

# Application of Logistic Regression Models in Measuring Corporate Investment and Financing Efficiency

Shengze Zhong  
Jinan New Channel-JUTES High School, Jinan, 250000, China  
chung88888888@foxmail.com

## ABSTRACT

Corporate investment and financing efficiency reflects the coordination between a company's financing capacity, capital allocation quality, and value creation capacity, and is an important research topic in corporate finance and corporate financial management. With the continuous accumulation of corporate financial, market, and governance data, the measurement of investment and financing efficiency has gradually shifted from single financial indicator analysis to a comprehensive method combining statistical models, efficiency evaluation models, and risk identification models. The Logistic Regression Model, due to its strong classification and identification capabilities, clear variable interpretation, and ease of converting results into probabilistic judgments, has high application value in identifying corporate financing constraints, judging investment efficiency, providing early warning of financial risks, and evaluating the investment and financing status. This paper analyzes the application of the Logistic Regression Model in measuring corporate investment and financing efficiency, focusing on the theoretical basis, efficiency status construction methods, explanatory variable selection, model setting path, result evaluation methods, and applicable scenarios. The study argues that the Logistic Regression Model should not be simply regarded as a tool for directly calculating efficiency values, but is more suitable as a method for identifying the status of corporate investment and financing efficiency and analyzing influencing factors. Future research should further improve the model in terms of efficiency label construction, dynamic data modeling, nonlinear variable handling, and practical application in enterprises.

**Keywords:** Logistic Regression; Investment and Financing Efficiency; Financing Constraints; Investment Efficiency; Corporate Financial Risk

## 1. INTRODUCTION

### 1.1 Research Background

Corporate investment and financing activities are the foundation for business expansion, technological upgrading, and value creation. Financing activities address the source of funds, while investment activities determine the direction of fund usage; both jointly influence the efficiency of corporate resource allocation. If a company's financing costs are too high, its financing channels are limited, or its capital structure is unbalanced, even with good investment opportunities, it may fail to invest due to insufficient funds. Conversely, if a company's internal governance is ineffective, its management is overly expansive, or its free cash flow is insufficiently constrained, funds may flow to low-yield projects, resulting in over-investment. Therefore, corporate investment and financing efficiency is not a problem that can be fully explained by a single financial indicator, but rather the result of the combined effects of capital structure, information asymmetry, financing environment, governance mechanisms, and investment decisions.

Modern corporate finance theory provides a foundation for the study of corporate investment and financing efficiency. The capital structure theory proposed by Modigliani and Miller, starting from the relationship between capital cost and corporate value, laid the theoretical starting point for financing decision analysis<sup>[1]</sup>. Myers and Majluf further explained the corporate financing sequence selection from the perspective of information asymmetry, pointing out that the information difference between external investors and corporate managers will affect corporate financing costs and financing methods<sup>[2]</sup>. Stiglitz and Weiss's research on credit rationing shows that under the condition of incomplete information in the financial market, companies may not be able to obtain funds due to insufficient project quality, but rather due to adverse selection and moral hazard leading to an imbalance in credit resource allocation<sup>[3]</sup>. These theories collectively illustrate that the key to corporate investment and financing efficiency lies in whether the ability to obtain funds matches the quality of fund use.

## 1.2 Significance of the Research

In corporate management practice, the measurement of investment and financing efficiency has direct application value. For enterprises, accurately identifying financing constraints and investment deviations helps optimize capital structure, control debt risk, and improve the quality of capital budgeting. For financial institutions, investment and financing efficiency evaluation can serve as an important reference for credit approval, credit rating, and risk monitoring. For regulatory authorities, identifying corporate fund mismatches and inefficient resource allocation helps enhance the capital market's ability to serve the real economy.

Traditional methods often assess a company's investment and financing efficiency using metrics such as debt-to-equity ratio, return on equity, cost of capital, investment cash flow sensitivity, DEA efficiency, or SFA efficiency. While these methods have different advantages, a more explicit classification model is still needed to identify whether a company is inefficient, has limited financing, is underinvested, or overinvested. Logistic regression models can transform the issue of corporate investment and financing efficiency into a probabilistic judgment problem. They can be used to identify a company's efficiency status and to explain the key factors influencing efficiency changes, thus exhibiting strong methodological adaptability.

## 1.3 Core Issues and Article Structure

This paper focuses on three key questions: First, how can corporate investment and financing efficiency be transformed into a dependent variable that can be handled by a Logistic regression model? Second, which financial, governance, and external environmental variables are suitable for inclusion in the model? Third, how to evaluate the application effect of the Logistic model in measuring investment and financing efficiency? The paper is structured as follows: Part 1 explains the research background and significance; Part 2 outlines the literature sources and selection criteria; Part 3 focuses on analyzing the theoretical foundation, measurement path, model construction, and application scenarios; Part 4 summarizes current shortcomings and future development directions; and Part 5 presents the conclusions.

## 2. EXPLANATION OF LITERATURE SOURCES AND SCREENING

The literature in this article was primarily sourced from Web of Science, ScienceDirect, JSTOR, Wiley Online Library, Oxford Academic, EconPapers, Google Scholar, CNKI, and Wanfang Database. English search keywords included "logistic regression," "corporate investment efficiency," and "financing." "constraints," "financial distress prediction," "data envelopment analysis," "stochastic frontier analysis", etc.; Chinese search keywords include "Logistic regression", "corporate investment and financing efficiency", "financing constraints", "investment efficiency", "financing efficiency measurement", "financial risk warning", etc.

The literature review primarily covers the period from 1957 to 2026. Literature on classical theories and fundamental methods is not subject to time restrictions, while applied research will prioritize findings from the last two decades. Selection criteria mainly include: clear source, relevance to the research topic, complete methodological description, and representative conclusions. Literature that only presents viewpoints, has unclear citations, or incomplete information will not be considered.

## 3. APPLICATION ANALYSIS OF LOGISTIC REGRESSION MODEL IN MEASURING CORPORATE INVESTMENT AND FINANCING EFFICIENCY

### 3.1 Definition of the Connotation of Corporate Investment and Financing Efficiency

Corporate investment and financing efficiency can be understood as a company's ability to obtain funds at a reasonable cost and allocate those funds to effective investment projects. As shown in Figure 1, financing efficiency emphasizes whether the fund acquisition process is efficient, mainly focusing on financing costs, financing structure, financing speed, financing stability, and financing risks; investment efficiency emphasizes whether the fund utilization process is reasonable, mainly focusing on whether investment projects align with the company's growth opportunities, whether investment expenditures deviate from reasonable levels, and whether investment returns cover the cost of capital.

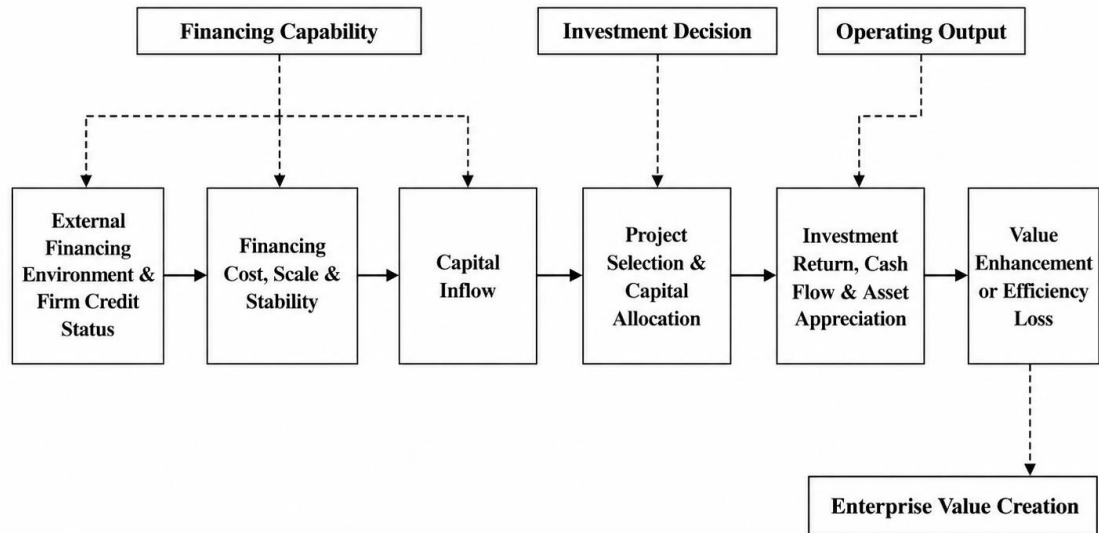


Figure 1. Formation Mechanism of Corporate Investment and Financing Efficiency.

Fazzari, Hubbard, and Petersen linked corporate investment with internal cash flow, proposing that financing constraints affect corporate investment expenditures, providing an important analytical framework for subsequent research on the relationship between corporate financing capacity and investment behavior (Figure 2)<sup>[4]</sup>. This idea suggests that corporate investment and financing efficiency should not only be judged by the scale of financing or investment, but should also focus on whether the company misses investment opportunities due to financing constraints or forms inefficient investments due to excessive liquidity.

From a practical perspective, corporate investment and financing efficiency encompasses at least three levels: First, financing efficiency, i.e., whether a company can obtain appropriate funds at a lower cost; second, investment efficiency, i.e., whether the scale of corporate investment matches growth opportunities and operational capabilities; and third, investment and financing matching efficiency, i.e., whether the funds obtained through financing are allocated to projects that can generate incremental value. The advantage of the Logistic Regression model lies precisely in its ability to transform these complex states into identifiable and interpretable categorical variables.

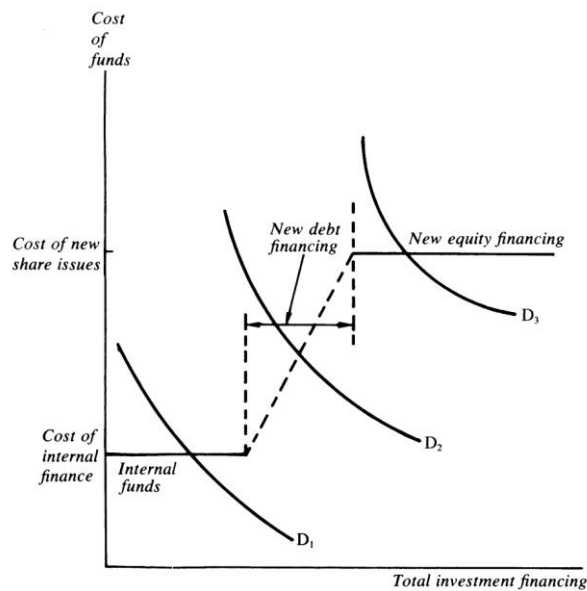


Figure 2. Cost of Funds and Investment Financing Choice under Financing Constraints.

### 3.2 Methods for Constructing Efficiency States

Logistic regression models require the dependent variable to have discrete categorical characteristics. Therefore, in measuring investment and financing efficiency, the primary issue is not directly building a model, but rather constructing appropriate efficiency status labels. There are three main common approaches.

The first type is based on thresholds for financial indicators. Researchers can classify companies into high-efficiency and low-efficiency groups based on indicators such as return on assets, cost of capital, operating cash flow, return on investment, and cost of debt. This method is simple to operate and suitable for studies with large sample sizes but simple data structures. However, this method is easily affected by industry differences, company size, and accounting policies. If the threshold settings lack a theoretical basis, the classification results may be distorted.

The second category is labels generated based on efficiency models. Farrell's concept of production efficiency laid the foundation for input-output efficiency evaluation<sup>[5]</sup>. After Charnes, Cooper, and Rhodes proposed the DEA model, relative efficiency evaluation under multiple input and multiple output conditions was widely used<sup>[6]</sup>. Banker, Charnes, and Cooper further distinguished between technical efficiency and scale efficiency, making the decomposition of enterprise efficiency more detailed<sup>[7]</sup>. The SFA method proposed by Aigner, Lovell, and Schmidt provides a measurement tool for efficiency frontier analysis by separating random error and inefficiency terms<sup>[8]</sup>. In the study of enterprise investment and financing efficiency, the efficiency value can be calculated first using DEA or SFA, and then enterprises can be divided into high-efficiency and low-efficiency states according to the industry median, sample quantiles, or theoretical thresholds. Logistic regression can then be used to explain the efficiency differences.

The third category is labeling based on the degree of investment deviation. Corporate investment efficiency is often measured by the deviation between actual and expected investment. If actual investment is significantly higher than the reasonable investment level, it can be considered overinvestment; if actual investment is significantly lower than the reasonable investment level, it can be considered underinvestment; if the deviation is small, it can be considered moderate investment. This method is suitable for further analysis using binary logistic or multinomial logistic models to determine the probability of a firm entering different investment states.

### 3.3 The Logic of the Logistic Regression Model

Logistic regression models are suitable for research scenarios where the dependent variable is binary or multi-class. McFadden's research on conditional logit models provides an important methodological basis for discrete choice behavior analysis<sup>[9]</sup>. In the measurement of corporate investment and financing efficiency, the model usually sets whether the enterprise is in an inefficient state, whether it is restricted in financing, whether it is underinvested or overinvested as the dependent variable, and explains its probability of occurrence through financial characteristics, governance structure and external environmental variables.

Hosmer, Lemeshow, and Sturdivant systematically summarized the modeling process of Logistic regression, including variable selection, model estimation, goodness-of-fit test, classification threshold setting, and prediction effect evaluation<sup>[10]</sup>. Based on research on corporate investment and financing efficiency, the basic model can be set as follows:

Corporate investment and financing efficiency status =  $f$  (debt repayment ability, profitability, cash flow, growth potential, financing cost, capital structure, equity structure, governance level, industry environment, macro-financial conditions).

The dependent variable is typically set as a 0-1 variable. For example, if a company is in a state of low investment and financing efficiency, it is assigned a value of 1; otherwise, it is assigned a value of 0. Explanatory variables can be divided into five categories based on theoretical mechanisms: First, solvency variables, such as debt-to-equity ratio, current ratio, and interest coverage ratio; second, profitability variables, such as return on equity, return on total assets, and operating profit margin; third, cash flow variables, such as operating cash flow, free cash flow, and cash holding level; fourth, growth variables, such as revenue growth rate, total asset growth rate, and Tobin's Q; and fifth, governance and external environment variables, such as equity concentration, board independence, audit quality, company nature, regional financial development level, and industry competition level.

The key to interpreting logistic regression models lies not in simply determining significance, but in analyzing the probabilistic impact of variables on a firm's efficiency level. For example, a significantly positive debt-to-equity ratio may indicate that higher debt pressure makes a firm more prone to inefficiency; significantly negative operating cash flow may suggest that ample cash flow reduces the likelihood of financing constraints or underinvestment; and significantly negative equity concentration may indicate that effective oversight helps curb inefficient investment. However, these interpretations

must be considered in conjunction with the firm's industry, life cycle, and financing environment, and cannot be applied mechanically.

### 3.4 Application in measuring financing efficiency

In the measurement of financing efficiency, Logistic regression is mainly used to identify whether a company has limited financing, excessive financing costs, or an unbalanced financing structure. Ohlson built a corporate bankruptcy probability prediction model based on financial ratios, proving that financial status can be effectively identified through probability models<sup>[11]</sup>. This idea can be transferred to the study of financing efficiency, that is, to judge whether a company is in a state of inefficient financing by using corporate financial indicators and market variables.

The dependent variable for financing efficiency can be set according to different research objectives. For example, companies with financing costs significantly higher than the industry average can be classified as financing inefficient companies; companies with long-term cash shortages, excessive debt burdens, and weak external financing capabilities can be classified as financing-restricted companies. Explanatory variables may include debt-to-equity ratio, short-term debt ratio, interest expense ratio, cash flow fluctuation, proportion of collateralized assets, credit rating, ownership structure, and bank relationships. If the model results show that the proportion of short-term debt, interest expense ratio, and cash flow fluctuation have a significant positive impact on the probability of financing inefficiency, it indicates that the decline in corporate financing efficiency mainly stems from an unreasonable debt structure and unstable cash flow.

Shumway's dynamic risk modeling approach suggests that corporate financial risk is not a static event, but rather something that accumulates gradually over a continuous operating cycle<sup>[12]</sup>. Therefore, the measurement of corporate financing efficiency should not rely solely on data from a single year. A more reasonable approach is to construct a panel logistic model to observe changes in a company's financing status over several years and identify the key triggering factors that cause a company to shift from smooth financing to restricted financing.

### 3.5 Application in Investment Efficiency Measurement

In the measurement of investment efficiency, Logistic regression is mainly used to determine whether a firm is underinvesting or overinvesting. Richardson identified the problem of overinvestment by using the expected investment model and linked free cash flow with agency conflict<sup>[13]</sup>. This method provides a clear path for constructing the state of investment efficiency: first, estimate the reasonable investment level of the firm, then calculate the deviation between actual investment and expected investment, and finally convert the deviation into a categorical variable.

In practical modeling, investment efficiency can be categorized into three states: underinvestment, moderate investment, and overinvestment. Using a binary logistic model, separate models can be constructed for "underinvestment" and "overinvestment." Using a multinomial logistic model, the probabilities of firms entering these three investment states can be compared simultaneously. Explanatory variables may include free cash flow, growth opportunities, debt-to-equity ratio, management shareholding ratio, equity concentration, internal control quality, and external oversight level.

Biddle, Hilary, and Verdi found that high-quality financial reporting helps improve corporate investment efficiency<sup>[14]</sup>. This indicates that information quality is an important variable affecting investment efficiency. The higher the quality of corporate information disclosure, the easier it is for external investors and creditors to judge the true operating conditions of the company, the lower the company's financing costs may be, and the space for management to make inefficient investments will be compressed. Chen, Hope, Li, and Wang further found in emerging market private companies that the impact of financial reporting quality on investment efficiency is affected by the institutional environment<sup>[15]</sup>. Therefore, when using logistic regression to analyze investment efficiency, information disclosure quality, audit quality, and institutional environment variables should be included in the model.

### 3.6 Application in measuring the efficiency of investment and financing matching

Investment and financing matching efficiency emphasizes whether the financing results are coordinated with the investment results. If, after obtaining funds, the investment returns can cover the cost of capital and generate stable cash flow and enterprise value growth, it indicates a high degree of investment and financing matching. Conversely, if the enterprise expands its financing scale but investment returns decline, debt pressure increases, or cash flow deteriorates, it indicates that the enterprise may have a mismatch of funds.

Logistic regression models can identify the matching status of investment and financing by constructing composite labels. For example, companies with "high financing costs and low investment returns" can be defined as inefficient companies in investment and financing, while companies with "reasonable financing costs and stable investment returns" can be

defined as efficient companies in investment and financing. In addition to traditional financial indicators, the explanatory variables of the model should also include variables such as investment opportunities, industry prosperity, corporate life cycle, governance structure, and macroeconomic financial environment. This can avoid evaluating corporate efficiency in isolation from the financing or investment side.

Kaplan and Zingales pointed out that investment cash flow sensitivity is not always a reliable measure of financing constraints<sup>[16]</sup>. This view has important implications for the measurement of investment and financing efficiency: research cannot simply assume that the greater the impact of cash flow on investment, the stronger the corporate financing constraints. Whited and Wu constructed a financing constraint index, providing a more comprehensive indicator design approach for measuring corporate external financing constraints<sup>[17]</sup>. Therefore, in Logistic modeling, multidimensional indicators should be used as much as possible to construct efficiency labels, rather than relying on a single financial variable.

### 3.7 Model Evaluation and Result Interpretation

When applying logistic regression models to measure corporate investment and financing efficiency, both explanatory and predictive capabilities should be considered. Explanatory capability is mainly reflected in the direction of variable coefficients, significance, and odds ratio; predictive capability is mainly reflected in classification accuracy, AUC, recall, precision, and confusion matrix. If the research objective leans towards theoretical explanation, the focus should be on analyzing the mechanism by which variables affect efficiency status; if the research objective leans towards corporate early warning, the focus should be on the model's ability to identify inefficient enterprises.

Lessmann, Baensens, Seow, and Thomas, after comparing various credit scoring classification algorithms, pointed out that complex classification models may outperform traditional models in terms of prediction performance<sup>[18]</sup>. However, in the study of corporate investment and financing efficiency, prediction accuracy is not the only standard. The advantage of Logistic regression lies in its transparent model, stable results, and clear interpretation path, making it suitable for academic research, corporate diagnosis, and risk assessment by financial institutions. Brown and Mues' research on class imbalanced data also reminds that when there are few samples of inefficient enterprises, the overall accuracy should not be reported alone, otherwise the model may tend to predict the majority of categories, thereby masking the truly risky enterprises that need to be identified<sup>[19]</sup>. Figure 3 shows the basic process of classification models from data partitioning, model training, validation set prediction to test set evaluation, reflecting the general path for comparing the performance of single classification models such as Logistic regression with ensemble classification models. For the measurement of corporate investment and financing efficiency, this process can provide a methodological reference for identifying inefficient enterprises, judging the state of financing constraints, and classifying investment deviations. Therefore, the Logistic regression model can not only be used to explain the influencing factors of corporate efficiency status, but also serve as a classification benchmark model for comparing the prediction performance with other machine learning methods.

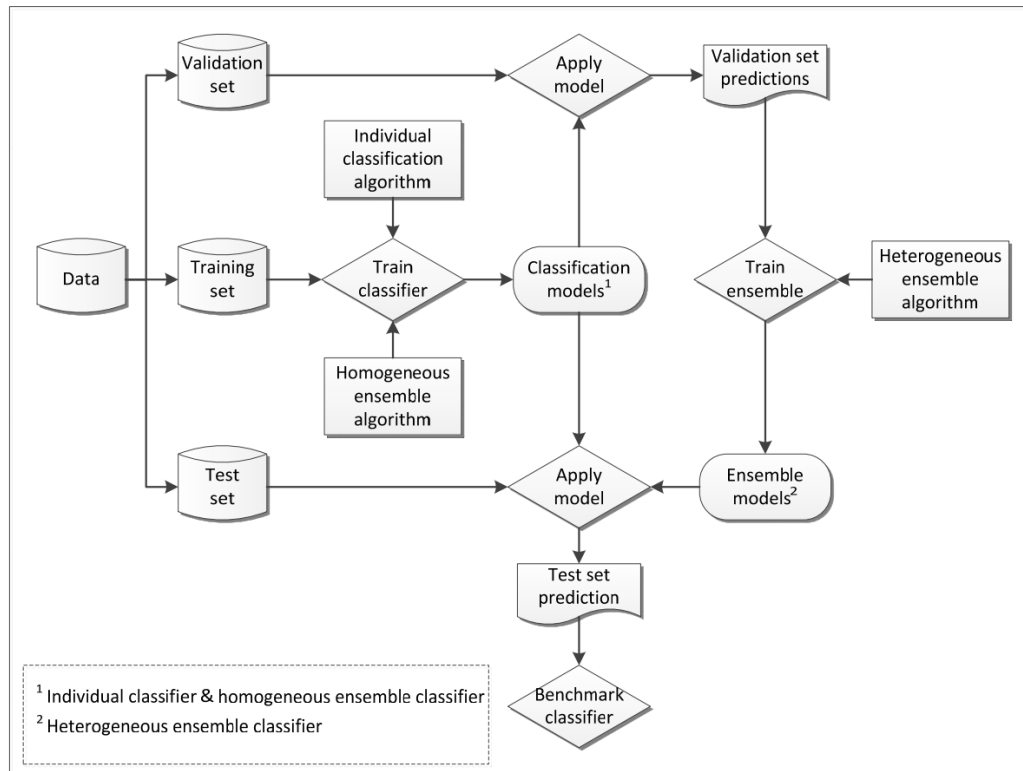


Figure 3. Training, Ensemble and Benchmarking Procedure for Classification Models <sup>[18]</sup>.

Therefore, a more reasonable evaluation method is as follows: First, report the overall fit of the model; second, examine whether the direction of the core variables conforms to theoretical expectations; third, use AUC, recall, and precision to evaluate classification ability; fourth, conduct sensitivity analysis on misclassification results under different thresholds; and fifth, conduct grouped tests in conjunction with industry and enterprise life cycles. Only when both theoretical rationality and predictive effectiveness are satisfied can the conclusions of the Logistic Regression model in measuring corporate investment and financing efficiency have high credibility.

## 4. CURRENT RESEARCH DEFICIENCIES AND DEVELOPMENT TRENDS

### 4.1 Current research is insufficient

Existing research has four main shortcomings. First, the construction of efficiency status labels is not rigorous enough; some studies directly classify efficiency as high or low based on the sample mean or median, lacking sufficient theoretical basis. Second, variable selection leans towards traditional financial indicators, with insufficient consideration for governance structure, information disclosure quality, financing channels, industry cycles, and regional financial environment. Third, models mostly use static cross-sectional data, making it difficult to reflect the dynamic changes in corporate investment and financing efficiency. Fourth, model evaluation indicators are not comprehensive enough; some studies only report accuracy rates, ignoring class imbalance and the cost of misjudgment.

### 4.2 Development Trends

Future research should focus on four key areas. First, constructing more reliable efficiency labels, which can combine DEA, SFA, investment deviation models, and financing constraint indices to form composite classification standards. Second, introducing panel logistic, dynamic logistic, and multinomial logistic models to characterize the continuous changes in corporate efficiency. Third, expanding data sources by combining financial statements, annual reports, audit opinions, ESG information, supply chain data, and financial market data. Fourth, comparing and integrating logistic regression with machine learning methods to improve predictive performance while maintaining interpretability.

## 5. CONCLUSION

The core value of the Logistic Regression model in measuring corporate investment and financing efficiency lies in transforming the complex investment and financing performance of enterprises into an interpretable problem of efficiency status identification. It does not directly replace DEA, SFA, or investment deviation models, but is more suitable for analyzing the probability and influencing factors of enterprises entering a state of high efficiency, low efficiency, financing constraints, underinvestment, or overinvestment after efficiency labels have been formed.

From an application perspective, a more reasonable approach is to first clarify the connotation of corporate investment and financing efficiency, then construct dependent variables based on financial indicators, efficiency models, or the degree of investment deviation. Subsequently, variables such as solvency, profitability, cash flow, growth potential, governance structure, and external environment should be selected to build a Logistic regression model. Finally, the effectiveness of the model should be judged through coefficient interpretation, advantage ratio analysis, and classification effect evaluation. Compared with complex machine learning models, although Logistic regression has shortcomings in capturing nonlinearity, it has strong transparency, stability, and economic interpretability, making it particularly suitable for corporate financial diagnosis, financing risk identification, and investment efficiency evaluation.

Overall, the measurement of corporate investment and financing efficiency is shifting from single-indicator evaluation to multi-method integrated analysis. Future research should place greater emphasis on defining efficiency status, dynamic data modeling, sample heterogeneity control, and practical scenario transformation, so that Logistic regression models can more effectively serve corporate capital allocation optimization and financial risk management.

## REFERENCES

- [1] Modigliani, F., Miller, MH The Cost of Capital, Corporation Finance and the Theory of Investment. *American Economic Review*, 48(3), 261-297(1958).
- [2] Myers, SC, Majluf, NS Corporate Financing and Investment Decisions When Firms Have Information That Investors Do Not Have. *Journal of Financial Economics*, 13(2), 187-221(1984).
- [3] Stiglitz, JE, Weiss, A. Credit Rationing in Markets with Imperfect Information. *American Economic Review*, 71(3), 393-410(1981).
- [4] Fazzari, SM, Hubbard, RG, Petersen, BC Financing Constraints and Corporate Investment. *Brookings Papers on Economic Activity*, 1988(1), 141-206(1988).
- [5] Farrell, MJ The Measurement of Productive Efficiency. *Journal of the Royal Statistical Society Series A*, 120(3), 253-290(1957).
- [6] Charnes, A., Cooper, WW, Rhodes, E. Measuring the Efficiency of Decision Making Units. *European Journal of Operational Research*, 2(6), 429-444(1978).
- [7] Banker, RD, Charnes, A., Cooper, WW Some Models for Estimating Technical and Scale Inefficiencies in Data Envelopment Analysis. *Management Science*, 30(9), 1078-1092(1984).
- [8] Aigner, D., Lovell, CAK, Schmidt, P. Formulation and Estimation of Stochastic Frontier Production Function Models. *Journal of Econometrics*, 6(1), 21-37(1977).
- [9] McFadden, D. Conditional Logit Analysis of Qualitative Choice Behavior. In Zarembka, P. ed. *Frontiers in Econometrics*. New York: Academic Press, 105-142(1974).
- [10] Hosmer, DW, Lemeshow, S., Sturdivant, RX *Applied Logistic Regression*. 3rd ed. Hoboken: John Wiley & Sons (2013).
- [11] Ohlson, JA Financial Ratios and the Probabilistic Prediction of Bankruptcy. *Journal of Accounting Research*, 18(1), 109-131(1980).
- [12] Shumway, T. Forecasting Bankruptcy More Accurately: A Simple Hazard Model. *Journal of Business*, 74(1), 101-124(2001).
- [13] Richardson, S. Over-Investment of Free Cash Flow. *Review of Accounting Studies*, 11(2-3), 159-189(2006).
- [14] Biddle, GC, Hilary, G., Verdi, RS How Does Financial Reporting Quality Relate to Investment Efficiency? *Journal of Accounting and Economics*, 48(2-3), 112-131(2009).
- [15] Chen, F., Hope, OK, Li, Q., Wang, X. Financial Reporting Quality and Investment Efficiency of Private Firms in Emerging Markets. *The Accounting Review*, 86(4), 1255-1288(2011).

- [16] Kaplan, SN, Zingales, L. Do Investment-Cash Flow Sensitivities Provide Useful Measures of Financing Constraints? Quarterly Journal of Economics, 112(1), 169-215(1997).
- [17] Whited, TM, Wu, G. Financial Constraints Risk. Review of Financial Studies, 19(2), 531-559(2006).
- [18] Lessmann, S., Baesens, B., Seow, HV, Thomas, LC Benchmarking State-of-the-Art Classification Algorithms for Credit Scoring: An Update of Research. European Journal of Operational Research, 247(1), 124-136(2015).
- [19] Brown, I., Mues, C. An Experimental Comparison of Classification Algorithms for Imbalanced Credit Scoring Data Sets. Expert Systems with Applications, 39(3), 3446-3453(2012).