

Multi-Source Data Fusion-Driven Cross-Border Fund Flow Behavior Identification and Risk Governance

Manlu Xu

Fortune Max International Asset Management Limited, Hong Kong, China

Abstract: With the rapid expansion of Chinese enterprises' overseas investment and cross-border financing, cross-border capital movements have become increasingly complex, involving multiple entities, jurisdictions, transaction structures and financial channels. Although China has established a relatively comprehensive regulatory framework covering Outward Direct Investment (ODI) management, foreign exchange supervision, customer due diligence and Anti-Money Laundering (AML) monitoring, fragmented multi-source data continue to constrain integrated fund behavior analysis and dynamic risk identification. From a data-driven risk management perspective, this study proposes a multi-source data fusion-driven framework for cross-border fund behavior identification. By integrating entity information, transaction records, ownership relationships, fund flow paths, relationship networks, and external risk factors, the framework constructs dynamic fund behavior profiles based on the analytical dimensions of entity–relationship–fund flow–time and identifies potential deviations between actual fund activities and expected business purposes. The study extends cross-border financial risk research from traditional entity-based assessment to fund behavior analysis, providing a data-driven analytical framework for continuous monitoring, risk early warning, and digital risk governance in cross-border investment and financing activities.

Keywords: cross-border investment and financing; outward direct investment (ODI); financial risk governance; digital risk management

1. Introduction

With the deepening of economic globalization and the increasing internationalization of Chinese enterprises, cross-border investment and financing activities have become an important channel for global resource allocation. Meanwhile, the associated capital movements have become increasingly complex, involving multiple entities, jurisdictions, transaction structures, and financial channels. Investment models have gradually evolved from traditional equity investment to cross-border mergers and acquisitions, overseas industrial deployment, cross-border funds investment and diversified financing arrangements, making fund flows more interconnected and dynamic. As a result, cross-border financial risks have shifted from isolated transaction compliance toward behavioral risks embedded in cross-border fund movements.[1]

China has established a relatively comprehensive regulatory framework covering ODI, foreign exchange administration, customer due diligence and Anti-Money Laundering (AML) supervision. Financial institutions routinely implement Know Your Customer (KYC), Ultimate Beneficial Owner (UBO) identification, ongoing due diligence and transaction monitoring to support regulatory compliance. However, increasing transaction complexity has exposed the limitations of fragmented risk management. Data related to entity information, transactions, ownership structures, regulatory filings, and suspicious activities remain dispersed across different institutions and systems, limiting the ability to construct integrated fund behavior profiles and identify abnormal fund behavior patterns. Therefore, the primary challenge is no longer the availability of regulatory data, but the effective integration of multi-source heterogeneous information for behavioral analysis and dynamic risk governance.

To address this challenge, this study investigates how multi-source data fusion can improve cross-border fund behavior identification and strengthen risk governance. Specifically, a data-driven analytical framework integrating identity information, transaction records, relationship networks, and external risk factors is proposed. The framework constructs unified fund behavior profiles, identifies abnormal fund behavior patterns, and supports continuous monitoring and dynamic risk governance throughout the investment lifecycle.

This study makes three contributions. First, it extends cross-border financial risk research from traditional entity-based assessment to fund behavior analysis, emphasizing the importance of capital flow paths, transaction patterns and relationship networks in risk identification. Second, it proposes a multi-source data fusion framework that integrates heterogeneous information to enhance behavioral analysis and risk identification under complex cross-border investment scenarios. Third, it provides practical implications for financial institutions, regulatory authorities, and multinational enterprises by supporting

data-driven risk monitoring, compliance management, and dynamic governance of cross-border fund behavior.

2. Literature Review and Research Gaps

2.1 Research on Cross-Border Investment and Financing Risks

Existing studies on cross-border investment and financing risks mainly focus on investment environments, institutional differences, and external uncertainties affecting overseas investment decisions and performance. Buckley et al. (2007), from the perspective of internationalization theory, argued that firms need to consider host-country institutional differences, market conditions, and risk factors when conducting overseas investment[2]. Kolstad and Wiig (2012) further emphasized the importance of political institutions and governance quality in influencing firms' ODI performance[3]. In addition, exchange rate volatility, policy changes, and compliance requirements are widely recognized as important sources of cross-border investment risks.

However, existing studies mainly examine risks related to investment environments and project performance, while paying limited attention to behavioral changes during cross-border fund movements. Traditional risk assessment frameworks often struggle to identify potential anomalies in fund disbursement, changes in fund utilization, related-party transactions, and capital repatriation. Therefore, a risk identification mechanism from the perspective of fund behavior analysis is needed.

2.2 Research on Anti-Money Laundering (AML) and Financial Risk Identification

The AML-related research has long focused on abnormal transaction identification, developing risk detection approaches based on rule-based models, statistical analysis, and machine learning algorithms. Chen et al. (2018) systematically reviewed the application of machine learning in suspicious transaction identification and found that transaction feature extraction can improve traditional rule-based models[4]. Jullum et al. (2020) further applied machine learning methods to transaction data for automated detection of potential money laundering activities. In recent years, graph network analysis has been introduced to capture relationships among transaction entities[5]. Savage et al. (2017) combined network structure analysis with supervised learning to identify money laundering networks, demonstrating that transaction relationships can complement traditional account-level analysis[6]. Colladon et al. (2017) further verified the effectiveness of social network analysis in anti-money laundering risk prevention[7]. Deprez et al. (2025) further showed that network analytics can improve AML models' ability to identify complex fund relationships[8].

Nevertheless, existing AML research mainly focuses on banking transactions and payment scenarios, with limited attention to the full-process analysis of investment entities, overseas enterprises, fund flow paths, and transaction behaviors in cross-border investment and financing activities. Due to their long investment cycles, multiple participants, and complex capital chains, cross-border investment activities cannot be fully assessed through transaction rules or bank records alone.

2.3 Research on Data Fusion and Intelligent Risk Management

With the development of artificial intelligence and big data technologies, multi-source data fusion has become an important approach in financial risk management[9]. Existing studies demonstrate that integrating transaction records, entity information, relationship networks, and external environmental data can improve risk identification and dynamic monitoring capabilities. However, current research mainly focuses on credit risk, market risk, and banking compliance management, while multi-dimensional data fusion mechanisms for cross-border fund behavior analysis remain underexplored.

In summary, existing studies provide important foundations for financial risk identification but have not yet established a fund behavior analytical framework suitable for cross-border investment and financing scenarios. This study therefore proposes a multi-source data fusion-driven mechanism for fund behavior profiling and anomaly identification throughout the investment lifecycle, extending risk management from traditional project-level assessment toward dynamic fund behavior governance.

3. Analysis of Practical Challenges in Cross-Border Fund Flow Behavior Identification

With the expansion of Chinese enterprises' international operations, cross-border investment and financing activities increasingly involve complex entities, fund flows, and transaction structures. China has gradually established a multi-layered regulatory framework covering ODI administration, Ministry of Commerce (MOFCOM) filing, foreign exchange registration, bank transaction authenticity verification, AML, KYC, and UBO identification. Therefore, the key challenge in current cross-border risk governance is no longer the lack of regulatory mechanisms, but how to effectively integrate increasing amounts of heterogeneous data for correlation analysis and dynamic risk identification, thereby shifting risk

governance from traditional compliance review toward behavior-oriented analysis.

3.1 Abundant Multi-Source Data yet Lack of Unified Correlation Mechanisms

Cross-border investment and financing activities involve multiple participants, processes, and markets. Effective risk identification requires the integration of entity information, ownership structures, fund flow paths, transaction relationships, and external risk factors. Currently, China’s regulatory system has accumulated substantial data resources related to cross-border fund behavior, including:

Table 1. Classification and Core Functions of Multi-Source Data for Cross-Border Fund Supervision

Data Type	Core Function
Entity Identity Information	Identify participating parties such as enterprises, individuals and financial institutions, and verify the authenticity of transaction entities.
UBO Information	Trace ultimate control relationships, identify hidden associated entities, and reveal complex ownership structures.
Foreign Exchange Administration Data	Track cross-border capital inflows and outflows, analyze cross-border fund flow chains.
AML Transaction Data	Identify abnormal transaction patterns and uncover potential money laundering and illicit fund risks.
Transaction Relationship Data	Construct entity relationship networks among enterprises, accounts, and institutions, and analyze fund flow chains.
External Risk Data	Support risk assessment by incorporating country risk, industry risk, market fluctuations, and other external factors.

Nevertheless, these datasets are generated across different administrative systems and business processes, resulting in differences in data standards, update frequencies, and application purposes. For example, corporate registration data provide basic entity information but cannot directly reveal fund flow paths; foreign exchange records capture cross-border fund movements but lack sufficient interpretation of ownership relationships and business backgrounds; AML data identify abnormal transaction patterns but mainly focus on individual or local transaction analysis.

Therefore, the main challenge is not data scarcity, but the lack of a unified mechanism for integrating heterogeneous information around “entity identity — control relationships — fund flow paths — transaction networks — external risks”. When risks are embedded across multiple dimensions and transaction stages, individual data sources cannot provide comprehensive risk assessment.

3.2 Traditional Regulatory Models Focus on Node-Level Identification and Fail to Detect Complex Risk Patterns

Current cross-border investment and financing supervision primarily relies on compliance review, focusing on investment qualifications, fund usage, transaction documentation, and transaction authenticity. Such approaches are effective in identifying isolated irregularities, including fraudulent transactions, incomplete documentation, and unqualified entities.

However, as cross-border transaction structures become increasingly complex, fund risks have gradually shifted from isolated violations to combined risks formed through interactions among multiple entities and transaction links. For example, a single capital flow may involve the following chain:



Figure 1. Typical Circulation Chain of Complex Cross-Border Capital Risk Under Indirect Overseas Investment Structure

Although each individual entity or transaction may appear compliant, including completed overseas investment filings, valid overseas entity registrations, and supporting transaction documents, holistic analysis of the fund network may reveal potential risks, such as fund flows inconsistent with business operations, frequent related-party transactions, and abnormal circulation patterns.

Therefore, supervision based on individual entities and transactions has limitations in identifying complex fund behaviors across multiple entities, jurisdictions and accounts. Future risk governance should move from assessing whether individual transactions comply with regulations toward evaluating whether overall fund behavior is consistent with commercial logic, identifying potential risks through behavioral patterns and relationship networks.

3.3 Lack of Dynamic Profiling Mechanisms Tailored to Fund Flow Behaviors

Existing cross-border investment and financing risk management mainly relies on static entity-based assessments, including corporate profiles, investment scale, ownership structures, transaction documents and historical compliance records. These factors describe what an enterprise is, but cannot continuously capture how fund behaviors evolve over time.

In practice, cross-border fund risks have strong dynamic characteristics. Normal business-related fund flows usually follow relatively stable commercial patterns, while abnormal behaviors are reflected in changes in fund flow paths, transaction frequency, relationship networks and deviations between actual fund usage and declared purposes. For example, enterprises with limited history of physical investment may suddenly conduct large cross-border transfers without reasonable commercial explanations; newly established overseas subsidiaries may engage in frequent large-value transactions; or fund flows may significantly deviate from business scope and investment objectives.

Such risks are difficult to identify through one-time compliance reviews. Therefore, cross-border risk governance requires a transition from static entity screening to dynamic behavior identification. By constructing dynamic fund behavior profiles through multi-source heterogeneous data fusion, risk management can continuously analyze how funds circulate, why such patterns occur, whether behaviors conform to historical patterns and whether they deviate from genuine commercial purposes.

4. Construction of the Cross-Border Fund Flow Behavior Profile and Risk Identification Model

To address the challenges of fragmented data sources, insufficient information correlation and limited dynamic analysis capability in cross-border fund behavior identification, this paper proposes a Multi-Source Data Fusion-based Cross-Border Fund Behavior Identification Framework (CFBIF). The framework does not replace existing regulatory mechanisms for cross-border investment and financing, but provides a data-driven enhancement approach based on existing overseas investment administration, foreign exchange supervision, bank transaction verification, and AML mechanisms. By integrating entity information, control relationships, fund flow paths, transaction behaviors and external risk factors, the framework constructs dynamic fund behavior profiles and promotes the transition of cross-border risk management from static compliance review toward behavior-oriented identification and risk governance.

The framework consists of four interconnected layers: data fusion, behavior profiling, risk identification, and governance application. The data fusion layer integrates heterogeneous information from different sources; the behavior profiling layer characterizes corporate cross-border fund activity patterns and their dynamic changes; the risk identification layer detects potential risks through behavior deviation analysis; and the governance application layer supports continuous monitoring, risk warning and due diligence optimization. The framework aims to improve risk identification capability through enhanced data utilization while maintaining the existing regulatory framework.

4.1 Multi-Source Data Fusion Mechanism: Construction of a Four-Dimensional Analytical Framework (Entity–Relationship–Fund Flow–Time)

Multi-source data fusion provides the foundation for cross-border fund flow behavior identification. Different from traditional data integration approaches that focus on simple data aggregation, the proposed mechanism establishes logical connections around fund behaviors through four dimensions: entity, relationship, fund flow and time.

The entity dimension focuses on identifying participants and control relationships involved in cross-border investment and financing activities. By integrating corporate information, operational data, and UBO information, the framework extends analysis beyond individual legal entities and identifies potential associated entities and actual control relationships. This helps overcome the limitations of evaluating enterprises solely based on registered identities.

The relationship dimension analyzes connections among different entities. Cross-border fund risks are often embedded in complex relationship networks rather than individual transactions. By incorporating ownership structures, control relationships, historical transactions, and business connections, the framework constructs relationship networks and identifies hidden associations during capital circulation.

The fund flow dimension focuses on the movement and utilization of capital throughout the transaction process. By

integrating cross-border payment records, account transaction information, and transaction relationships, the framework analyzes capital sources, transfer paths, transaction nodes, and final usage, thereby evaluating whether fund movements are consistent with declared investment purposes and commercial logic.

The time dimension captures dynamic changes in fund behavior patterns. Since normal corporate capital activities usually follow relatively stable patterns, deviations from historical behaviors may indicate potential risks. Through comparison between current activities and historical patterns, the framework identifies abnormal changes in fund flow behaviors.

Through the above four-dimensional mechanism, fragmented cross-border fund data can be transformed into interconnected behavioral networks, providing the basis for subsequent behavior profiling and risk identification.

4.2 Construction of Cross-Border Fund Flow Behavior Profiles

Based on multi-source data fusion, this paper constructs cross-border fund behavior profiles to characterize enterprises' investment and financing activities. Unlike traditional risk assessment approaches that mainly focus on static entity information, fund behavior profiles emphasize dynamic patterns of capital movement and changes in business relationships throughout the cross-border investment lifecycle.

The entity profile describes fundamental characteristics of enterprises involved in cross-border activities, including industry background, operational scale, ownership structure, control relationships, and overseas investment arrangements. It provides contextual information for risk assessment and avoids evaluating individual transactions separately from enterprise characteristics.

The fund flow path profile focuses on capital movement trajectories, including fund destinations, transfer routes, intermediary entities and potential deviations from investment purposes. Through analysis of capital circulation paths, the framework evaluates whether fund movements are consistent with genuine commercial activities.

The transaction behavior profile analyzes transaction characteristics, including transaction scale, frequency, timing, and counterparties. Significant deviations from historical transaction patterns or industry characteristics may provide important signals for potential risk identification.

The relationship network profile examines connections among enterprises and related entities through shared controllers, affiliated companies, and historical transaction networks. It helps identify potential risks hidden within complex transaction structures and multi-layer ownership arrangements.

Through these behavior profiles, cross-border fund activities are transformed from fragmented transaction records into dynamic behavioral representations, enabling more comprehensive risk analysis.

4.3 Risk Identification Mechanism Based on Behavior Deviation Analysis

Based on the constructed fund behavior profiles, this paper proposes a risk identification mechanism based on behavior deviation analysis. It should be emphasized that the mechanism does not directly determine whether enterprises violate laws or regulations; rather, it identifies inconsistencies between fund behaviors and reasonable commercial logic to provide analytical support for risk governance.

Cross-border financial risks are often gradual and difficult to detect through isolated transactions. Instead, they may emerge through accumulated changes in fund flow patterns, transaction behaviors, and relationship structures. Therefore, risk identification focuses on detecting deviations between actual fund activities and expected business purposes.

The proposed mechanism integrates entity characteristics, fund flow paths, transaction behaviors, relationship networks, and external risk factors to support dynamic risk assessment. By continuously monitoring changes in fund behaviors, it can assist financial institutions in improving customer due diligence, support enterprises in strengthening overseas investment capital management, and provide data-driven support for regulatory analysis.

In summary, the CFBIF framework extends cross-border financial risk analysis from traditional entity-oriented assessment toward fund behavior-oriented governance. Through the “entity–relationship–fund flow–time” analytical framework, it transforms fragmented information into dynamic behavioral insights. Rather than modifying existing cross-border investment and financing regulatory mechanisms, the framework enhances data utilization and risk identification capabilities within the current institutional environment, providing a new analytical pathway for intelligent cross-border fund risk governance.

5. Research Implications and Future Prospects

Addressing the increasing complexity of cross-border capital flows, fragmented data resources and dynamic risk identification challenges, this study proposes a cross-border fund behavior identification framework based on multi-source data fusion. The study suggests that improving cross-border risk governance depends not on additional regulatory procedures,

but on better utilization of existing data resources. By integrating entity information, control relationships, fund flow paths, transaction behaviors and external risk factors, the framework enables dynamic analysis and identification of cross-border fund behavior risks.

From a practical perspective, the framework supports the transition of cross-border investment and financing risk management from static compliance review toward dynamic behavior-based identification. It can assist financial institutions in enhancing due diligence and risk assessment, help enterprises strengthen capital management during overseas investment activities and provide regulatory authorities with data-driven support for cross-border fund monitoring and risk decision-making.

Future research can be further developed in three directions. First, data integration can be expanded by incorporating corporate operations, international trade, industry development and external risk factors. Second, artificial intelligence, machine learning, and graph computing technologies can be explored for complex fund network analysis. Third, model validation based on real business scenarios can further examine applications in cross-border banking services, corporate risk management, and regulatory technology.

Overall, multi-source data fusion and intelligent analysis provide a new pathway for cross-border investment and financing risk governance, supporting the transition from ex-post risk response toward proactive identification and continuous monitoring.

References

- [1] XU H, MIAO W, ZHANG S. Cross-border capital inflows and systemic financial risks[J]. *International Review of Economics & Finance*, 2025, 99: 104049.
- [2] Buckley P.J., Clegg L.J., Cross A.R., et al. The determinants of Chinese outward foreign direct investment[J]. *Journal of International Business Studies*, 2007.
- [3] Kolstad I., Wiig A. What determines Chinese outward FDI?[J]. *Journal of World Business*, 2012.
- [4] Chen Z., Khoa L.D.V., Teoh E.N., et al. Machine learning techniques for anti-money laundering (AML) solutions in suspicious transaction detection: a review[J]. *Knowledge and Information Systems*, 2018.
- [5] Jullum M., Løland A., Huseby R.B., et al. Detecting money laundering transactions with machine learning[J]. *Journal of Money Laundering Control*, 2020.
- [6] Savage, D., Wang, Q., Chou, P., Zhang, X., & Yu, X. (2017). Detection of money laundering groups: Supervised learning on small networks. In *AAAI Workshop on AI for Financial Services* (pp. 1–6). AAAI Press.
- [7] Colladon, A. F., & Remondi, E. (2017). Using social network analysis to prevent money laundering. *Expert Systems with Applications*, 67, 49–58.
- [8] Deprez B., Vanderschueren T., Baesens B., et al. Network Analytics for Anti-money Laundering—A Systematic Literature Review and Experimental Evaluation[J]. *INFORMS Journal on Data Science*, 2025.
- [9] WANG G, CHEN G, ZHAO H M, et al. Leveraging multisource heterogeneous data for financial risk prediction: A novel hybrid-strategy-based self-adaptive method[J]. *MIS Quarterly*, 2021, 45(4): 1949-1998.