Proceedings of ICFTBA 2025 Symposium: Global Trends in Green Financial Innovation and Technology
DOI: 10.54254/2754-1169/2025.BJ30101

The Role of Behavioral Biases in Algorithmic Trading: A Comprehensive Review of Evidence from Global Equity Markets

Jinghao Yang

Hong Kong Polytechnic University, Hong Kong, China
22099878D@connect.polyu.hk

Abstract. Behavioral finance explores the psychological influences and cognitive biases that affect investor behavior and financial decision-making, including herding, the disposition effect, overconfidence, and others. Algorithmic trading is a method that uses computer programs to automatically execute buy and sell orders based on predefined mathematical models and trading strategies. With the continuous development of modern technology, the advent of the Web3 era, and the gradual evolution of artificial intelligence, algorithmic trading is becoming increasingly prevalent and garnering significant attention. While algorithmic trading is automated and may seem immune to human cognitive biases, the opposite is often true. This study aims to review the main findings of existing research from the perspective of the stock market, exploring the interactive relationship between behavioral finance and algorithmic trading and how cognitive biases such as herding and the disposition effect can influence algorithm performance. The results emphasize the importance of behavioral finance in both the research and practice of algorithmic trading, while also proposing the potential for using machine learning techniques to advance the field of behavioral finance. By integrating existing theories, this study contributes to a deeper understanding of the relationship between behavioral finance and algorithmic trading and offers new perspectives for its future development.

Keywords: Cognitive Bias, Behavioral Finance, Algorithmic Trading, Big Data, Machine Learning

1. Introduction

Algorithmic trading utilizes complex computational algorithms and high-frequency data to automatically execute buy and sell orders based on pre-set strategies and rules, identifying and capturing trading opportunities within milliseconds [1]. These systems rely on objective market data and quantitative analysis to reduce the impact of human emotions and significantly increase trading frequency. However, despite this, automated trading based on algorithms cannot completely eliminate human influence. Even though algorithmic programs operate automatically based on set parameters, cognitive biases can inevitably be incorporated into the algorithm's design process,

© 2025 The Authors. This is an open access article distributed under the terms of the Creative Commons Attribution License 4.0 (https://creativecommons.org/licenses/by/4.0/).

DOI: 10.54254/2754-1169/2025.BJ30101

thereby affecting the final trading outcome [2]. Algorithms are created by humans, and the potential cognitive biases inherent to humans should be considered.

Traditional academic finance emphasizes theories such as Modern Portfolio Theory and the Efficient Market Hypothesis, but the emerging field of behavioral finance studies the psychological and sociological issues that influence the decision-making processes of individuals, groups, and organizations [3]. Studying behavioral finance may help better understand these cognitive biases present in algorithmic trading. Common cognitive biases include herding, the disposition effect, overconfidence, loss aversion, etc., all of which can lead decision-makers to make irrational choices. For example, a consequence of heuristic simplification (overconfidence) is that people tend to think they are better than they actually are [4]. This may lead algorithm models to exhibit overreaction or underreaction to outcomes, as designers might cause pre-set parameters and models to be insufficiently rational, preventing the achievement of maximum efficiency.

This paper aims to investigate the interactive relationship between behavioral biases and algorithmic trading. By reviewing existing research findings, we integrate theoretical research and suggest potential future trends. Starting from traditional behavioral finance, we analyze how it studies investor decision-making, then explain why algorithmic trading cannot escape the influence of behavioral biases. Subsequently, we discuss how behavioral finance and algorithmic trading influence each other, combining the latest research theories and directions to hypothesize about potential future research avenues. Therefore, elucidating the interplay between behavioral biases and algorithmic trading not only facilitates a more accurate interpretation of price fluctuations and risk drivers in real markets but also provides a crucial theoretical foundation for developing the next generation of intelligent trading systems that are more adaptive and better equipped to circumvent human frailties.

2. Literature review

In recent decades, the interconnected fields of behavioral finance and artificial intelligence (AI) have attracted increasing attention from scholars, financial institutions, and regulators worldwide. The relationship between human behavioral biases and algorithmic trading systems has evolved into a prominent topic at the intersection of finance, technology, and psychology. Understanding these biases is crucial for developing more accurate and adaptive algorithmic models. This chapter comprehensively reviews key literature to clarify how behavioral finance principles interact with algorithmic trading, identify current theoretical gaps, and outline possible directions for future research.

Ritter's seminal work in behavioral finance demonstrated that human thinking patterns often produce systematic, predictable errors in financial decision-making [5]. These biases, such as overconfidence, loss aversion, representativeness, and confirmation bias, distort investor behavior and market outcomes. For instance, representativeness bias leads investors to underestimate long-term average returns while disproportionately focusing on recent performance, additionally exhibiting strong tendencies to make intuitive judgments about securities, companies, or market conditions based on salient representative characteristics rather than comprehensive statistical analysis, resulting in suboptimal decisions. These systematic errors appear deeply rooted in human psychology and information processing limitations, suggesting that no market participant, regardless of sophistication or experience, can maintain perfect rationality across all decisions and market conditions.

Chan, Jegadeesh, and Lakonishok provided early empirical evidence that algorithmic trading, though grounded in quantitative principles, remains susceptible to biases transmitted through human

DOI: 10.54254/2754-1169/2025.BJ30101

developers, historical data, and market interactions [6]. These researchers highlighted the unconscious incorporation of overconfidence and anchoring biases into algorithmic models, challenging the assumption that automation guarantees objectivity. Their findings marked a turning point in understanding how behavioral tendencies could indirectly shape supposedly rational trading systems.

DeLong, Shleifer, Summers, and Waldmann's classic study on noise traders revealed that many traders base decisions primarily on subjective cognitions, personal experiences, and emotional reactions rather than rigorous fundamental analysis or rational valuation models [7]. In such informationally inefficient markets characterized by significant noise trader participation, algorithmic trading systems must necessarily analyze, interpret, and respond to these irrational trends and behavioral patterns present in market data. An exclusively rational analytical approach to fundamentally irrational market phenomena can easily lead algorithmic systems to overlook or misinterpret these anomalous factors, potentially resulting in erroneous decisions, unexpected losses, or contribution to the amplification of existing behavioral biases throughout financial markets. This complex interaction between rational algorithms and irrational market participants creates fascinating dynamics that warrant continued investigation.

The conceptual and methodological relationship between behavioral finance and algorithmic trading continues to deepen and evolve rapidly. Yang advanced this field by explicitly integrating behavioral biases into algorithmic models to better capture and exploit market anomalies [8]. Using advanced analytical methods, their research modeled how biases such as overreaction and underreaction affect trading behavior. The study demonstrated that incorporating behavioral insights could enhance algorithmic adaptability and improve risk-adjusted returns, underscoring the potential synergy between behavioral and quantitative design to generate superior risk-adjusted returns compared to traditional approaches.

In more recent cutting-edge research, Jing identified that the increasing complexity of algorithmic models makes their decision logic difficult to interpret or validate [9]. This opacity creates significant challenges for incorporating the intuitive and nuanced judgments that frequently dominate human trading decisions, particularly under market stress. These subtle psychological factors prove exceptionally difficult to quantify, parameterize, and integrate into mathematical models, representing a fundamental challenge in current algorithmic trading system design. While researchers and practitioners increasingly attempt to leverage behavioral finance knowledge to refine existing algorithmic frameworks, progress in the crucial quantification and implementation stages remains limited and incremental. Furthermore, given the inherent complexity, context-dependence, and limited predictability of human behavior, computer models struggle to generate accurate predictions about complex financial futures based exclusively on numerical data and statistical relationships. Contemporary models contain numerous difficult-to-quantify variables that resist straightforward interpretation or incorporation, suggesting that the ambitious conjecture of systematically utilizing behavioral finance theory to perfect algorithmic trading models continues to face substantial theoretical, methodological, and practical difficulties that will require interdisciplinary collaboration to overcome.

3. Discussion

3.1. Human cognitive biases as the fundamental source of algorithmic biases

Despite claims of objectivity and rationality, algorithmic trading systems remain fundamentally rooted in human cognition. Behavioral biases such as overconfidence, loss aversion,

DOI: 10.54254/2754-1169/2025.BJ30101

representativeness bias, confirmation bias, and availability heuristic become inevitably embedded during design and calibration stages. During the initial design and specification phase, the underlying logic, parameter choices, risk limits, and operational rules comprising algorithmic trading strategies are ultimately determined and implemented by human developers, researchers, and financial engineers. The developers' behavioral biases, blind spots, and subjective preferences frequently become unconsciously integrated into system architecture; for instance, establishing stop-loss points, profit targets, or position sizing rules might implicitly incorporate loss aversion tendencies, disposition effects, or overconfidence in predictive accuracy. During the critical data training and model calibration phase, if the historical market data utilized for training algorithmic systems already contains substantial irrational behaviors, cognitive biases, and emotional reactions of market participants (such as herding behavior, momentum chasing, or overreaction to news), then the resulting algorithms inevitably learn, internalize, and potentially amplify these biased patterns rather than discovering fundamental economic relationships. This creates a paradoxical situation where algorithms designed to eliminate human irrationality may instead perpetuate and systematize it through different mechanisms.

3.2. Algorithmic trading systems as powerful "accelerators" and "systematic amplifiers" of market biases

Once algorithmic trading systems containing embedded behavioral biases become deployed in live market environments, they typically not only fail to eliminate or reduce existing biases but may dramatically amplify them through higher execution speeds, larger transaction volumes, and coordinated behavior across multiple systems. The reinforcement of pre-existing behavioral biases represents a significant concern: as empirically suggested by research from Chan et al., and DeLong et al., algorithmic systems can rapidly identify, analyze, and systematically chase market trends initially driven by noise traders, retail investors, or other potentially irrational market participants, thereby transforming localized, small-scale behavioral biases into widespread, intense market volatility affecting entire sectors or markets [6, 7]. Furthermore, the widespread adoption of algorithmic trading may generate completely novel "algorithmic biases" that did not previously exist in human-dominated markets: multiple algorithmic systems operating according to similar logical frameworks, data sources, or risk models can lead to emergent, systematic phenomena such as "algorithmic herding," "flash crashes," "liquidity black holes," or "quantitative contagion," whose astonishing speed, coordination, and impact magnitude vastly exceed capabilities of traditional human traders. These phenomena highlight a transformation: algorithmic systems evolve from passive executors of human bias to active agents shaping market dynamics, raising profound implications for financial stability and regulatory oversight.

3.3. Behavioral finance as a theoretical tool for optimizing algorithms

After recognizing the above issues, behavioral finance has transformed from the "root of the problem" into the "key to the solution." Researchers have begun actively using behavioral finance theory to correct and enhance algorithmic trading:

(1) Explaining Market Anomalies: Using behavioral biases (like overreaction and underreaction) to explain why market prices deviate from fundamentals, thereby providing new, more effective profit opportunities for algorithmic trading (e.g., Yang's 2019 research). (2) Improving Model Design: Quantifying and incorporating behavioral factors (such as investor sentiment, and attention)

DOI: 10.54254/2754-1169/2025.BJ30101

into algorithmic models aims to make algorithms more adaptable to real, not perfectly efficient markets, thereby improving their robustness and performance.

3.4. The dilemma and future direction of deep integration

Despite promising progress, achieving deep integration between behavioral finance and algorithmic trading remains challenging. A key difficulty lies in quantifying intuition and subjective judgment, dimensions that remain elusive to formal modeling. As Jing observed, factors such as intuition and fuzzy judgment in human trading, which are difficult to quantify, are currently hard for algorithmic models to capture [9]. The high complexity of models also makes their internal logic difficult to understand and control, like a "black box." Overcoming this may require advanced AI paradigms such as deep learning and reinforcement learning, capable of capturing non-linear behavioral dynamics and simulating human decision complexity. Gupta and Rao further advocate leveraging large-scale AI infrastructures such as distributed computing, cloud infrastructure, and sophisticated machine learning algorithms to systematically analyze, model, and predict investor behavioral biases and their market impact, thereby harnessing technological advancement to accelerate the development of behavioral finance as an empirical discipline and its practical applications in investment management [10].

The findings generated by this comprehensive research study carry numerous implications for practical application within financial institutions, investment firms, and regulatory bodies. For instance, systematically combating overconfidence bias in algorithmic systems could involve implementing more conservative position sizing, incorporating broader uncertainty estimates, or increasing the weight assigned to recent performance metrics and regime changes. Conversely, effectively addressing loss aversion tendencies might require designing more dynamic, context-aware risk management rules that adjust exposure based on market volatility, portfolio drawdowns, or correlation structures. Alternatively, financial engineers could employ advanced statistical techniques and machine learning methods to conduct rigorous comparative performance analyses between algorithms containing specific behavioral biases and theoretically neutral benchmarks, meticulously analyzing key performance indicators, risk metrics, and behavioral patterns to precisely determine the quantitative impact magnitude of particular biases on trading performance, risk-adjusted returns, and portfolio characteristics across different market environments.

4. Conclusion

This paper argues that algorithmic trading, despite its apparent rationality and automation, remains deeply intertwined with human behavioral biases. Algorithms are both the product of behavioral biases and their amplifier, while behavioral finance, in turn, provides the key theoretical tools for optimizing algorithms.

The main content of the paper indicates that algorithms, being designed by humans and trained on historical data, are inherently embedded with the cognitive biases of their designers (such as overconfidence and loss aversion) as well as the irrational patterns present in the market data. When these algorithms with "biased genes" enter the market, they can reinforce and even amplify existing behavioral biases and on a massive scale, triggering algorithm-driven herding or market flash crashes, thus evolving from passive execution tools into active sources of market disruption. However, this predicament also points out: behavioral finance theory can be actively used to correct and optimize algorithmic trading. By quantifying and integrating behavioral biases like overreaction into models, algorithms can better explain market anomalies and enhance performance in inefficient

DOI: 10.54254/2754-1169/2025.BJ30101

markets. Ultimately, the paper reveals that the core challenge currently lies in the problem of "quantifying intuition"—how to translate the fuzzy intuition of human decision-making into data processable by algorithms, and how to understand the increasingly complex algorithmic "black box."

Nonetheless, as a literature-based analysis, this paper has some limitations. Its main constraint lies in its focus on theoretical discussion and lack of empirical support, failing to visually demonstrate the quantitative impact of behavioral biases on trading performance through specific algorithmic case backtesting. In terms of its literature foundation, the discussion of applications of cutting-edge AI technologies, particularly deep learning and reinforcement learning, in behavioral finance remains relatively weak, making the paper's connection to the "AI era" less forceful.

Looking ahead, future research could focus on using explainable AI techniques to crack the algorithmic black box, making the algorithm's decision-making process (e.g., based on which behavioral features it makes judgments) transparent and trustworthy. Simultaneously, by integrating multimodal data such as news and social media, and applying natural language processing techniques, models will be able to more accurately capture hard-to-quantify market sentiment and intuition. Furthermore, Agent-based modeling methods will provide a powerful "digital laboratory," allowing researchers to test the complex outcomes of interactions between different behavioral biases and algorithms in a simulated market, thereby foreseeing and mitigating potential risks at the design stage. Ultimately, developing intelligent algorithms capable of adapting to market states and autonomously switching strategies between rational and irrational environments will be the frontier direction for achieving true unification of behavioral finance and algorithmic trading.

References

[1] Aldridge, I., & Krawciw, S. (2017) Real-time risk: What investors should know about FinTech, high-frequency trading, and flash crashes. John Wiley & Sons.
[2] Das, T., Jeelani, S., Das, S., Giri, P., & Chatterjee, A. (2024) Human Bias in Algorithmic Trading: Evaluating Behavioral Finance Impacts on Automated Systems. In 2024 Second International Conference Computational and Characterization Techniques in Engineering & Sciences (IC3TES) (pp. 1-5). IEEE.
[3] Ricciardi, V., & Simon, H. K. (2000) What is behavioral finance?. Business, Education & Technology Journal, 2(2): 1-9.
[4] Trivers, R. (1991) Deceit^ ind Self-Deception The Relationship.
[5] Ritter, J. R. (2003) Behavioral finance. Pacific-Basin finance journal, 11(4): 429-437.
[6] Chan, L. K., Jegadeesh, N., & Lakonishok, J. (1999) The profitability of momentum strategies. Financial Analysts Journal, 55(6): 80-90.
[7] De Long, J. B., Shleifer, A., Summers, L. H., & Waldmann, R. J. (1990) Noise trader risk in financial markets. Journal of political Economy, 98(4): 703-738.
[8] Li, X., Yang, Q., and Wang, X. (2019) Incorporating behavioral biases into trading strategies: Evidence from high-frequency trading. Quantitative Finance, 19(3): 407–423.
[9] Jing, Z. W. (2025) Algorithmic Supervision of Market Manipulation Behaviors in Programmatic Trading. Open Journal of Legal Science, 13: 1189.
[10] Gupta, S., & Rao, M. V. S. AI in Behavioral Finance: Understanding Investor Bias Through Machine Learning.