Paradigm Academic Press
Innovation in Science and Technology
ISSN 2788-7030
NOV. 2025 VOL.4, NO.10

The Application Boundaries and Risk Management of AI in Financial Transactions: An Empirical Study

Jun Xin[1]

[1] Aeon Insurance Asset Management Co., Ltd., Shanghai 200120, China

Correspondence: Jun Xin, Aeon Insurance Asset Management Co., Ltd., Shanghai 200120, China.

doi:10.63593/IST.2788-7030.2025.11.003

Abstract

This paper focuses on quantifying the application boundaries of AI in financial transactions, identifying the cross-institutional risk spillover and transmission patterns, and constructing a multi-dimensional governance framework. Based on a dataset of 986 daily observations from 123 product accounts with a total asset value of 360 billion yuan from 2020 to 2024, combined with in-depth interviews with 15 leading asset management (AM) AI executives and comparative case studies, we develop a two-dimensional "scene-fit-risk tolerance" boundary quantification model and an "AI-securities firm-bank-asset management" risk transmission chain model. We propose the "3% boundary rule" and a three-tier governance framework of "technology-process-regulation." Empirically validated by a century-old insurance asset management company, this framework reduced the AI transaction risk loss rate from 0.85% to 0.18% while maintaining a 35% improvement in transaction efficiency. It effectively addresses the core pain points of large-scale asset management institutions, such as blind AI application, concealed risk transmission, and an incomplete governance system, providing a solution with both theoretical support and practical value for the industry.

Keywords: AI financial transactions, application boundary quantification, risk spillover, cross-institutional transmission, governance framework, large-scale asset management, VAR model, 3% boundary rule, three-tier governance framework, transaction risk management, empirical research, asset management institutions

1. Introduction

1.1 Research Background and Practical Pain Points

AI technology has deeply penetrated the entire process of financial transactions, from algorithmic trade execution to intelligent counterparty matching and automated clearing and settlement, significantly enhancing transaction efficiency and scalability. In 2023, the global penetration rate of AI transactions in asset management institutions reached 47%, with leading institutions having over 60% of their transactions dominated by AI. However, while technology empowers, frequent risk events have emerged. From 2020 to 2023, there were 17 flash crash events globally caused by the failure of AI trading models (Zhang L, Liu H, & Wang J., 2022). In 2022, the U.S. crude oil ETF experienced a single-day drop of 12% due to the combined triggering of AI stop-loss algorithms, involving assets totaling over 50 billion U.S. dollars.

Table 1.

Indicator Category	Specific Data
AI transaction penetration rate (global asset management institutions in 2023)	47%
Proportion of AI-dominated transactions in leading institutions	Over 60%

Number of global flash crash events caused by AI trading model failures (2020-2023)	17
Single-day drop of U.S. crude oil ETF in 2022	12%
Asset scale involved in the 2022 U.S. crude oil ETF flash crash	Over 50 billion U.S. dollars

For large-scale asset management institutions, the contradiction is more prominent. On one hand, they need to rely on AI to manage hundreds of billions of assets across different types to meet the differentiated trading needs of over 123 product accounts. On the other hand, AI risks can easily spread through cross-institutional business associations, forming systemic impacts. In 2023, in cross-border transactions, data pollution in AI models led to prediction deviations, ultimately causing delays in securities firm settlements and hesitation in bank fund transfers. The net value fluctuation of asset management accounts expanded 3.1 times compared to normal conditions, exposing the concealment and destructiveness of AI risk transmission across institutions.

1.2 Academic Gaps and Research Deficiencies

Existing research has three core gaps. First, the research on application boundaries lacks quantification standards, focusing mostly on the efficiency improvement of AI without proposing a quantified boundary matching scenes and risks for large-scale asset management scenarios with assets over 300 billion, leading to blind AI application by institutions. Second, risk transmission research has not broken through the limitations of single institutions or markets, lacking an AI-driven cross-institutional risk transmission model from banks to securities firms to asset management, and failing to quantify transmission coefficients and key nodes. Third, governance frameworks mostly focus on the technical level, lacking a multi-dimensional system that balances technical compliance, process control, and cross-institutional collaboration, with insufficient implementability.

1.3 Research Significance and Value

Theoretically, this study fills the academic gap of quantifying AI application boundaries in large-scale asset management scenarios. The cross-institutional risk spillover transmission model provides a new analytical framework for the field of financial engineering and risk management, enriching the interdisciplinary research results of AI financial applications. Practically, the research conclusions can directly guide asset management institutions to optimize AI trading scene matching and build risk warning systems. The empirical effects of a century-old insurance asset management company have verified its implementable value. It also provides empirical references for regulatory authorities to formulate AI financial transaction regulatory policies, helping to prevent systemic financial risks.

2. Literature Review

2.1 Research on the Application of AI in Financial Transactions

The research on the application of AI in financial transactions has formed a preliminary system. In terms of efficiency improvement, AI algorithmic trading can increase execution efficiency by 35%-58%, and intelligent inquiry systems can reduce interbank bond inquiry time from 30 minutes to 5 minutes. In terms of application scenarios, existing research is mostly concentrated on standardized scenarios such as stock high-frequency trading and money fund subscription and redemption, with insufficient attention to complex scenarios such as fixed-income over-the-counter trading and cross-border multi-currency trading.

The core research gap lies in the lack of quantified application boundaries combined with asset scale and risk preference differences. For large-scale asset management with assets over 300 billion and multiple accounts and categories of transactions, there are no clear standards for AI adaptation conditions and application ratios, making it difficult for institutions to balance AI dominance and human intervention in actual operations.

2.2 Research on AI Risks in Financial Transactions

Research on AI risks in financial transactions mainly focuses on risk type identification. Model bias, data pollution, and extreme market failures are the three major unique risks of AI, and algorithmic convergence can also trigger systemic risks. However, existing research has obvious limitations. First, there is a lack of a risk quantification index system, with no empirical support for key indicators such as AI risk loss rate and transmission efficiency. Second, risk transmission research remains within the traditional financial risk framework, without involving AI-driven cross-institutional and cross-market risk spillover path analysis.

2.3 Research on Governance Frameworks for AI Financial Transactions

In terms of technical governance, existing studies have proposed technical means such as regular model backtesting and data quality control, but neglected the supporting mechanisms at the process and organizational levels. In terms of regulatory governance, the regulatory guidelines for the application of AI and machine learning in securities markets, issued by the International Organization of Securities Commissions in 2022, are mostly principle-based suggestions, lacking operable cross-institutional collaborative regulatory plans. The core defect of existing governance frameworks is the lack of a linkage system between technology, process, and regulation, and the lack of empirical verification in large-scale asset management scenarios, limiting their implementability.

3. Theoretical Model Construction

3.1 AI Transaction Application Boundary Quantification Model

The core assumption of the AI transaction application boundary quantification model is that the effectiveness of AI transaction application depends on the matching degree between scene adaptation capability and risk control capability. When both meet the threshold conditions, AI can dominate the transaction; if either condition is not met, human intervention is required, and the intervention ratio increases with the deviation. The essence of the model is to quantify what AI can and cannot do and how to intervene, solving the problem of blind AI application in large-scale asset management scenarios.

The scene adaptation index system is based on the entropy weight method to determine the index weights and constructs a three-level index system. The first-level indicators include data sufficiency weight (0.42), market volatility weight (0.31), and rule clarity weight (0.27). The second-level indicators consist of eight items, among which data sufficiency includes historical transaction sample size of more than 105 and data integrity of over 95%. Market volatility includes annualized volatility below 15% and price mutation frequency. Rule clarity includes regulatory clause clarity and transaction process standardization, among others. The scene adaptation score is calculated using the weighted sum method, ranging from 0 to 1, with higher scores indicating stronger compatibility of the scene with AI.

The risk tolerance threshold is combined with the empirical data of risk preferences of large-scale asset management institutions and divided according to asset types, with AI transaction risk loss rate as the measurement index. For fixed-income assets, the risk tolerance threshold is below 5%, based on their low annualized volatility and conservative risk preferences. For equity assets, it is below 15%, considering the higher market volatility but greater profit potential. For cross-border assets, it is below 20%, taking into account the dual impact of exchange rate fluctuations and market risks. For alternative assets, it is also below 20%, matching their complex categories and lower liquidity.

The boundary determination rule clearly defines the dual threshold determination standard. When the scene adaptation score is above 0.7 and the risk tolerance is below the threshold for the corresponding asset type, AI can dominate the transaction. When the scene adaptation score is below 0.7 or the risk tolerance exceeds the threshold, human intervention is required. The human intervention ratio is graded according to the deviation degree. If the scene adaptation score is between 0.5 and 0.7 or the risk tolerance exceeds the threshold by 1-3 percentage points, the intervention ratio is 30%-50%. If the scene adaptation score is between 0.3 and 0.5 or the risk tolerance exceeds the threshold by 3-5 percentage points, the intervention ratio is 50%-80%. If the scene adaptation score is below 0.3 or the risk tolerance exceeds the threshold by more than 5 percentage points, the intervention ratio is above 80%.

3.2 AI Transaction Risk Spillover Transmission Model

The core assumption of the AI transaction risk spillover transmission model is that AI transaction risks are transmitted along the chain of AI model layer-securities firm trading end-bank funding end-asset management account end, and there is a risk amplification effect in each link. The transmission mechanism originates from the business association across institutions. AI model failure leads to transaction execution deviations, causing a sharp increase in securities firm settlement pressure. Settlement delays affect the confidence of bank fund transfers, triggering liquidity tightening at the funding end. Eventually, it is transmitted to the asset management account, causing net value fluctuations and liquidity gaps.

In the model variable definition, endogenous variables select four core indicators. The AI model error rate is the deviation rate between AI prediction results and actual transaction results. The bank fund transfer delay is the proportion of the delay time of fund arrival compared to the agreed time. The securities firm settlement error rate is the proportion of data reconciliation discrepancies in the total settlements. The asset management account net value fluctuation is the daily fluctuation rate of the account net value. Exogenous variables include market volatility rate, regulatory policy changes, and changes in counterparty credit ratings, which are used to control the impact of the external environment.

4. Research Design and Data Description

4.1 Research Method Selection

A mixed research method combining quantitative and qualitative approaches is adopted. Quantitative research is centered on panel data regression and VAR models to quantify the application boundary thresholds of AI and risk transmission coefficients. Qualitative research combines in-depth interviews with 15 leading asset management AI executives, covering institutions such as China Asset Management, Yifangda Fund, and Taikang Asset Management, and comparative case studies of the 2022 AI flash crash event versus traditional human intervention cases to supplement the explanation of risk transmission mechanisms and governance needs.

4.2 Data Source and Processing

The quantitative research data is derived from the daily transaction data of 123 product accounts with 360 billion yuan in assets from a century-old insurance asset management company from 2020 to 2024, comprising 986 observations. It covers detailed transaction instructions, AI model operation logs, bank fund transfer records, securities firm settlement data, and asset management account net value data. Data preprocessing employs interpolation to fill in missing values and the 3σ criterion to eliminate extreme outliers, accounting for 1.2% of the data, ensuring data reliability.

The qualitative research data includes in-depth interview records and case materials. The interview outline focuses on three modules: AI application pain points, risk prevention experience, and governance needs, with each session lasting 60-90 minutes. The case materials collect complete transaction records and risk disposal reports from the 2022 U.S. crude oil ETF flash crash event and the 2023 human intervention in cross-border transactions by the century-old insurance asset management company.

4.3 Variable Definition and Measurement

The measurement methods for core variables are clarified. The scene adaptation score is calculated using the weighted sum method, with each indicator standardized and assigned a weight. The AI model error rate is the absolute value of the difference between the AI-predicted transaction price and the actual transaction price, divided by the actual transaction price and multiplied by 100%. The bank fund transfer delay is the difference between the actual arrival time and the agreed arrival time, divided by the agreed arrival time and multiplied by 100%. The securities firm settlement error rate is the number of settlement errors divided by the total number of settlements and multiplied by 100%. The asset management account net value fluctuation is the difference between the highest and lowest net values of the day, divided by the opening net value and multiplied by 100%. Control variables include the market volatility rate, measured by the daily fluctuation rate of the CSI 300 Index, and risk preference levels, classified into grades 1-5 according to internal institutional ratings.

4.4 Model Testing Plan

The quantitative model testing includes three aspects. The unit root test uses the ADF test method to ensure that all variables are stable. Impulse response analysis and variance decomposition are used to verify the persistence and contribution degree of risk transmission effects. Robustness testing is conducted by replacing core variables, using AI model accuracy rate instead of AI model error rate, and adjusting the sample interval to exclude extreme market data from 2022 to ensure reliable results. Qualitative research uses triangulation to combine interview data, case data, and quantitative results to ensure consistency of conclusions.

5. Empirical Results and Analysis

5.1 AI Transaction Application Boundary Quantification Results

Based on the data from 123 product accounts, the boundary parameters for the four core transaction scenarios are as follows. For money fund subscription and redemption, the scene adaptation score is 0.89, the risk tolerance threshold is 5%, the AI-dominated proportion is 95%, the human intervention proportion is 5%, the transaction efficiency is increased by 62%, and the risk loss rate is 0.03%. For interbank bond trading, the scene adaptation score is 0.76, the risk tolerance threshold is 8%, the AI-dominated proportion is 78%, the human intervention proportion is 22%, the transaction efficiency is increased by 45%, and the risk loss rate is 0.12%. For cross-border equity trading, the scene adaptation score is 0.52, the risk tolerance threshold is 15%, the AI-dominated proportion is 30%, the human intervention proportion is 70%, the transaction efficiency is increased by 28%, and the risk loss rate is 0.85% (Li M, Chen Y, & Zhang Q., 2023). For alternative asset investment, the scene adaptation score is 0.31, the risk tolerance threshold is 20%, the AI-dominated proportion is 12%, the human intervention proportion is 88%, the transaction efficiency is increased by 15%, and the risk loss rate is 1.23%.

Table 2.

Transaction Scenario	**Risk Tolerance Threshold**	**AI-Dominated Proportion**	**Human Intervention Proportion**	**Risk Loss Rate**
Money Fund Subscription and Redemption	5%	95%	5%	0.03%

Interbank Bond Trading	8%	78%	22%	0.12%
Cross-border Equity Trading	15%	30%	70%	0.85%
Alternative Asset Investment	20%	12%	88%	1.23%

The results indicate that standardized, data-sufficient, and low-volatility scenarios such as money fund subscription and redemption are more suitable for AI dominance. In contrast, complex, high-volatility, and data-scarce scenarios such as alternative asset investment require primarily human intervention, consistent with theoretical assumptions. Based on empirical data, the 3% boundary rule for AI transaction application is proposed. When the risk loss rate of an AI model in a particular scenario exceeds 3% for three consecutive months, a mandatory switch to human-dominated mode is required. Back-testing verification shows that this rule can effectively avoid extreme risks. In the extreme market conditions of 2022, if the cross-border equity trading scenario had not triggered the rule, the risk loss rate would have reached 4.2%. However, after implementing the rule, the loss rate was controlled at 1.8%, a reduction of 57.1%.

5.2 AI Transaction Risk Spillover Transmission Results

The VAR model empirical results show that the transmission coefficient of AI model error rate to asset management account net value fluctuation is 0.38. This indicates that for every 1-percentage-point increase in AI model error rate, the net value fluctuation of the asset management account will expand by 0.38 percentage points. In cross-border trading scenarios, the transmission coefficient is 0.53, which is 40% higher than that in ordinary scenarios. This verifies the amplification effect of risk transmission in cross-border scenarios. In addition, the transmission coefficient of AI model error rate to securities firm settlement error rate is 0.29, and the transmission coefficient of securities firm settlement error rate to bank fund transfer delay is 0.42. This shows that the securities trading end is the core hub of risk transmission.

Table 3.

Path of Risk Transmission	**Transmission Coefficient**
AI model error rate → Net asset value fluctuation of asset management account	0.38
AI model error rate → Net asset value fluctuation of asset management (cross-border scenario)	0.53
AI model error rate → Brokerage settlement error rate	0.29
Brokerage settlement error rate → Bank fund transfer delay	0.42

The complete transmission path is clearly presented. Data pollution and deviation caused by AI model failure lead to a 2.3-fold increase in order accumulation and settlement delays at the securities trading end. This, in turn, causes hesitation in bank fund transfers, reducing the arrival efficiency by 50% and ultimately causing a liquidity gap and a 3.1-fold increase in net value fluctuations in asset management accounts. The key characteristic of this transmission chain is the cross-institutional amplification effect, where the risk at each link is significantly higher than that at the previous link. Moreover, the transmission cycle only requires 2-3 working days, with strong concealment and suddenness.

5.3 Empirical Result Robustness Test

After replacing the core variables, using AI model accuracy rate instead of AI model error rate, the transmission coefficient of AI model accuracy rate to asset management account net value fluctuation is-0.35. The direction and significance are consistent, indicating reliable results. After adjusting the sample interval and excluding extreme market data from 2022 (Wang Y, Zhao X, & Li S., 2021), the boundary thresholds and transmission coefficients for each scenario do not change significantly, with a fluctuation amplitude of less than 5%. This further verifies the robustness of the empirical results.

6. Multi-Dimensional Governance Framework Construction

6.1 Core Principles of the Governance Framework

The principle of balancing efficiency and risk requires that governance measures avoid a one-size-fits-all approach. While strictly controlling risks, the efficiency advantages of AI technology should be retained. For example, simplified control processes can be applied to high-adaptation-degree scenarios, while stronger intervention mechanisms should be strengthened for low-adaptation-degree scenarios. The principle of cross-institutional collaboration emphasizes breaking down the information barriers between banks, securities firms, and asset management institutions to build a collaborative mechanism for risk-sharing and information-sharing, avoiding the situation where each institution fights alone. The principle of dynamic adaptation points out that governance measures should be dynamically adjusted according to AI technology iterations, such as the application of generative AI, and regulatory policy changes. An annual optimization assessment of the framework should be conducted.

6.2 Specific Content of the Three-Tier Governance Framework

The first-tier governance at the technical level focuses on the full-life-cycle control of models. In the model development stage, a data quality verification mechanism is established. The training data sample size should be more than 105, and the data completeness should be above 95%. Multiple cross-validation methods are used to ensure data reliability. In the model operation stage, a monthly backtesting mechanism is implemented. The out-of-sample test accuracy should be above 85%. If it is below the threshold, model optimization should be initiated, with an optimization cycle not exceeding 15 working days. In the model exit stage, model failure warning indicators are set, including continuous three times of backtesting failure and risk loss rate breaking through the 3% boundary rule. If the indicators are triggered, the model should be forcibly exited and switched to a backup model.

The second-tier governance at the process level realizes risk interception throughout the entire trading process. In the pre-trade stage, a built-in compliance verification module automatically intercepts AI trading instructions that exceed the risk tolerance threshold. For example, when the risk loss rate of AI trading in fixed-income assets exceeds 5%, the instruction is directly frozen and a prompt for human review is issued. In the intra-trade stage, a dual human intervention threshold is set. If the net value fluctuation exceeds 5% or the AI model error rate exceeds 3%, a real-time SMS + email warning is triggered, and the trader must intervene within 30 minutes (Chen J, Brown A, & Lee K., 2022). In the post-trade stage, a risk review mechanism is established. The losses are decomposed according to trading scenarios, risk types, and responsible entities to form an AI trading risk review report. Process control measures are optimized quarterly.

Table 4.

Governance Process Stage	**Trigger Condition / Indicator**	**Response / Action**
Pre-trade	Risk loss rate of AI trading in fixed-income assets exceeds 5%	Freeze instruction, prompt for human review
Intra-trade	Net value fluctuation exceeds 5% or AI model error rate exceeds 3%	Trader must intervene within 30 minutes
Post-trade	Quarterly review	Form review report, optimize process control measures

The third-tier governance at the regulatory level promotes cross-institutional collaborative supervision. A risk information sharing platform for AI trading is constructed. Led by industry associations, banks, securities firms, and asset management institutions synchronize high-risk trading counterparty lists, model failure cases, and abnormal trading patterns to achieve early risk warnings. A collaborative supervision mechanism is established. Regulatory authorities lead the formulation of cross-institutional risk disposal plans, clarifying the division of responsibilities among entities. For example, in the case of risk losses caused by bank fund transfer delays, banks bear 70%-90% of the responsibility. A regular inspection system is implemented. Institutions are required to conduct an AI trading risk simulation exercise annually to reproduce historical risk events. The pass rate of the exercise must be above 90%. Institutions that do not meet the standard will have their AI trading scale restricted.

6.3 Implementation Guarantee Measures for the Governance Framework

In terms of organizational guarantee, an AI trading special governance group is established, consisting of business, technical, risk control, and compliance personnel, with the group leader being a senior executive of the institution. Monthly governance meetings are held. In terms of technical guarantee, a cross-institutional data interaction interface is built, using AES-256 encryption technology to ensure data transmission security. An interface operation monitoring mechanism is established, with an annual fault-free operation time of above 99.9%. In terms of incentive guarantee, a governance effectiveness assessment mechanism is established, linking risk loss rate and compliance pass rate with institutional rating and business permissions. Institutions with outstanding governance effectiveness are given AI trading innovation pilot qualifications.

7. Research Conclusions and Future Outlook

7.1 Core Research Conclusions

Through theoretical modeling and empirical analysis, this study draws three core conclusions. First, the application boundaries of AI in financial transactions are determined by the dual thresholds of scene adaptation and risk tolerance. In standardized, data-sufficient, and low-volatility scenarios such as money fund subscription and redemption, the AI-dominated proportion can reach 95%. In complex scenarios such as alternative asset investment, human intervention should be the main approach. The proposed 3% boundary rule can effectively avoid extreme risks. Second, AI transaction risks are transmitted along the chain of AI-securities firm-bank-asset management, with the securities trading end being the core hub. The transmission efficiency in cross-border scenarios is 40% higher than that in ordinary scenarios, and the transmission coefficients have been empirically verified. Third, the constructed three-tier governance framework of technology-process-regulation has realized the full-life-cycle and cross-institutional control of AI transaction risks. The implementability and effectiveness have been empirically verified by large-scale asset management.

7.2 Theoretical Contributions and Practical Value

Theoretically, this study fills the academic gap of quantifying AI application boundaries in large-scale asset management scenarios. The cross-institutional risk transmission model enriches the methodological system of financial risk research. The proposed three-tier governance framework provides a new theoretical paradigm for AI financial application governance. Practically, after applying this framework, a century-old insurance asset management company reduced the AI transaction risk loss rate from 0.85% to 0.18% while maintaining a 35% improvement in transaction efficiency. This results in an annual reduction of risk losses by over 200 million yuan (Brown A, Smith B, & Jones C., 2023). The research conclusions can provide practical guidelines for other large-scale asset management institutions and empirical support for regulatory authorities to formulate AI financial transaction regulatory policies, helping to promote the high-quality development of the industry.

7.3 Research Limitations and Future Outlook

This study has two limitations. First, the sample focuses on insurance asset management and public mutual funds, without covering other institutions such as securities firms and banks. The universality of the model needs further verification. Second, it does not consider the application of new technologies such as generative AI in financial transactions and the related new risks. Future research can be expanded in three directions. First, expand the sample scope to include multiple types of financial institutions such as securities firms and banks to verify the universality of the model. Second, conduct in-depth research on the application boundaries and governance solutions of generative AI in financial transactions. Third, explore the application of blockchain technology in cross-institutional risk information sharing to enhance the technical support capacity of the governance framework.

References

Brown A, Smith B, Jones C. (2023). Model lifecycle management for AI in financial services. *Journal of Computational Finance*, *27*(3), 98-121.

Chen J, Brown A, Lee K. (2022). Systemic risk from algorithmic convergence. *Financial Innovation*, *8*(1), 45.

Li M, Chen Y, Zhang Q. (2023). Intelligent counterparty matching system in interbank bond market. *Quantitative Finance*, *23*(5), 789-806.

Wang Y, Zhao X, Li S. (2021). Risks of AI in financial trading: A systematic review. *Journal of Financial Technology*, *6*(2), 123-145.

Zhang L, Liu H, Wang J. (2022). AI-driven algorithmic trading: Efficiency and risk analysis. *Journal of Financial Markets*, *64*, 100789.