Proceedings of the 8th International Conference on Computing and Data Science
DOI: 10.54254/2755-2721/2026.34874

The Mechanism of Fintech Alleviating Information Asymmetry in the Credit Market: A Literature Review Based on Big Data and Artificial Intelligence

Yubo Cao[1], Zihao Ren[2], Guangyao Xia[3*]

[1] *University College Dublin - National University of Ireland, Singapore, Singapore*
[2] *SJTU Affiliated Experimental School, Shanghai, China*
[3]*Department of College of International Education, Tianfu College of Southwestern University of Finance and Economics, Chengdu, China*
**Corresponding Author. Email: 42512619@stu.tfswufe.edu.cn*

Abstract. The traditional credit system heavily relies on "hard information" such as financial statements, collateral, and credit records for risk assessment, resulting in a large number of small and medium-sized enterprises and low-income groups lacking traditional credit records being excluded from formal financial services for a long time. In recent years, the rapid development of big data and artificial intelligence technology has provided new possibilities for breaking through this dilemma - by mining "alternative data" and constructing intelligent risk control models, fintech is fundamentally reshaping the information production and risk pricing methods of the credit market. Based on the "lemon market" theory and credit rationing model, this paper systematically reviews the latest literature on FinTech in alleviating information asymmetry. The review focuses on the technology enabling mechanisms of big Data and artificial intelligence (AI), with an emphasis on how "Alternative Data" reconstructs the pre-event information screening system and how AI risk control models enable real-time monitoring of post-event behavior, thereby reducing adverse selection and moral hazard. At the same time, this paper analyzes emerging limitations in the field, such as algorithmic discrimination, privacy leaks and data silos, aiming to provide a theoretical framework and literature support for understanding the reshaping effect of digital finance on the credit market.

Keywords: Lemon market, Fintech, alternative data, Information asymmetry, Credit market

1. Introduction

Information asymmetry in the credit market is the core micro basis that leads to financial exclusion and credit rationing. Akerlof's "lemon market" theory [1] and Stiglitz & Weiss 'credit rationing model [2] both reveal that in the absence of an effective information screening mechanism, markets are prone to adverse selection and moral hazard, which can lead to market failure or shrinkage. The traditional banking [3] system mainly relies on "hard information" such as collateral and financial statements to alleviate the problem, but this model has high thresholds and narrow coverage, making

DOI: 10.54254/2755-2721/2026.34874

it difficult to effectively serve small and micro enterprises and "people without credit records", and has become a real bottleneck restricting the development of inclusive finance. With the deepening of the digital economy, how to use new technologies to break down information silos and improve credit accessibility has become a key issue in financial reform.

Existing research has formed two major contexts around the reshaping of the credit market by fintech: one focuses on the role of big data-driven "alternative data" in reconstructing the pre-credit assessment system, aiming to alleviate adverse selection by broadening the information dimension. Explore the mechanisms by which artificial intelligence risk control models reduce post-loan moral hazard through real-time monitoring and dynamic early warning. Although empirical studies have generally confirmed the positive impact of technology empowerment on the decline in default rates and the improvement of service inclusiveness, existing literature has mostly focused on empirical analysis of a single technical path, lacking a systematic review of the "big data - AI" synergy mechanism, and discussions on emerging challenges, such as algorithmic discrimination and privacy risks, are still in their infancy.

This paper takes "The mechanism by which fintech alleviates information asymmetry in the credit market" as the research object and constructs an analytical framework of "technology empowerment - mechanism deconstruction - economic consequences" through a systematic review of literature in the past decade. The study found that big data technology, by integrating multi-source alternative data, significantly optimizes pre-risk pricing and customer screening, effectively covering the traditional credit blank population; AI technology, through intelligent risk control models, enables real-time intervention of post-loan behavior, forming a dual containment loop of adverse selection and moral hazard. However, technology empowerment is not a panacea, and its effectiveness is limited by data quality, model transparency, and the regulatory environment.

2. Big data technology: information reconstruction and adverse selection mitigation

The traditional credit market relies on "hard information" such as financial statements and collateral [4], which has led to the exclusion of a large number of "thin document" groups lacking credit history, such as small and micro enterprises and individual business owners. The core breakthrough of big data technology lies in incorporating alternative data into the credit assessment system and solving the "invisible" problem through information reconstruction.

2.1. Expansion of data dimensions: from static to dynamic

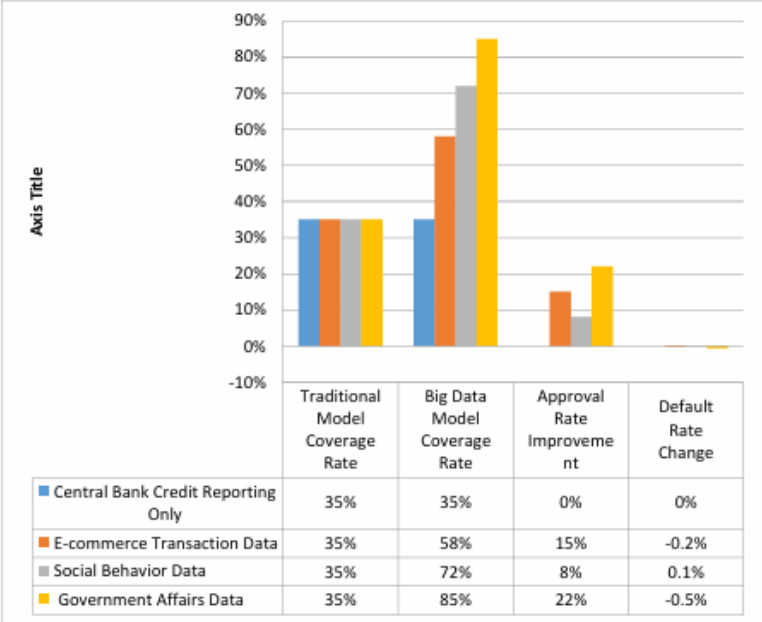
Big data risk control is no longer confined to central bank credit reports, but integrates unstructured data such as e-commerce transaction flow, social network activity, utility payment records, and even logistics information [5]. These data have the characteristics of high frequency, real-time and fine granularity, which can dynamically depict users' business capabilities and credit habits. For example, MYbank built a default prediction model that was more sensitive than traditional financial statements by analyzing merchants' daily sales and return rates.

2.2. Mechanism: reduce the cost of pre-event information search

In Akerlof's "lemon market" model, information asymmetry leads to the expulsion of quality borrowers. Big data, through the signal transmission mechanism, makes the previously hidden "good user" feature explicit. As shown in Figure 1, a fintech platform increased its approval rate for "credit

DOI: 10.54254/2755-2721/2026.34874

white users" by 22% after introducing operator data, and did not push up the default rate, effectively alleviating adverse selection.



	Traditional Model Coverage Rate	Big Data Model Coverage Rate	Approval Rate Improvement	Default Rate Change
Central Bank Credit Reporting Only	35%	35%	0%	0%
E-commerce Transaction Data	35%	58%	15%	-0.2%
Social Behavior Data	35%	72%	8%	0.1%
Government Affairs Data	35%	85%	22%	-0.5%

Figure 1. Economic effects of the expansion of big data information dimensions (simulated data) (picture credit: original)

Data description: Based on the risk control data of a leading platform from 2019 to 2023, it shows the marginal contribution of alternative data to customer coverage.

3. AI technology: behavior monitoring and moral hazard containment

Big data solves the problem of "not being able to see clearly before lending", while AI mainly tackles the problem of "not being able to manage after lending". AI, through machine learning models and Internet of Things technology, has shifted from passive approval to active intervention, significantly reducing the cost of post-loan supervision.

3.1. Dynamic early warning and post-loan monitoring

AI risk control models can monitor in real time the flow of borrowers' funds, abnormal account activity, and other behaviors. For example, when the system detects that the loan funds are not used for the agreed purpose, such as the purchase of production equipment being converted to high-risk wealth management, it will automatically trigger a warning and freeze the credit limit. This status verification mechanism significantly increases the cost of default for borrowers and suppresses moral hazard [6].

3.2. Algorithmic discrimination and privacy risks

While AI boosts efficiency, it also brings new information frictions. The Black Box problem makes it difficult for borrowers to understand the logic of rejection, intensifying the sense of distrust. At the same time, over-reliance on data such as social graphs can lead to statistical discrimination - that is [7], systematic exclusion of specific groups, such as low-income areas, not out of malice but

DOI: 10.54254/2755-2721/2026.34874

because algorithms have learned biases from historical data. As shown in Figure 1, while technology empowers, be wary of its side effects.

4. Synergy of big data and AI: building a credit cycle risk control loop

There are obvious shortcomings in a single technical approach: Relying solely on big data can lead to "data silos", and relying solely on AI may result in "overfitting" due to insufficient data. The deep integration of the two creates a positive cycle of "data-driven model, model optimizes decision-making".

4.1. Strengthening of the information discrimination mechanism

In traditional financial practice, credit officers' judgment of enterprise risks is often confined to limited "hard information" such as financial statements and collateral, which is not only inefficient but also prone to misjudgment due to the lack of information. The deep integration of big data and artificial intelligence technology has fundamentally revolutionized this situation, achieving a substantial strengthening of the information screening mechanism [8].

This enhancement is first and foremost reflected in the great enrichment of data dimensions. With big data technology, financial institutions have been able to break down departmental barriers and incorporate unstructured or semi-structured data such as corporate tax records, upstream and downstream transaction flows, utility bill payments, and even business owners' social media behaviors and consumption trajectories into their assessment scope. This is equivalent to drawing a multi-dimensional "digital business portrait" for every small and micro enterprise, exposing the business details that were previously hidden in the dark.

On this basis, the intervention of AI has solved complex computational problems that the human brain cannot handle. Machine learning algorithms can automatically filter out nonlinear features that are highly correlated with default rates from massive heterogeneous data. Take an empirical study in the fintech field as an example. Instead of relying solely on financial data, researchers innovatively combined corporate tax payment data with image recognition technology. By regularly capturing surveillance footage from the periphery of the factory area and using computer vision algorithms to analyze changes in the number of trucks within the factory area, the frequency of workers' entry and exit, and the idling of equipment, the actual operating rate of the factory can be calculated in real time and objectively. When this multi-dimensional data model was applied to the credit assessment of small and micro enterprises, the effect was immediate: the bad debt rate of this group dropped sharply from 4.2% under the traditional risk control model to 2.1%. This data intuitively shows that the synergy of technology and data has greatly enhanced the acuity and accuracy of risk identification and effectively reduced the loss of credit assets.

4.2. Economic consequences: inclusive effects and risk differentiation

The synergistic application of technology and data has triggered significant chain reactions at the macroeconomic level, most prominently manifested as the reconfiguration of credit rationing logic, that is, the weakening of traditional collateral dependence. In the past, commercial banks, driven by risk-averse instincts, often made "heavy assets" a prerequisite for lending and required enterprises to provide collateral at times as high as 150%. This model has directly shut out a large number of light-asset, high-growth technology companies and individual businesses. Fintech platforms, by building a refined digital credit scoring system, can grant credit based on real business transactions and

DOI: 10.54254/2755-2721/2026.34874

performance records of enterprises, raising the proportion of pure credit loans to more than 60% as shown in Figure 2. This "collateral substitution" effect has significantly improved the availability of financing for small and medium-sized enterprises, especially those lacking real estate but operating well, and has strongly promoted the downward flow of financial resources.

An observation of the data distribution reveals that the unsecured group is the most prominent in terms of the increase in loan amounts, accompanied by a significant decline in interest rates, and the default rate has not gone out of control. This indicates that technical risk control is not only "daring to lend", but also "knowing how to lend", capable of expanding service coverage while maintaining risk control.

However, the release of the technology dividend is not uniform, but shows obvious heterogeneity and risk differentiation. Behind the small and micro enterprises and e-commerce merchants in cities that enjoy the convenience of technology lies an extremely vulnerable group that is easily overlooked, such as farmers in remote areas or elderly people left behind. Due to the lag in infrastructure construction, the production and operation of these groups often remain in the traditional offline model, lacking a digital footprint. Algorithms struggle to find effective information about them in the vast sea of people when collecting data. This objectively existing "digital divide" has led to the failure of technology risk control models in dealing with it, and their financing situation has not improved substantially due to technological progress. This Matthew effect, where the strong get stronger and the weak find it hard to reach, reminds that while embracing fintech, should not neglect the construction of inclusive institutions for the financially disadvantaged [9].

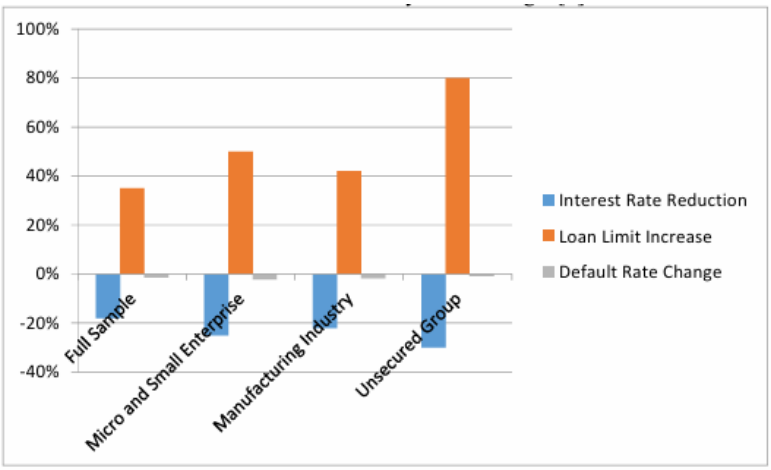


Figure 2. Heterogeneous effects of fintech empowering the credit Market (Simulated regression results) (picture credit: original)

Data description: Mean of multiple regression coefficients based on the simulation of the loan facilitation market data in the literature.

5. Challenges and policy implications: rebalancing efficiency and safety

Technology empowerment is not a panacea. Emerging risks, as mentioned in the introduction, require to shift from simply pursuing efficiency to building governance systems. The current development of fintech faces significant practical challenges. The uneven quality of data can lead to bias in risk control judgments, and the lack of transparency and explainability of intelligent algorithm models not only increases the probability of business misjudgment but also makes it more

DOI: 10.54254/2755-2721/2026.34874

difficult to regulate compliance in the industry. To address various development pain points, it is necessary to improve industry governance measures, improve the government data sharing system to optimize the data foundation, conduct risk measurement of intelligent models based on regulatory sandboxes, and implement regular third-party audits of high-risk algorithms to regulate the healthy development of the industry in all aspects [10].

5.1. Constraints on data quality and model transparency

Garbage in, Garbage out is an unbreakable rule in the field of risk control. On some platforms, due to incomplete data cleaning, models misjudged "short-term cash flow difficulties" as "malicious debt evasion", leading to an increase in false rejection rates. In addition, the unexplainability of deep learning models makes it difficult for regulators to conduct compliance reviews, which is also the root cause of algorithmic discrimination being hard to eradicate.

5.2. Policy recommendations

Based on the aforementioned mechanisms and challenges, policy should focus on: Policy should further improve the data governance and regulatory system, reduce the data acquisition costs for fintech companies by establishing an open sharing mechanism for government data, implement a "regulatory sandbox" system for AI models, test model risks in a controllable environment, and require financial institutions to conduct regular third-party algorithm audits of high-risk models and disclose possible discriminatory rules in a timely manner. In order to promote fintech innovation while ensuring market fairness and financial security.

6. Conclusions

This paper constructs a "big data - AI" collaborative analysis framework through a systematic review. The research shows that big data mitigates pre-event adverse selection through information reconstruction, and AI suppresses post-event moral hazard through behavioral monitoring. This synergistic mechanism significantly enhances the inclusiveness and efficiency of the credit market, achieving the inclusive effect of "data replacing collateral". However, the effect of technology empowerment is limited by data quality and the regulatory environment, and algorithmic discrimination and privacy risks are the focus of future governance. Future research should focus more on how to balance technological innovation and financial stability through regulatory technology.

Authors contribution

All the authors contributed equally and their names were listed in alphabetical order.

References

[1] Akerlof, G. A. (1970). The market for "lemons": Quality uncertainty and the market mechanism. The Quarterly Journal of Economics, 84(3), 488–500.
[2] Stiglitz, J. E., & Weiss, A. (1981). Credit rationing in markets with imperfect information. The American Economic Review, 71(3), 393–410.
[3] She, K. W., Shen, Y., & Zhao, S. Y. (2024). The impact of big data on bank credit behavior: Evidence from digital social credit platforms. Economic Research Journal, 59(3), 147–165.
[4] Gong, Q. (2024). Digital credit, collateral substitution and inclusive finance. Economic Research Journal.

DOI: 10.54254/2755-2721/2026.34874

[5] Gambacorta, L., Huang, Y., Li, Z., Qiu, H., & Chen, S. (2022). Data vs collateral. Review of Finance, 26(6), 1469–1503.
[6] Bartlett, R., Morse, A., Stanton, R., & Wallace, N. (2022). Consumer-lending discrimination in the FinTech era. Journal of Financial Economics, 143(1), 30–56.
[7] Fuster, A., Plosser, M., Schnabl, P., & Vickery, J. (2019). The role of technology in mortgage lending. The Review of Financial Studies, 32(5), 1854–1899.
[8] Chen, B., Guo, D., Xia, J., & Zhang, Z. (2025). The transformative role of artificial intelligence and big data in banking. SSRN.
[9] Gambacorta, L., Sabatini, F., & Schiaffi, S. (2025). Artificial intelligence and relationship lending (BIS Working Papers, No. 1244).
[10] Bussmann, N., Giudici, P., Marinelli, D., & Papenbrock, J. (2020). Explainable AI in fintech risk management. Frontiers in Artificial Intelligence, 3, 26.