WEP Warwick Evans Publishing

Transactions on Economics, Business and Management Research
ISSN: 2960-1789, eISSN: 2960-2254 | Volume 14
AFTEM 2024

Research on Influencing Factors of corporate financial risk based on LightBM-SHAP model: Differentiation analysis under different corporate nature

Renguang Cao [1, #], Jing Cao [2, #], Ting Huang [3, #, *]

[1] School of Economics and Management, Henan institute of science and technology, Xinxiang, China, 453003

[2] School of Economics and Management, Shanxi University, Taiyuan, China, 030000

[3] School of Accountancy, Guangzhou Xinhua University, Dongguan, China, 523133

* Corresponding author: a5268992024@163.com

[#] These authors contributed equally.

Abstract. With the deepening of economic globalization and the increasing complexity of enterprise operating environment, it is of great theoretical and practical significance to study the key factors affecting enterprise financial risk and their mechanisms for improving enterprise risk management level and market competitiveness. Based on the annual data of China's A-share listed companies from 2019 to 2023, this paper constructs A comprehensive evaluation index system of financial risk, and adopts entropy method to conduct a quantitative evaluation of corporate financial risk. Through the introduction of LightGBM model and Shapley value method, the paper conducts feature screening on the multi-dimensional factors that affect the financial risk of enterprises, and identifies the core indicators that are highly correlated with the financial risk. Then, from the Angle of enterprise nature, the paper discusses the mechanism of different influence of these indicators on financial risk in different enterprise types. The empirical results show that both financial and non-financial indicators have significant effects on the financial risk of enterprises, but the importance of these indicators is significantly different in different natures of enterprises, especially the influence of non-financial indicators in non-state-owned enterprises is more prominent. Based on the above research conclusions, this paper puts forward the corresponding policy suggestions, which provide a guiding reference for the financial risk management of enterprises.

Keywords: financial risk; firm nature; feature selection; causal inference.

1. Introduction

With the acceleration of the globalization process, the business environment faced by enterprises is becoming increasingly complex, and the market competition is becoming more and more fierce [1]. In this context, the business activities of enterprises are always accompanied by risks, resulting in the future is full of uncertainty. If the enterprise fails to effectively control the risk, it is very likely to fall into financial crisis [2]. As an inevitable part of enterprise operation, financial risk not only directly affects the sustainable development of the enterprise, but also concerns its long-term value creation. Therefore, in-depth exploration of the influencing factors of financial risk and its mechanism can not only help enterprises improve their risk management capabilities, but also enhance their competitive advantages in the market, and provide strong support for the strategic decision-making of enterprise management.

The research of early scholars mostly focused on the construction of comprehensive evaluation system through financial indicators. For example, Li Xinhui et al. (2010) [3] built a financial risk assessment model for public hospitals based on factor analysis, CAI Xinglin and Zhang Yaoga (2019) evaluated and warned the financial risk status of listed sporting goods companies [4], Wang Zhuquan et al. (2020) [5] analyzed the defects of traditional short-term financial risk analysis indicators, and then, The comprehensive short-term financial risk assessment system with both stock and flow is reconstructed to provide support for financial institutions and rating agencies to optimize risk assessment and early warning models. Shen Qingyuan (2021) [6] constructs a financial risk early warning model applicable to the automobile industry based on the fuzzy comprehensive evaluation method. These studies show that financial indicators play an important role in financial risk assessment. However, traditional approaches often fail to capture the full complexity and dynamics of the enterprise [7].

With the rapid development of big data analysis technology, more and more researches begin to introduce mathematical models and machine learning algorithms and apply them to financial distress prediction, financial fraud prediction and stock market prediction and quantification [8]. Machine learning has the ability to automatically extract patterns and knowledge from data [9], has higher flexibility in the process of data fitting, and has stronger explanatory ability and prediction accuracy in the study of nonlinear and non-directly observed causality [10], which significantly improves the depth and breadth of analysis of economic and financial phenomena. Help scholars overcome the limitations of data collection and highly complex correlation prediction problems [11]. For example, Ke Jian and Wang Qiqi (2023) [12] studied the financial risk early warning of R&D manufacturing based on machine learning, and expanded the application of machine learning in specific industries. In addition, Xingli Z. et al. (2024) [13] optimized the financial risk model in the enterprise management decision system based on deep learning in the context of digital transformation, further improving the intelligence level of risk assessment. By combining macroeconomic environment and corporate financial data, Jiang Fuwei et al. (2023) [7] applied machine learning technology to give early warning to corporate bond default risk and further verified the importance of financial indicators in revealing corporate financial risk.

At the same time, the role of non-financial indicators has been paid more and more attention. Guo Yaxin (2021) [14] uses financial indicators to build a financial risk evaluation index system, uses fuzzy comprehensive evaluation method to calculate the comprehensive evaluation value of financial risks, and introduces the influence of non-financial indicators on financial risks, which has reference significance for smes to evaluate their own financial risks. Wang Chong and Liu Yali (2024) [2] used Benford's Law and XGBoost model to analyze the financial risks of Shanghai and Shenzhen A-share listed companies in depth, and found that non-financial indicators such as corporate governance structure and industry characteristics also had A significant impact on financial risks. Gu et al. (2024) [15], when studying the impact of green credit policies on the financial risks of heavily polluting enterprises, emphasized the important role of non-financial indicators such as environmental performance in financial risk assessment. These studies not only expand the field of financial risk research, but also highlight the specific risks that companies face in the process of sustainable development.

To sum up, the research on enterprise financial risk assessment and management at home and abroad has made remarkable progress, especially with the introduction of mathematical models and machine learning technology, which has greatly improved the accuracy and efficiency of financial risk assessment. However, most of the current research focuses on the function analysis of a single indicator or simple combination application, while the research on the impact mechanism of financial indicators and non-financial indicators on financial risk under different enterprise nature background is relatively insufficient. Therefore, this paper explores how financial and non-financial indicators affect the financial risk of a firm, whether the relationship between them is positive or negative, and whether these relationships vary with the nature of the firm. The clarification of these issues will help to deepen the research in the field of financial risk and better explain the key factors affecting financial risk and their mechanisms.

In this paper, the comprehensive evaluation index system of financial risk is constructed, and the entropy method is used to quantify the enterprise financial risk. By using LightGBM algorithm and Shapley value method, the key indicators that are significantly related to financial risk are screened out from a large number of potential impact factors, and then the impact mechanism of these indicators on financial risk is discussed based on different enterprise nature backgrounds. The empirical results show that both financial indicators and non-financial indicators have a significant impact on the financial risk of enterprises, but there are significant differences in the intensity and importance of their impact under different enterprise nature backgrounds, especially in non-state-owned enterprises, the role of non-financial indicators is particularly prominent. Based on these findings, this paper further puts forward relevant policy and management recommendations.

The innovation of this paper is mainly reflected in two aspects: First, machine learning method (LightGBM and Shapley values) is introduced to identify and quantify the key indicators affecting financial risk from a large number of enterprise data. This data-driven approach not only improves the accuracy and efficiency of financial risk assessment, but also reveals the specific role and priority of financial and non-financial indicators in the formation of financial risks, providing a new technical path for quantitative analysis of financial risks. Second, it discusses the influence of financial index and non-financial index on financial risk from the perspective of enterprise nature, breaking through the limitation of traditional financial risk research. By taking corporate ownership structure into account, this study provides a more comprehensive analytical framework and reveals the unique characteristics and causes of financial risk in different types of firms.

2. Theory and methodology

2.1. Research hypothesis

In recent years, academics have conducted in-depth discussions on the impact of non-financial performance indicators on corporate financial risk, with a particular focus on the differences in performance in the context of different corporate natures. Wang and Liu (2019) [16], studying the technology industry, point out that non-financial performance indicators play an important role in enhancing financial stability and that the nature of the firms triggers significant differences among them. Li and Chen (2020) [17] further investigate the impact of corporate governance, a non-financial factor, on financial risk and find that the governance structure is crucial to financial risk management, and point out that differences in governance mechanisms between state-owned and non-state-owned enterprises may lead to different financial risk management effects. Similarly, Chen and Jiang (2021) [18] studied the impact of financial indicators on financial risk under different ownership structures, emphasizing the differences in financial risk management strategies between SOEs and private firms. Based on these findings, this paper proposes the following research hypotheses.

First, financial indicators are the core indicators reflecting the profitability, solvency and capital structure of an enterprise, and changes in the financial situation are directly related to the financial risk of an enterprise. Financial indicators not only provide investors with key data on the operational efficiency and financial health of enterprises, but also help management make more accurate decisions. Based on this, this paper proposes hypothesis H1:

H1: Financial indicators have a significant effect on corporate financial risk.

Second, non-financial indicators, such as employee satisfaction, customer relations, corporate governance and corporate social responsibility, although not directly reflecting the financial performance of an enterprise, have a significant impact on its long-term stability and market reputation. These factors usually act indirectly on financial risk by affecting firms' intangible assets and internal management. Therefore, this paper proposes hypothesis H2:

H2: Non-financial performance indicators play an important role in corporate financial risk.

Finally, there are significant differences in enterprises (e.g. SOEs vs. private enterprises) in terms of capital structure, market objectives, regulatory environment and policy support, which may lead to different performance of financial and non-financial indicators in enterprises of different natures. SOEs usually benefit from stronger policy support and financing advantages, so their financial risk management may rely more on financial indicators. On the other hand, private enterprises rely more on market mechanisms and autonomous decision-making, and therefore may rely more on non-financial indicators, such as customer relationship and employee performance, in their financial risk management. Based on this, this paper proposes hypothesis H3:

H3: Financial and non-financial indicators have different levels of influence on financial risk under different business natures.

2.2. Research Principles

2.2.1. Comprehensive financial risk evaluation system.

In constructing the comprehensive evaluation index system of financial risk, this paper adopts Altman (1968) [19] Z-score model to judge the occurrence of financial risk. The calculation of each index of Z-score model and the meaning of the indexes are shown in Table.1.

Table 1. Z-score model indicator calculation and indicator meaning

Indicator	Indicator Meaning
Working Capital/Total Assets	The higher the value of this indicator, the better the liquidity of the company's assets.
Retained Earnings/Total Assets	The higher the value of this indicator, the higher the degree of accumulation of undistributed profits.
EBIT/Total Assets	The higher the value of this indicator, the more stable and well-funded the enterprise is.
Market Value of Ownership Interest/Total Liabilities	The higher the value of this indicator, the more stable is the long-term solvency of the enterprise.
Total Sales/Total Assets	The higher the value of this indicator, the more stable the enterprise's own operation is.

After establishing the indicator system, the entropy value method is applied to calculate the weights of the selected indicators. Entropy value method is a weight allocation technique based on information theory, and its core principle is to determine the relative importance of each indicator in the overall evaluation by quantifying its information entropy. Specifically, the smaller the entropy value of the indicator, the more concentrated its information, reflecting the importance of the indicator in the financial risk assessment. Referring to Wang Shan et al. (2024) [20], the assignment steps are as follows:

(1) Constructing the original matrix

Construct the original data matrix $R = \left(X_{ij}\right)n \times m$, where X_{ij} is the indicator value of the ith listed company under the jth indicator, n is the number of listed companies, and m is the number of indicators.

(2) Data standardization

$$a_{ij} = \frac{X_{ij} - Min(X_{ij})}{Max(X_{ij}) - Min(X_{ij})} \tag{1}$$

(3) Calculation of indicator weights

$$P_{ij} = \frac{a_{ij}}{\sum_{i=1}^{n} a_{ij}} \tag{2}$$

(4) Calculate the information entropy of each indicator

$$e_j = -\frac{1}{\ln n} \sum_{n=1}^{n} (P_{ij} * \ln P_{ij}) \tag{3}$$

(5) Calculation of the coefficient of variation for each indicator

$$g_j = 1 - e_j \tag{4}$$

(6) Calculation of the weights of the indicators

$$W_j = \frac{g_j}{\sum_{j=1}^{m} g_j} \tag{5}$$

Through a series of calculations we obtain the weights of each indicator, which are summed to obtain the Z-indicator score for each sample. Subsequently, the probability density distribution of the Z-indicator scores is plotted, and firms with scores in the top 20% of SOEs and private firms, respectively, are defined as having financial risk.

2.2.2. Selection and Testing of Significant Influential Factors Based on LightGBM and Shapley Values.

In the data processing stage, this paper utilizes the SMOTE algorithm to process the unbalanced data derived from 2.2.1 using oversampling and undersampling methods to avoid the problem of biased results due to sample imbalance. In terms of feature selection, this paper uses LightGBM model combined with Shapley study.

LightGBM is an efficient and superior gradient boosting tree algorithm that can effectively handle large-scale datasets. The objective function of LightGBM consists of two parts: the loss function and the regularization term. It is known that the training set data$T = \{(x_1, y_1), (x_2, y_2), \cdots, (x_n, y_n)\}$, the loss function $l(y_i, \hat{y}_i)$, and the regularization term $\Omega(f_k)$, then the objective function of LightGBM can be represented as

$$L^t(\phi) = \sum_i^n l(y_i, \hat{y}_i) + \sum \Omega(f_k) \tag{6}$$

Where n refers to the total number of samples, y_i is the actual value of the ith sample, $\hat{y}_i$ is the predicted value of the ith sample, f_k refers to the kth tree model, and $\sum \Omega(f_k)$ can represent the complexity of the k trees. Then the objective function can be transformed into the following form

$$(\phi) = \sum_{i=1}^{n} l(y_i, \hat{y}_i^{t-1} + f_t(x_i)) + \Omega(f_t) + \text{Const} \tag{7}$$

Here Const is a constant term, and the final objective function can be obtained by expanding the second-order Taylor term and removing the constant term, regularizing the expansion and removing the constant term, and combining the coefficients of the primary term and the coefficients of the quadratic term. Second order Taylor expansion of the objective function

$$L^t(\phi) = \sum_{i=1}^{n} [l(y_i, \hat{y}_i^{t-1}) + g_i f_t(x_i) + \frac{1}{2} h_i f_t^2(x_i)] + \Omega(f_t) + \text{Const} \tag{8}$$

Here $g_i = \partial \hat{y}^{(t-1)} l\left(y_i, \hat{y}_i^{(t-1)}\right), h_i = \partial^2 \hat{y}^{(t-1)} l\left(y_i, \hat{y}_i^{(t-1)}\right)$. Next, the objective function is further simplified

$$\begin{aligned} L^t(\phi) &= \sum_{i=1}^{n} \left[g_i f_t(x_i) + \frac{1}{2} h_i f_t^2(x_i)\right] + \Omega(f_t) \\ &= \sum_{i=1}^{T} \left[\left(\sum_{i \in I_j} g_j\right) w_j + \frac{1}{2}\left(\sum_{i \in I_j} h_i + \lambda\right) w_j^2\right] + \gamma T \end{aligned} \tag{9}$$

Let $G_j = \sum_{i \in I_j} g_j, H_j = \sum_{i \in I_i} h_j$, then the objective function is

$$L^t(\phi) = \sum_{j=1}^{T} [G_j w_j + \frac{1}{2}(H_j + \lambda) w_j^2] + \gamma T \tag{10}$$

The objective function for each leaf node j is

$$f(w_j) = G_j w_j + \frac{1}{2}(H_j + \lambda) w_j^2 \tag{11}$$

Equation (11) is a quadratic function about w_j. When the objective value obj is smaller, the structure of the tree is also better, obj is the smallest, that is, the optimal solution of the objective function. The objective equation of each leaf node of the objective function is independent of each other, so when each leaf node reaches the maximum value, the whole objective function also reaches the maximum value. So that $w_j = -\frac{G_j}{H_j + \lambda}$ ubstituting into the formula (11), can get the optimal form of the objective function

$$L^t(\phi) = -\frac{1}{2} \sum_{j=1}^{r} \frac{G_j^2}{H_j + \lambda} + \gamma T \tag{12}$$

The SHAP method measures the extent to which individual features in the model and their interaction terms influence the model output by calculating the marginal contributions of these features. These marginal contributions are referred to as Shapley values and enable an effective assessment of the role of different features in the model's prediction results. In this way, SHAP provides an interpretation of the prediction results of the black-box model and helps the user to understand the specific impact of each feature. For feature i in the feature set S, the Shapley value is calculated as follows:

$$\Phi_i = \sum_{S \subseteq N \setminus \{i\}} \frac{|S|!(|N| - |S| - 1)!}{|N|!} (v(S \cup \{i\}) - v(S)) \tag{13}$$

Where N denotes the set of all features; S is any subset of features that does not contain feature i; $|S|$ is the number of features in the set S; $v(S)$ is the feature set S contributes to the predicted output of the model; $v\,(S \cup \{i\})$ is the contribution of the feature set $S \cup \{i\}$ containing feature i to the predicted output of the model.

Through the experiment, we will be able to analyze in depth the difference in importance of each financial and non-financial indicator between state-owned and non-state-owned enterprises, so as to identify the key factors affecting the financial risk of enterprises. The results obtained are subsequently tested using XGBoost and Random Forest combined with Shapley's method.

2.2.3. Causal analysis - two-way fixed effects model.

In order to explore the relationship between the characteristic variables screened in 2.2.2 and financial risk, this paper will use the two-way fixed effects model to analyze. The two-way fixed-effects model can effectively control individual heterogeneity, i.e., the influence of those time-independent characteristics on the dependent variable, so as to more accurately reveal the intrinsic correlation between the independent variables and the dependent variable. The general form of the fixed effects model can be expressed as:

$$Y_{it} = \alpha + \beta_1 X_{1it} + \beta_2 X_{2it} + \cdots + \beta_k X_{kit} + \mu_i + \epsilon_{it} \tag{14}$$

Where Y_{it} represents the dependent variable, specifically the firm's level of financial risk, which is usually quantified by the score of the Z indicator. X_{kit} represents the independent variables, including the key financial and non-financial indicators obtained through the LightGBM model and the Shapley value analysis in 2.2.2. α is a constant term; β_k is the regression coefficient of the respective variable, which captures the degree of influence of the independent variable on the dependent variable; μ_i is an individual fixed effect to capture firm characteristics that do not change over time; ϵ_{it} is a random error term reflecting fluctuations in the dependent variable that cannot be explained by the independent variables. To ensure the relevance and validity of the study, this paper will construct the fixed-effects model for state-owned enterprises and non-state-owned enterprises separately. Through this grouping analysis, it can reveal more clearly the differences in the role of different enterprise nature on financial and non-financial indicators in financial risk.

3. Empirical analysis

3.1. Data sourcing and preprocessing

In this study, we utilized all A-share companies as the research sample and gathered relevant data on 20 financial and non-financial indicators associated with financial risk from the Guotai An database for the years 2019 to 2023. The dataset comprised 25,527 observations. To mitigate the impact of noisy data on result accuracy, we initially conducted data cleansing and specifically processed the sample data as follows: (1) excluded samples from financial industry enterprises; (2) removed samples from ST-class enterprises; (3) eliminated samples with missing variables; and (4) addressed sample bias issues resulting from imbalance using over-sampling and under-sampling methods through the SMOTE algorithm. Ultimately, we obtained 3,375 observation records.

3.2. The results of the comprehensive evaluation of financial risks

The article adopts the modified Z-score financial indicator proposed by Altman as a quantitative measure of corporate financial risk [19]. The specific calculation formula is as follows:

$$Z - score = \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4 x_4 + \beta_5 x_5 \quad (15)$$

Among them, where x_1 represents working capital/total assets, x_2 represents retained earnings/total assets, x_3 represents EBIT/total assets, x_4 represents market value of equity/total assets, and x_5 represents operating income/total assets. The coefficients β_1 to β_5 correspond to the variables x_1 to x_5. Each indicator has different weights, with x_4 having the highest weight and x_2 having the lowest weight. To accurately analyze the impact of each indicator on financial risk, we use the entropy method in comprehensive evaluation to calculate the weights of various influencing factors on financial risk. The specific weights are shown in Table 2 below:

Table 2. The weight of each indicator of financial risk

x_1	x_2	x_3	x_4	x_5
0.056	0.053	0.159	0.535	0.198

Based on the calculated weights for each indicator, we obtain specific scores for Z-score and plot a probability density distribution graph for Z-score, taking the lower 20% as indicating financial risk. Z-score is a binary variable ranging from 0 to 1, with 0 indicating financial risk. The larger the Z-score, the smaller the financial risk. The specific model process diagram for entropy method is shown in Figure below:

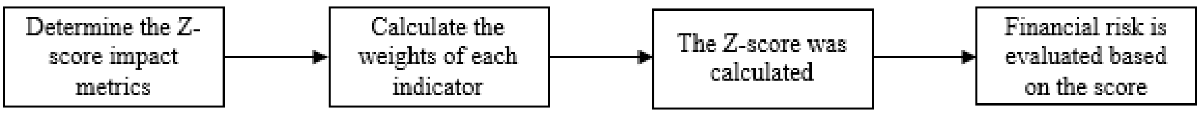

Figure 1. Financial risk comprehensive evaluation model

3.3. Important influencing factor selection results

Due to the fact that corporate financial risk is not only related to financial indicators such as ownership ratio, debt ratio, return on total assets, fixed asset ratio, leverage level, current ratio, quick ratio, company size etc, but also to non-financial indicators such as executive shareholding proportion, independent director proportion, customer concentration degree, ESG score and equity concentration degree. To analyze more factors that influence corporate financial risk in this study, we comprehensively refered to the practices of scholars like Liu Yunjing [21], Zhang Qian et al. [22] and Zhou Chen et al. [23]. Finally, we selected a total of 20 influencing factors from two categories: financial indicators and non-financial indicators. The specific details are shown in Table 3 below:

Table 3. Type of financial indicator

Indicator category	Indicator name	Indicator symbols	Description of indicators
Financial Indicators	Financial risk	Z-score	Calculated from equation (15)
	Financial risk	O-score	Sourced from the CSMAR
	Net interest rate on total assets	Roa	Net margin/average total assets
	Gearing ratio	Tdr	Total liabilities/total assets
	Company size	Size	Total assets
	Current ratio	Liqui	Current assets/current liabilities
	Quick ratio	Qr	Quick assets/current liabilities
	Combined leverage	Dtl	Operating leverage factor * financial leverage factor
	Total asset turnover rate	Tat	Sales revenue/total assets
	Ratio of intangibles	Ia	Intangible/total assets
	Tangible assets ratio	Tan	Tangible/total assets
	Fixed asset ratio	Fixed	Fixed/total assets
	Equity ratio	Cpr	Total Liabilities/owners' equity
	Gross operating income growth rate	Growth	Current year's operating income growth / Previous year's operating income growth
	Cash flow ratio	Cfr	Cash flows from operating activities/current liabilities
Nonfinancial indicators	Percentage of independent directors	Dire	Number of independent directors/number of directors
	Executive Shareholding Ratio	Mo	Number of shares held by executives/total shares
	ESG scores	Esg	Huazheng ESG Ratings
	Shareholding concentration	Share	Shareholding ratio of the largest shareholder
	Customer concentration	Cc	Proportion of sales to total sales from top five customers of the enterprise

To explore the impact of different categories of financial indicators on the financial risk of state-owned enterprises (SOEs) and private enterprises (PEs), we divided companies into SOEs and PEs for separate analysis. Moreover, due to numerous indicators affecting corporate financial risk, this study used LightGBM+Shapley method to select important indicators that affect corporate financial risk. The specific results are shown in the following figure (left side for SOEs and right side for PEs):

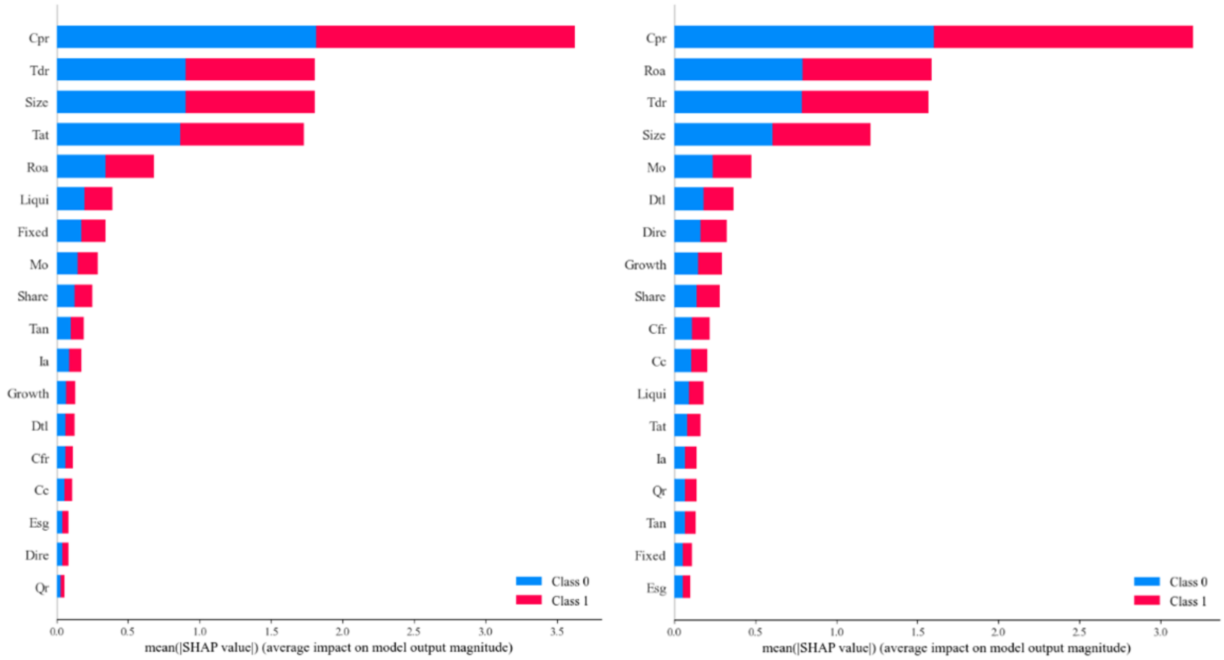

Figure 2. LightGBM+Shapley prediction results

To increase the credibility of prediction results, we used XGBoost+Shapley value method and random forest + Shapley value method to predict important indicators affecting financial risks, and then compared them with the prediction results obtained by LightGBM+Shapley method. The specific illustration is shown below (left side for SOEs and right side for PEs):

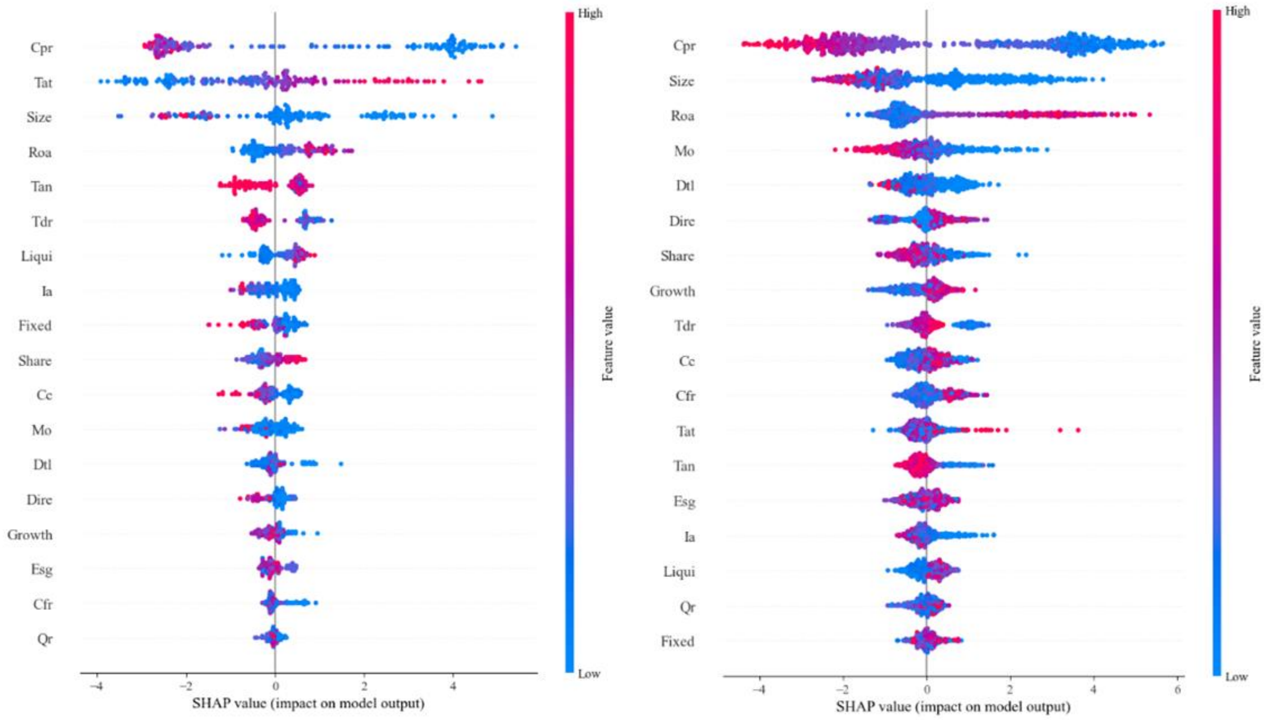

Figure 3. XGBoost+Shapley prediction results

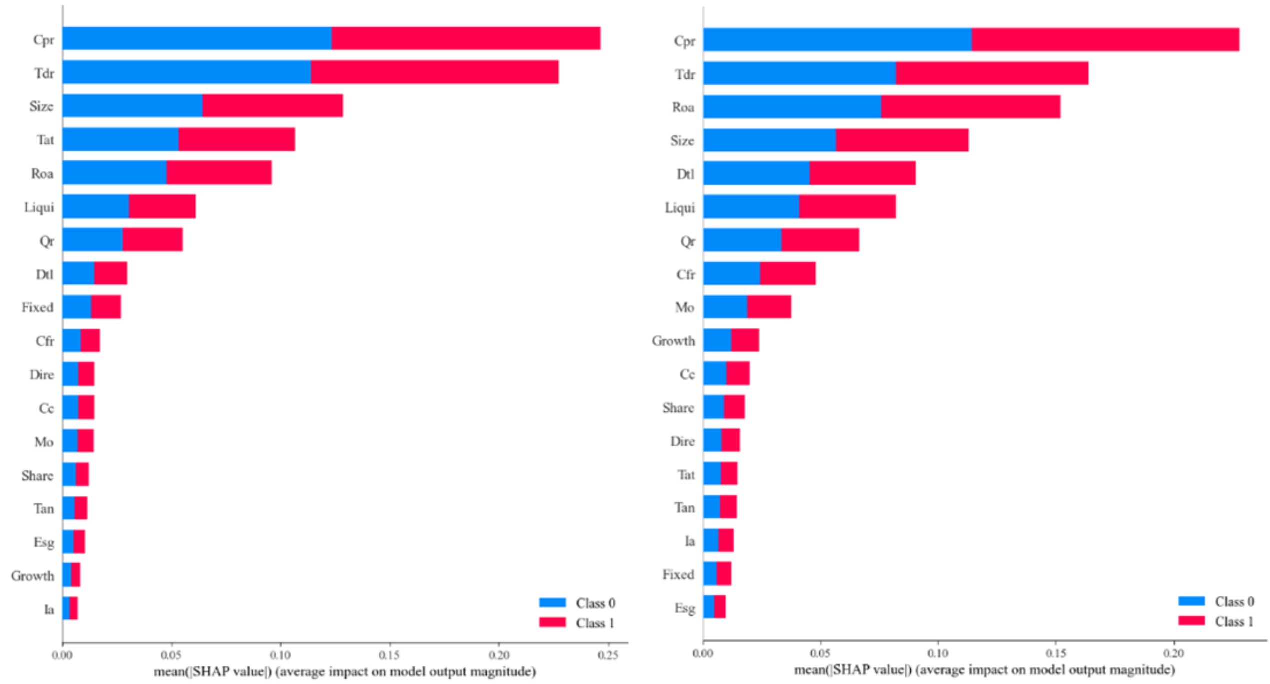

Figure 4. RF + Shapley prediction results

Comparing the prediction results of LightGBM+Shapley with XGBoost+Shapley and Random Forest+Shapley, it was found that the accuracy of all prediction results is above 90%. Therefore, the LightGBM+Shapley model used in this study can accurately select important indicators that affect corporate financial risk. Among them, Precision (also known as accuracy) represents the proportion of true positive samples among predicted positive samples; Recall (also known as sensitivity) represents the proportion of true positive samples among all actual positive samples; F1 score is calculated by assigning equal weights to Precision and Recall [21]. The different model's prediction results are shown in Table 4-5 below:

Table 4. SOE forecast results

	precision	recall	F1-score
LightGBM+Shapley	0.93	0.93	0.93
XGBoost+Shapley	0.93	0.92	0.92
RF+Shapley	0.93	0.92	0.92

Table 5. PE forecast results

	precision	recall	F1-score
LightGBM+Shapley	0.94	0.94	0.94
XGBoost+Shapley	0.94	0.94	0.94
RF+Shapley	0.93	0.93	0.93

3.4. Results of the analysis of influencing factors

Based on the prediction results of the LightGBM+Shapley model, we selected the top 7 indicators respectively as important factors influencing financial risk for state-owned enterprises and private enterprises. The indicators affecting state-owned enterprises are ownership ratio, debt-to-assets ratio, company size, net profit margin on total assets, total asset turnover rate, current ratio and fixed asset ratio; while for private enterprises they are ownership ratio, debt-to-assets ratio, company size, net profit margin on total assets, executive shareholding percentage, comprehensive leverage and independent director proportion.

Tables 6 and 7 show descriptive statistical results for major indicators affecting financial risk for state-owned enterprises and private enterprises respectively. Among them, zscore and oscore measure enterprise financial risk with similar but opposite directions. There is a significant difference between maximum and minimum values of enterprise financial risk, and there is a large variation in levels of financial risk among sample companies which aligns with our expectations. The descriptive statistical results for other indicators also generally meet our expectations.

Table 6. Descriptive statistics on key indicators of State-owned enterprises

Variable	Obs	Mean	Std. Dev.	Min	Max
oscore	726	-9.069	1.884	-17.687	-4.786
zscore	726	2.843	4.323	0.107	37.12
ttm	726	0.769	0.745	0.036	8.634
liqui	726	2.01	1.438	0.179	16.113
fixed	726	0.194	0.168	0.003	0.8
cpr	726	1.217	1.339	0.064	14.975
tdr	726	0.459	0.187	0.06	0.937
size	726	6.101e+10	2.606e+11	6.700e+08	2.900e+12
roa	726	0.054	0.042	-0.004	0.295

Table 7. Descriptive statistics on key indicators of private enterprises

Variable	Obs	Mean	Std. Dev.	Min	Max
oscore	2649	-9.7	2.235	-26.978	-5.51
zscore	2649	3.94	5.794	0.097	86.496
mo	2649	17.857	15.217	0.001	72.45
dtl	2649	2.374	19.579	0.722	913.802
dire	2649	0.385	0.054	0.25	0.714
cpr	2649	0.724	0.612	0.022	5.798
roa	2649	0.07	0.058	-0.012	0.88
tdr	2649	0.369	0.162	0.021	0.853
size	2649	1.055e+10	3.266e+10	3.700e+08	6.800e+11

By comparing the prediction results of LightGBM+Shapley models for state-owned enterprises and private enterprises, it was found that the ownership ratio, asset-liability ratio, company size, and net profit margin of total assets are common indicators that affect corporate financial risk. By using non-common indicators as explanatory variables and conducting regression analysis separately for state-owned enterprises and private enterprises, it was found that both were statistically significant at the 1% level, which indicates that financial and non-financial indicators have a significant impact on corporate financial risk, and Hypothesis 1 and Hypothesis 2 are confirmed. However, only financial indicators have a significant impact on the financial risk of state-owned enterprises, while non-financial indicators such as executive shareholding ratio also have a significant impact on the financial risk of private enterprises at the 1% level. Based on this, Hypothesis 3 is supported. The difference in the nature of the indicators that affect the financial risk of enterprises may be determined by the special status of state-owned enterprises in the national economy and the internal demand of private enterprises to enhance their core competitiveness. In addition, the significant levels of the common indicators for state-owned enterprises and private enterprises are the same, so the conclusion of the study does not have a heterogeneity problem. The specific results are shown in Table 8-9 below:

Table 8. The return of state-owned enterprises to the results

	(1)	(2)
	zscore	zscore
cpr	0.379***	0.348**
	(2.771)	(2.527)
tdr	-10.266***	-7.997***
	(-7.492)	(-4.998)
size	-0.000	-0.000
	(-0.559)	(-0.597)
roa	8.361***	8.724***
	(2.796)	(2.690)
ttm		0.705
		(1.611)
liqui		0.293**
		(2.485)
fixed		5.539***
		(3.005)
_cons	6.703***	3.476***
	(10.934)	(3.255)
N	726	726
R^2	0.121	0.143
F	19.551	13.465

***p<0.01, **p<0.05, *p<0.10

Table 9. The return of private enterprises to the results

	(1)	(2)
	zscore	zscore
cpr	4.542***	4.684***
	(7.043)	(7.280)
tdr	-28.687***	-29.603***
	(-12.278)	(-12.653)
size	-0.000**	-0.000**
	(-2.134)	(-2.406)
roa	9.634***	9.957***
	(6.056)	(6.247)
mo		-0.054***
		(-4.609)
dtl		-0.001
		(-0.311)
dire		0.599
		(0.207)
_cons	10.684***	11.646***
	(19.602)	(9.441)
N	2649	2649
R^2	0.109	0.118
F	64.400	40.180

***p<0.01, **p<0.05, *p<0.10

To address endogeneity issues that may arise from omitted variables, reverse causality, selection bias, etc., instrumental variable tests were conducted based on previous research. Lagged variables were used as instrumental variables for testing the endogeneity of financial indicators, with the example of asset-liability ratio as a common variable. Non-financial indicators, such as executive shareholding percentage, were tested for endogeneity using the square of executive shareholding percentage as an instrumental variable. The regression analysis results indicate that there are no significant endogeneity problems in both state-owned enterprises and private enterprises, ensuring high reliability of the findings. The specific results are shown in Table 10-11:

Table 10. Endogeneity test of financial indicators

	(1)	(2)
	tdr	zscore
ltdr	0.305***	-3.928***
	(15.496)	(-3.285)
cpr		-0.745***
		(-3.543)
roa		10.826***
		(6.780)
size		-0.000
		(-1.562)
_cons	0.273***	5.307***
	(35.508)	(11.274)
N	2700	2700
R^2	0.106	0.040
F	240.135	20.837

***p<0.01, **p<0.05, *p<0.10

Table 11. Endogeneity test of non-financial indicators

	(1)	(2)
	mo	zscore
mo1	0.015***	-0.001***
	(123.376)	(-4.792)
dtl		-0.001
		(-0.317)
dire		0.469
		(0.162)
cpr		4.690***
		(7.292)
roa		9.889***
		(6.210)
tdr		-29.751***
		(-12.711)
size		-0.000**
		(-2.344)
_cons	9.715***	11.272***
	(125.602)	(9.249)
N	2649	2649
R^2	0.879	0.119
F	15221.532	40.455

***p<0.01, **p<0.05, *p<0.10

Drawing on existing research findings, this paper substitutes O-score for Z-score as a measure of corporate financial risk for robustness testing [22]. In this case, the larger the O-score, the higher the corporate financial risk. The results show that financial indicators have a significant impact on the corporate financial risk of both state-owned enterprises and private enterprises, while non-financial indicators also have a significant impact on the financial risk of private enterprises, further verifying the accuracy of Hypothesis 1, Hypothesis 2, and Hypothesis 3. Therefore, the research conclusion has high robustness. The specific results are shown in Tables 12-13 below:

Table 12. The return of stability to state-owned enterprises

	(1)	(2)
	oscore	oscore
cpr	-0.130***	-0.088***
	(-4.289)	(-3.012)
tdr	8.708***	7.305***
	(28.604)	(21.449)
size	-0.000**	-0.000**
	(-2.304)	(-2.264)
roa	-15.021***	-16.057***
	(-22.605)	(-23.258)
ttm		-0.022
		(-0.233)
liqui		-0.202***
		(-8.035)
fixed		-0.257
		(-0.654)
_cons	-12.045***	-10.928***
	(-88.433)	(-48.068)
N	726	726
R^2	0.769	0.793
F	472.484	309.534

***p<0.01, **p<0.05, *p<0.10

Table 13. The regression results of the stability of private enterprises

	(1)	(2)
	oscore	oscore
cpr	-1.107***	-1.128***
	(-8.118)	(-8.270)
tdr	13.565***	13.698***
	(27.467)	(27.627)
size	-0.000*	-0.000*
	(-1.842)	(-1.668)
roa	-14.248***	-14.248***
	(-42.374)	(-42.179)
mo		0.007***
		(2.763)
dtl		0.001
		(1.456)
dire		-0.378
		(-0.615)
_cons	-12.879***	-12.894***
	(-111.785)	(-49.321)
N	2649	2649
R^2	0.645	0.647
F	956.729	550.008

***p<0.01, **p<0.05, *p<0.10

4. Conclusion and suggestion

With the acceleration of economic globalization, corporate financial risk has an important impact on its sustainable development and value creation. This paper takes the annual data of China's A-share listed companies from 2019 to 2023 as samples, adopts comprehensive evaluation method to construct financial risk index system, and carries out quantitative evaluation by entropy method. Further, the LightGBM algorithm and Shapley value method are used to select the characteristics of the multi-dimensional factors affecting financial risk, identify the core indicators, and explore the differential impact mechanism of these indicators on financial risk under different corporate nature. The empirical results show that: (1) Both financial and non-financial indicators have significant influence on the financial risk of enterprises; (2) Among enterprises of different natures, non-financial indicators have a more significant impact on the financial risks of non-state-owned enterprises. Due to the flexible governance structure and market adaptability of non-state-owned enterprises, non-financial indicators have a greater impact on financial risks in governance and decision-making efficiency. Based on this, the following suggestions are put forward:

(1) Internal management of enterprises: State-owned enterprises and non-state-owned enterprises should optimize the capital structure, reduce financial leverage, reduce reliance on debt financing, and enhance financial robustness. State-owned enterprises should strengthen risk management, especially risk assessment for large-scale projects, improve the efficiency of asset use, and strengthen cost control and operation management. Non-state-owned enterprises should design reasonable executive incentive mechanisms, increase the proportion of independent directors, enhance the transparency of corporate governance, and reduce internal fraud and risks.

(2) External management: Enterprises should establish a risk-sharing mechanism to reduce financing costs, share resources and diversify risks. State-owned enterprises should explore introducing more market mechanisms to improve their competitiveness. Non-state-owned enterprises should adapt to market changes through innovation and business model adjustment, enhance communication with investors, and reduce market fluctuations caused by information asymmetry.

(3) Government level: Governments should coordinate macroeconomic policies to provide a stable economic environment and reduce the impact of volatility on corporate financial risk. Strengthening the supervision of the financial activities of state-owned and non-state-owned enterprises, controlling systemic risks, promoting the reform of state-owned enterprises, reducing policy burdens and enhancing market competitiveness; at the same time, creating a level playing field for private enterprises, encouraging innovation, promoting stable revenue growth and reducing financial risks.

These suggestions aim to jointly reduce corporate financial risks by optimizing internal corporate management, strengthening external cooperation and innovation, and improving government policy support.

References

[1] Gu lisi. Research on financial risk early warning of listed companies in China's pharmaceutical manufacturing industry--Based on Stacking integration algorithm [J]. Small and Small and Medium-sized Enterprises Management and Technology,2023, (08): 164 - 166.

[2] WANG Chon, LIU Yali. A study on financial risk early warning by Benford and XGBoost model fusion - based on the data analysis of A-share listed companies in Shanghai and Shenzhen from 2000 to 2021 [J]. Research on Finance and Accounting, 2024, (01): 35 - 46.

[3] LI Xinhui, DONG Siping, XIAO Jingjing, et al. Research on financial risk assessment model of public hospitals based on factor analysis [J]. China Health Economics, 2010, 29 (10): 88 - 90.

[4] CAI Xinglin, Zhang Yaoga. Research on financial risk assessment and early warning of listed Chinese sporting goods companies: Based on financial data from 2009 to 2016 [J]. Journal of Xi 'a University of Physical Education, 2019, 36 (03): 282 - 289.

[5] WANG Zhuquan, SONG Xiaobin, WANG Yuanzhu. Evaluation and judgement of short-term financial risk of China's real economy--Comprehensive assessment and early warning of short-term financial risk considering stock and flow [J]. Management World, 2020, 36 (10): 156 - 170+216 - 222.

[6] Shen Qingyuan. Research on financial risk early warning management of automobile industry based on fuzzy comprehensive evaluation method--Taking JAC as an example [J]. Modern Business,2019, (35): 53 - 55.

[7] JANG Fuwei, LIN Yihao, MA Tian. Corporate bond default risk in the context of ‘de-rigging’: an investigation of machine learning early warning and economic mechanisms [J]. Financial Research,2023, (10): 85 - 103.

[8] Gepp A, Linnenluecke M K, O'Neill T, et al. Big Data Techniques in Auditing Research and Practice: Current Trends and Future Opportunities[J]. Social Science Electronic Publishing, 2017, 40: 102 - 115.

[9] ZHAO Qi, Xu Weijun, Ji Yucheng, Liu Guifang, Zhang Weiguo. Review on the application of Machine learning in financial asset price prediction and allocation [J]. Journal of Management, 20, 17 (11): 1716 - 1728.

[10] Mullainathan S, Spiess J. Machine Learning: An Applied Econometric Approach [J]. Journal of Economic Perspectives, 2017, 31(2): 87 - 106.

[11] Wang Fang, Wang Xuanyi, Chen Shuo. Machine Learning in economics research: Review and Prospect [J]. Journal of Quantitative and Technical Economics, 2019, 37 (04): 146 - 164.

[12] Ke Jian, Wang Qiqi. Research on early warning of financial risk in R&D-based manufacturing industry based on machine learning [J]. Management Accounting Research, 2023, (05): 22 - 32.

[13] Xingli Z, Wenjie W, Guochao L, et al. Optimizing Financial Risk Models in Digital Transformation-Deep Learning for Enterprise Management Decision Systems [J]. Journal of Organizational and End User Computing (JOEUC), 2024, 36 (1): 1 - 19.

[14] Guo Yaxin. Research on comprehensive evaluation of financial risk of small and medium-sized enterprises--Based on fuzzy comprehensive evaluation method [J]. Times Economy and Trade, 2021, 18 (09): 68 - 71.

[15] Gu X, Wang Y, Xing X, et al. The impact of green credit policy on heavily polluting enterprises’ financial risk: evidence from China [J]. Clean Technologies and Environmental Policy, 2024, 1 - 19.

[16] Wang, T., & Liu, X. (2019). "The Effects of Non-Financial Performance Indicators on Financial Risk: Evidence from China’s Technology Sector." Technological Forecasting and Social Change, 146, 620 - 628.

[17] Li, H., & Chen, Q. (2020). "Corporate Governance and Financial Risk: The Role of Non-Financial Factors." Corporate Governance: An International Review, 28 (6), 467 - 485.

[18] Chen, K., & Jiang, Y. (2021). "Financial Indicators and Their Impact on Financial Risk in Different Ownership Structures: Evidence from China." Emerging Markets Finance and Trade, 57 (8), 2255 - 2268.

[19] Altman, E, 1968, Financial Ratios, Discriminant Analysis and the Prediction of Corporate Bankruptcy0, Journal of Finance 23 (Sept.), pp. 589 - 609.

[20] WANG Shan, et al. Design and research on comprehensive evaluation model of water quality based on entropy value method [J]. Automation Application, 2024, 65 (12): 242 - 243+246.

[21] Liu Yunjing, Wu Bin, Zhang Min. Can independent directors identify the risk of corporate financial fraud - a study of predicting financial fraud based on machine learning [J]. Accounting Research, 2023, (09): 21 - 34.

[22] ZHANG Qian, MEI Yali, WANG Kui. Does ESG Performance Affect Corporate Financial Risk - Empirical Evidence Based on A-share Listed Companies [J]. Friends of Accounting, 2023, (21): 105 - 114.

[23] ZHOU Chen, ZHANG Feiteng, HAN Chen. Research on the impact of customer relationship on financial risk of listed companies [J]. Friends of Accounting, 2024, (07): 68 - 75.