Strategic Finance and Economics Insights
ISSN(Online): 3104-6827
Volume 2, Issue 3, May-Jun 2026, Page 9-15
Original Research Article
Received: 12-05-2026| Accepted: 15-06-2026| Published: 20-06-2026

Generative Artificial Intelligence and Investor Decision-Making

Amit Kumar *

Abstract
Generative artificial intelligence (GenAI) tools, most visibly ChatGPT, have moved from a novelty to a routine input in how individual investors gather information, interpret news, and decide whether to buy or sell. This paper reviews recent empirical and experimental research on the role of GenAI in investor decision-making, drawing on studies published in accounting, finance, and behavioral economics outlets between 2015 and 2026. Three strands of evidence are examined. First, large language models show a measurable capacity to extract sentiment from financial text and to anticipate short-run stock price reactions, although this predictive edge appears to narrow as adoption spreads and markets adjust. Second, archival and survey data confirm that a substantial share of retail investors already rely on GenAI chatbots to screen and interpret company information, with usage shifting toward more sophisticated tasks as familiarity grows. Third, long-standing psychological barriers, particularly algorithm aversion and overconfidence, continue to limit how much weight investors place on machine-generated advice, even when that advice is demonstrably useful. The review also considers risks that are specific to generative systems, including fabricated outputs and an emerging "illusion of ability" that interactive tools can create in novice users. The paper closes with implications for financial educators, platform designers, and regulators, and it identifies gaps that future longitudinal and cross-market research should address.

Keywords
Generative Artificial Intelligence, ChatGPT, investor behavior, robo-advisors, algorithm aversion, behavioral finance

Independent Research

INTRODUCTION
The public release of ChatGPT in November 2022 marked a turning point in how ordinary people interact with artificial intelligence, and financial markets were among the first arenas to feel the effect. Within months, brokerage platforms began embedding chatbots into their apps, financial commentators started asking language models to explain earnings calls, and academics began testing whether a system trained to predict the next word in a sentence could also say something useful about the next day's stock return. That question has since generated a fast-growing body of research spanning accounting, finance, and behavioral economics.

Generative artificial intelligence differs from the earlier wave of algorithmic and robo-advisory tools in an important way. Traditional robo-advisors apply a fixed set of rules to a client's stated risk tolerance and produce a portfolio recommendation. Generative systems instead hold something closer to a conversation: an investor can ask a follow-up question, request a plainer explanation, or probe the reasoning behind a suggestion. This shift from static output to interactive exchange raises the stakes of two older debates in behavioral finance. The first concerns whether investors trust and use algorithmic advice at all, a question studied for decades under the heading of algorithm aversion. The second concerns what happens to judgment quality when a tool feels more capable and more approachable than it actually is.

This paper brings together three strands of recent literature that are usually discussed separately. The first strand treats generative models as analytical instruments and asks how well they process financial text and forecast returns. The second strand documents how retail investors actually adopt and use these tools once they are available. The third strand returns to behavioral finance and asks how trust, bias, and the design of the interface shape whether AI-generated insight ever translates into better decisions. Bringing these strands together clarifies where the

***Corresponding Author:** Amit Kumar

technology delivers a genuine informational advantage and where the limiting factor is not the model but the investor.

LITERATURE REVIEW

Generative AI as an Analytical and Forecasting Tool

The earliest and most cited empirical result in this literature comes from Lopez-Lira and Tang (2023), who fed tens of thousands of post-training news headlines into ChatGPT and asked it to classify each as good or bad news for the mentioned firm. GPT-4's classifications predicted roughly ninety percent of the next trading day's directional portfolio returns for the immediate reaction, and its scores also forecast the slower drift that follows, particularly among small firms and negative news. The authors interpret this as evidence that sophisticated return forecasting is an emergent capability of large models rather than something explicitly programmed, and they show that a long-short strategy built on the model's scores would have earned a meaningful premium over their sample period. Importantly, they also find that this edge shrinks as more market participants adopt similar tools, which is consistent with the classical view that information advantages erode once they become common knowledge.

Follow-on work has asked whether this text-processing skill extends to portfolio construction rather than single-stock prediction. Ko and Lee (2024) tested ChatGPT's ability to select and weight assets for diversified portfolios and found an asymmetry: the model was reasonably effective at choosing which assets to include but noticeably weaker at assigning the weights that would make the portfolio efficient. Performance improved substantially once the model's asset picks were paired with a conventional optimization step, suggesting that generative AI is currently better suited to the qualitative, language-heavy parts of investment analysis than to the quantitative optimization that portfolio theory demands. A related study by Chen et al. (2023) linked ChatGPT-derived sentiment from front-page financial news to both stock returns and macroeconomic indicators, reinforcing the idea that the model's main comparative advantage lies in compressing large volumes of text into a usable signal rather than in numerical reasoning.

Adoption and Usage Patterns Among Retail Investors

A separate line of research moves away from testing what the models can do in principle and instead asks what investors actually do with them. Blankespoor et al. (2026) offer the most direct evidence to date, combining an archive of more than four hundred thousand queries submitted to a major brokerage's GenAI chatbot with a survey of over two thousand retail investors. Close to half of the surveyed investors reported already using GenAI, most often to interpret earnings releases and market-moving news rather than to receive explicit buy-or-sell recommendations. The archival data show that usage patterns mature over time: new users lean on the tool for broad stock screening, while more experienced users shift toward detailed monitoring of specific holdings. The same study finds that concerns about answer quality and data privacy remain the main obstacles for investors who have not yet adopted the technology.

Cheng et al. (2025) approach adoption from a different angle by treating brief ChatGPT service outages as a natural experiment. Trading volume fell by roughly ten percent relative to average levels during outage periods, and the drop was largest for stocks with corporate news released just before or during the disruption and for stocks with heavier ownership by short-horizon institutional investors. Price impact and return variance also declined while ChatGPT was unavailable, a pattern consistent with less informed trading taking place. Over the longer run, the study associates GenAI-assisted trading with improved stock price informativeness, suggesting that the technology's aggregate effect on markets may be to speed up, rather than distort, the incorporation of public information into prices.

Adoption is not uniform across investor populations, and financial literacy appears to be one of the more important moderating factors. Ullah et al. (2024) surveyed retail investors in

Pakistan and found that the relationship between how investors use ChatGPT and the investment decisions they subsequently make is not fixed; it strengthens or weakens depending on the investor's underlying financial literacy. This finding complicates a simple story in which GenAI levels the playing field for less sophisticated investors. Niszczota and Abbas (2023) add a related caution: in their assessment, GPT-3.5 answered sixty-five percent of a standard financial literacy test correctly against a thirty-three percent guessing baseline, while GPT-4 reached a near-perfect ninety-nine percent, indicating that financial literacy has become an emergent property of the more advanced models. Yet their companion survey of laypeople taking advice from the tool suggests that people with weaker economic knowledge are, if anything, more inclined to trust its output, which raises the possibility that the investors least equipped to catch an error are also the most likely to accept one.

Trust, Algorithm Aversion, and Behavioral Biases

Even a technically capable tool is only useful if investors are willing to act on what it says, and this is where a much older behavioral finding becomes relevant. Dietvorst, Simmons, and Massey (2015) showed that people frequently abandon an algorithm after watching it make even a single mistake, even when the algorithm continues to outperform human judgment on average. This tendency, labeled algorithm aversion, has proven durable in financial settings. Filiz et al. (2022) placed participants in a laboratory diversification task in which a robo-advisor consistently produced the option with the highest expected payoff. Despite this visible advantage, participants delegated the decision to the robo-advisor in only about forty percent of trials, and their reluctance measurably reduced their own earnings.

Design choices appear to soften, though not eliminate, this resistance. Kulkarni et al. (2025) surveyed 461 retail investors and found that robo-advisors perceived as personalized, interactive, transparent about their algorithms, and appropriately autonomous were associated with lower overconfidence and loss-aversion bias, which in turn predicted greater willingness to use the tool. Structural assurance, meaning formal guarantees about the platform's reliability, was the one design factor that did not show a significant effect in their sample. Taken together with the laboratory evidence on algorithm aversion, this literature suggests that the barrier to AI-assisted investing is rarely a lack of technical performance; it is much more often a matter of how much the investor understands and trusts the process that produced the recommendation.

Risks Specific to Generative Systems

Generative tools introduce at least one risk that the older robo-advisor literature did not have to confront: the conversational format itself can change how capable an investor feels, independent of whether their judgment actually improves. Croom (2026) tested this directly and found that interacting with a generative AI tool tended to inflate investors' confidence in their own analytical ability beyond what their actual decision quality would justify, an effect the study describes as an illusion of ability. Because the tool answers follow-up questions fluently and in natural language, users may come away crediting themselves with insight that in fact came largely from the model, which could make them more willing to trade on borrowed confidence rather than borrowed information.

A second, more familiar risk is fabrication. Language models can produce plausible-sounding but inaccurate financial figures or citations, and Niszczota and Abbas (2023) note that simply rephrasing or repeating a query can reduce, though not eliminate, this kind of error. Neither risk is captured well by the earlier robo-advisor research, which generally assumed that the algorithm's output was internally consistent even when investors distrusted it. With generative tools, the more pressing concern is sometimes the opposite: investors may trust an output that is not, in fact, reliable.

Synthesis of Reviewed Evidence

Table 1 summarizes the ten core studies reviewed above, organized by their primary focus, method, and central finding. The studies cluster naturally into the three strands described earlier:

computational tests of forecasting and text-processing ability, archival and survey evidence on adoption, and experimental work on trust and behavioral bias.

Table 1: *Summary of Core Studies on Generative AI and Investor Decision-Making*

Study	Focus	Method / Sample	Key Finding
Lopez-Lira & Tang (2023)	Return forecasting from news text	50,000+ headlines; GPT-3.5/GPT-4	GPT-4 scores predict ~90% of next-day directional reaction; edge narrows with adoption
Ko & Lee (2024)	Portfolio construction	ChatGPT-guided asset selection and weighting	Strong at asset selection, weaker at optimal weighting; improves when paired with optimization models
Blankespoor et al. (2026)	Retail adoption and usage	400,000+ chatbot queries; survey of 2,000+ investors	Nearly half of investors use GenAI; usage shifts from screening to detailed monitoring over time
Cheng et al. (2025)	Market-level effects of GenAI use	ChatGPT outages as natural experiment	Trading volume falls ~10% during outages; long-run gains in price informativeness
Ullah et al. (2024)	Financial literacy as moderator	Survey of retail investors, Pakistan	Financial literacy moderates the link between ChatGPT usage and investment decisions
Niszczota & Abbas (2023)	Model literacy and advice-taking	Financial literacy test; survey of 184 laypeople	GPT-4 scores 99% on literacy test; less financially literate users show greater trust in advice
Dietvorst et al. (2015)	Algorithm aversion	Behavioral experiments	People abandon algorithms after seeing an error, even when the algorithm outperforms on average
Filiz et al. (2022)	Delegation to robo-advisors	Laboratory diversification task	Only ~40% of decisions delegated to a consistently optimal robo-advisor
Kulkarni et al. (2025)	Robo-advisor design and bias	Survey of 461 retail investors (SEM)	Personalization, interactivity, and transparency lower overconfidence and loss aversion
Croom (2026)	Illusion of ability	Experimental study of investor judgment	GenAI interactivity inflates perceived ability beyond actual judgment quality

Note. All figures and percentages are as reported in the cited primary studies. SEM = structural equation modeling.

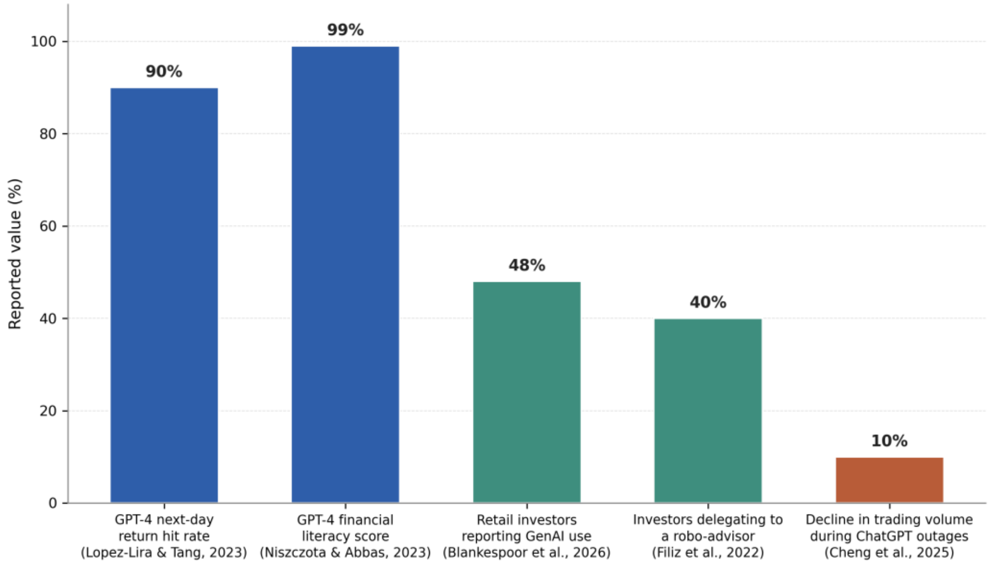

***Figure 1**. Reported Quantitative Findings From Selected Studies on Generative AI in Investment Contexts*
Note. Bars display single reported statistics drawn from the indicated studies and are not directly comparable across studies because sample, method, and time period differ. GenAI = generative artificial intelligence.

DISCUSSION AND IMPLICATIONS

Read together, these studies suggest that generative AI has already cleared the bar of being useful for certain investing tasks, particularly the rapid interpretation of text-heavy information such as headlines, earnings releases, and disclosures. What remains uneven is the human side of the equation: whether investors trust that usefulness enough to act on it, and whether the confidence that a fluent chatbot inspires is actually earned. This has practical implications for at least three audiences.

For individual investors and financial educators, the literacy-contingent pattern documented by Ullah et al. (2024) and Niszczota and Abbas (2023) argues for pairing GenAI access with baseline financial education rather than treating the tool as a substitute for it. An investor who cannot independently sanity-check a chatbot's output is poorly positioned to catch the fabricated figures or overconfident framing that these systems can occasionally produce.

For platform designers, the robo-advisor literature offers a fairly consistent recipe: transparency about how a recommendation was generated, personalization to the investor's situation, and a degree of investor autonomy all reduce algorithm aversion and increase constructive use of the tool. Croom's (2026) illusion-of-ability finding complicates this recipe, however, since the same interactive features that build trust may also be the ones that inflate a user's sense of their own competence. Interface choices that build appropriate trust without also manufacturing false confidence are likely to be an active area of product design in the coming years.

For regulators and market infrastructure providers, Cheng et al.'s (2025) outage study is a useful reminder that GenAI adoption is now large enough to be visible in aggregate trading data. A single company's service disruption measurably changed trading volume and price dynamics across the market, which suggests that concentration risk in AI infrastructure is now a market-structure question and not only a technology question. Disclosure standards for AI-assisted advice, along with clearer accountability rules for fabricated or misleading outputs, would address a gap that the current robo-advisor regulatory framework was not built to cover.

LIMITATIONS AND FUTURE RESEARCH

This review has several boundaries worth noting. Most of the evidence summarized here comes from U.S. equity markets and English-language models, with the Pakistani sample in Ullah et al. (2024) a useful but isolated exception; whether these patterns generalize to other markets, languages, and regulatory regimes is largely untested. The literature is also still young: several of the studies cited are forthcoming or recently published, and effect sizes reported in early work, particularly around return predictability, may shift as more investors adopt similar tools and prices adjust accordingly, a dynamic Lopez-Lira and Tang (2023) themselves anticipate. Finally, almost all of the reviewed studies examine GenAI in an advisory or informational role. As agentic systems capable of executing trades with limited human oversight become more common, the behavioral and market-structure questions raised here will need to be revisited for a setting where the human is no longer the final decision-maker. Longitudinal studies that follow the same investors as models improve, along with cross-market replications, would meaningfully extend what is currently a fast-moving but still narrow evidence base.

CONCLUSION

Generative artificial intelligence has moved quickly from an experimental curiosity to a routine part of how many retail investors process financial information. The evidence reviewed here indicates that the technology's analytical capability is real but narrowing as it becomes more widely adopted, that investor uptake is uneven and shaped strongly by financial literacy, and that the psychological barriers documented in decades of algorithm-aversion research have not disappeared simply because the algorithm can now hold a conversation. If anything, the conversational format introduces a new and still poorly understood risk: an investor may leave an interaction with a chatbot more confident, without necessarily being more correct. Addressing that gap, through better financial education, more thoughtfully designed interfaces, and clearer regulatory guardrails, is likely to determine whether generative AI ultimately improves investor decision-making or simply changes the form that old biases take.

REFERENCES

1. Blankespoor, E., Croom, J., & Grant, S. M. (2026). Generative AI and investor processing of financial information. *Journal of Accounting and Economics*, *82*(2), forthcoming.
2. Chen, J., Tang, G., Zhou, G., & Zhu, W. (2023). *ChatGPT, stock market predictability and links to the macroeconomy* (SSRN Working Paper No. 4660148). https://doi.org/10.2139/ssrn.4660148
3. Cheng, Q., Lin, P., & Zhao, Y. (2025). Does generative AI facilitate investor trading? Early evidence from ChatGPT outages. *Journal of Accounting and Economics, 80*, Article 101821. https://doi.org/10.1016/j.jacceco.2025.101821
4. Croom, J. (2026). Interactivity and illusions of ability: How using generative AI affects investor judgments. *Journal of Accounting Research*. Advance online publication. https://doi.org/10.1111/1475-679X.70017
5. Dietvorst, B. J., Simmons, J. P., & Massey, C. (2015). Algorithm aversion: People erroneously avoid algorithms after seeing them err. *Journal of Experimental Psychology: General, 144*(1), 114–126. https://doi.org/10.1037/xge0000033
6. Filiz, I., Judek, J. R., Lorenz, M., & Spiwoks, M. (2022). Algorithm aversion as an obstacle in the establishment of robo advisors. *Journal of Risk and Financial Management, 15*(8), Article 353. https://doi.org/10.3390/jrfm15080353
7. Ko, H., & Lee, J. (2024). Can ChatGPT improve investment decisions? From a portfolio management perspective. *Finance Research Letters, 64*, Article 105433. https://doi.org/10.1016/j.frl.2024.105433
8. Kulkarni, M. S., Patil, K. P., & Pramod, D. (2025). The role of robo-advisors in behavioural finance, shaping investment decisions. *Cogent Economics & Finance, 13*(1), Article 2571403. https://doi.org/10.1080/23322039.2025.2571403
9. Lopez-Lira, A., & Tang, Y. (2023). *Can ChatGPT forecast stock price movements? Return*

predictability and large language models (SSRN Working Paper No. 4412788). *Journal of Financial Economics*, advance online publication. https://doi.org/10.2139/ssrn.4412788

10. Niszczota, P., & Abbas, S. (2023). GPT has become financially literate: Insights from financial literacy tests of GPT and a preliminary test of how people use it as a source of advice. *Finance Research Letters, 58*, Article 104333. https://doi.org/10.1016/j.frl.2023.104333
11. Ullah, R., Ismail, H. B., Khan, M. T. I., & Zeb, A. (2024). Nexus between ChatGPT usage dimensions and investment decisions making in Pakistan: Moderating role of financial literacy. *Technology in Society, 76*, Article 102454. https://doi.org/10.1016/j.techsoc.2024.102454

Conflict of Interest: The authors declare no conflicts of interest
Source of Funding: Author(s) Funded the Research
How to Cite: Kumar, A. (2026). Generative Artificial Intelligence and Investor Decision-Making. *Strategic Finance and Economics Insights, 2*(3), 9-15.

.