dimension of AI adoption is closely tied to trading activity, bid-ask spreads, and lottery-like return distributions.

2 Hypotheses

In markets with information frictions, tools such as AI that lower production costs or alter how readers process content can affect discussion patterns and market outcomes. Motivated in part by the model of Stiglitz and Ventura-Bolet (2025), we argue that the magnitude and direction of these effects are critically shaped by the institutional and behavioral environment of a platform. The sharp contrast between Seeking Alpha and WallStreetBets in their governance and user base provides a natural setting to test this idea. We first outline these institutional differences and then develop hypotheses regarding AI adoption motives, user responses, and associated capital market outcomes.

2.1 Institutional Background

Seeking Alpha and WallStreetBets differ fundamentally in governance, user sophistication, and content norms. Seeking Alpha, founded in 2004, is a platform for its crowd-sourced and curated equity research.[6] Its contributors include a network of stock analysts, traders, economists, academics, financial advisors and industry experts. As of 2021, the platform had more than 17,000 individual contributors and over 350 contributing firms. An in-house team of over 60 editors reviews every article submission, providing iterative feedback to authors until the content meets the platform's standards—a process that typically takes 10-24 hours for weekday submissions. This moderation extends to the comments section, which is actively monitored to maintain substantive discussion. Seeking Alpha has more than 10 million registered users and over 17 million monthly visitors, with an average visit duration double that of Morningstar and Yahoo Finance, and four times that of The Economist, Barron's, and the Wall Street Journal.

This structured governance of Seeking Alpha is reinforced by content guidelines for contributors. The platform encourages "investment-oriented", "original", and "fundamental analysis", and explicitly prioritizes articles that help an investor in their decision making. Furthermore, authors must disclose any business relationships with or personal holdings in the companies they cover. Seeking Alpha writers are paid a fixed amount for each article and a variable amount based on views by paid subscribers. Despite allowing pseudonyms, the platform requires private identity disclosure and prohibits users from posting under multi-

[6]The information presented in this subsection for Seeking Alpha is based on content from its official website, for example, https://static.seekingalpha.com/uploads/pdf_income/sa_media_kit_2021_final.pdf.

ple accounts.[7]

Seeking Alpha's user base also reflects this professional focus. The average retail user is a 46-year-old male (80% of the user base) with a reported household income of $321,000 and $1.5 million in investable assets. Behaviorally, they are active investors, conducting an average of 11.6 transactions annually. Moreover, a large majority are considered "affluencers" (80%) and act as the lead financial decision-maker in their household (76%). The platform's average professional client manages over $165 million in assets (AUM) and possesses deep industry experience, with 93% having worked in finance for at least a decade. Their practice is typically focused on an elite clientele, as 80% of these professionals managing assets for high-net-worth individuals.

In contrast, WallStreetBets (WSB) is a decentralized Reddit community, created on January 2012, known for its meme-centric discussions with a particular focus on highly speculative trading strategies. As the largest finance-related subredit with over 13 million subscribers, its substantial growth is largely attributable to the GameStop (GME) short squeeze event, during which the forum expanded from 500,000 users in July 2018 to 10.7 million by June 2021 (Bradley et al. 2023).[8] The conventional view is that WallStreetBets' user base consists of a less affluent and less sophisticated retail investor demographic. Data from a 2020 survey indicates that the user base is overwhelmingly male (90.2%), young (72% below the age of 29), and of modest financial means, with nearly half (47.1%) earning less than $50,000 annually.[9]

WSB posts appear instantly without any pre-publication review, and there is only limited post-hoc moderation for spam and abuse. Posts often highlight high-risk, speculative trades. In line with its speculative culture, WallStreetBets encourages users to share screenshots of high-risk holdings and trades through designated "YOLO" and "Gains/Losses" tags

[7]See https://about.seekingalpha.com/policy-on-pseudonymous-analysts.

[8]New members who joined WSB in early 2021 tend to emphasize coordinated buy-and-hold strategies for a handful of meme stocks, with little regard to the company's fundamentals. For example, John Coffee of Columbia Law School characterized WSB users as a "mob of uninformed, unsophisticated retail traders" (Detrixhe 2021). Also see Asarch (2021) for the anecdotal evidence on new WSB members.

[9]See https://docs.google.com/presentation/d/1ozj-S3eIwSa6ZERs0kTdE1LiMYcN1kwBUGIDVcVlzLg.

(known as 'flairs' on Reddit).[10] The highlighting of trading gains versus losses is likely to trigger reference-dependent thinking (Shefrin and Statman (1985)) , a form of imperfectly rational cognition that has been heavily documented in behavioral economics and finance. There is no direct monetary compensation for posting on WSB. Instead, the reward is social capital, which includes karma points (a quantitative score reflecting a user's reputation and contributions, earned through upvotes from other members), and collective recognition and sense of identity as a community member. WSB permits anonymity and there are no restrictions on using multiple accounts.

Taken together, these structural differences in governance, content guidelines, and user sophistication and incentives create distinct environments for AI deployment and its relationship with markets. In light of our central hypothesis—that AI lowers the cost of generating both informative and misleading content, with the balance shaped by a platform's institutional and behavioral context—these differences are key determinants of how AI affects each platform's opinion ecosystem.

SA fosters reputation-building through in-depth analysis, whereas WSB rewards engagement through memes and high-risk/high-gain narratives. This fundamental dichotomy is the foundation for our analysis. At the broadest level, we are interested in whether contributors adopt AI for rational, information-provision purposes, such as processing and generating signals that help market participants better assess the fundamentals of the assets, or whether adoption is instead driven by engagement-driven motives such as entertainment, identity reinforcement, or persuasion. In the latter case, contributors may pursue glory as influencers and use AI primarily to generate more vivid, accessible, and emotionally resonant content that boosts engagement, rather than to enhance the informational content of their posts.

[10] See https://www.reddit.com/r/wallstreetbets/wiki/linkflair/. For example, users can select from a list of "flairs" to add to their post, where "YOLO" corresponds to trades with minium value of risk at $10,000 in options, or $25,000 in equity, "gain" and "loss" show off a solid winning or losing trade with minimum gain of $2,500 for options and $10,000 for shares. The WSB content guidelines focus primarily on combating spam and abuse with rules such as "no crypto," "no penny stocks," "no paper trading," "no pump-and-dumps," "no market manipulation," and "no social begging." Moderators also intervene to remove posts that misuse 'flairs.'

2.2 Adoption of AI in Content Creation

We begin with general motivation. AI reduces the cost of producing both meaningful and misleading content. Its net effect therefore hinges on whether adoption is channeled toward informational enhancement or informational distortion. Under the *Information Enhancement Hypothesis*, AI strengthens a platform's role in knowledge discovery by helping creators extract, synthesize, and communicate value-relevant signals. Under the *Information Distortion Hypothesis*, AI contributes instead to the generation of vivid, affectively charged, or persuasive narratives that diverge from fundamental information, thereby degrading the informational environment.

Based on this, we develop specific predictions regarding the adoption of AI in financial content creation.

The first mechanism is an information-enhancement channel: AI reduces the cost of producing fundamental analysis and other value-relevant information. Recent studies show that AI substantially reduces the cost and time required to generate professional-quality text, with benefits documented for customer-support agents, professional writers, financial analysts, and retail investors, especially when users lack prior domain expertise.[11] Under this motive for adoption, AI is a complement to information acquisition. Creators adopt AI to produce value-relevant analysis when the unassisted cost of such analysis is high. This occurs when authors cover unfamiliar firms or industries, or when public news flow is thin and independent information discovery is more costly. We therefore expect AI adoption to be more frequent in first-time coverage settings and for firms with lower levels of news availability.

[11] For example, Cao et al. (2024) and Bertomeu et al. (2025) find evidence suggesting that AI helps financial analysts generate more accurate forecasts. Experimental evidence in Noy and Zhang (2023) indicates that access to ChatGPT enables authors to complete professional writing tasks more quickly. Brynjolfsson, Li, and Raymond (2025) find that AI assistance in a Fortune 500 company's customer support operation significantly enhances worker productivity. A 2024 survey of retail investors indicates widespread adoption of GenAI, primarily for gathering and interpreting financial or market data (Blankespoor, Croom, and Grant, 2024). Wang et al. (2025) conduct experiments to compare human and AI agents across essential work-related skills such as data analysis, engineering, computation, writing, and design and find that "agents deliver results 88.3% faster and cost 90.4-96.2% less than humans."

The second mechanism—information distortion—applies when creators use AI to generate narratives that are systematically inaccurate, noisy or misleading. Such narratives may emerge because creators themselves hold biased or misinformed beliefs, or because creators strategically deploy inaccurate narratives to exploit viewers' behavioral biases. In either case, AI enables the production of emotionally resonant, sensational, or identity-affirming content designed to shape beliefs or mobilize attention. This can be without regard to accuracy or with deliberate intent to foster inaccuracy. Experimental evidence shows that LLMs outperform humans in persuasion tasks (Schoenegger et al. 2025) and can be turned into persuasion tools with minimal prompting (Hackenburg et al. 2025). We refer to this broad class of persuasive, non-informational uses as *persuasion engineering*, which includes both attention-grabbing visual content (e.g., images and memes) and crafted narratives that we term *rhetorical engineering* (Matz et al. 2024; Cong et al. 2024). Such uses are especially valuable on lightly moderated platforms with less sophisticated users who are more susceptible to emotions and social influence.

We therefore hypothesize the following:

Hypothesis 1a *The probability of AI adoption for creating financial content increases with information production costs, such as when authors address unfamiliar topics or operate in information-scarce environments.*

Hypothesis 1b *The probability of AI adoption for creating financial content is positively associated with recent surges in retail buying activity.*

We argue that the relevance of these two motives for AI adoption depends critically on the platform context. On curated platforms such as Seeking Alpha, where editorial screening ensures baseline quality standards and users are relatively sophisticated, AI adoption is more likely to to reflect the information-enhancement motive. The platform's incentive structure rewards well-researched fundamental analysis, making AI a natural complement to value-relevant information production, consistent with Hypothesis 1a. In contrast, on lightly moderated and speculative platforms such as WallStreetBets, AI adoption is more likely to reflect information distortion—oriented motives—arising either from creators' own biased beliefs or from attempts to shape or exploit the behavioral biases of others through emotionally charged, entertaining, or persuasive content. In such settings, the use of AI for persuasion engineering becomes particularly valuable during periods of retail-driven trading disconnected from fundamentals, consistent with Hypothesis 1b. We therefore expect Hypothesis 1a to be most relevant for SA, and Hypothesis 1b for WSB.

2.3 User Responses to AI-Generated Content

Having established that AI adoption in content creation can reflect either the *information enhancement* or the *information distortion* motive, we next consider how users respond to AI-generated financial content. We hypothesize that if AI is used to synthesize and disseminate value-relevant information, it will support rational information processing; when it is used to craft distorted or affectively charged narratives, it will amplify behavioral biases and sentiment propagation.

Specifically, if AI adoption is driven by the information-enhancement motive, then AI-generated content will help structure complex information into clearer and more accessible formats, thereby reducing investor uncertainty about fundamentals. Lower uncertainty leaves less room for divergent interpretations, implying that AI-generated posts will tend to facilitate convergence in beliefs. Our key prediction about consensus-building is that this effect will be especially pronounced on curated platforms such as Seeking Alpha, where editorial oversight and a sophisticated user base orient discussion toward fundamentals.

In contrast, when AI adoption is driven by the information-distortion motive—whether due to creators' own biased beliefs or to persuasion engineering—the use of AI generates stylistic, emotionally charged, or sensational narratives rather than accurate value-relevant analysis. Such content can reinforce behavioral biases and shift beliefs away from fundamentals, especially on lightly moderated platforms with less sophisticated users such as WallStreetBets.

These dynamics of information degradation align with existing theories of social transmission. For example, in the model of Han, Hirshleifer, and Walden (2022), the social transmission of a salient signal about a stock, such as a large positive return, can induce naïve receivers to overweight this signal in forming beliefs about future performance. Alternatively, in the model of Hirshleifer (2020b), social transmission can systematically cause signals about a stock to be biased upward, which can induce overoptimism among receivers. Furthermore, there are models in which optimism versus pessimism about a stock spreads via social network (Burnside, Eichenbaum, and Rebelo 2016,Hirshleifer 2020b, Pedersen 2021). In our context, if AI enables an author to make a post more salient and persuasive, unsophisticated viewers are likely to overreact, thereby amplifying sentiment contagion and producing tighter alignment between viewer comments and the sentiment expressed in the initial post.

Based on these considerations, we propose the following two hypotheses about the relationship between AI-generated content and user engagement.

Hypothesis 2 *AI-generated content is associated with greater user consensus, as reflected in lower sentiment dispersion in comments.*

Hypothesis 3 *AI-generated content is associated with greater sentiment contagion, where the sentiment of viewer comments more closely aligns with the sentiment of the initial post.*

We expect Hypothesis 2 to be more applicable to Seeking Alpha, where AI supports value-relevant information processing, and Hypothesis 3 to be more applicable to WallStreetBets, where AI amplifies distortion-oriented narratives and behavioral contagion.

2.4 AI-Generated Content and Capital Market Outcomes

Finally, we consider how AI-generated content maps into capital market outcomes. Consistent with earlier discussion, AI deployment can either enhance information efficiency by facilitating genuine information production or amplify speculative dynamics by reinforcing behavioral distortions.

On curated platforms such as Seeking Alpha, where editorial oversight and user sophistication channel AI toward value-relevant analysis, we expect AI deployment to help investors accurately evaluate firms. If these signals are not fully and immediately incorporated into prices, attentive investors—particularly rational retail traders facing high information costs—may earn higher abnormal returns. Consistent with evidence that aggregate retail buying predicts positive future returns (Kaniel, Saar, and Titman 2008; Barrot, Kaniel, and Sraer 2016; Boehmer et al. 2021), AI-generated content should further strengthen this informativeness by lowering information frictions for these rational retail participants.

Furthermore, the information-enhancement channel implies improvements in information discovery and market efficiency. By reducing uncertainty and mitigating adverse selection (see, e.g., Glosten and Milgrom 1985; Glosten 1989; Glosten and Harris 1988; Huang and Stoll 1997; Stoll 1989), AI-generated value-relevant content results in narrower bid-ask spreads. More broadly, if AI facilitates the incorporation of fundamental information into prices, it should oppose price bubbles and other deviations from intrinsic value.

In contrast, on lightly moderated and speculative platforms such as WallStreetBets, AI adoption is expected to more often derive from information-distortion motives. This results in opposite implications for information discovery in capital markets. As discussed in Hypothesis 3, AI-generated narratives that reinforce creators' own biases or exploit the behavioral tendencies of others can heighten sentiment contagion. Such content encourages noise trading (Black 1986, Barber and Odean 2000), resulting in higher speculative trading volume and greater excess volatility. Consistent with Cookson, Engelberg, and Mullins (2023), who find that echo chambers stimulate trading activity, AI deployment may intensify these dynamics by making persuasive or emotionally charged posts more salient.

Distortionary AI usage can also amplify social-transmission bias (Hirshleifer 2020a), in which social interactions systematically shift ideas or signals in a particular direction as they pass from person to person. The interaction of social transmission bias with short-selling constraints—common in retail-dominated markets—can increase the probability of sharp price run-ups unanchored from fundamentals (e.g., Bali et al. 2025). AI usage is then predicted to generate lottery-like return distributions characterized by occasional extreme upside movements and lower expected returns.

The effect of AI on bid-ask spreads in speculative settings is theoretically ambiguous.

Noise trading may reduce spread by reducing adverse selection (e.g., Glosten and Milgrom 1985; Kyle 1985). However, recent evidence suggests the opposite when retail herding is concentrated.[12] Thus, the relationship between AI content and spreads is ultimately an empirical question, shaped by the relative strength of the adverse-selection channel versus the retail-clustering channel.

Based on these considerations, we propose the following two hypotheses linking AI-generated content to market outcomes:

Hypothesis 4 *AI-generated content associated with reduced uncertainty and improved price discovery, as reflected in narrower bid-ask spreads, more informative retail order flows, and lower likelihood of price bubbles.*

Hypothesis 5 *AI-generated content is associated with more speculative trading, as reflected in higher subsequent abnormal trading volume and return volatility, and more positively skewed ("lottery-like") short-horizon return distributions. The relationship between such content and bid-ask spread could be ambiguous.*

As with user responses, we expect that Hypothesis 4 will be more relevant for Seeking Alpha, where AI supports information-enhancement motives, while Hypothesis 5 will be more relevant for WallStreetBets, where AI amplifies speculative dynamics and sentiment contagion.

3 Data

This section details the data sources and variable construction for our empirical analysis. We begin by introducing our primary data from two social media platforms, Seeking Alpha (SA) and Reddit's WallStreetBets (WSB). We then describe our methodology for identifying AI-generated content. Finally, we define the key variables used in our tests and present summary statistics.

[12] Barber et al. (2022) document substantial clustering in retail demand, and Eaton et al. (2022) find that such clustering reduces liquidity and increases volatility by raising inventory risk for market makers (e.g., Ho and Stoll 1981; Grossman and Miller 1988; Hendershott and Menkveld 2014).

3.1 Social Media Financial Content

Our primary data consist of the textual content collected from two online platforms: Seeking Alpha (SA) and Reddit's r/WallStreetBets forum (WSB). The sample period spans from December 2022, immediately after the launch of ChatGPT in November 2022, to December 2024.

For the Seeking Alpha sample, we collect 83,654 articles published between December 1, 2022, and December 31, 2024, within the Analysis section, excluding earnings call transcripts and corporate presentations to focus on original analytical work. Following past literature, we further restrict the sample to include (1) single-ticker articles (Chen et al. 2014b), and (2) multi-ticker articles for which the author specifies a primary ticker identifying the focal firm (Campbell, DeAngelis, and Moon 2019). 57,581 (or 69%) of the SA articles in our sample period meet this criterion.

For WallStreetBets, we consider all submissions (original posts) and subsequent comments, with submissions accounting for 10.35% of the total messages. The raw dataset contains 969,747 posts and comments. We apply two filters. First, to enable reliable detection of AI-generated content, messages must exceed 50 words which eliminates trivial remarks (about 250 characters, or 40-45 words), leaving 82,523 posts and comments. Second, messages must reference a single valid stock symbol, excluding spurious ticker matches; this yields 64,848 posts and comments.[13]

We process individual articles with GPTZero Model 3.1b (Base build, January 9, 2025) with default settings. Recent work applying GPTZero to a corpus of general-purpose, human-written, and AI-generated articles finds that the False Positive Rate (FPR)—the proportion of human-written text incorrectly classified as AI—is 0.01 or lower for medium to long passages, which corresponds to the article lengths used in our analysis. Similarly, the False Negative Rate (FNR)—the proportion of AI-generated text incorrectly classified as human—is 0.05 or

[13] Section 7 presents robustness checks with alternative samples that include submissions only, or that include all WSB messages regardless of length. Bradley et al. (2023) examines due diligence posts, which are relatively rare in our sample; there are 1,617 such posts (0.17%), and due to their small sample size, we do not analyze them separately.

lower for these passages (Jabarian and Imas 2025).[14]

A natural question is whether it performs well on social-media financial texts that contain domain-specific jargon and analysis. To assess whether GPTZero has substantial resolution ability for financial text, we conduct a validation study using 1,000 pre-ChatGPT human-written SA and WSB articles, along with AI-generated counterparts created from pre-ChatGPT *Wall Street Journal* news articles using multiple LLMs prompted to mimic SA and WSB styles. As shown in Appendix Table A2, GPTZero achieved an accuracy of 98% and an F1 score of 0.99, outperforming mainstream open-source detectors (e.g., Wu et al. 2025) in distinguishing AI from human content in this setting.[15]

Figure 1 plots the monthly fraction of AI content on each platform. On SA, adoption rises sharply after the release of ChatGPT, peaking at 13.7% by late 2023 before dropping to 3% by the end of 2024. This drop coincides with an editorial prohibition on AI use (Bradshaw et al. 2025). On WSB, adoption is slower but steady: AI content remains under 2% until mid-2023 and then climbs gradually to around 7% by December 2024. The slower uptake is likely associated with the short, slang-heavy, and meme-driven format of WSB, where human spontaneity is already low-cost, reducing the relative advantage of early AI tools.

3.2 Variables Construction

We collect financial market data from a number of sources. Daily stock returns, trading volume, and pricing information are from the Center for Research in Security Prices (CRSP); firm-level accounting data are from Compustat; quarterly institutional ownership is from Thomson Reuters Institutional Holdings (13F) database; and analyst coverage is from Insti-

[14] Jabarian and Imas (2025) evaluate four commercial and open-source AI-text detectors using FNR and FPR metrics on a 1,992-passage text corpus spanning everyday genres (news, blogs, consumer reviews, novels, restaurant reviews, and resumes). Verified human-generated texts are paired with AI-generated text using GPT 4o, Claude 3.5 Sonnet, and an open-source model LLama 3.3. The authors find that commercial detectors, including GPTZero, achieve relatively low FNR and FPR, substantially outperforming open-source alternatives, and that these results are stable across AI models, document lengths, and topical domains.

[15] Accuracy is the share of documents correctly classified as AI-generated or human-written. The F1 score is the harmonic mean of precision and recall, providing a single measure that balances between false positively and false negatives.

tutional Brokers' Estimate System (IBES). To construct microstructure measures, we use 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). Here, we focus on the key AI variables and defer the detailed definitions of all other commonly used variables to Appendix Table B.

For AI-content detection, GPTZero analyzes each article or message and outputs the probabilities that it belongs to one of three categories: AI-ONLY," "HUMAN-ONLY," or "MIXED" (human with AI assistance). These probabilities sum to one. For example, an article may have probabilities of 0.58, 0.02, and 0.40 for being "AI-ONLY," "MIXED," and "HUMAN-ONLY," respectively. Based on these outputs, we define a continuous AI probability score, *AIProb*, as the sum of the "AI-ONLY" and "MIXED" probabilities, and use *AIProb* for our article-level analysis. In this example, the article's *AIProb* would be 0.60.[16]

The API also outputs a categorical label for each article, determined by the category with the highest probability. In the same example, the article would be classified as "AI-ONLY" because its probability is the highest among the three categories. At the stock-day level, we define $AIDay^{SA} = 1$ if at least one SA article about a given stock on day t is classified as "AI-ONLY" or "MIXED," and 0 otherwise. Similarly, we define $AIDay^{WSB}$ for WSB posts. These platform-specific AI indicators serve as our key measures of AI presence in the stock-level analysis.

Other article-level variables include length (word count); *Sentiment*, measured as the ratio

$$\frac{N_{\text{positive}} - N_{\text{negative}}}{N_{\text{positive}} + N_{\text{negative}}},$$

where N_{positive} and N_{negative} are the counts of positive and negative words; textual complexity, measured by the proportion of complexity words (Loughran and McDonald 2024); the Fog

[16]We include both AI-ONLY and MIXED for this definition to make sure we capture all potential AI-generated contents. The mean probability of a "MIXED" classification is 2.38% for SA and 0.10% for WSB. For robustness, Appendix Table A4 reports results where *AIProb* is constructed using only the "AI-ONLY" category.

Index, a readability measure defined as

$$0.4 \times \frac{\text{Words}}{\text{Sentences}} + 100 \times \frac{\text{Complex Words}}{\text{Words}},$$

where complex words are have three or more syllables; and indicators for quantitative content (count of numbers) and graphical content (count of images). Viewer disagreement is measured as the standard deviation of sentiment scores across comments posted within the 10-day window beginning on the publication date, calculated using platform-specific sentiment dictionaries.

We align publication times with market trading: articles posted before 4:00 p.m. are assigned to the same trading day, while those posted after 4:00 p.m. or on non-trading days are assigned to the next trading day. Outcome variables are measured after publication, typically over $t+1$ to $t+5$: cumulative abnormal returns (CAR, cumulative buy-and-hold returns adjusted by the market return); abnormal trading volume (AVOL, mean daily log dollar volume over $t+1$ to $t+5$ minus the mean over $t-41$ to $t-11$); effective bid-ask spreads (Spread, mean daily volume-weighted effective bid-ask spread); and realized volatility (Vol, standard deviation of the daily stock returns).

Table 1 presents the summary statistics for our sample. Panel A reports article characteristics. SA articles are substantially longer than that of WSB (1,454 vs. 114 words), more positive in tone, more complex, and contain more quantitative elements and figures. SA articles also attract more comments (11.3), whereas WSB posts draw significantly fewer comments (3.85). These patterns align with the emphasis of SA on in-depth analysis relative to the shorter, meme-centric style of WSB. SA articles have nearly three times the AI probability (10.6% vs. 3.6%), suggesting that authors on SA are more likely to use AI as flagged by GPTZero.

Panel B summarizes author characteristics. Our sample includes 1,998 unique authors on SA and 28,563 on WSB. Of these, 457 (23%) SA authors and 19,025 (67%) WSB authors publish only once; these one-time contributers account for 0.79% of SA and 22.8% of WSB articles, respectively. On average, a SA author publishes 28.78 articles and covers 16.24 stocks across 10.24 four-digit SIC industries. In comparison, the average WSB author publishes 2.27

articles and covers 1.64 stocks across 1.60 industries. Overall, these statistics indicate that SA authors are less numerous but more prolific, covering a broader range of firms, while the vast number of WSB authors contributes less frequently on an individual basis.

Panel C reports the characteristics of the stock-day sample. SA covers more stocks than WSB (3,193 vs. 1,351), resulting a larger stock-day sample. However, for the stocks covered, WSB exhibits higher coverage intensity, generating 9.43 articles per 100 stock-days compared to 3.59 for SA. AI days occur in 0.24% of SA stock–days and 0.18% of WSB stock–days. Firm characteristics indicate that, compared to WSB, SA coverage tilts toward smaller firms with higher book-to-market ratios, lower news coverage, and less analyst coverage. Liquidity profiles also differ: stocks discussed on SA exhibit wider spreads (0.41% vs. 0.29%). Taken together, Table 1 establishes SA and WSB as distinct ecosystems with different informational, engagement, and liquidity environments.

Table 2 describes the association between AI usage and the textual attributes of posts. Columns 2, 4, and 6 including author fixed effects for SA and WSB posts, respectively; columns 5 and 6 also add an indicator for initial submissions to distinguish them from replies. The result indicates that AI content is generally more positive in sentiment, higher in complexity, and harder to read (as measured by the Fog Index). On SA, AI articles also contain fewer images and numbers, whereas on WSB they tend to include fewer images but more numbers—though these associations become insignificant once author fixed effects are included. Notably, WSB submissions exhibit a 193% (6.964%/3.6%) higher AI probability than comments, indicating that users are more likely to rely on AI for creating original posts than for replies.

4 AI Adoption in Content Production

In this section, we empirically investigate the determinants of AI adoption for the creation of financial content. We test whether observed AI deployment is more consistent with the *information enhancement* channel (Hypothesis 1a), in which AI helps overcome information costs, or with the *information distortion* channel (Hypothesis 1b), in which AI use comoves with episodes of intense retail trading and speculative activity. To measure the information costs faced by creators, we proxy their need for assistance using indicators of firm- or industry-level unfamiliarity and the richness of the firm's information environment (firm size, analyst coverage, and news flow). To capture periods conducive to distortion-driven adoption, we identify episodes of elevated retail speculation by examining days with unusually high retail net buying.

4.1 Information Costs

To test Hypothesis 1a, which asserts that AI adoption is used to obtain useful information at lower cost, we examine the relationship between AI probability and proxies for an author's inexperience in covering a specific firm or industry. We estimate the following article-level panel regression:

$$AIProb_{ijt} = \beta_1 Unfamiliarity_{ijt} + Controls + FE + \epsilon_{ijt}, \tag{1}$$

where the dependent variable $AIProb_{ijt}$ is the AI probability score article j covering stock i on day t. $Unfamiliarity_{ijt}$ is an indicator variable that equals one if the author/user covers the stock i for the first time in the past six months.[17] *Controls* include stock size, book-to-market ratio and institutional ownership, as well as the author's past six-month activity (number of articles, number of AI articles, and average comments per article in 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 two-way clustered by stock and day.

The results in Table 3 support Hypothesis 1a on the Seeking Alpha platform. On SA (column 1), first-time firm coverage is associated with a 2.32 percentage points (pp) higher AI probability ($t = 9.42$). This effect is economically significant relative to a mean of 10.65 pp on SA: when venturing into new domains, authors are 22% more likely to use AI. To ad-

[17] We re-estimate using first-time industry coverage and find similar results. See Appendix Table A5.

dress unobserved author characteristics—such as writing style, technological familiarity, or risk preferences—that may influence adoption, column 2 includes author fixed effects. We find that the effect of first-time firm coverage remains positive and robust, though smaller in magnitude.

We also find that firm size is negatively related to AI adoption: SA authors are more likely to use AI when writing about smaller firms, which typically suffer from limited information availability. For author characteristics, those with fewer articles in the past six months, those with prior experience using AI, and those whose articles generate more engagement (measured by comments per article) are all more likely to adopt AI. These associations persist even after controlling for author fixed effects. These patterns are consistent with the information-enhancement channel on SA: contributors turn to AI when traditional research is more costly, in line with Hypothesis 1a.

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 indicates that, unlike on SA, the adoption of AI by WSB users is not primarily associated with efforts to overcome information frictions. Firm size is weakly and negatively correlated with AI adoption, and book-to-market ratio and institutional ownership show no systematic relationship. For WSB author characteristics, column 3 shows higher AI adoption among users with fewer past articles, prior AI use, and lower comment engagement, but these associations weaken once author fixed effects are introduced. These results suggest that on WSB, AI adoption is not systematically tied to learning about new firms or filling gaps in traditional information sources. This tends to oppose the information-enhancement hypothesis on this platform.

We further test Hypothesis 1a by examining whether AI adoption rises during periods when firm-specific information is more scarce. 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 article-level panel regression:

$$AIProb_{ijt} = \beta_1 News_{it} + \beta_2 Analyst_{it} + Controls + FE + \epsilon_{ijt}, \quad (2)$$

where $News_{it}$ is the logarithm of one plus the number of highly relevant Dow Jones news articles (with a Ravenpack relevance score above 90 on firm i in the last three days), and $Analysts$ is the logarithm of one plus the number of analysts following the stock i in the most recent quarter. Control variables, fixed effects, and standard error clustering are identical to the specification in Table 3.

Results in Table 4 are also consistent with Hypothesis 1a on SA. On the SA sample (columns 1-2), AI adoption is significantly less likely when news or analyst coverage is abundant. 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 indicate that SA contributors are more likely to use AI when traditional information sources are limited. In contrast, no such relationship is observed on WSB (columns 3-4), and in some specifications analyst coverage is even positively associated with AI adoption. This divergence highlights that information scarcity is not a primary driver of AI use on WSB.

Overall, these findings are consistent with the idea that AI flattens the learning curve in SA, lowering the marginal cost of research when authors lack deep prior knowledge. However, these patterns are absent on WSB, underscoring that AI adoption differs across platforms, with SA usage aligning with the information-enhancement motive and WSB usage instead consistent with an information distortion motive, to which we now turn.

4.2 Intensity of Retail Buying

We next test Hypothesis 1b, which posits that AI adoption on lightly moderated and speculative platforms such as WSB is associated with retail-driven speculation. Under this hypothesis, users turn to AI when retail interest is elevated, using it to craft persuasive or attention-grabbing narratives that resonate with peers and reinforce community dynamics. In this information-distortion view, AI amplifies sentiment and narrative transmission around retail buying surges, rather than primarily serving to reduce information frictions.

We therefore test whether AI adoption is associated with intense retail investor buying. We estimate the following article-level regression.

$$AIProb_{ijt} = \beta_1 IntenseRetailBuy_{i[t-3,t-1]} + Controls + FE + \epsilon_{ijt}, \tag{3}$$

where $IntenseRetailBuy_{i[t-3,t-1]}$ equals one if retail net buying during days [t-3, t-1] falls in the top 10% of the cross-sectional distribution. Following Boehmer et al. (2021), retail net buying is defined as the number of shares bought minus the number of shares sold by retail investors, scaled by total retail trades. A spike in retail net buying therefore serves as a proxy for strong retail buying pressure. Control variables, fixed effects, and standard error clustering are identical to the specification in Table 3.

The results, presented in Table 5, are consistent with Hypothesis 1b. For messages on WSB (columns 3–4), the coefficient on *IntenseRetailBuy* is positive and highly significant, both statistically and economically. A surge in retail buying predicts a 1.59 to 1.99 percentage point increase in a content's AI probability, which corresponds to a 44% to 55% rise relative to the platform's sample average. This indicates that WSB users are substantially more likely to adopt AI when a stock is experiencing intense retail buying.

In contrast, this relationship is not observed on Seeking Alpha. As shown in columns 1 and 2, the coefficient for *IntenseRetailBuy* is statistically insignificant for SA articles. This divergence suggests that while AI adoption on SA is more closely associated with information frictions and cost-saving (information enhancement), AI adoption on WSB is more strongly linked to retail-driven surges and speculative dynamics (information distortion).

5 User Responses to AI-Generated Contents

We now turn to the association between AI-generated content and user responses. In our framework, AI can either support information enhancement by reducing disagreement and helping users converge on fundamentals (Hypothesis 2), or support information distortion by amplifying sentiment contagion and the social transmission of biased narratives (Hypothesis 3). We first test whether AI is associated with lower disagreement among users, measured as the standard deviation of comment sentiment. This test evaluate the prediction of Hypothesis 2 that AI-generated content fosters greater user consensus. We then test whether AI is related to sentiment contagion between original posts and subsequent user comments to evaluate Hypothesis 3.

5.1 AI Content and Subsequent User Disagreement

We measure disagreement as the standard deviation of sentiment scores in the comments received within ten days of the initial post, standardized to zero mean and unit variance. The regression specification, as detailed in Table 6, is:

$$Disagreement_{ij[t,t+10]} = \beta_1 AIProb_{ijt} + Controls + FE + \epsilon_{ijt}, \tag{4}$$

where $Disagreement_{ij[t,t+10]}$ 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, $AIProb_{ijt}$ is the AI probability of the initial article. *Controls* include the article characteristics used in Table 2, firm characteristics, and author past activity and performance. Stock, day, and author fixed effects are included. Standard errors are two-way clustered by stock and day.

The results in Table 6 support Hypothesis 2 on SA, indicating that the enhanced information acquisition associated with AI-generated content lowers user disagreement. On SA (column 1–2), the coefficient on AI-related probability is consistently negative and statistically significant. With day, stock, and author fixed effects, the coefficient is −0.083, implying that AI-generated articles are associated with a 0.083 standard-deviation reduction in disagreement, which is a 10% decline relative to the mean. These findings are consistent with Hypothesis 2 (information-enhancement) that AI-assisted articles help structure and clarify fundamental information, reducing the scope for divergent interpretations. In contrast, on WSB (columns 3-4), there is no significant association between AI and disagreement. This contrast suggests that while AI fosters user consensus on SA, no such pattern emerges on WSB.[18]

5.2 Sentiment Contagion

As WSB users are less sophisticated and more susceptible to social influence than SA users, WSB users may be more subject to the information distortion effects of AI. We next examine whether the persuasive capabilities of AI (Schoenegger et al. 2025; Hackenburg et al. 2025) amplify sentiment contagion on either platform. Specifically, we test whether AI-generated content is associated with stronger sentiment contagion between articles and subsequent user comments using the following specification:

$$\begin{aligned} CommentSentiment_{ij[t,t+10]} = &\beta_1 AIProb_{ijt} \times Sentiment_{ijt} + \beta_2 AIProb_{ijt} + \\ &\beta_3 Sentiment_{ijt} + Controls + FE + \epsilon_{ijt} \end{aligned} \tag{5}$$

where $CommentSentiment_{ij[t,t+10]}$ is the average sentiment of comments following article j about stock i over the window $[t, t+10]$, $Sentiment_{ijt}$ is the sentiment of the article itself, and $AIProb_{ijt}$ is the AI probability score. The interaction term β_1 captures whether AI content creates stronger sentiment transmission from articles to comments. Other control variables, fixed effects, and standard error clustering are identical to the specification in Table A6.

Table 7 presents the results. For Seeking Alpha, we find no significant association be-

[18] Bradshaw et al. (2025) find that AI articles on SA is linked with lower engagement compared to human articles. Motivated by this, We modify equation 4 by replacing the dependent variable with log number of comments received for a content between $[t, t+10]$. Table A6 presents the results. On Seeking Alpha (columns 1–2), an AI-generated article ($AIProb = 1$) is associated with 6–7% fewer comments compared to a human-generated article ($AIProb = 0$), consistent with Bradshaw et al. (2025). On WSB, there is a similar negative relationship between AI content and follow-up comments without author fixed effects (column 3) but the association becomes insignificant after author fixed effects (column 4).

tween AI and sentiment contagion. Columns 1 and 2 show that the interaction coefficients are -0.022 and -0.046 and insignificant, so there is no evidence that AI-generated content is associated more strongly than human-written content with sentiment contagion. The baseline sensitivity of comment sentiment to article sentiment remains strong at 0.217, but AI does not appear to strengthen this relationship.

In contrast, on WallStreetBets (columns 3 and 4), consistent with sentiment contagion, these interactions coefficients are statistically significant.The interaction coefficients are 0.090 and 0.135, respectively. Given the baseline sensitivity of 0.080 and 0.067, this represents a doubling of sentiment influence when content is AI-generated. Consistent with Hypothesis 3, these results suggest that the persuasive capabilities of AI can amplify the social transmission bias modeled in Hirshleifer (2020b) and Han, Hirshleifer, and Walden (2022). Our evidence suggests that AI, when deployed on speculative platforms, may serve as a powerful amplifier of such dynamics.

The social transmission bias model further predicts that such AI-assisted amplification of positive sentiment through social networks contributes to price bubbles and the emergence of "lottery-like" stocks. Motivated by this prediction, we next turn to the asset pricing patterns linked to AI-generated content.

6 AI-Generated Content and Capital Market Outcomes

In this section, we investigate the capital market implications of AI-generated content. We examine whether its adoption is associated with improved market quality and price discovery (Hypothesis 4, consistent with information enhancement) or with increased speculative trading and lottery-like return patterns (Hypothesis 5, consistent with information distortion). We first conduct a direct test of informativeness by examining the return predictability of AI article sentiment. Next, we analyze the association between AI content and broad indicators of market quality: abnormal trading volume, volatility, and bid-ask spreads. We then assess whether AI enhances the informativeness of retail order flows. Finally, we test for evidence of speculative price dynamics by investigating whether AI content predicts "lottery-like" return events.

6.1 Informativeness of AI Articles

We proxy for an article's fundamental informativeness by examining the extent to which its sentiment aligns with subsequent stock returns. Under the information-enhancement hypothesis, if an article conveys fundamentally accurate signals, and if this information is not fully and immediately incorporated into market price, then we expect a positive association between article sentiment and future returns. That is, when an article expresses bullish sentiment, future returns should, on average, be positive, and vice versa. By the same token, if bullish sentiment is followed by neutral or negative returns, the article is less likely to contain fundamentally accurate information.

Again, this prediction is premised on such information not being fully and immediately incorporated into market prices. This is consistent with limited attention and trading frictions, and with a large empirical literature documenting underreaction to earnings announcements and analyst forecast revisions (e.g., Bernard and Thomas 1989; Hirshleifer, Lim, and Teoh 2009; Stickel 1991; Chan, Jegadeesh, and Lakonishok 1996).[19]

To test these implications, we estimate the following article-level regression:

$$CAR_{i[t+1,t+5]} = \beta_1 AIProb_{ijt} \times Sentiment_{ijt} + \beta_2 AIProb_{ijt} + \beta_3 Sentiment_{ijt} + Controls + FE + \epsilon_{ijt}, \tag{6}$$

where $CAR_{i[t+1,t+5]}$ is the cumulative market-adjusted abnormal return over the subsequent five trading days, $AIProb_{ijt}$ is the AI probability score assigned to article j about firm i on day t, and $Sentiment_{ijt}$ is the article's sentiment score. The coefficient on *Sentiment* captures the baseline effect—the extent to which the sentiment of human-authored articles align with

[19] An article whose bullish sentiment systematically predicts negative subsequent returns can still be informative for sophisticated traders who profit by trading against its views. Our definition of informativeness here is narrower: we focus on whether the article contains fundamentally accurate information.

future returns. The key variable of interest, the coefficient β_1 on the interaction term, which measures the incremental tendency of AI-generated content relative to human-generated articles to predict higher future returns. *Controls* include the article characteristics in Table 2 and the stock controls as in Equation (8). *FE* corresponds to firm and day fixed effects.

Table 8 presents the results. Columns 1–2 show results for SA articles and columns 3–4 for WSB posts. In column 1, the coefficient on *Sentiment* is −0.149 (significant at the 1% level), indicating that the sentiment of human-written SA articles *negatively* predicts returns in our sample. Hence, our finding, estimated with the December 2022 to December 2024 sample, while different from earlier work in which SA articles were found to be informative (Chen et al. 2014b), is consistent with recent work documenting the decay of the value of social media signals following the 2021 GameStop event (Bradley et al. 2023).[20]

Crucially, the interaction term, *AIProb* × *Sentiment*, is positive and statistically significant in predicting the next five-day returns. Economically, for a one-standard-deviation increase in (bullish) sentiment, a fully AI-generated article (*AIProb* = 1) predicts 35 basis points higher one-week-ahead returns than does a human-generated article (*AIProb* = 0). This result is consistent with the sentiment of AI articles containing useful information about future fundamental prospects.

We next test whether these return patterns simply reflect temporary sentiment-driven price pressure. If AI article sentiment is associated with short-lived mispricing, its initial return predictability should reverse as the distortion corrects. Column 2 finds no reversal in the subsequent $[t+6, t+10]$ window—the interaction coefficient is insignificant—indicating that AI article sentiment is not followed by correction. This reinforces the interpretation that AI-generated content on SA conveys value-relevant information rather than noise, consistent with the information-enhancement hypothesis.

For r/WallStreetBets posts (columns 3-4), however, the interaction coefficients are insignificant across both time windows. This suggests that AI content on this platform is not

[20]The coefficient remains negative and significant even if *AIProb*-related variables are excluded from the regression.

informative about future returns.[21] This divergence between the findings for SA and WSB aligns with our earlier findings that the influence of AI on returns differs across platforms.

6.2 Informativeness of Retail Orders

The return predictability of AI articles on SA suggests that such articles contain value-relevant information not immediately incorporated into prices. This raises a natural question: can such signals help attentive investors—particularly rational retail traders facing high information costs—earn higher returns?

As shown by Kaniel, Saar, and Titman (2008) and Boehmer et al. (2021), retail net buying is associated with positive subsequent abnormal returns, which indicates that net retail buying is informative or is associated with liquidity provision. Our question here is whether the presence of AI-generated content strengthens this relationship.

If AI adoption reflects rational information acquisition (Hypothesis 4), retail net buying should be more predictive of future returns on days when AI-generated content is present. Alternatively, under the information-distortion channel, AI deployment should not enhance—and may even weaken—the link between retail order imbalance and subsequent performance.

To test these predictions, we next turn to stock-level analysis and run the following stock-day panel regression to predict returns:

$$\begin{aligned} CAR_{i[t+1,t+5]} = {} & \beta_1 AIDay_{it} \times OIB_{it} + \beta_2 HumanDay_{it} \times OIB_{it} \\ & + \beta_3 AIDay_{it} \times Sentiment_{it} + \beta_4 HumanDay_{it} \times Sentiment_{it} \\ & + \beta_3 AIDay_{it} + \beta_3 HumanDay_{it} + \beta_4 OIB + \beta_5 Sentiment_{it} \\ & + \beta_6 Attention_{it} + Controls + FE + \epsilon_{it}, \end{aligned} \tag{7}$$

where $\mathrm{CAR}_{i[t+1,t+5]}$ is the market-adjusted 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

[21] The higher R-squared for the WSB sample than the SA sample likely reflects differences in firm coverage: as shown in Table 1, SA incentivizes coverage of smaller, under-followed stocks, whereas WSB posts tend to focus on larger firms.

AI article for stock i on day t, and OIB_{it}, the retail order imbalance calculated as the ratio between the net retail buys and total retail volume (in shares) following Boehmer et al. (2021). The coefficient on this term (β_1) captures whether retail order flows are more informative on AI days.

We also include $HumanDay_{it}$, an indicator for firm-days with at least one human-generated article, where human articles are identified by GPTZero as having the highest probability of being "HUMAN-ONLY" content. We further control for the total number of articles ($Attention_{it}$) as a proxy for social media attention and average sentiment of all content for that stock on that day (*Sentiment*). *Controls* is a vector that includes size, book-to-market ratio, institutional ownership, analyst coverage, past news, an earnings announcement dummy, and past one-month return and volatility. We further include platform-specific interactions between $AIDay_{it}$ and platform-level sentiment. Standard errors are two-way clustered by stock and day.

Table 9 presents the results. Across all specifications, we find a positive and statistically significant coefficient on the $AIDay^{SA} \times OIB$ interaction in the SA sample. The coefficient of 0.431 in column 2 indicates that on days with AI content on Seeking Alpha, a one-standard-deviation greater OIB predicts an additional 43.1 basis points in cumulative abnormal returns over the subsequent five trading days, compared to days without AI content. This association is economically substantial and statistically significant. In contrast, the interaction between $HumanDay^{SA}$ and OIB is much smaller (0.053) and statistically insignificant, indicating that human articles do not enhance the return predictability of retail order flows. These findings are consistent with Hypothesis 4: AI content on SA appears to enhance the informativeness of retail trading, in line with the information-enhancement channel.

In contrast, the interaction coefficient for the WSB sample ($AIDay^{WSB} \times OIB$) is positive but not statistically significant (columns 3–4), suggesting no detectable association between AI presence and the strength of the OIB-return relationship on WSB. The $HumanDay^{WSB} \times OIB$ interaction is similarly insignificant (around 0.070), indicating that neither AI nor human content on WSB enhances the informativeness of retail order flows. This difference between platforms may reflect the distinct nature of their user bases and content types, with the more analytical focus of Seeking Alpha being associated with improved price discovery from AI-assisted information synthesis. The combined model (columns 5–6) shows the same pattern. The SA interaction remains significant while the WSB interaction remains insignificant.

Overall, these results suggest that AI deployment on SA is consistent with Hypothesis 4: AI content is associated with more informative retail order flows, consistent with AI-generated contents helping rational and attentive retail investors on this platform better synthesize and act upon complex information. The contrast with human content—for which the interaction effect is much smaller and insignificant—indicates that this finding reflects AI-specific mechanisms rather than general content effects. On WSB, neither AI nor human content appears to enhance the informativeness of retail trading.

6.3 Volume, Volatility, and Bid-ask Spreads

We next examine how the presence of AI-generated content is associated with market quality measures such as trading behavior and stock liquidity. We estimate the following stock-day panel regression using all trading days:

$$\begin{aligned} MarketQuality_{i[t+1,t+5]} = {} & \beta_1 AIDay_{it} + \beta_2 HumanDay_{it} + \beta_3 Attention_{it} + \beta_4 Sentiment_{it} \\ & + Controls + FE + \epsilon_{it}. \end{aligned} \tag{8}$$

MarketQuality refers to the following trading activity variables measured over the subsequent five days: abnormal trading volume, return volatility, and the bid-ask spread. All dependent variables are standardized, so the coefficients represent changes in standard deviations. The key independent variable is $AIDay_{it}$, an indicator for firm-days with at least one AI-generated article for a given platform. Since bid-ask spreads are highly persistent (Chordia, Sarkar, and Subrahmanyam 2005), we also control for the average daily bid-ask spread in the past one month when *MarketQuality* is measured by bid-ask spread. For the other explanatory variables and fixed effects included, definitions follow those provided in Table

7.

We estimate this regression separately for the SA and WSB samples. The SA sample includes all stocks that are covered by at least one SA article during our sample period, while the WSB sample includes all stocks that appear in at least one WSB message. For the pooled sample analysis, we use the union of these two samples—that is, all stocks that appear on either SA or WSB during our sample period. On days when a stock has no content on a particular platform, we set the *AIDay*, *Sentiment*, and *Attention* variables to zero for that platform, following Cookson et al. (2024).

The findings, presented in Table 10, indicate important differences in the relation of AI to trading and capital market outcomes across platforms. Columns 1–3 examine Seeking Alpha: over the next five trading days, AI presence is associated with a 0.018 standard deviation (SD) decrease in effective bid-ask spreads (or a 5.3% decrease relative to the sample mean). This pattern of tighter spreads is consistent with the information-enhancement channel and Hypothesis 4: AI-generated content on SA appears to reduce uncertainty and adverse selection costs faced by liquidity providers. The coefficients on abnormal volume and volatility are also negative although insignificant.

On WSB, shown in columns 4–6, we observe a markedly different pattern for AI content. AI presence is associated with a 0.086 SD increase in abnormal volume, a 0.074 SD increase in volatility (significant at the 10% level).[22] The divergent relationship between AI content and market outcomes on the SA and WSB platforms underscores the critical role of the institutional and behavioral environment in shaping AI's market influence.

In the combined model (columns 7–9), where we pool observations from both platforms, the opposing effects across platforms persist. For SA, AI content is associated with a 0.016 SD decrease in spreads (significant at the 1% level), while the coefficients on volume (0.015 SD) and volatility (0.022 SD) are positive but not statistically significant.

Overall, the across-platform comparison suggests that the relationship between AI with

[22]The coefficient on $HumanDay^{WSB}$ for volatility is −0.066 (significant at the 1% level). The negative coefficient represents the effect of human content conditional on total attention. In an alternative specification omitting the $Attention_{it}^{WSB}$ control, $HumanDay^{WSB}$ becomes positive and significant, indicating that the unconditional association of human posts with volatility is positive, similar to AI content.

market variables depend on institutional context. On SA, where editorial review ensures content quality and with relatively sophisticated user base, AI is associated with value-relevant content and greater liquidity, consistent with Hypothesis 4. On WSB, where anyone can instantly post AI-generated narratives, the same technology is linked to higher volatility and abnormal volume in subsequent periods, consistent with Hypothesis 5 and the information-distortion channel.

6.4 Lottery-like Return Events

We have seen that the combination of elevated trading volume and volatility following AI-generated content on WSB is consistent with speculative trading. An alternative explanation is that these patterns are instead driven by large fundamental information shocks. To differentiate these possibilities, we next test whether AI content predicts extreme, lottery-like return events—sharp price run-ups. Such patterns of sharp price run-up followed by return reversals have been identified by past studies as characteristic of overvaluation (e.g., Kumar, Page, and Spalt 2011; Bali, Cakici, and Whitelaw 2011).

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-day 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 less likely to be associated with fundamental information arrival. We estimate the following stock-day logistic regressions to predict the probability of these events:

$$\begin{aligned} MAX/Lottery_{i[t+1,t+5]} = {} & \beta_1 AIDay_{it} + \beta_2 HumanDay_{it} + \beta_3 Attention_{it} + \beta_4 Sentiment_{it} \\ & + Controls + FE + \epsilon_{it}, \end{aligned} \tag{9}$$

where the dependent variables are indicators for whether stock i experiences a MAX or lottery event either during the subsequent five-day window $[t+1, t+5]$. The explanatory variables, fixed effects, and standard error clustering are identical to the specification in Table 10.

Table 11 presents striking results. On WSB, AI content predicts the occurrence of both types of extreme events. For MAX events, AI presence is associated with an increase in odds by 18.0% ($e^{0.166}-1$, column 3) in the next five days. The lottery event results are even more dramatic: AI content is associated with an increase in odds by 44.2% ($e^{0.366}-1$, column 4) for the next five days. In the combined sample (columns 5–6), the WSB AI Day coefficients remain positive and significant, with odds increases of 18.6% ($e^{0.171}-1$) for MAX events and 49.7% ($e^{0.403}-1$) for lottery events. In contrast, $HumanDay^{WSB}$ shows no significant association with either type of extreme event, suggesting that the lottery-like dynamics are specific to AI-generated content rather than reflecting general posting activity on WSB.

On SA, we find the opposite pattern. AI presence is associated with a *reduction* in lottery events: the coefficient of −0.336 (marginally significant at the 10% level, column 2) implies that AI content is associated with a 28.5% decrease in the odds of lottery events ($1-e^{-0.336}$). Human-generated content shows a stronger negative association: $HumanDay^{SA}$ is associated with a 30.6% decrease in lottery event odds ($1-e^{-0.365}$, significant at the 5% level). The coefficients for MAX events are also negative for both AI Day (−0.027) and Human Day (−0.065) on SA, though not statistically significant. In the combined model (columns 5–6), the negative associations for SA content persist, though they are attenuated and no longer reach conventional significance levels.

Extreme right-tail events are arguably symptomatic of price bubbles, i.e., speculative overvaluation, rather than of fundamental news. So the finding that AI presence predicts these return episodes on WSB supports the relevance of Hypothesis 5 and the information-distortion channel on that platform. Conversely, the absence—and in some cases reduction—of lottery-like events on SA is consistent with Hypothesis 4 and the information-enhancement channel.

7 Further Analaysis and Robustness

This section provides additional analysis and robustness checks for our main findings. First, the results remain when we restrict the WSB sample to original posts (submissions). Second, our inferences are not sensitive to including or excluding very short WSB messages, for which AI-content detection is likely noisy.[23]

7.1 WSB Submissions Only

A possible concern is that our WSB findings could be influenced by pooling high-effort original posts (submissions) with subsequent comments. It is possible that original posts and comments play different roles. To address this, we restrict the WSB sample to submissions only and replicate Table 3 to 6. Results are presented in Appendix Table A8.

Despite the dramatically reduced sample size, our key findings remain consistent with those obtained from the full WSB sample. Columns 1-6 show that proxies for information cost, such as covering a firm for the first time (Unfamiliarity) or number of news articles (News), remain insignificant in explaining AI probability. Conversely, the coefficient on Intense Retail Buy remains significantly positive. The probability of using AI in crafting WSB submissions increases by 13.41–17.53 percentage points after extreme retail buying, relative to an average AI probability of 12.54 percentage points for submissions on WSB. Columns 7-10 show that, after controlling for author fixed effects, AI content is not statistically associated with the number of follow-up comments or the level of disagreement.

[23]The use of AI on Seeking Alpha declined substantially following when the platform started to impose a ban on AI-generated content around the end of 2023 (see Figure 1). The drop in the fraction of AI-generated content started on December 2023, and then stabilizes to around 3% after March 2024. As an additional robustness check, we examined outcomes before and after this ban. However, the sharpness of this reduction implies that the post-ban sample of AI articles may be too small to deliver strong statistical power. We confirm this in the data. We extend the main analyses from Table 3 through Table 6 by interacting the key variables with indicators for the pre- and post-ban periods (omitting the intercept). Appendix Table A7 shows that the documented patterns are concentrated in the pre-ban period, whereas post-ban the associations are generally insignificant. Seeking Alpha introduced its own "Analyze with AI" feature around May 2025 as part of an update to its Premium subscription's Advanced Charting tool, which use Generative AI to "compose an up-to-date easy-to-read company report."

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 and including an indicator variable for messages exceeding 50 words in all specifications.

We replicate Table 3 to 6 using this sample and report the results in Appendix Table A9. The findings on the determinants of AI adoption and user engagement remain robust. The proxies for information-production cost remain insignificant. Furthermore, *IntenseRetailBuy* remains positive and marginally significant, confirming that this result is not an artifact of our sample filtering.

8 Conclusion

This paper provides some of the first evidence on how generative AI is reshaping financial discourse across online platforms, and the association of generative AI use with investor behavior and market outcomes. Motivated by the idea that AI can either enhance information production or distort it—as the technology simultaneously lowers the cost of generating value-relevant analysis and persuasive, affectively charged narratives—we compare two prominent platforms with sharply contrasting institutional environments: the curated Seeking Alpha, where contributors are primarily semi-professional analysts, and the largely unmoderated Reddit's r/WallStreetBets, dominated by retail traders with limited sophistication. This comparison allows us to explore how platform governance, user sophistication, and incentive structures are related to differences in how AI's lowered content-production costs manifest in financial discourse.

Our analysis uses a state-of-the-art AI detection tool which has been found to achieve high accuracy across various genres of text while minimizing false positives (Jabarian and Imas 2025). We further verify that the detection is highly effective in distinguishing AI- from human-generated text in the context of online investment commentary..

We find that AI adoption surged following ChatGPT's launch, but with sharply divergent patterns. On Seeking Alpha, adoption is consistent with the information enhancement channel: authors rely on AI most when information acquisition is costly—such as when covering unfamiliar firms or operating in information-scarce environments. In contrast, on WallStreetBets, AI usage spikes during episodes of intense retail buying, consistent with the information distortion channel in which AI facilitates emotionally resonant, persuasive, or bias-amplifying narratives rather than value-relevant analysis.

User responses display an analogous split: on Seeking Alpha, AI-generated posts are associated with lower disagreement among commenters, consistent with AI reducing uncertainty and fostering consensus. On WallStreetBets, by contrast, AI posts is linked to stronger sentiment contagion, consistent with AI augmenting the salience and transmissibility of speculative or affectively charged narratives.

Market reactions similarly suggest that the role of AI differs substantially across the two platforms. On Seeking Alpha, AI article sentiment positively aligns with future returns, and following AI deployment spreads narrow and retail order flows become more informative. This evidence is consistent with improved information transmission and enhanced price discovery. On WallStreetBets, AI deployment is followed by higher trading volume, greater volatility, and a higher incidence of lottery-like return outcomes, consistent with AI amplifying speculative dynamics and sentiment-driven trading rather than improving the incorporation of fundamental information into prices.

These findings bear upon policy debates about whether AI deployment enhances or undermines the information ecosystem of financial markets. Our findings suggest that the same tool may facilitate price discovery on a curated platform while amplifying exuberance on an unmoderated one. In essence, the nature of AI-assisted thought transmission—whether it serves to disseminate rational analysis or to engineer persuasive, speculative narratives—is not inherent in the technology itself. Instead, it is critically shaped by the governance, incentives, and community norms of the platform on which it operates.

This highlights a key tension in democratizing financial discourse. While AI can lower barriers to entry, our evidence from WSB suggests that, absent appropriate governance and investor education, there is a risk that AI degrades rather than improves market quality. This risk is particularly acute in environments that function as “product market traps,” wherein users participate due to social pressure or a “feel of missing out” (FOMO), even to their own detriment (Bursztyn et al. 2025a,b). Although these studies analyze social dynamics in non-financial settings—including decisions to use platforms such as TikTok and Instagram and the demand for AI-based learning tools—the underlying mechanisms of peer pressure, conformity, and contagion are closely related. Our findings suggest that within investment social media generative AI can similarly act as an accelerant of such mechanism. Understanding how AI-mediated transmission of ideas interacts with social pressures can help illuminate the design of platform protocols and governance.

References

Acemoglu, D., A. Ozdaglar, and J. Siderius. 2025. AI and social media: A political economy perspective. Working paper, Massachusetts Institute of Technology.

Asarch, S. 2021. Reddit's wallstreetbets is facing a culture divide as new users flood the forum. *Business Insider*, https://www.businessinsider.com/wallstreetbets-reddit-forum-divided-as-new-users-flood-subreddit-2021-2.

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., and T. Odean. 2000. Trading is hazardous to your wealth: The common stock investment performance of individual investors. *Journal of Finance* 55:773–806.

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.

Barrot, J.-N., R. Kaniel, and D. Sraer. 2016. Are retail traders compensated for providing liquidity? *Journal of Financial Economics* 120:146–68.

Bernard, V. L., and J. K. Thomas. 1989. Post-earnings-announcement drift: Delayed price response or risk premium? *Journal of Accounting Research* 27:1–36.

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* 80:101782.

Black, F. 1986. Noise. *Journal of Finance* 41:529–43.

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

Blankespoor, E., E. deHaan, and Q. Li. 2025. Generative ai in financial reporting. Research Paper 4986017, Stanford University Graduate School of Business.

Boehmer, E., C. M. Jones, X. Zhang, and X. Zhang. 2021. Tracking retail investor activity. *Journal of Finance* 76:2249–305.

Bradley, D., J. Hanousek, Jan, R. Jame, and Z. Xiao. 2023. Place your bets? The value of investment research on reddit's wallstreetbets. *Review of Financial Studies* 37:1409–59.

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.

Burnside, C., M. Eichenbaum, and S. T. Rebelo. 2016. Understanding booms and busts in housing markets. *Journal of Political Economy* 124:1088–147.

Bursztyn, L., B. R. Handel, R. Jiménez-Durán, and C. Roth. 2025a. When product markets become collective traps: The case of social media. *American Economic Review* Forthcoming.

Bursztyn, L., A. Imas, R. Jiménez-Durán, A. Leonard, and C. Roth. 2025b. Social dynamics of ai adoption. NBER Working Paper 34488, National Bureau of Economic Research.

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.

Cao, S., W. Jiang, J. Wang, and B. Yang. 2024. From man vs. machine to man + machine: The art and AI of stock analyses. *Journal of Financial Economics* 160:103910.

Chan, L. K. C., N. Jegadeesh, and J. Lakonishok. 1996. Momentum strategies. *Journal of Finance* 51.

Chang, A., X. Dong, X. Martin, and C. Zhou. 2023. Ai democratization and trading inequality. *Olin Business School Center for Finance & Accounting Research Paper No. 2024/19* .

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

———. 2014b. Wisdom of crowds: The value of stock opinions transmitted through social media. *Review of Financial Studies* 27:1367–403.

Chen, H., and B.-H. Hwang. 2022. Analysts' and the crowd's earnings forecasts. *The Accounting Review* 97:153–78.

Cheng, Q., P. Lin, and Y. Zhao. 2025. Does generative ai facilitate investor trading? early evidence from chatgpt outages. *Journal of Accounting and Economics* 80:101821.

Chordia, T., A. Sarkar, and A. Subrahmanyam. 2005. An empirical analysis of stock and bond market liquidity. *Review of Financial Studies* 18:85–129.

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.

Detrixhe, J. 2021. Good luck proving reddit traders did anything illegal by pumping gamestop. *Quartz*, https://qz.com/1965494/are-wallstreetbets-reddit-traders-manipulating-gamestop-shares.

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

Dim, C. 2024. The role of social media in transmitting information across firms. *Journal of Financial and Quantitative Analysis* 59:1126–58.

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.

Farrell, M., T. C. Green, R. Jame, and S. Markov. 2022. The democratization of investment research and the informativeness of retail investor trading. *Journal of Financial Economics* 145:616–41.

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

Glosten, L. R. 1989. Insider trading, liquidity, and the role of the monopolist specialist. *Journal of Business* 62:211–35.

Glosten, L. R., and L. E. Harris. 1988. Estimating the components of the bid/ask spread. *Journal of Financial Economics* 21:123–42.

Glosten, L. R., and P. R. Milgrom. 1985. Bid, ask and transaction prices in a specialist market with heterogeneously informed traders. *Journal of Financial Economics* 14:71–100.

Grossman, S. J., and M. H. Miller. 1988. Liquidity and market structure. *Journal of Finance* 43:617–33.

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

Hackenburg, K., B. M. Tappin, L. Hewitt, E. Saunders, S. Black, H. Lin, C. Fist, H. Margetts, D. G. Rand, and C. Summerfield. 2025. The levers of political persuasion with conversational AI. *arXiv:2507.13919v1* .

Han, B., D. Hirshleifer, and J. Walden. 2022. Social transmission bias and investor behavior. *Journal of Financial and Quantitative Analysis* 57:390–412.

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

Hendershott, T., and A. J. Menkveld. 2014. Price pressures. *Journal of Financial Economics* 114:405–23.

Hirshleifer, D. 2020a. Presidential Address: Social Transmission Bias in Economics and Finance. *Journal of Finance* 75:1779–831.

———. 2020b. Presidential address: Social transmission bias in economics and finance. *Journal of Finance* 75:1779–831.

Hirshleifer, D., S. S. Lim, and S. H. Teoh. 2009. Driven to distraction: Extraneous events and underreaction to earnings news. *Journal of Finance* 64:2289–325.

Ho, T., and H. R. Stoll. 1981. Optimal dealer pricing under transactions and return uncertainty. *Journal of Financial Economics* 9:47–73.

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.

Huang, R. D., and H. R. Stoll. 1997. The components of the bid-ask spread: A general approach. *Review of Financial Studies* 10:995–1034.

Jabarian, B., and A. Imas. 2025. Artificial writing and automated detection. Working Paper, University of Chicago.

Kaniel, R., G. Saar, and S. Titman. 2008. Individual investor sentiment and stock returns. *Journal of Finance* 63:273–310.

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

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

Kogan, S., A. L. Kostyuk, and M. Solovyeva. 2023. Linguistic complexity and market reactions to analyst reports. *Journal of Financial Economics* 150:103708.

Kumar, A., J. K. Page, and O. G. Spalt. 2011. Religious beliefs, gambling attitudes, and financial market outcomes. *Journal of Financial Economics* 102:671–708.

Kyle, A. S. 1985. Continuous auctions and insider trading. *Econometrica* 53:1315–35.

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.

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]* .

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

Pedersen, L. H. 2021. Game on: Social networks and markets. Working Paper, Copenhagen Business School.

Schoenegger, P., F. Salvi, J. Liu, X. Nan, R. Debnath, B. Fasolo, E. Leivada, G. Recchia, F. Günther, et al. 2025. Large language models are more persuasive than incentivized human persuaders. *arXiv:2503.01844* .

Shefrin, H., and M. Statman. 1985. The disposition to sell winners too early and ride losers too long: Theory and evidence. *Journal of Finance* 40:777–90.

Stickel, S. E. 1991. Common stock returns surrounding earnings forecast revisions. *Accounting Review* 66:402–16.

Stiglitz, J. E., and M. Ventura-Bolet. 2025. The impact of ai and digital platforms on the information ecosystem. NBER Working Paper 34318, National Bureau of Economic Research.

Stoll, H. R. 1989. Inferring the components of the bid-ask spread: Theory and empirical tests. *Journal of Finance* 44:115–34.

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:2306.05540* .

Wang, Z. Z., Y. Shao, O. Shaikh, D. Fried, G. Neubig, and D. Yang. 2025. How do ai agents do human work? comparing ai and human workflows across diverse occupations.

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* .

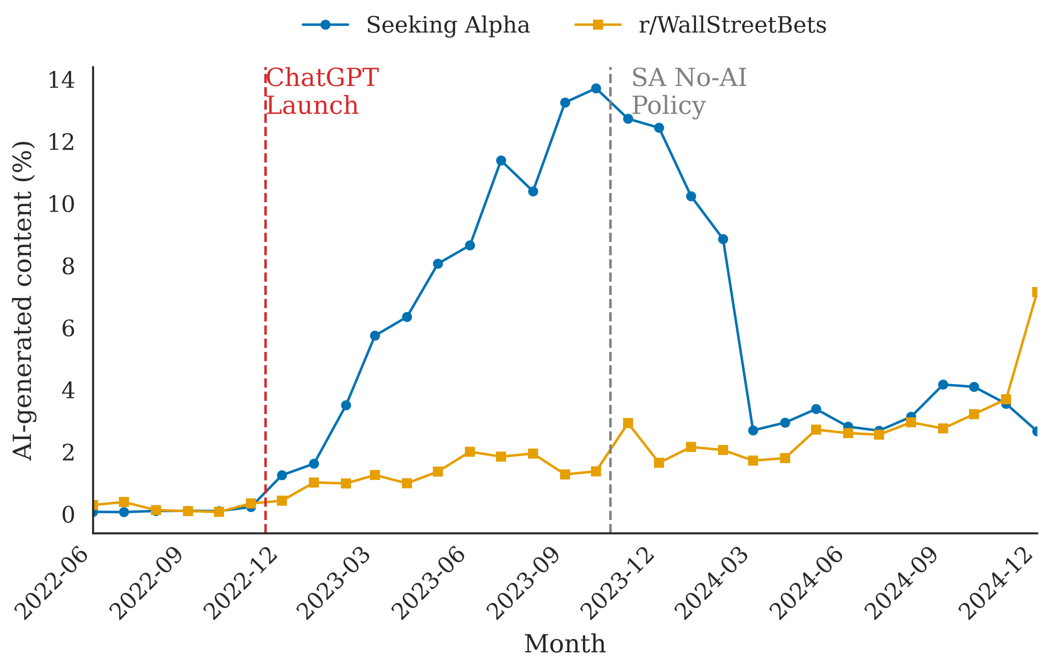
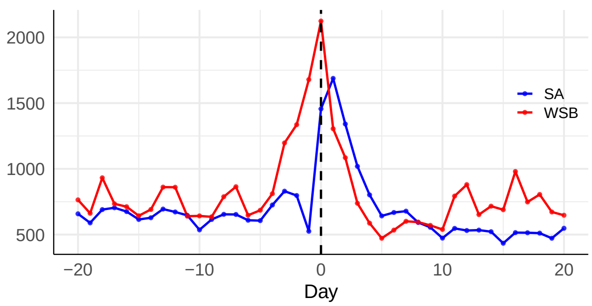
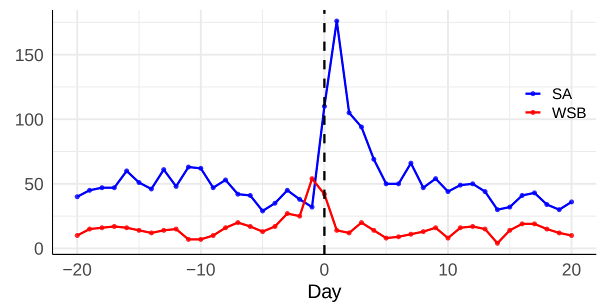


Figure 1: AI-generated Content Over Time

This figure plots the monthly share of AI-generated content on two platforms–Seeking Alpha and Reddit's r/WallStreetBets– from March 2022 through December 2024. The AI Proportion for each platform is computed from GPTZero's classification of all sampled articles/posts on that platform. The red vertical dashed line marks November 30, 2022, the launch date of ChatGPT. The grey vertical dashed line marks November 30, 2023, when Seeling Alpha started the no-AI policy.

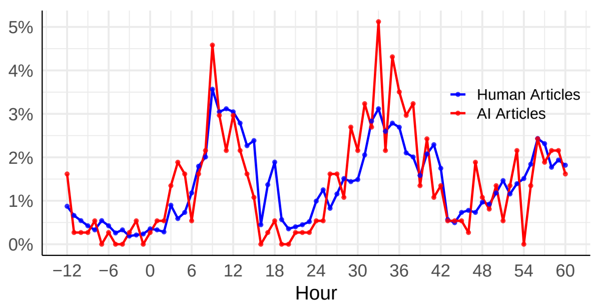


(a) Human Articles

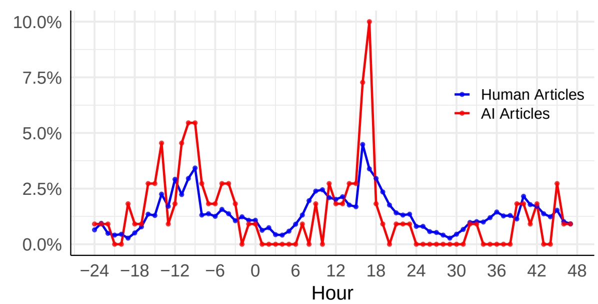


(b) AI Articles

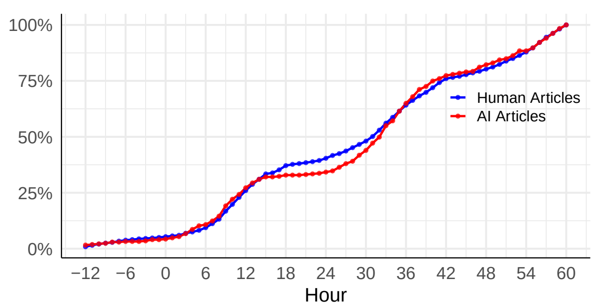
Figure 2: **Articles Around Earnings Announcements.** This figure plots the daily number of articles posted on Seeking Alpha (SA) and WallStreetBets (WSB) around an earnings announcement. Day 0 corresponds to the announcement day.



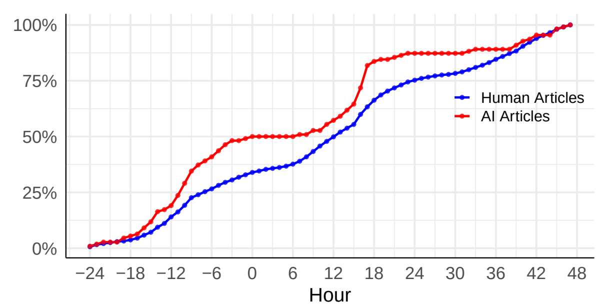
(a) Hourly Share on SA



(b) Hourly Share on WSB

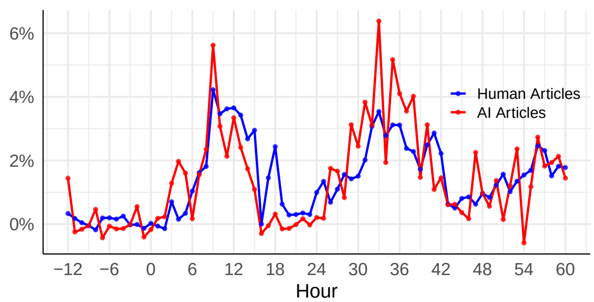


(c) Cumulative Hourly Share on SA

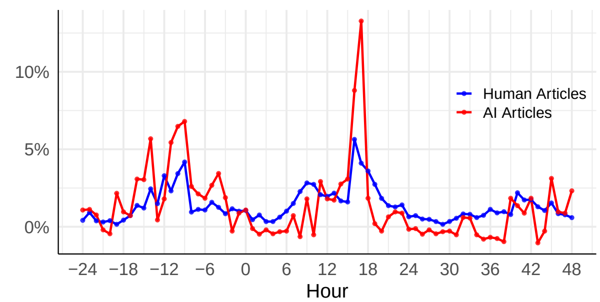


(d) Cumulative Hourly Share on WSB

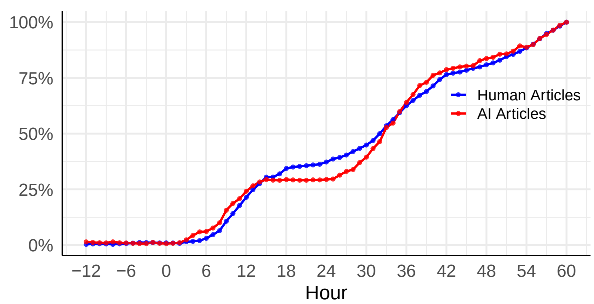
Figure 3: **Hourly Distribution of Articles Around Earnings Announcements.** This figure plots the hourly distribution of articles posted on Seeking Alpha (SA) and WallStreetBets (WSB) in the three-day window around an earnings announcement. Hour 0 corresponds to the first hour of the announcement day. Hourly share is calculated as the number of articles published in a given hour divided by the total number of articles in the three-day window. Because Seeking Alpha articles undergo an editorial review process that typically takes 10–14 hours, we assume an average delay of 12 hours and shift all SA plots by 12 hours. WSB posts are published instantaneously, so no adjustment is applied.



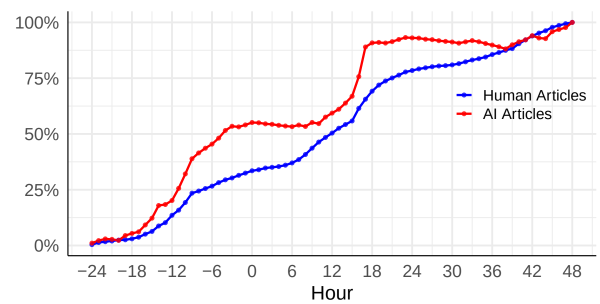
(a) Hourly Share on SA



(b) Hourly Share on WSB



(c) Cumulative Hourly Share on SA



(d) Cumulative Hourly Share on WSB

Figure 4: **Hourly Distribution of Abnormal Articles Around Earnings Announcements.** This figure plots the hourly distribution of articles posted on Seeking Alpha (SA) and WallStreetBets (WSB) in the three-day window around an earnings announcement. Hour 0 corresponds to the first hour of the announcement day. Hourly share is calculated as the number of articles published in a given hour divided by the total number of articles in the three-day window. Because Seeking Alpha articles undergo an editorial review process that typically takes 10–14 hours, we assume an average delay of 12 hours and shift all SA plots by 12 hours. WSB posts are published instantaneously, so no adjustment is applied.

Table 1. Summary statistics

This table presents summary statistics of article, author and firm characteristics from Seeking Alpha and Reddit's r/WallStreetBets during the sample period (December 2022–December 2024). Panel A reports summary statistics on average article characteristics; Panel B, on total author activity; and Panel C, on variables at the stock-day level. Detailed definitions of all variables are provided in Appendix Table A1.

Panel A: Article Characteristics

	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	3.85	36.67	0	1	2	7.47***
Disagreement	57,581	0.04	0.05	0.00	0.04	0.07	64,848	0.02	0.04	0.00	0.00	0.02	0.05***

Panel B: Author Characteristics

	SA						WSB					
	N	Mean	SD	p25	p50	p75	N	Mean	SD	p25	p50	p75
Author Article Count	1,998	28.78	80.88	2	5	18	28,563	2.27	5.07	1	1	2
# Covered Stocks	1,998	16.24	36.90	1	4	13	28,563	1.64	2.09	1	1	2
# Covered Industries	1,998	10.24	18.92	1	4	10	28,563	1.60	1.77	1	1	2

Panel C: Stock-Day Sample Characteristics

	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
News	1,484,486	2.59	1.28	1.79	2.71	3.43	653,497	2.92	1.32	2.20	3.09	3.76
Earnings Day	1,484,486	0.01	0.12	0	0	0	653,497	0.02	0.12	0	0	0
$Ret_{[t-21,t-1]}$ (%)	1,484,486	1.60	23.80	-6.94	0.30	7.95	653,497	1.72	20.35	-6.56	0.47	7.75
$Volatility_{[t-21,t-1]}$ (%)	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. Textual Attributes of AI-Generated Articles

This table examines the association between GPTZero's AI probability score and textual attributes of articles/posts at the article level using the following regression:

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

where $AIProb_{ijt}$ is the AI probability score for content j about firm i, posted on day t. $\boldsymbol{X}_{ijt}$ is 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 results for posts on Reddit's r/WallStreetBets. Standard errors are two-way clustered by stock and day, with t-statistics 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. Author Unfamiliarity and AI Adoption

This table examines the association between author unfamiliarity and the likelihood of AI use at the article level using the following panel regression specification:

$$AIProb_{ijt} = \beta_1 Unfamiliarity_{ijt} + Controls + FE + \epsilon_{ijt},$$

where $AIProb_{ijt}$ is the AI probability score for content j about firm i, posted on day t. $Unfamiliarity_{ijt]}$ is an indicator equal to one if the author has not covered firm i in the prior six months. *Controls* include firm characteristics and measures of the author's prior activity/performance (each standardized to zero mean and unit variance). Columns 1–3 report results for articles published on Seeking Alpha; Columns 4–6 report results for posts on Reddit's r/WallStreetBets. Standard errors are two-way clustered by stock and day; t-statistics are in parentheses. *p <.1; **p <.05; ***p <.01.

	SA		WSB	
	(1)	(2)	(3)	(4)
Unfamiliarity	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. Information Scarcity and AI Adoption

This table examines the association between information scarcity–proxied by recent news flow and analyst coverage-and article-level AI usage using the following regression specification:

$$AIProb_{ijt} = \beta_1 News_{it} + \beta_2 Analyst_{it} + Controls + FE + \epsilon_{ijt},$$

where $AIProb_{ijt}$ is the AI probability score for content j about firm i, posted on day t. $News_{it}$ is the log of one plus the number firm-specific news items in the prior three days and $Analysts_{it}$ is the log of one plus the number of analysts covering the firm in the most recent quarter. *Controls* include firm characteristics and measures of the author's prior activity/performance (each standardized to zero mean and unit variance). Columns 1–3 report results for articles published on Seeking Alpha; Columns 4–6 report results for posts on Reddit's r/WallStreetBets. Standard errors are two-way clustered by stock and day; *t*-statistics are 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. Intense Retail Buying and AI Adoption

This table examines the association between recent intense retail buying and article-level AI usage using the following panel regression specification:

$$AIProb_{ijt} = \beta_1 IntenseRetailBuy_{i[t-3,t-1]} + Controls + FE + \epsilon_{ijt},$$

where $AIProb_{ijt}$ is the AI probability score for content j about firm i, posted on day t. *IntenseRetailBuy* is an indicator equal to one if retail order imbalance over [t-3:t-1] ranks in cross-sectional top 10%. *Controls* include firm characteristics and measures of the author's prior activity/performance (each standardized to zero mean and unit variance). Columns 1–2 report results for articles published on Seeking Alpha; Columns 3–4 report results for posts on Reddit's r/WallStreetBets. Standard errors are two-way clustered by stock and day; t-statistics are in parentheses. *p <.1; **p <.05; ***p <.01.

	SA		WSB	
	(1)	(2)	(3)	(4)
IntenseRetailBuy	-0.186	-0.460	1.588**	1.992***
	(-0.31)	(-1.10)	(2.53)	(3.19)
Size	-0.775	-1.844**	-1.548**	-0.713
	(-0.72)	(-2.23)	(-2.00)	(-0.68)
BM	0.949***	0.347	0.227	-0.226
	(2.85)	(1.00)	(1.47)	(-1.12)
IO	1.095**	0.440	0.427	-0.565
	(1.99)	(1.13)	(1.07)	(-1.37)
Author Article	-3.719***	-1.135***	-0.606***	-0.043
	(-21.86)	(-8.04)	(-9.01)	(-0.49)
Author AI Article	8.499***	2.510***	0.757***	-0.362**
	(36.76)	(18.87)	(5.97)	(-2.52)
Author Comments	0.428***	0.493***	0.031	0.071
	(2.80)	(3.83)	(0.39)	(0.61)
Submissions			8.812***	4.320***
			(10.62)	(5.50)
Day FE	X	X	X	X
Stock FE	X	X	X	X
Author FE		X		X
Obs.	52,695	52,695	60,473	60,473
Adj. R^2 (%)	18.8	60.0	9.6	57.9

Table 6. AI-Generated Content and Comment Disagreement

This table examines the association between an article's AI probability score and the dispersion of follow-up comment sentiment over the subsequent ten days using the following regression:

$$Disagreement_{ij[t,t+10]} = \beta_1 AIProb_{ijt} + Controls + FE + \epsilon_{ijt},$$

where $Disagreement_{ij[t,t+10]}$ is the standard deviation of comment sentiment following content j about firm i over $[t, t+10]$ (standardized to zero mean and unit variance). $AIProb_{ijt}$ is the AI probability score for content j about firm i, posted on day t. *Controls* include article characteristics, firm characteristics, and measures of the author's prior activity/performance (standardized to zero mean and unit variance). Columns 1–2 report results for articles published on Seeking Alpha; Columns 3–4 report results for contents on Reddit's r/WallStreetBets. Standard errors are two-way clustered by stock and day; *t*-statistics are in parentheses. **p* <.1; ***p* <.05; ****p* <.01.

	SA		WSB	
	(1)	(2)	(3)	(4)
AIProb	-0.083***	-0.082***	-0.002	0.001
	(-4.05)	(-3.18)	(-0.07)	(0.17)
Length	0.072***	0.063***	0.037***	0.056***
	(13.48)	(8.34)	(7.68)	(6.64)
Sentiment	0.005	-0.014**	-0.019***	-0.022***
	(1.05)	(-2.45)	(-5.83)	(-3.94)
Complexity	-0.027***	-0.012**	-0.002	0.001
	(-4.96)	(-2.45)	(-0.46)	(0.14)
Fog Index	-0.026***	0.014*	-0.007**	-0.001
	(-6.46)	(1.81)	(-1.97)	(-0.17)
Graphical	0.051***	0.018**	0.012***	0.023**
	(10.77)	(2.54)	(2.78)	(2.44)
Quantitative	0.003	0.003	0.043***	0.043***
	(0.52)	(0.44)	(11.57)	(7.67)
Size	0.132*	0.157***	0.039	-0.020
	(1.92)	(2.78)	(0.78)	(-0.22)
BM	-0.012	-0.008	-0.001	-0.013
	(-0.61)	(-0.45)	(-0.07)	(-0.55)
IO	0.002	-0.017	-0.023*	-0.052
	(0.10)	(-0.71)	(-1.82)	(-1.57)
News	0.044***	0.034***	0.022***	0.020
	(6.58)	(5.15)	(3.45)	(1.62)
Analysts	-0.005	-0.008	0.007	0.009
	(-0.47)	(-0.78)	(0.65)	(0.46)
Author Article	-0.055***	-0.014*	0.020***	0.015
	(-8.79)	(-1.83)	(4.67)	(1.45)
Author AI Article	0.025***	0.013**	0.008	0.007
	(4.86)	(2.39)	(1.41)	(0.63)
Author Comments	0.091***	0.005	-0.005	-0.024***
	(14.56)	(0.73)	(-0.95)	(-3.21)
Submissions			1.536***	1.539***
			(73.97)	(42.06)
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 (%)	23.8	29.2	24.1	24.9

Table 7. AI-Generated Content and Sentiment Alignmment

This table examines the association between an article's AI probability score and the alignment between article sentiment and subsequent comment sentiment using the following regression:

$$CommentSentiment_{ij[t,t+10]} = \beta_1 AIProb_{ijt} \times Sentiment_{ijt} + \beta_2 AIProb_{ijt} + \beta_3 Sentiment_{ijt} + Controls + FE + \epsilon_{ijt}$$

where $Comment\ Sentiment_{ij[t,t+10]}$ is the mean sentiment of comments following content j about firm i over $[t, t+10]$. $AIProb_{ijt}$ is the AI probability score for content j about firm i, posted on day t. $Sentimentijt$ is the main text sentiment score. *Controls* 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 results for posts on Reddit's r/WallStreetBets. Standard errors are two-way clustered by stock and day; *t*-statistics are in parentheses. **p* <.1; ***p* <.05; ****p* <.01.

	SA		WSB	
	(1)	(2)	(3)	(4)
AIProb × Sentiment	-0.022	-0.046	0.090**	0.135**
	(-0.40)	(-0.83)	(2.38)	(2.22)
AIProb	-0.001	-0.001	-0.004**	-0.008**
	(-0.80)	(-0.50)	(-2.49)	(-2.43)
Sentiment	0.227***	0.217***	0.080***	0.067***
	(13.49)	(11.85)	(12.93)	(7.27)
Length	0.003***	0.002***	0.000	0.000
	(14.10)	(7.17)	(-1.61)	(0.66)
Complexity	0.000*	0.000	0.000	0.000
	(-1.95)	(-0.06)	(-0.06)	(-0.17)
Fog Index	0.000	0.001***	0.000	0.000
	(-1.41)	(3.46)	(-1.58)	(0.29)
Graphical	0.002***	0.001***	0.000	0.000
	(8.79)	(4.32)	(0.63)	(0.03)
Quantitative	0.001***	0.000	0.001***	0.001**
	(2.81)	(0.69)	(3.98)	(2.20)
Size	0.002	0.003	-0.001	0.000
	(0.85)	(1.25)	(-0.16)	(-0.03)
BM	-0.001	0.000	-0.001	0.000
	(-0.93)	(-0.66)	(-1.59)	(-0.11)
IO	0.001	0.001	0.000	0.001
	(0.94)	(0.57)	(0.06)	(0.37)
News	0.000	0.000	0.000	0.000
	(0.38)	(-0.11)	(-0.62)	(-0.49)
Analysts	0.000	0.000	0.000	0.002
	(0.20)	(0.00)	(0.93)	(1.57)
Author Article	-0.001***	0.000	0.000	0.000
	(-3.46)	(-1.32)	(0.24)	(-0.24)
Author AI Article	0.000	0.000	0.000	0.000
	(1.14)	(0.01)	(-0.20)	(0.18)
Author Comments	0.001***	0.000	0.000	0.000
	(5.11)	(1.01)	(1.32)	(-0.61)
Submissions			0.021***	0.019***
			(33.19)	(11.23)
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 (%)	8.3	10.1	3.4	0.1

Table 8. AI Content and Return Predictability

This table examines whether AI content can better predict future returns using the following article-level regression:

$$CAR_{i[t+1,t+5]} = \beta_1 AIProb_{ijt} \times Sentiment_{ijt} + \beta_2 AIProb_{ijt} + \beta_3 Sentiment_{ijt} + Controls + FE + \epsilon_{ijt}.$$

The dependent variable is cumulative abnormal returns over $[t+1, t+5]$ in columns 1 and 3, and over $[t+6, t+10]$ in columns 2 and 4. AIProb is the AI probability score assigned to each article. Sentiment is the article-level sentiment score. All continuous control variables are standardized. Standard errors are two-way clustered by date and firm. t-statistics are in parentheses. *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

	SA		WSB	
	$CAR_{[t+1,t+5]}$ (1)	$CAR_{[t+6,t+10]}$ (2)	$CAR_{[t+1,t+5]}$ (3)	$CAR_{[t+6,t+10]}$ (4)
AIProb × Sentiment	0.350**	-0.122	-0.179	0.007
	(2.114)	(-0.851)	(-0.793)	(0.034)
AIProb	-0.327*	-0.093	-0.259	0.077
	(-1.675)	(-0.579)	(-1.009)	(0.262)
Sentiment	-0.149***	-0.075*	0.019	-0.020
	(-3.398)	(-1.891)	(0.500)	(-0.451)
Length	0.070*	-0.005	0.006	0.022
	(1.806)	(-0.145)	(0.167)	(0.681)
Complexity	-0.019	-0.002	-0.019	-0.013
	(-0.663)	(-0.075)	(-0.490)	(-0.365)
Fog Index	0.014	0.021	-0.100**	-0.039
	(0.399)	(0.645)	(-2.355)	(-1.099)
Graphical	0.087**	0.059*	-0.061*	-0.046
	(2.440)	(1.952)	(-1.903)	(-1.214)
Quantitative	0.035	0.026	-0.054	-0.056
	(1.112)	(1.024)	(-1.607)	(-1.343)
Size	-3.549***	-2.927***	-2.857*	-3.197**
	(-3.959)	(-3.716)	(-1.801)	(-2.020)
BM	-0.484	-0.329	0.327	-0.138
	(-1.610)	(-1.300)	(0.707)	(-0.380)
Analysts	-0.124	-0.041	0.319	0.310
	(-1.144)	(-0.366)	(0.626)	(0.990)
IO	-1.002**	-0.624**	-0.120	0.376
	(-2.195)	(-2.443)	(-0.260)	(0.914)
News	-0.089	-0.034	-0.299	-0.091
	(-0.759)	(-0.343)	(-0.680)	(-0.191)
Earnings Day	0.156	-0.119	-0.619	-0.211
	(0.675)	(-0.436)	(-1.073)	(-0.344)
$Ret_{[t-21,t-1]}$	-0.189	-0.309***	-0.806**	0.305
	(-0.905)	(-2.792)	(-2.155)	(0.629)
$Volatility_{[t-21,t-1]}$	-0.010	0.202	0.247	0.057
	(-0.058)	(1.620)	(0.595)	(0.135)
Submissions			0.205	0.250
			(1.580)	(1.280)
Day FE	Yes	Yes	Yes	Yes
Stock FE	Yes	Yes	Yes	Yes
Obs.	54,070	54,070	62,785	62,785
Adj. R^2 (%)	8.6	9.1	20.0	19.3

Table 9. AI Day and Retail Order Flow Informativeness

This table examines the association between "AI Day" and the relationship between retail order flow and subsequent returns, using the following stock-day panel regression specification:

$$\begin{aligned} CAR_{i[t+1,t+5]} &= \beta_1 AIDay_{it} \times OIB_{it} + \beta_2 HumanDay_{it} \times OIB_{it} \\ &+ \beta_3 AIDay_{it} \times Sentiment_{it} + \beta_4 HumanDay_{it} \times Sentiment_{it} \\ &+ \beta_3 AIDay_{it} + \beta_3 HumanDay_{it} + \beta_4 OIB + \beta_5 Sentiment_{it} \\ &+ \beta_6 Attention_{it} + Controls + FE + \epsilon_{it}, \end{aligned}$$

where $CAR_{[t+1,t+5]}$ is the market-adjusted cumulative abnormal return over $[t+1, t+5]$. $OIB_{i,t}$ is the net retail order imbalance (net retail buys divided by total retail share volume) for stock i on day t. $AIDay_{it}^{SA}$ is an indicator equal to one if at least one AI-generated article/post about stock i appears on day t on SA. $AIDay_{it}^{WSB}$ is defined similarly for WSB. $HumanDay_{it}^{SA}$ is an indicator equal to one if at least one Human article/post about stock i appears on day t on SA. $Sentiment_{it}^{SA}$ is the average platform-level sentiment across all content about stock i on day t on SA. $Attention_{it}^{SA}$ is the logarithm of one plus the number of articles about stock i on day t on SA. All regressions include firm-characteristic controls and 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	
	(1)	(2)	(3)	(4)	(5)	(6)
$\text{AIDay}^{SA} \times \text{OIB}$	0.403*	0.431**			0.407*	0.435**
	(1.940)	(2.069)			(1.929)	(2.057)
$\text{HumanDay}^{SA} \times \text{OIB}$	0.053	0.053			0.050	0.050
	(0.795)	(0.793)			(0.749)	(0.747)
AIDay^{SA}	-0.294	-0.577			-0.297	-0.573
	(-1.549)	(-1.565)			(-1.596)	(-1.564)
HumanDay^{SA}	0.032	0.086			0.027	0.081
	(0.180)	(0.456)			(0.156)	(0.433)
Sentiment^{SA}	-0.021	0.012			-0.021	0.012
	(-1.593)	(0.305)			(-1.559)	(0.306)
Attention^{SA}	0.019	0.015			0.019	0.015
	(0.595)	(0.456)			(0.621)	(0.476)
$\text{AIDay}^{SA} \times \text{Sentiment}^{SA}$		0.025				0.023
		(0.398)				(0.380)
$\text{HumanDay}^{SA} \times \text{Sentiment}^{SA}$		-0.041				-0.040
		(-0.927)				(-0.914)
$\text{AIDay}^{WSB} \times \text{OIB}$			0.371	0.325	0.478	0.478
			(0.998)	(0.884)	(1.061)	(1.060)
$\text{HumanDay}^{WSB} \times \text{OIB}$			0.070	0.069	0.090	0.090
			(0.954)	(0.945)	(0.995)	(0.991)
AIDay^{WSB}			0.064	-0.097	-0.086	-0.085
			(0.195)	(-0.269)	(-0.235)	(-0.234)
HumanDay^{WSB}			0.183	0.194	0.189	0.188
			(1.235)	(1.317)	(1.251)	(1.241)
Sentiment^{WSB}			-0.002	-0.139	-0.090	-0.091
			(-0.224)	(-1.094)	(-1.070)	(-1.072)
Attention^{WSB}			-0.028	-0.029	-0.019	-0.019
			(-0.773)	(-0.797)	(-0.753)	(-0.741)
$\text{AIDay}^{WSB} \times \text{Sentiment}^{WSB}$				0.121	0.077	0.077
				(1.004)	(0.953)	(0.954)
$\text{HumanDay}^{WSB} \times \text{Sentiment}^{WSB}$				0.136	0.090	0.090
				(1.082)	(1.065)	(1.067)
OIB	0.004	0.004	0.022	0.022	-0.003	-0.003
	(0.174)	(0.174)	(1.021)	(1.024)	(-0.132)	(-0.132)
Size	-2.943***	-2.943***	-2.945***	-2.945***	-3.002***	-3.002***
	(-10.493)	(-10.493)	(-9.177)	(-9.179)	(-10.600)	(-10.600)
BM	-0.295***	-0.295***	-0.307**	-0.307**	-0.281***	-0.281***
	(-2.721)	(-2.721)	(-2.312)	(-2.313)	(-2.621)	(-2.620)
IO	-0.046	-0.046	-0.056	-0.056	-0.007	-0.007
	(-1.060)	(-1.060)	(-0.632)	(-0.632)	(-0.101)	(-0.101)
Analysts	-0.166***	-0.166***	-0.095*	-0.095*	-0.172***	-0.172***
	(-3.026)	(-3.026)	(-1.831)	(-1.830)	(-3.128)	(-3.128)
News	-0.075	-0.075	-0.145**	-0.145**	-0.087*	-0.087*
	(-1.616)	(-1.617)	(-2.381)	(-2.381)	(-1.893)	(-1.893)
Earnings Day	-0.103	-0.102	-0.080	-0.081	-0.120	-0.119
	(-0.918)	(-0.913)	(-0.664)	(-0.668)	(-1.064)	(-1.059)
$\text{Ret}_{[t-21,t-1]}$	-0.397***	-0.397***	-0.342***	-0.342***	-0.433***	-0.433***
	(-9.301)	(-9.300)	(-5.117)	(-5.118)	(-9.613)	(-9.613)
$\text{Volatility}_{[t-21,t-1]}$	0.060	0.060	0.076	0.076	0.056	0.056
	(0.695)	(0.695)	(0.727)	(0.727)	(0.716)	(0.716)
Day FE	X	X	X	X	X	X
Stock FE	X	X	X	X	X	X
Obs.	1,370,181	1,370,181	611,883	611,883	1,391,849	1,391,849
Adj. R^2 (%)	3.2	3.2	4.4	4.4	3.2	3.2

Table 10. AI Day and Subsequent Volume, Volatility, and Bid-Ask Spreads

This table examines the association between platform-level "AI Day" and subsequent trading metrics over $[t+1, t+5]$, using the following stock-day panel regression specification:

$$MarketQuality_{i[t+1,t+5]} = \beta_1 AIDay_{it} + \beta_2 HumanDay_{it} + \beta_3 Attention_{it} + \beta_4 Sentiment_{it} + Controls + FE + \epsilon_{it},$$

$MarketQuality_{i[t+1,t+5]}$ measures trading activity for stock i, standardized to zero mean and unit variance. Columns (1), (4), and (7) report abnormal daily volume; columns (2), (5), and (8) report return volatility; and columns (3), (6), and (9) report the effective bid-ask spread. $AIDay_{it}^{SA}$ is an indicator equal to one if at least one AI-generated article/post about stock i appears on day t on SA. $HumanDay_{it}^{SA}$ is an indicator equal to one if at least one Human article/post about stock i appears on day t on SA. $Sentiment_{it}^{SA}$ is the average platform-level sentiment across all content about stock i on day t on SA. $Attention_{it}^{SA}$ is the logarithm of one plus the number of articles about stock i on day t on SA. $AIDay_{it}^{WSB}$, $HumanDay_{it}^{WSB}$, $Sentiment_{it}^{WSB}$, and $Attentionit^{WSB}$ are defined similarly for WSB. All regressions include firm-characteristic controls and 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		
	AVOL (1)	Volatility (2)	Spread (3)	AVOL (4)	Volatility (5)	Spread (6)	AVOL (7)	Volatility (8)	Spread (9)
AIDay^{SA}	-0.023	-0.019	-0.018***				0.015	0.022	-0.016***
	(-1.017)	(-0.923)	(-5.710)				(0.721)	(1.160)	(-5.453)
HumanDay^{SA}	0.006	0.017	-0.009***				0.046**	0.060***	-0.007***
	(0.270)	(0.993)	(-3.081)				(2.191)	(3.452)	(-2.646)
Attention^{SA}	0.020***	0.013***	0.001**				0.011***	0.003	0.001
	(4.901)	(4.000)	(2.153)				(2.810)	(1.077)	(1.510)
Sentiment^{SA}	-0.004***	-0.012***	0.000				-0.003**	-0.011***	0.000
	(-2.809)	(-8.737)	(-0.297)				(-2.367)	(-8.606)	(-0.168)
AIDay^{WSB}				0.086***	0.074*	0.000	0.088***	0.078*	0.004
				(2.816)	(1.865)	(-0.033)	(2.958)	(1.959)	(0.851)
HumanDay^{WSB}				-0.017	-0.066***	-0.004	-0.014	-0.062***	-0.004
				(-0.816)	(-2.957)	(-1.301)	(-0.674)	(-2.754)	(-1.406)
Attention^{WSB}				0.038***	0.045***	0.002***	0.024***	0.028***	0.001***
				(6.499)	(6.423)	(2.926)	(5.685)	(6.153)	(2.595)
Sentiment^{WSB}				0.006***	0.003**	0.000	0.004***	0.002***	0.000
				(5.324)	(2.548)	(-1.545)	(5.336)	(2.595)	(-0.978)
Size	-0.192***	-0.178***	-0.111***	-0.165***	-0.193***	-0.100***	-0.198***	-0.180***	-0.105***
	(-6.370)	(-6.797)	(-8.666)	(-4.574)	(-5.451)	(-4.691)	(-6.780)	(-7.023)	(-8.480)
BM	-0.007	-0.042***	0.005	-0.025*	-0.022*	-0.001	-0.009	-0.037***	0.006
	(-0.712)	(-3.990)	(1.152)	(-1.837)	(-1.655)	(-0.117)	(-0.931)	(-3.740)	(1.243)
IO	-0.018***	-0.006	-0.005*	-0.008***	-0.003	-0.007	-0.016**	-0.004	-0.007*
	(-2.625)	(-0.663)	(-1.723)	(-3.026)	(-1.103)	(-1.173)	(-2.562)	(-0.486)	(-1.818)
Analysts	-0.036***	-0.003	-0.002	-0.036***	-0.003	-0.003	-0.035***	-0.003	-0.002
	(-7.155)	(-0.614)	(-0.969)	(-4.985)	(-0.479)	(-1.319)	(-7.132)	(-0.506)	(-1.088)
News	0.063***	-0.057***	-0.024***	0.032***	-0.080***	-0.014***	0.063***	-0.056***	-0.024***
	(8.230)	(-8.321)	(-12.973)	(3.146)	(-11.034)	(-5.226)	(7.945)	(-8.231)	(-12.770)
Earnings Day	0.430***	0.140***	0.003	0.465***	0.114***	0.005**	0.413***	0.134***	0.002
	(29.524)	(9.244)	(1.251)	(28.626)	(7.818)	(2.136)	(29.622)	(9.323)	(0.752)
$\text{Ret}_{[t-21,t-1]}$	0.255***	0.020***	-0.022***	0.316***	0.018**	-0.029***	0.258***	0.021***	-0.022***
	(4.240)	(2.821)	(-3.117)	(25.742)	(2.143)	(-10.308)	(4.398)	(2.919)	(-3.218)
$\text{Volatility}_{[t-21,t-1]}$	0.028	0.047	-0.011***	0.116***	0.099***	-0.021***	0.036	0.051	-0.012**
	(0.762)	(1.508)	(-2.853)	(3.829)	(5.083)	(-3.655)	(0.893)	(1.590)	(-2.572)
$\text{Spread}_{[t-21,t-1]}$			0.644***			0.605***			0.643***
			(41.360)			(15.778)			(41.713)
Day FE	X	X	X	X	X	X	X	X	X
Stock FE	X	X	X	X	X	X	X	X	X
Obs.	1,472,523	1,472,523	1,343,895	648,295	648,295	594,458	1,496,823	1,496,823	1,365,371
Adj. R^2 (%)	17.6	39.0	88.6	20.8	42.1	90.6	17.3	38.9	88.6

Table 11. Lottery-Like Return Events Following AI Days

This table examines the association between "AI Day" and the incidence of large-return events (MAX and Lottery) using the following the following logistic stock-day panel regression specification:

$$MAX/Lottery_{i[t+1,t+5]} = \beta_1 AIDay_{it} + \beta_2 HumanDay_{it} + \beta_3 Attention_{it} + \beta_4 Sentiment_{it} + Controls + FE + \epsilon_{it},$$

where the dependent variables, $MAX_{i[t+1,t+5]}$ and $Lottery_{i[t+1,t+5]}$, are indicators equal to one if stock i experiences a MAX or Lottery event in the $[t+1, t+5]$ window, and zero otherwise. A MAX event occurs when the daily return is the highest over the prior 20 trading days. A lottery event is a MAX event for which the corresponding maximum return also ranks in the top decile across all stocks on that day. $AIDay_{it}^{SA}$ is an indicator equal to one if at least one AI-generated article/post about stock i appears on day t on SA. $AIDay_{it}^{WSB}$ is defined similarly for WSB. *HumanDay*, *Sentiment* and *Attention* measures are define as in Table 10. All regressions include firm-characteristic controls and 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	
	$MAX_{[t+1,t+5]}$ (1)	$Lottery_{[t+1,t+5]}$ (2)	$MAX_{[t+1,t+5]}$ (3)	$Lottery_{[t+1,t+5]}$ (4)	$MAX_{[t+1,t+5]}$ (5)	$Lottery_{[t+1,t+5]}$ (6)
$AIDay^{SA}$	-0.027	-0.336*			0.020	-0.115
	(-0.374)	(-1.659)			(0.284)	(-0.542)
$HumanDay^{SA}$	-0.065	-0.365**			-0.014	-0.106
	(-1.116)	(-2.237)			(-0.250)	(-0.606)
$Attention^{SA}$	0.040***	0.148***			0.029***	0.091***
	(4.184)	(6.070)			(3.179)	(3.356)
$Sentiment^{SA}$	-0.017***	-0.010			-0.016***	-0.007
	(-3.801)	(-0.661)			(-3.612)	(-0.457)
$AIDay^{WSB}$			0.166**	0.366**	0.171**	0.403**
			(2.120)	(2.009)	(2.193)	(2.243)
$HumanDay^{WSB}$			0.002	0.139	0.017	0.190
			(0.030)	(0.982)	(0.332)	(1.389)
$Attention^{WSB}$			0.047***	0.133***	0.028***	0.079***
			(3.649)	(4.176)	(3.299)	(3.862)
$Sentiment^{WSB}$			0.009***	0.008	0.006***	0.006
			(2.693)	(0.724)	(2.702)	(0.905)
Size	-0.448***	-0.729***	-0.498***	-0.701***	-0.443***	-0.747***
	(-9.120)	(-8.776)	(-7.696)	(-6.188)	(-9.027)	(-9.126)
BM	-0.090***	-0.125***	-0.094***	-0.088	-0.086***	-0.115***
	(-4.337)	(-3.300)	(-3.678)	(-1.639)	(-4.281)	(-3.179)
IO	-0.005	-0.098	0.001	-0.091	-0.003	-0.083*
	(-0.474)	(-1.551)	(0.189)	(-1.544)	(-0.337)	(-1.761)
Analysts	-0.010	-0.040	-0.009	-0.055	-0.010	-0.045
	(-1.028)	(-1.331)	(-0.694)	(-1.121)	(-1.022)	(-1.474)
News	-0.122***	-0.154***	-0.164***	-0.175***	-0.126***	-0.157***
	(-11.021)	(-6.092)	(-11.495)	(-4.099)	(-11.534)	(-6.347)
Earnings Day	1.066***	1.543***	1.124***	1.645***	1.055***	1.502***
	(37.404)	(28.929)	(34.852)	(24.581)	(37.524)	(28.501)
$Ret_{[t-21,t-1]}$	-0.928***	-0.136***	-0.805***	-0.128***	-0.931***	-0.151***
	(-37.523)	(-7.245)	(-26.364)	(-5.033)	(-37.734)	(-8.249)
$Volatility_{[t-21,t-1]}$	-1.682***	0.008	-1.282***	0.012	-1.695***	0.009
	(-37.949)	(0.438)	(-22.789)	(0.600)	(-36.558)	(0.449)
Day FE	X	X	X	X	X	X
Stock FE	X	X	X	X	X	X
Obs.	1,424,072	680,696	635,342	274,574	1,446,735	700,185

Appendix

A AI Detection Method Comparison

A.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 by 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—GPT-4o, Claude 3.5 Sonnet, and Llama 3.3, -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.

A.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.

A.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.

B Examples of AI and Human Contents

Table A3 compares human-authored content with suspected AI-generated content across both Seeking Alpha and Reddit's r/WallStreetBets. The human-written samples often exhibit a conversational, idiosyncratic tone. The suspected AI-generated texts typically display mechanical transitions, a lack of personal voice, and a rigid summary structure. For each example, we provide the excerpt context, the associated stock, the GPTZero classification probabilities, and the stylistic cues highlighted by GPTZero.

Table A1. Description of Variables

Variable	Definition
Article-level	
AIProb_{ijt}	The sum of probability scores of content j for the AI-ONLY and MIXED categories, assigned by GPTZero.
$\text{AI-Only Prob}_{ijt}$	The probability scores of content j for the AI-ONLY categories, assigned by GPTZero.
Class_{ijt}	The classification (i.e., AI-ONLY, MIXED, HUMAN) of content j, 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 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: $$\text{Complexity}_{ijt} = \frac{\text{Words in Dictionary}}{\text{Total Words}}$$ Where #Words in Dictionary are the count of the specific complexity words identified in Loughran and McDonald (2024).
$\text{Quantitative}_{ijt}$	Total count of numbers mentioned divided by word count of a article, winsorized at the 1st and 99th percentiles.
Graphical_{ijt}	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 content j; Sentences represents the total number of sentences in 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.
Submissions_{ijt}	An indicator that identifies whether content j in Reddit r/WallStreetBets is a submission (top-level post).
$\text{Unfamiliarity}_{ijt}$	An indicator that identifies whether an author or user is addressing a specific firm i in their published content j for the first time within a recent six-month period.

Variable	Definition
First Time Industry$_{ijt}$	This variable is an indicator that identifies whether an author or user is addressing the specific industry of the firm i in their published content for the first time within a recent six-month period.
News$_{ijt}$	Natural logarithm of one plus the number of news articles concerning firm i mentioned in content j in the past three days ($[t-2,t]$), sourced from RavenPack.
Analysts$_{ijt}$	Natural logarithm of one plus the number of analysts covering firm i mentioned in content j, sourced from Institutional Brokers' Estimate System (IBES).
IntenseRetailBuy$_{i[t-3,t-1]}$	An indicator variable that equals 1 if firm i's retail order imbalance ranks in the preceding three days ranks in the top 10% across-sectionally. Order imbalance is calculated as in Boehmer et al. (2021).
Comments$_{ij[t,t+10]}$	Natural logarithm of number of user comments received on an article from its publication day (t) through 10 days post-publication ($t+10$).
Disagreement$_{ij[t,t+10]}$	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.
CommentSentiment$_{ij[t,t+10]}$	Mean sentiment of comments following content j from its publication day (t) through 10 days post-publication ($t+10$).
Size$_{ijt}$	Natural logarithm of the market capitalization for firm i mentioned in content j.
BM$_{ijt}$	Natural logarithm of book-to-market ratio for firm i mentioned in content j.
IO$_{ijt}$	Institutional ownership, calculated as the percentage of common shares outstanding of firm i mentioned in content j, held by institutional investors as of the last quarter ending before day t.
Author Articles$_{ijt}$	Natural logarithm of one plus the number of articles published by a given author in the past six months.
Author AI Articles$_{ijt}$	Natural logarithm of one plus the number of AI-ONLY or MIXED articles by a given author in the past six months.
Author Comments$_{ijt}$	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.
Author level	
Author Article Count$_a$	Number of articles that author/user a published during the sample period.
# Covered Stocks$_a$	Number of distinct stocks that author/user a covered at least once during the sample period.
# Covered Industries$_a$	Number of distinct CRSP four-digit SIC industries across the set of tickers covered by author/user a during the sample period.
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.
Article$_{i,t}$	The number of stock-specific articles about firm i published on day t across the two platforms.

Variable	Definition
$\text{AVOL}_{i[t+1,t+5]}$	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.
$\text{Volatility}_{i[t+1,t+5]}$	Standard deviation of the daily stock returns for firm i over the 5 trading days from $t+1$ to $t+5$.
$\text{Spread}_{i[t+1,t+5]}$	Average daily volume-weighted effective bid-ask spread for firm i over the 5 trading days from $t+1$ to $t+5$.
$\text{MAX}_{i[t+1,t+5]}$	A MAX event is when the daily return is the highest over the prior 20 trading days.
$\text{Lottery}_{i[t+1,t+5]}$	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.
$\text{CAR}_{i[t+1,t+5]}$	Cumulative buy-and-hold market-adjusted returns over the 5 trading days from $t+1$ to $t+5$.
OIB_{it}	Retail order imbalance for stock i on day t, computed as $$OIB_{i,t} = \frac{\text{\#retail buy trades}_{i,t} - \text{\#retail sell trades}_{i,t}}{\text{\#retail buy trades}_{i,t} + \text{\#retail sell trades}_{i,t}}$$ using marketable retail trades identified following Boehmer et al. (2021).
Sentiment_{it}	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, following Cookson et al. (2024)
Attention_{it}	Natural logarithm of one plus the number of articles on SA (or posts on WSB) of firm i on day t.
Size_{it}	Natural logarithm of the market capitalization for firm i.
BM_{it}	Natural logarithm of book-to-market ratio for firm i.
IO_{it}	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.
Analysts_{it}	Natural logarithm of one plus the number of analysts covering firm i, sourced from Institutional Brokers' Estimate System (IBES).
News_{it}	Natural logarithm of one plus the number of news articles concerning firm i in the past 21 trading days, sourced from RavenPack.
$\text{Ret}_{i[t-21,t-1]}$	Cumulative returns over the preceding 21 trading days from $t-21$ to $t-1$.
$\text{Volatility}_{i[t-21,t-1]}$	Standard deviation of the daily stock returns for firm i over the preceding 21 trading days from $t-21$ to $t-1$.
$\text{Spread}_{i[t-21,t-1]}$	Average daily volume-weighted effective bid-ask spread for firm i over the preceding 21 trading days from $t-21$ to $t-1$.

Table A2. AI Detection Method Comparison

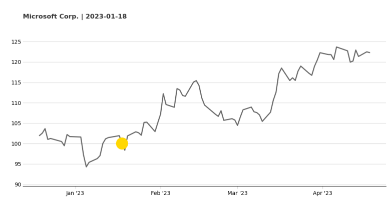
This table compares AI text detection methods on our benchmark dataset, reporting overall accuracy and F1 scores. Accuracy is the proportion of documents whose provenance (AI-generated or human-written) is correctly identified, and the F1 score is the harmonic mean of precision and recall, providing a balanced assessment of classification performance. Higher values indicate better performance on both metrics. 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. **Examples of Suspected AI-generated and Human-written Content**

Seeking Alpha Article Example 1

Title: Microsoft Doubling Down On OpenAI
Ticker: MSFT (Microsoft Corp)
Date: Jan. 18, 2023
User: User: Ivy Global Insights
Link: https://seekingalpha.com/article/4570487-microsoft-stock-doubling-down-openai

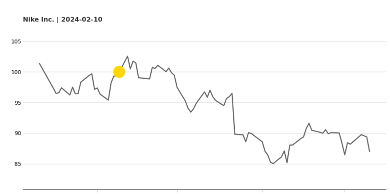


Excerpt: "Everyone knows ChatGPT. It is a smart coder, an intelligent writer, and an elegant Chat Bot. We all get used to search engines that pops up random pages with key words, and we dig into those links on the first or second page of search results in hoping to locate what we want... For the first time, search becomes a peaceful and delightful human-like dialogue."

GPTZero Output: Human (99.25%), Mixed (0.20%), AI-Only (0.54%)

Seeking Alpha Article Example 2

Title: Nike Is Running Well, But Its Valuation Is A Hurdle
Ticker: NKE (Nike Inc.)
Date: Feb. 10, 2024
User: Khen Elazar
Link: https://seekingalpha.com/article/4669281-nike-is-running-well-but-its-valuation-is-a-hurdle

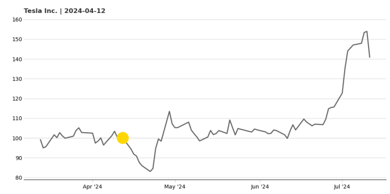


Excerpt: "These risks, while manageable, necessitate strategic planning and operational flexibility to mitigate potential impacts on revenue and profitability. Given the current market valuation, the shares are overvalued. Taking into account NIKE's solid foundation, growth prospects, and the outlined risks, the verdict leans towards a HOLD."

GPTZero Output: AI-Only (97.75%), Mixed (0.87%), Human (1.37%)
Why GPTZero think it's AI: Mechanical Transitions, Task-Oriented, Formulaic, Lacks Creativity.

WallStreetBets Message Example 1

Title: TSLA earnings play
Ticker: TSLA (Tesla Inc.)
Date: Apr. 12, 2024
User: u/Fragrant_Swing9987
Link: https://www.reddit.com/r/wallstreetbets/comments/1c2fztr/tsla_earnings_play/

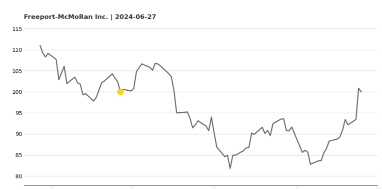


Excerpt: "Hello regards. Today I thought of a potential TSLA earnings call play. Before you guys bomb me with hate notice that the earning expectations are very low for TSLA, with a EPS estimate of 0.36(compared to 0.6 Q4 2023). Thus, I think a potential outperformance can occur at TSLA earnings shooting up the price. What are your thoughts ? (Not financial advice, my brain is smooth)"

GPTZero Output: Human (98.71%), Mixed (0%), AI-Only (1.28%)

WallStreetBets Message Example 2

Title: DD FCX
Ticker: FCX (Freeport-McMoRan Inc.)
Date: June, 27, 2024
User: u/Age_Specialist
Link: www.reddit.com/r/wallstreetbets/comments/1dq1vio/dd_fcx/



Excerpt: "Executive Summary Freeport-McMoRan Inc. (FCX) presents a compelling investment opportunity driven by the surging demand for copper, which is a critical component in the technology and renewable energy sectors. As a leading global copper producer, FCX is well-positioned to capitalize on the growing needs of AI companies, tech startups, and the broader electrification trend."

GPTZero Output: AI-Only (96.59%), Mixed (3.42%), Human (0%)
Why GPTZero think it's AI: Speculative Focus, Lacks Complexity, Artificial Simplicity, Predictable Rhythm.

WallStreetBets Message Example 3

Title: NKE puts for all you regards that lost money through TSM
Ticker: NKE (Nike Inc.)
Date: March, 21, 2024
User: u/Ready_Confidence5326
Link:
https://www.reddit.com/r/wallstreetbets/comments/1bjz3ut/nke_puts_for_all_you_regards_that_lost_money/

Excerpt: "For all you regards who lost all your money this week, gather round cuz you can make it all back with puts on $NKE.
First off, these earnings coming up? They ain't lookin' pretty. Analysts getting cold feet and slashing those EPS and revenue estimates like there's no tomorrow. Revenue estimated to be down 1% YoY, operating income down 7%, EPS down 6%...I'm talking fiscal Q3 '24 lookin' like a dumpster fire."

GPTZero Output: AI-Only (99.98%), Mixed (0.02%), Human (0%)
Why GPTZero think it's AI: Speculative Focus, Lacks Complexity, Artificial Simplicity, Predictable Rhythm.

WallStreetBets Message Example 3

Title: 40K+ in Intel INTC ! I love USA chip PART 2
Ticker: INTC (Intel Inc.)
Date: December, 10, 2024
User: u/Intelligent-Cellist6
Link: https://www.reddit.com/r/wallstreetbets/comments/1hbakxw/40k_in_intel_intc_i_love_usa_chip_part_2/

Excerpt:
Going **all in** on $INTC!
If you've seen my prior post, you know I crushed it with LEAPS, riding Intel from **$18** to **$26**. Now, I'm back for round two. Currently sitting on **2K shares** and planning to **load up another 3K with margin**-yes, I'm that bullish.
Here's why I'm doubling down:
nosep **CEO announcement incoming** - This could be the game-changer the market's waiting for.
nosep **Massive grants secured** - Perfect timing for growth and innovation.
nosep **Undervalued gem** - At these levels, $INTC is a straight-up steal.

I see huge upside potential here-like staring into the abyss of tendies. Who's riding with me?"

GPTZero Output: AI-Only (99.98%), Mixed (0.02%), Human (0%)
Why GPTZero think it's AI: Promotional Tone, Artificial Simplicity, Speculative Focus.
Sentiment: 0.09 (Higher than 98.8% of Messages in our sample)

Table A4. Robustness: AI Content Measure Based on "AI-Only" Category

This table re-examines the main article-level analyses in Tables 3-6 using an alternative AI measure that relies only on the "AI-only" category from GPTZero. Panels A and B report results for Seeking Alpha and r/WallStreetBets, respectively. Columns 1—6 examine article-level AI use. Columns 7—8 and 9—10 examine the association between AI content and subsequent comment volume and disagreement, respectively. Standard errors are two-way clustered by stock and day, and *t*-statistics are shown in parentheses. **p* <.1; ***p* <.05; ****p* <.01.

Panel A: SA Sample

	AI-Only Prob						Comments		Disagreement	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Unfamiliarity	2.352***	0.511***								
	(11.69)	(3.69)								
News			-0.381***	-0.129*			0.088***	0.072***	0.044***	0.034***
			(-3.18)	(-1.74)			(9.49)	(9.52)	(6.56)	(5.16)
Analysts			-0.591***	-0.265*			0.010	0.007	-0.005	-0.008
			(-3.28)	(-1.91)			(0.85)	(0.67)	(-0.49)	(-0.77)
IntenseRetailBuy					-0.080	-0.249				
					(-0.17)	(-0.73)				
AI-Only Prob							-0.154***	-0.081**	-0.127***	-0.076**
							(-5.56)	(-2.58)	(-4.82)	(-2.21)
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.	54,429	54,429	54,429	54,429	52,695	52,695	54,429	54,429	54,429	54,429
Adj. R^2 (%)	17.0	61.8	16.6	61.7	16.9	62.2	51.6	64.2	23.8	29.2

Panel B: WSB Sample

	AI-Only Prob						Comments		Disagreement	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Familiarity	-0.004	-0.027								
	(-0.03)	(-0.19)								
News			-0.056	-0.204			0.011^{**}	0.013	0.021^{***}	0.019
			(-0.43)	(-1.45)			(2.10)	(1.26)	(3.14)	(1.42)
Analysts			0.513^{**}	0.163			0.009	0.025^{*}	0.006	0.007
			(2.41)	(0.83)			(1.14)	(1.80)	(0.55)	(0.35)
IntenseRetailBuy					1.499^{**}	1.886^{***}				
					(2.43)	(3.09)				
AI-Only Prob							-0.145^{***}	-0.017	0.012	0.018
							(-5.01)	(-0.33)	(0.34)	(0.20)
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.	63,400	63,400	63,400	63,400	60,473	60,473	63,400	63,400	63,400	63,400
Adj. R^2 (%)	11.4	56.9	11.4	56.9	9.4	58.5	38.5	43.1	18.6	22.0

Table A5. Author Industry Unfamiliarity and AI Adoption

This table examines the association between author unfamiliarity with a firm's industry and article-level AI use, using the following panel regression specification:

$$AIProb_{ijt} = \beta_1 First\ Time\ Industry_{ijt} + Controls + FE + \epsilon_{ijt},$$

where $AIProb_{ijt}$ is AI probability score of content j about firm i posted on day t. $First\ Time\ Industry_{ijt]}$ is an indicator equal to one if the author has not covered firm i's industry in the prior six months. $\mathbf{X}_{ijt}$ includes firm characteristics and measures of the author's prior activity/performance (each standardized to zero mean and unit variance). Columns 1–3 report results for articles published on Seeking Alpha; Columns 4–6 report results for posts on Reddit's r/WallStreetBets. Standard errors are two-way clustered by stock and day; *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 A6. AI-Generated Content and Comment Volume

This table examines the association between AI probability score and the volume of viewer comments over the subsequent ten days, using the following regression:

$$\log(1 + Comments_{ij[t,t+10]}) = \beta_1 AIProb_{ijt} + Controls + FE + \epsilon_{ijt},$$

where $Comments_{ij[t,t+10]}$ is the number of comments received on content j about firm i between day t and $t+10$. $AIProb_{ijt}$ is the AI probability score for content j about firm i, posted on day t. *Controls* include article characteristics, firm characteristics, and measures of the author's prior activity/performance (standardized to zero mean and unit variance). Columns 1–3 report results for articles published on Seeking Alpha; Columns 4–6 report results for posts on Reddit's r/WallStreetBets. Standard errors are two-way clustered by stock and day; t-statistics are in parentheses. *p <.1; **p <.05; ***p <.01.

	SA		WSB	
	(1)	(2)	(3)	(4)
AIProb	-0.073***	-0.065**	-0.142***	-0.014
	(-3.13)	(-2.48)	(-5.13)	(-0.26)
Length	0.115***	0.114***	0.069***	0.088***
	(13.88)	(11.71)	(19.23)	(12.20)
Sentiment	-0.046***	-0.080***	-0.023***	-0.020***
	(-7.16)	(-11.74)	(-10.93)	(-5.16)
Complexity	-0.057***	-0.032***	-0.001	-0.004
	(-7.07)	(-5.15)	(-0.31)	(-0.75)
Fog Index	-0.067***	-0.033***	-0.013***	-0.014**
	(-13.89)	(-3.91)	(-6.30)	(-2.54)
Graphical	0.048***	-0.012	0.005	0.005
	(8.41)	(-1.65)	(0.91)	(0.80)
Quantitative	-0.011**	-0.042***	0.027***	0.032***
	(-2.01)	(-6.33)	(7.54)	(7.90)
Size	0.184*	0.199**	0.022	0.041
	(1.66)	(2.44)	(0.44)	(0.50)
BM	-0.012	-0.001	0.012	-0.006
	(-0.47)	(-0.06)	(1.00)	(-0.28)
IO	-0.026	-0.056**	-0.005	0.008
	(-0.82)	(-2.01)	(-0.47)	(0.33)
News	0.089***	0.072***	0.011**	0.013
	(9.53)	(9.52)	(2.10)	(1.26)
Analysts	0.010	0.007	0.009	0.025*
	(0.88)	(0.67)	(1.14)	(1.80)
Author Article	-0.106***	-0.015**	0.017***	0.014**
	(-15.41)	(-2.06)	(4.45)	(2.00)
Author AI Article	0.026***	0.006	0.014***	0.006
	(4.87)	(1.25)	(3.41)	(0.99)
Author Comments	0.195***	0.012*	0.004	-0.026***
	(20.25)	(1.84)	(0.76)	(-4.89)
Submissions			1.787***	1.930***
			(54.31)	(53.81)
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 (%)	51.6	64.2	38.5	43.1

Table A7. Analysis Surrounding the Seeking Alpha AI Ban

This table replicates the results from Table 3 to Table 6, showing separate interaction coefficients for key variables in the Pre-Ban and Post-Ban (After 2023-10-31) periods. The specifications control for the same variables as in the original tables. 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)
Unfamiliarity × Pre-Ban	4.197***	1.107***								
	(11.87)	(4.77)								
Unfamiliarity × Post-Ban	0.450	0.024								
	(1.52)	(0.10)								
News × Pre-Ban			-0.696***	-0.262*						
			(-3.40)	(-1.94)						
News × Post-Ban			0.040	-0.157						
			(0.21)	(-1.24)						
Analysts × Pre-Ban			-0.912***	-0.313*						
			(-3.83)	(-1.66)						
Analysts × Post-Ban			-0.454*	-0.346*						
			(-1.89)	(-1.82)						
Intense Rtl. Buy × Pre-Ban					-0.012	0.600				
					(-0.01)	(0.97)				
Intense Rtl. Buy × Post-Ban					0.382	0.321				
					(0.51)	(0.54)				
AIProb × Pre-Ban							-0.150***	-0.090***	-0.005***	-0.003**
							(-4.97)	(-2.69)	(-4.19)	(-2.14)
AIProb × Post-Ban							0.027	-0.037	-0.002*	-0.005***
							(0.87)	(-1.15)	(-1.96)	(-3.54)
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.	54,429	54,429	54,429	54,429	52,695	52,695	54,429	54,429	54,429	54,429
Adj. R^2 (%)	18.9	59.6	18.6	59.5	18.8	60.0	51.6	64.2	22.1	27.1

Table A8. Robustness: WSB Submissions-only Sample

This table presents replication results of Tables 3–6 for the r/WallStreetBets (WSB) submissions-only sample. Columns 1-6 examine article-level AI use; Columns 7-8 and 9-10 examine the association between AI content and user responses on the platform. Standard errors are two-way clustered by stock and day; *t*-statistics are shown in parentheses. **p* <.1; ***p* <.05; ****p* <.01.

	AIProb						Comments	Disagreement
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Unfamiliarity	0.532	-1.112						
	(0.69)	(-0.73)						
News			-0.653	-3.484			0.016	0.001
			(-0.80)	(-1.51)			(0.85)	(0.74)
Analysts			3.012**	-6.026			-0.009	0.003
			(2.53)	(-1.28)			(-0.27)	(1.21)
IntenseRetailBuy					13.408***	17.534*		
					(2.73)	(1.87)		
AIProb							-0.315***	-0.040
							(-5.38)	(-0.64)
Day FE	X	X	X	X	X	X	X	X
Stock FE	X	X	X	X	X	X	X	X
Author FE		X		X		X		
Obs.	6,487	6,487	6,487	6,487	6,127	6,127	6,487	6,487
Adj. R^2 (%)	10.4	23.1	10.5	23.8	11.5	25.9	8.1	4.7

Table A9. Robustness: Including Messages below 50 Words on WSB

This table presents replication results of Tables 3–6 for the full Reddit r/WallStreetBets message sample (including submissions and comments below 50 words). Columns 1-6 examine article-level AI use; Columns 7-8 and 9-10 examine the association between AI content and user responses on the platform. Standard errors are two-way clustered by stock and day; t-statistics are in parentheses. *p < .10; **p < .05; ***p < .01.

	AIProb						Comments		Disagreement	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Unfamiliarity	-0.008	0.017								
	(-0.56)	(1.53)								
News			-0.006	-0.005			0.015***	0.013***	0.025***	0.020***
			(-0.56)	(-0.61)			(6.88)	(6.37)	(9.27)	(7.68)
Analysts			0.039**	0.032***			0.002	0.001	0.000	0.000
			(2.53)	(3.33)			(0.65)	(0.35)	(0.07)	(0.08)
IntenseRetailBuy					0.159*	0.112*				
					(1.85)	(1.95)				
AIProb							-0.373***	-0.262***	0.130***	0.000
							(-5.38)	(-10.13)	(3.44)	(-0.00)
Word Count > 50	2.925***	2.675***	2.924***	2.677***	2.761***	2.572***	-0.488***	-0.510***	-0.056***	-0.024***
	(26.15)	(28.10)	(26.19)	(27.99)	(27.01)	(27.49)	(-45.88)	(-63.88)	(-8.65)	(-4.23)
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	1.4	1.0