Rethinking Risk Management in Robo-Advisory: Insights from Behavioural Finance and Algorithmic Transparency

Yanda Deng *

School of Business, Hong Kong Baptist University, Hong Kong, China

* Corresponding Author Email: 22257578@life.hkbu.edu.hk

Abstract. With the rapid development of financial technology, robo-advisors have become a pivotal tool in wealth management. However, the risk management mechanisms face dual challenges in practice: first, investor Behavioural biases lead to distorted model inputs; second, algorithmic opacity undermines user trust and regulatory transparency. To adress these two challenges, this article takes Behavioural finance and algorithmic transparency as entry points, systematically reviewing research progress in these areas and identifying key gaps in theoretical integration and practical application. Based on this, it proposes a dual-layer risk management framework of 'Behavioural awareness—transparent feedback', emphasising the crucial synergy between dynamic investor behaviour recognition and algorithmic decision-making explainability. This framework not only enhances platforms' responsiveness to irrational behaviour but also optimises risk control and regulatory alignment mechanisms under algorithmic transparency. At the same time, the establishment of this framework provides a new direction for the development of robot-advisors. The article holds significant theoretical and practical value for designing optimised robo-advisor systems, constructing trust mechanisms, and formulating regulatory strategies.

Keywords: Robo-Advisors, Risk Management, Behavioural Finance, FinTech Governance, Algorithmic Transparency.

1. Introduction

Global wealth management is being revolutionised by the development of financial technology (FinTech) rapidly. Robo-advisors are the famous tools that attract many people. Although robo-advisors increase productivity, there is growing scrutiny of their risk management procedures. The "black box" nature of complicated algorithms, the distortion of input quality brought on by investor behavioural biases, and the increasingly strict transparency and compliance requirements of regulators present the largest obstacles. Robo-advisor risk management combines computer science, Behavioural finance, and finance, and has attracted much attention in recent years. Behavioural finance provides a basic theoretical framework for understanding how psychological and emotional biases distort the decision-making process, and also provides expertise for analyzing irrational investor behavior. For example, Rad et al. used a decision tree regression model to show how Behavioural biases affect portfolio selection [1]. This means that robo-advisors need to constantly adjust their models to respond to market changes [2]. On the other hand, research on algorithm transparency highlights the dangers of opaque models. Arrieta et al. proposed an explainable artificial intelligence (XAI) framework to illustrate how the inability to understand how algorithms work exacerbates privacy issues and reduces user trust [3]. Anshari et al. also proposed digital twin technology as a tool to improve risk management and transparency [4].

This article combines existing research to explain the systematic nature of investor irrationality, its potential impact on algorithm feedback mechanisms, and its significant impact on user trust and market acceptance. In addition, this article proposes new framework that integrates Behavioural cognition and transparent feedback.

2. Evolution and Emerging Challenges of Robo-Advisors

FinTech has developed rapidly since the financial crisis, and robo-advisors have contributed to the surge in automated wealth management. From simple models based on risk questionnaires and passive investing, related platforms have improved into extremely dynamic investment platforms that use big data and artificial intelligence. In addition, this technological development has radically changed risk dynamics and investor behaviour.

But innovation also brings with it new difficulties. On the one hand, cognitive and emotional biases (overconfidence and loss aversion) continue to have a significant impact on investor behaviour, resulting in skewed risk inputs. However, deep learning and ensemble approaches have become increasingly popular due to the growing complexity of models. Interpretability has suffered as a result, leading to the emergence of unidentified "black box" systems. These elements worsen compliance risks and damage trust.

Through history, robo-advisors have consistently evolved in step with risk management innovations. Early systems, reliant on historical data and static user profiles, often overlooked the dynamic dimensions of investor behaviour, which resulted in late or weak responses to market volatility. As user needs and regulatory requirements increase, platforms must incorporate behavioural models and enhance algorithmic transparency to rebuild trust.

In this background, this article proposes a systematic exploration of the response strategies for intelligent investment advisors in the new risk environment, taking the two-dimensional integration of behavioural finance and algorithmic transparency as the entry point. Unlike previous research that examined these dimensions in isolation, we advocate constructing a unified analytical framework that links dynamic behavioural feedback with explainable algorithmic logic. This approach not only fills a theoretical void but also establishes a foundation for a credible robo-advisory systems.

Since regulatory policies vary from country to country, the robo-advisors studied in this article do not represent discretionary robo-advisors, but rather semi-discretionary investment advisor business models that are more admissible in each country.

3. The Impact of Behavioural Finance on Risk Management in Robo-Advisory Services

Behavioural finance offers a more realistic explanation of investor decision-making, overturning the assumption of perfectly rational agents in traditional economic models [5]. When applied to robo-advisory platforms, these theories reveal multiple ways in which user behavior can interfere with the effectiveness of risk management. In combination with existing research, this article explores the specific impact of Behavioural finance on the risk management mechanism from following four dimensions:

3.1. Cognitive distortions compromise the accuracy of risk assessments

Common cognitive biases among investors, such as overconfidence, anchoring, and loss aversion, directly impact the quality of their decisions when filling out risk preference questionnaires or making asset allocation decisions. These biases can distort the data captured by intelligent investment systems, leading to a mismatch between risk tolerance and strategic allocation. For example, investors may exhibit aggressive tendencies in a bull market but quickly redeem their investments during a market decline. This kind of input data bias caused by Behavioural bias will weaken the model's ability to identify the real risk, affecting the overall risk control effect from the source.

3.2. Shifting behaviours challenge model stability over time

Investor behaviour is not fixed. People's financial decisions tend to evolve as markets change or as personal circumstances shift. A risk profile established at the time of onboarding may lose relevance within weeks or even days. Research by Rad et al. highlights how Behavioural drift affects

asset allocation, and suggests that monitoring tools—such as tracking changes in asset selection frequency or abrupt reallocation—can help detect early signs of Behavioural volatility [6]. Robo-advisors that build in such mechanisms may be better positioned to recalibrate portfolios before discrepancies become damaging. Without this adaptability, risk models may lag behind user behaviour, rather than guiding it.

3.3. Herd behaviour can amplify systemic risks

Bias not only distorts individual decision-making, it can spread across groups, especially in stressful market environments. A sudden redemption by one investor may be rational, but if followed by thousands in a short period of time, it could overwhelm platform liquidity and disrupt asset rebalancing strategies. Behaviours such as panic selling, fear-driven asset switching, or a general disregard for long-term strategy, if unforeseen, will expose systemic weaknesses. If investors are viewed as isolated individuals, platforms may miss the early warning signs of group psychology. A more resilient strategy should group users based on Behavioural tendencies and identify the broader risks that collective action can trigger.

3.4. Investor psychology influences how risk messages are received

Even when systems accurately assess risk, their effectiveness depends on how users interpret and respond to the information provided. Many risk warnings or performance updates fail not because they are inaccurate, but because they are not tailored to the investor's emotional state or decision-making context. Behavioural research has shown that framing effects, timing, and presentation format (e.g., charts vs. text) all influence user response. Designs such as setting a temporary confirmation delay ("cooling-off period"), reminding of long-term goals, or providing personalized alerts can significantly change how users react when faced with market declines or volatility. These subtle design choices often determine whether a user stays invested or withdraws at the worst possible time.

In summary, Behavioural finance theory provides a theoretical basis and methodological path for intelligent investment platforms to build a more adaptive and forward-looking risk management mechanism by identifying and explaining the Behavioural deviations of investors in irrational states. By integrating Behavioural finance into risk control logic, not only is individual risk identification and strategy matching optimized, but the system's tolerance for Behavioural imbalances in extreme market conditions is also enhanced, thereby improving the overall stability of the platform and user trust.

4. The Impact of Algorithm Transparency on Risk Management in Robo-Advisory Services

In intelligent investment systems that rely heavily on data and models, algorithmic transparency is becoming a part of risk management that cannot be ignored. In the past, people paid more attention to the yield performance and technical accuracy of algorithms, but now they increasingly realize that an "unintelligible" algorithmic system, even if it is more powerful, can hardly win the real trust of investors and regulation.

4.1. Transparency as the Basis for User Trust

First of all, transparency is the premise of establishing user trust. In traditional finance, communication between people can establish trust through language, tone and context, while intelligent investment is completely dependent on algorithmic output. It is difficult for users to form a stable trust relationship without understanding the logic behind a vague "suggestion", and the Explainable Artificial Intelligence (XAI) model proposed by Arrieta et al. attempts to solve this problem: users not only know "what", but also understand "why" [3]. This understanding is critical to maintaining risk expectations and strategy execution, especially in the face of market volatility,

where users who lack confidence in the source of advice are more likely to engage in irrational redemptions and other behaviors.

4.2. Enhancing Risk Recognition through Visualized Logic

Transparency also determines whether risks can be clearly identified and predicted. In reality, many platforms only show the results of a simple recommendation, while the underlying data sources, variable weightings, strategy assumptions, etc. are all in a "black box" state. This puts users at an information disadvantage and may even misjudge risks. On the contrary, if the platform can present the basis for decision-making in a popular visual way, such as showing which factors affect the current recommendations, historical backtesting results of similar situations, etc., it can significantly enhance the user's recognition of the strategy and willingness to implement.

4.3. Regulatory Demands for Algorithmic Interpretability

The interpretability of algorithms is gradually becoming a rigid requirement for regulation. With the rising data compliance and AI ethical issues, regulations such as the EU GDPR have clearly put forward the auditable and interpretable requirements of algorithms. As an automated system for mass users, smart investment advisors are more likely to need to reflect the principle of transparency in their compliance design. For example, the "digital twin" model proposed by Anshari et al. can be used to conduct simulation audits of models, which can enhance the regulatory transparency and accountability of platforms [7]. The introduction of such technology not only prevents collective misinformation caused by algorithmic bias, but also lays the foundation for the future development of financial regulatory technology (RegTech). Also, for certain algorithmically generated alternative investments, such as insurance products, disclosure of the underlying algorithms is also necessary [8].

4.4. The Case for Layered Transparency

It is worth noting that, in some cases, transparency itself is not necessarily better. Excessive disclosure of model structure may lead to new problems, such as arbitrage, strategy leakage, and even misleading ordinary users. Therefore, transparency design needs to find a balance between technical complexity and user cognitive ability. The best solution is to achieve "layered explanation": provide ordinary investors with intuitive results and basic logical explanations, and provide professional investors and regulators with more detailed model descriptions and technical documents. Specialized research and development for specific companies can present specific data [9]. This layered mechanism can enhance trust and understanding while maintaining the sensitivity of algorithmic strategies.

In short, algorithmic transparency is not only the basis for users to understand and accept smart investment advisors, but also an important guarantee for compliance and risk control. From the cross-framework perspective of Behavioural finance, the transparency mechanism constitutes the key channel for the platform to "respond to user psychology and stabilize operational behavior." In the future, if smart investment advisory systems want to better serve investors, in addition to "strong algorithms", they must also add the difficult problem of "understandability".

5. Integrative Analysis: The Impact of Behavioural Finance × Algorithmic Transparency on Risk Management in Robo-Advisory Services

Algorithmic transparency and Behavioural finance are not separate entities. Instead, they are highly complementary and interdependent. While the latter focuses on a system's ability to make its decision-making process transparent and auditable, the former examines investor Behavioural tendencies and identifies deviations from rational decision-making. Combining the two not only makes conceptual sense but also help with two key challenges currently facing: poor decision quality and eroding user trust.

As the article mentioned, algorithmic transparency ensures the system's responsiveness, while Behavioural finance provides a powerful explanatory framework for understanding how investors manage uncertainty. Abnormal trading frequencies and portfolio allocations that deviate from risk preferences are among the problems they create [1]. Also, From a systemic perspective, if these trends are not recognized and addressed, risks may inadvertently increase [2]. Therefore, the integration of algorithmic transparency and Behavioural finance should not be limited to simply comparing the two mechanisms, but must achieve deep functional embedding and feedback loops [10].

Based on this synergy, this article proposes a two-layer risk management framework - Behavioural perception and transparent feedback, to achieve the coordinated optimization of Behavioural recognition mechanism and explainability mechanism. The framework consists of two interrelated functional layers:

Behaviour perception layer: This front-end layer builds a user behaviour model based on interactive behaviour, portfolio selection, and emotional response through real-time monitoring and data collection. Drawing on Behavioural finance theory, the system classifies users by Behavioural labels (such as "high volatility preference", "short-term trading mode", and "panic selling") [11]. Once irrational tendencies are detected, the system triggers dynamic portfolio rebalancing and risk reassessment to achieve pre-personalized risk management.

Transparent feedback layer: The back-end outputs explainable investment logic to users through a hierarchical interpretation strategy. Different user groups (retail investors, professionals, and regulators) obtain different levels of model details: retail users can view the basic logic and updated risk levels, professionals can obtain change weights and algorithm rules, and regulators can review the decision path and input data structure to facilitate audits and ensure compliance. The innovation of this framework lies in its construction of a closed-loop system, with Behavioural perception at the input end and transparent feedback at the output end.

In conclusion, the combination of algorithmic transparency and behavioural finance is not a simple "1+1=2" strategy. Instead, it leverages data-driven feedback and embedded mechanisms to systematically enhance risk management capabilities, allowing these two components to converge to form a new mechanism. This new mechanism, in turn, focuses on shifting risk response strategies from reactive defense to proactive Behavioural change.

6. Conclusion

The risk management philosophy employed by robo-advisors has also evolved, from a reliance on technical models to the integration of multifaceted strategies, and from basic passive asset allocation solutions to complex AI-based systems. During this evolution, algorithmic opacity and irrational investor behaviour have garnered widespread attention.

To address these two core issues, this article proposes a two-tiered architecture that combines Behavioural tendency sensing with clear explanations. On the input side, real-time detection and intervention improve the accuracy and speed of risk information responses. On the output side, by implementing explainable methods and a layered disclosure strategy, algorithmic recommendations become more understandable and reliable. This model provides a framework for creating more accountable and responsive robo-advisors that address both the technical and human aspects.

Its primary contributions are represented in two ways: first, it closes a gap in the literature by combining algorithmic explainability and behavioural finance into a single risk management paradigm, where system explainability and psychological insights support and enhance one another; second, it offers an operational implementation framework for creating robo-advisor systems that are more effective, user-friendly, and compliant.

The following avenues can be further investigated in future studies: (1) the two-layer architecture's flexibility and constraints in various regulatory contexts; (2) the best way to combine hierarchical transparency techniques with personalised intervention mechanisms; and (3) the privacy trade-off between model transparency and user behaviour data. Robot advisors risk control system must

develop further in tandem with the financial technology ecosystem in order to strike a dynamic balance between "technology-trust-regulation."

References

[1] Kamoune A., Ibenrissoul N. Traditional versus Behavioural Finance Theory - [La théorie de la finance traditionnelle contre la théorie de la finance comportementale]. IDEAS Working Paper Series from RePEc, 2022,3 (2-1): 282-294.

[2] Rad, D., Cuc, L. D., Croitoru, G., et al. Modeling Investment Decisions Through Decision Tree Regression—A Behavioural Finance Theory Approach. Electronics (Basel), 2025, 14 (8): 1505.

[3] Arrieta A., Díaz-Rodríguez N, Se J, et al. Explainable Artificial Intelligence (XAI): Concepts, Taxonomies, Opportunities and Challenges toward Responsible AI. Information Fusion, 2020, 58: 82–115.

[4] Anshari M., Almunawar M. N., Masri M. Digital twin: Financial Technology's next frontier of robo-advisor. Journal of Risk and Financial Management, 2022, 15 (4): 163.

[5] Kumar I. Evolution of financial advisory market with the advent of robo-advisors. Managerial Finance, 2025, 51 (5): 831–856.

[6] Dreyer, S., Egger, A., Püschel, L., Röglinger, M. Prioritising smart factory investments – A project portfolio selection approach. International Journal of Production Research, 2022, 60 (3): 999–1015.

[7] Anshari M., Almunawar M. N., Masri M. Digital twin: Financial Technology's next frontier of robo-advisor. Journal of Risk and Financial Management, 2022, 15 (4): 163.

[8] Anđelinović, M., Škunca, F. Optimizing insurers' investment portfolios: incorporating alternative investments. Zbornik Radova Ekonomskog Fakulteta u Rijeci, 2023, 41 (2): 361–389.

[9] Lee, Y. S., Kim, T., Choi, S., et al. When does AI pay off? AI-adoption intensity, complementary investments, and R&D strategy. Technovation, 2022, 118:102590.

[10] Darskuviene V, Lisauskiene N. Linking the Robo-advisors phenomenon and behavioural biases in investment management: An interdisciplinary literature review and research agenda. Organizations and Markets in Emerging Economies, 2021, 12 (2): 459–477.

[11] Romanko, O., Narayan, A., Kwon, R. H. ChatGPT-Based Investment Portfolio Selection. Operations Research Forum, 2023, 4 (4): 91.