J. COMBIN. MATH. COMBIN. COMPUT. 127b (2025) 507--522

Journal of Combinatorial Mathematics and Combinatorial Computing
www.combinatorialpress.com/jcmcc

Combinatorial Press

Machine Learning Strategies and Realization Paths for Internal Audit Quality Improvement Enabled by New Quality Productivity

Liying Zheng[1], Juanjuan Liang[1,✉]

[1] *School of Management Engineering, Guilin University, Guilin, Guangxi, 541000, China*

ABSTRACT

This paper takes 2020-2022 Shanghai main board listed companies as the research object, and empirically examines the relationship between factors such as the establishment of internal audit department and the quality of internal audit empowered by new quality productivity, with the effectiveness of the quality of internal control as the explanatory variable, the degree of separation of two powers and so on as the explanatory variable, and the corporate governance structure as the control variable to carry out a via gradient descent Logistic regression analysis optimized by gradient descent algorithm. On this basis, to address the problem that internal audit is prone to bias or falsehood due to management's self-interest, the fsQCA method is combined to analyze the influencing factors of the choice of auditing policy (capital item or expense item) for general R&D expenditures. It is found that there is a significant positive relationship between companies with an internal audit department and a higher hierarchical level of affiliation and obtaining a standard audit opinion, and the regression relationship holds at the 0.05 level of significance, with a positive correlation with a regression coefficient of 3.745, and an OR value of 40.099. However, the effect of the company's twofold separation of powers governance structure on the quality of the audit fails the significance test. Firms with lower profitability levels, higher R&D intensity, higher debt levels, lower tax benefits for R&D additions and deductions and lower external audit quality are more likely to capitalize R&D expenditures. The study uses cutting-edge algorithms to accurately analyze new quality productivity-enabling internal audit quality factors and innovate corporate compliance internal control paths.

Keywords: internal audit, equity concentration, logistic regression, gradient descent, fsQCA

1. Introduction

New quality productivity refers to the productivity based on new technologies, new modes and new ideas, which is the upgrading and expansion of traditional productivity [1]. With the continuous

✉ Corresponding author.
E-mail address: jjbango@glc.edu.cn (J. Liang).
Received 10 January 2024; Revised 25 May 2024; Accepted 20 October 2024; Published Online 16 April 2025.
DOI: 10.61091/jcmcc127b-029

progress and innovation of science and technology, the new quality of productivity continues to emerge, bringing more efficient, smarter and more sustainable production methods and methods for enterprises, prompting an unprecedented change in enterprise financial management. Mainly manifested in the promotion and application of financial shared service model, which will be scattered in different regions, different departments of the financial function centralized, standardized, professional processing methods to improve the efficiency of the internal audit work, and reduce the operating costs of enterprises, and further enhance the accuracy, transparency, timeliness of financial data [2-3]. However, the implementation of this management model has brought new challenges to the internal control and audit work of enterprises. Under the financial shared service model, the centralized and standardized processing of financial information within the enterprise provides a rich data base for intelligent auditing [4]. Intelligent auditing realizes rapid analysis and accurate identification of massive financial data through the use of big data, artificial intelligence and other advanced technologies, thus helping enterprises discover potential risks in a timely manner and improving audit efficiency and quality [5]. At the same time, intelligent auditing also helps to promote the transformation of enterprise internal auditing to the whole process of pre-warning, monitoring and analyzing afterwards, providing a strong guarantee for the healthy development of enterprises [6].

In the context of the new quality productivity, enterprises have a more comprehensive understanding of financial management, enterprise internal audit, etc., which promotes the establishment of a perfect audit system and standards, and further enhances the rigor and reliability of the audit work. Literature [7] uses machine learning algorithms to create intelligent models with efficient identification, assessment, and very low audit risk to provide a roadmap for corporate decision makers to optimize the audit risk management process, which avoids potential non-compliance issues in traditional audit risk management systems due to the complexity of corporate transactions. Literature [8] uses digital tools to propose responsive digital solutions for different stages of internal auditing in the petroleum industry to facilitate the achievement of strategic business objectives in the petroleum industry by assisting in the development of innovative business practices and processes through trend analysis and testing internal controls. Literature [9] emphasizes the importance of big data thinking in the context of the current era and proposes an enterprise internal audit system based on big data analysis technology, which not only overcomes the problems existing in the traditional enterprise internal audit system, but also significantly improves the efficiency and quality of internal auditing using digital technology. Literature [10] shows that the enterprise internal audit system is the optimal method for solving accounting conflicts in enterprises, while improving the digitalization ability of internal auditors through the development of intelligent tools based on artificial neural networks and machine training can effectively improve the efficiency of internal auditing and further enhance the efficiency of handling accounting conflicts. Literature [11] investigated an automated supervisory modeling design methodology aimed at reducing internal audit risks in enterprises by establishing an audit-related architecture that integrates supervisory machine learning to provide internal audit staff with automated aids that enhance their analytical skills and report more potential risks for decision makers. Literature [12] proposes the digital transformation of auditing based on machine learning artificial intelligence to solve the complexity and low fault-tolerance problems of internal auditing in enterprises, and experiments show that the intelligent internal auditing system effectively improves the auditing accuracy as well as work efficiency. Literature [13] discusses the specific functions of data analytics and machine learning capabilities linked to enterprise internal auditing, machine learning technology not only provides full coverage for internal auditing, but also introduces objectivity in the analysis of key areas, so that intelligent analytics can be used to promote the innovation of services in the various stages of the internal audit life cycle. Literature [14] clarifies that machine learning technology has powerful data processing and pattern recognition capabilities, which can effectively respond to the internal audit tasks of enterprises in complex and changing environments, and proposes an internal audit system based on hybrid intelligent optimization algorithms, which provides technical support for improving the level of audit intelligence, reducing audit costs, and improving the quality of audit. Literature [15] integrates process mining and machine learning into the internal control evaluation model of the audit process, which can not only discover the audit transactions that deviate from the standardized process, but also quantitatively measure the level of the transaction, so as to delineate the audit anomaly area with high risk.

Based on the principle of new quality productivity-enabled internal auditing, this paper puts forward relevant research hypotheses and selects the sample data of listed companies on the main board of the Shanghai market from 2020 to 2022, constructs the logistic regression model, and utilizes machine learning algorithms to mine the factors affecting the quality of internal auditing with new quality productivity-enabled internal auditing. For the loss function of logistic regression, the gradient descent method is used to solve the parameters. SPSS is used to carry out descriptive statistics for each dependent variable designed, and the sample is analyzed by regression. Finally, the key factors influencing companies to make accounting policy choices for R&D expenditures are analyzed.

2. New quality productivity-enabling internal audit assumptions and logistic regression modeling

2.1. Research hypothesis and definition of variables

2.1.1. Formulation of the original hypothesis

1) The higher the status of the internal audit of the standardized company, the stronger the independence, the more beneficial to play the role of corporate governance, and the strong independence of the internal audit, indicating that the company's system is standardized, the control system is sound, the existence of the risk of material misstatement is the smaller, the more beneficial to the certified public accountants to issue a high standard of audit opinion. The higher the level of the internal audit department, the stronger the independence of the company's internal audit organization, the more beneficial to the role of internal audit, the more conducive to improving the quality of internal audit [16]. Measured by company size and degree of standardization as quantitative indicators, denoted as CS. therefore the following hypotheses are proposed:

Hypothesis 1 Internal audit affiliation level and the establishment of internal audit department are positively related to internal audit quality, i.e., the higher the internal audit department is established and the higher the internal audit department's affiliation level is, the higher the audit quality is.

2) Chinese listed companies usually combine the chairman of the board with the general manager. The chairman of the board also serves as the general manager, which violates the basic principles of internal control of listed companies, thus adversely affecting the quality of internal auditing. When there is a situation in which the two positions are combined, it is easy to promote the motives and behaviors of managers to harm the interests of the company. Managers, for their own interests, will inhibit the internal audit department from exercising its supervisory function, limit the scope of responsibilities of the internal audit department, weaken the internal audit department's ability to identify the company's operational weaknesses, and is not conducive to the new quality productivity empowering the internal audit governance effect. The chairman of the board and the general manager represent the company's control and claiming power respectively, and when the two powers are separated to a large extent, the appropriation motive of the major shareholders and the fraud of the internal managers will form the resistance to the fact of internal control self-evaluation. The degree of separation of powers is measured by whether the chairman and general manager are united or separated, which is denoted as PS. therefore, the following hypotheses are proposed:

Hypothesis 2 The quality of internal audit when the two positions of chairman and general manager are held by different people is higher than the quality of internal audit when the two positions are combined with the new quality of productivity empowerment. That is, the higher the audit quality when the positions of chairman and managing director are not held simultaneously.

3) The independent director system is a supervisory system created to protect the interests of shareholders of listed companies. In the governance mechanism of a listed company, independent directors, because of their independence from the listed company, can be unaffected in the process of exercising their supervisory function over the behavior of the operator. However, China's independent director system is far from achieving the expected goal, showing the phenomenon of independent directors being assimilated by "insiders", and some of them are even in collusion with the major shareholders and management, accepting bribes from the management, and "acquiescing" in their irregularities and participating in the board meetings, Exercising the power of independent directors is often a mere formality. Measured by the proportion of independent directors who truly fulfill their functions, it is denoted as IP, so the following hypothesis is proposed:

Hypothesis 3 There is no significant correlation between the proportion of independent directors and internal audit quality.

4) Since China's state-owned holding companies, the ownership of which belongs to the national government, cannot directly manage the enterprise, they can only manage the enterprise by appointing government officials as the representatives of the owners of state-owned assets. However, for the implementation of the management of government officials, there is no residual claim to enterprise assets, the lack of enthusiasm for enterprise management, in order to pursue their own interests, will certainly be on the new quality of productivity-enabling internal audit work to obstruct, so that the role of internal audit can not be well played, thus reducing the quality of internal audit. Take whether it is state-owned control as the research variable, denoted as OP. based on this, the following hypotheses are proposed:

Hypothesis 4 The actual control category of listed companies, state-owned holding is negatively related to the quality of internal audit.

5) In the shareholding structure of the company, the more concentrated the shareholding, the majority of the shareholding is occupied by a few owners. An overly concentrated shareholding structure is not an effective mechanism by which the general meeting of shareholders is likely to be manipulated, and the majority shareholders are free to divert the company's resources for their own interests, at the expense of small and medium shareholders [17]. The possibility of whitewashing financial statements and fraud increases, which in turn affects the audit opinion. The shareholding percentage of the top ten shareholders is taken as the research variable and the mode of influence is positive, which is denoted as CR.

Hypothesis 5 Equity concentration is negatively related to the quality of internal audit of new quality productivity empowerment.

6) The degree of equity checks and balances refers to the degree to which the proportion of shares held by each shareholder in the shareholding structure of a listed company constrains each other. If the proportion of shares held by each shareholder is relatively close to each other, a stronger equity check and balance can be formed, and the more likely the listed company will be fair and equitable in corporate governance. Therefore, internal audit, as an important part of corporate governance, will also be affected by the level of equity checks and balances. The higher the degree of equity checks and balances, the more likely the shareholders will pay attention to the work of the internal audit, so as to ensure their own economic interests and improve the quality of internal audit. The number of the board of directors and the average shareholding of shareholders are taken as variables, denoted as BS.

Hypothesis 6 The degree of equity checks and balances is positively related to internal audit quality.

7) Audit costs include internal audit costs and external audit costs. In order to improve audit quality, listed companies are bound to invest corresponding human and material resources, and if necessary, may also seek expert advice to make suggestions on the design of internal control, and audit costs to a certain extent reflect the necessary financial resources invested in improving audit quality. The existence of internal control consulting and service fees is taken as the research variable, denoted as

CC. therefore the following hypothesis is proposed:

Hypothesis 7 Audit fees are positively related to internal audit quality.

In addition another financial risk and profitability are set as reference variables, and debt ratio and return on net worth are taken as values, denoted as Z and SS, respectively.

2.1.2. Research Sample and Data Sources. The Shanghai Stock Exchange has stipulated in the Guidelines on Internal Audit Work of Listed Companies on Small and Medium-sized Enterprises Board that listed companies in small and medium-sized enterprises must set up an internal audit department, which is responsible to the Audit Committee and reports to the Audit Committee, and Article 8 stipulates that listed companies shall, in accordance with the size of the company, the characteristics of production and operation as well as the relevant regulations, allocate full-time staff to engage in the work of internal auditing, and that the full-time staff should be not less than 3 persons. For main board listed companies, it is stipulated that enterprises should set up an audit committee under the board of directors, and set up internal organizations according to the needs of business characteristics, which shows that it is not mandatory for main board listed companies to set up a new quality productivity-enabling internal auditing department, therefore, this paper chooses main board listed companies in Shanghai as the object of the study, and selects the sample data from 2020 to 2022.

2.2. Logistic regression modeling

2.2.1. Logistic regression model derivation process. The logistic regression model is a generalized linear model, and we abbreviate the linear regression model as equation (1). The model shown in equation (2) is called a "generalized linear model", and the function is called a "link function".

$$y = \omega^T x + b \tag{1}$$

$$y = g^{-1}(\omega^T x + b) \tag{2}$$

Considering the binary classification task, the output is $y \in \{0,1\}$ and the linear regression model produces the predicted values as shown in equation (3):

$$z = \omega^T x + b \tag{3}$$

The values are real and need to be converted to 0/1 values, ideally as a unit step function, as shown in equation (4):

$$y = \begin{cases} 0, z < 0; \\ 0.5, z = 0; \\ 1, z > 0. \end{cases} \tag{4}$$

In the unit step function, if the predicted value is greater than 0, it is judged as a positive case; if the predicted value is less than 0, it is judged as a negative case, and when the predicted value is a critical value of 0, it can be judged arbitrarily [18].

Since the unit step function is not continuous, we hope that we can find the cost function that can approximate the unit step function to some extent.

The logistic regression function, as shown in equation (5):

$$y = \frac{1}{1 + e^{-z}} \tag{5}$$

This happens to be a common alternative function, and it can be seen that the logistic regression function is a "Sigmoid" function, which converts a value of z to a value y close to 0 or 1, and the output value varies very steeply around $z = 0$.

Bringing the logistic regression function into the linear model yields:

$$y = \frac{1}{1+e^{-(\omega^T x + b)}} \tag{6}$$

Eq. (6) can be transformed into:

$$\ln \frac{y}{1-y} = \omega^T x + b \tag{7}$$

The objective function of the logistic regression solution is a convex function that is integrable of any order, and y in Eq. (7) can be regarded as a class of a posteriori probabilities to estimate $p(y=1 \mid x)$. The logarithmically changed Eq. is then rewritten as:

$$\ln \frac{p(y=1 \mid x)}{p(y=0 \mid x)} = \omega^T x + b \tag{8}$$

According to Eq. (8) the probability of belonging to a positive case for the input vector x can be obtained as shown in Eq. (9);

$$p(y=1 \mid x) = \frac{e^{\omega^T x + b}}{1+e^{\omega^T x + b}} = \sigma(\omega^T x + b) \tag{9}$$

The probability of being a counterexample is shown in equation (10):

$$p(y=0 \mid x) = \frac{1}{1+e^{\omega^T x + b}} = 1 - \sigma(\omega^T x + b) \tag{10}$$

The unknown parameters ω and b can be estimated by the method of maximum likelihood. Given dataset $\{(x_i, y_i)\}_{i=1}^n$, the logistic regression model maximizes the "log likelihood":

$$\ell(\omega, b) = \sum_{i=1}^{n} \ln p(y_i \mid x_i; \omega, b) \tag{11}$$

Order:

$$\begin{aligned} &\beta = (\omega; b) \\ &\hat{x} = (x; l) \\ &p_1(\hat{x}; \beta) = p(y=1 \mid \hat{x}; \beta) \\ &p_0(\hat{x}; \beta) = p(y=0 \mid \hat{x}; \beta) = 1 - p_1(\hat{x}; \beta) \end{aligned} \tag{12}$$

Then:

$\omega^T x + b$ becomes $\beta^T \hat{x}$, the likelihood term in Eq. (11) can be rewritten in the form of Eq. (13):

$$p(y_i \mid x_i; o, b) = y_i p_1(\hat{x}_i; \beta) + (1-y_i) p_0(\hat{x}_i; \beta) \tag{13}$$

Maximizing the likelihood function (11), which is equivalent to minimizing Eq:

$$\ell(\beta) = \sum_{i=1}^{n} (-y_i \beta^T \hat{x}_i + \ln(1 + e^{\beta^T \hat{x}_i})) \tag{14}$$

The above equation is a higher order derivable continuous convex function with respect to β. Classical numerical optimization algorithms such as Newton's method and gradient descent method can find its optimal solution, $\beta^* = \arg\min_n \ell(\beta)$.

2.2.2. Gradient descent algorithm. Through the initial point in each iteration process to select the corresponding gradient direction to find new iteration points, and then the iteration point as the current point to continue to find new iteration points, and then to change the parameters that need to be modified until the optimal solution is found under the given conditions, the specific process is shown below:

First, an initial point β_0 is selected, and the iteration process is based on equation (15):

$$d_i = -\frac{\partial}{\partial \beta} L(\beta) | \beta_i \tag{15}$$

The negative gradient direction of the loss function is solved to update the value of the requested parameter, and the step parameter of the parameter update is specified as α. The iteration for β_{i+1} is performed according to $\beta_{i+1} = \beta_i + \alpha d_i$ until the termination condition is satisfied.

For the loss function $\ell(w,b)$ of logistic regression, the parameters are solved using the gradient descent method as shown in Eqs. (16) and (17):

$$\nabla \omega_j (\ell_\beta) = -\frac{1}{n} \sum_{i=1}^{n} \left(y_i - \frac{e^{\beta^T \hat{x}_i}}{1 + e^{\beta^T \hat{x}_i}} \right) x_{ij} \tag{16}$$

$$\nabla b_j (\ell_\beta) = -\frac{1}{n} \sum_{i=1}^{n} \left(y_i - \frac{e^{\beta^T \hat{x}_i}}{1 + e^{\beta^T \hat{x}_i}} \right) \tag{17}$$

where x_{ij} denotes the j rd component of sample x_i. When $\omega_0 = b, x_0 = 1$ is taken, the gradient formula (18) and the parameter iteration formula (19) are obtained:

$$\nabla \omega_j (\ell_\beta) = -\frac{1}{n} \sum_{i=1}^{n} (y_i - \frac{e^{\beta^T \hat{x}_i}}{1 + e^{\beta^T \hat{x}_i}}) x_{ij} \tag{18}$$

$$\omega_{j+1} = \omega_j + \alpha \nabla \omega_j (\ell_\beta) \tag{19}$$

The training of the logistic regression model and the solution of the weights as well as the biases in the model can be achieved through the above process.

Advantages of the logistic regression model (LR): the results of the model are easy to understand, the coefficients of the independent variables are directly linked to the weights, which can be interpreted as the magnitude of the predictive value of the independent variables on the target variable; the model training is fast and efficient, and the related algorithms are easy to implement.

Disadvantages of the logistic regression model (LR): the number of samples corresponding to each category in the target variable should be sufficient to support the modeling: care should be taken to exclude the problem of covariance in the white variables; outliers can cause a great deal of interference in the model: the model itself cannot deal with missing values, and it needs to be preprocessed for missing values.

3. Empirical analysis and pathways to realization

3.1. Analysis of new quality productivity-enabled internal audit quality improvement

3.1.1. Descriptive analysis of continuous variables. SPSS was first utilized to perform descriptive statistics for each of the dependent variables in this paper, and in accordance with the assumptions made above, no hypothetical anomalies were found to be present. The descriptive results are shown in Table 1. The CS and PS means were relatively high, and the rest of the items had means below 10.0.

Table 1. Descriptive Statistics

CE		N	Min	Max	Mean	SD
0	CS	48	19.4785	26.6105	22.324	1.2888

	SS	48	0.9656	8.074	3.4691	1.1725
	Z	48	0.021	0.8954	0.4129	0.1763
	CR	48	0.2493	0.91	0.4713	0.1295
	BS	48	7.0259	15.127	9.7927	1.9442
	IP	48	0.297	0.5494	0.3738	0.0178
	PS	48	0.0199	35.7543	9.5933	9.4304
	Valid N(Listwise)	48				
1	CS	66	21.1947	27.4172	23.701	1.5994
	SS	66	3.001	9.0784	5.3685	1.8157
	Z	66	0.2523	3.5667	0.9249	0.6229
	CR	66	0.0364	0.8287	0.5187	0.1636
	BS	66	5.0064	15.1191	9.0356	1.9113
	IP	66	0.3017	0.6673	0.4219	0.0659
	PS	66	0.0306	35.7438	4.2489	8.0233
	Valid N(Listwise)	66				

3.1.2. Correlation analysis. Correlation analysis of categorical variables, two categorical variables CC (whether internal control consulting) and OP (nature of the ultimate controller) were correlated with 2×2 column table using SPSS respectively, and the results are shown in Table 2.

As can be seen from the above table, the correlation coefficient of CE and CC is 0.418, which is correlated at 0.001 level of significance, and the correlation coefficient of CE and OP is 0.251, which is correlated at 0.01 level of significance.

Table 2. Correlation analysis

CE * CC Crosstabulation						
		CC		Total	r	Sig
		Unconsulted	Consult			
CE	0	38	12	50	0.418	0
	1	15	48	63		
Total		53	60	113		
CE *OP Crosstabulation						
		CC		Total	r	Sig
		Unconsulted	Consult			
CE	Ineffectiveness	20	28	48	0.251	0.007
	In effect	12	56	68		
Total		32	84	116		

The correlations between the dependent variable and other continuous variables were analyzed by Spearman rank correlation and the results are shown in Table 3.

As can be seen from the table, the correlations between CE and CS.Z.IP.PS.SS all hold at the 0.05 level of significance; where there is a negative correlation with PS and a positive correlation with the other variables.

Table 3. Spearman

		CE	CS	Z	CR	BS	IP	PS	SS
CE	Correlation Coefficient	1							

CS	Correlation Coefficient	0.391**	1						
Z	Correlation Coefficient	0.547**	0.292**	1					
CR	Correlation Coefficient	0.07	0.04	0.075	1				
BS	Correlation Coefficient	-0.156	0.112	-0.148	-0.045	1			
IP	Correlation Coefficient	0.248*	0.007	0.124	0.008	-0.17	1		
PS	Correlation Coefficient	-0.415**	-0.104	-0.108	-0.348**	-0.134	-0.145	1	
SS	Correlation Coefficient	0.541**	0.244**	0.349**	0.037	-0.157	0.07	-0.188*	1

3.1.3. Model profile analysis. The chi-square value, coefficient of determination and p-value of the model were tested using SPSS and the results are shown in Table 4. The value of the chi-square is 114.238, the p-value is less than 0.05, it can be considered that the whole model is set up reasonably, and then look at the results of goodness-of-fit, the COX coefficient of determination is 0.664, 66.4% of the dependent variable can be explained by regression relationship with by the independent variable, and the model fit is good.

Table 4. Logistic regression

COX	Chi-square	P
0.664	114.238	0.000

3.1.4. Logistic regression. Combining the above influencing factors, the sample was analyzed by regression and the regression results are shown in Table 5. Based on the analysis results, the following conclusions are drawn:

1) The regression relationship between internal control audit quality effectiveness and internal control consulting holds at 0.05 level of significance, with a regression coefficient of 3.745, a positive correlation, and an OR value of 40.099, which means that internal control consulting positively affects internal control quality effectiveness.

2) The regression relationship between internal control quality effectiveness and company size is established at 0.1 level of significance, with a regression coefficient of 0.584 and a positive relationship, i.e., company size positively affects internal control quality effectiveness.

3) The regression relationship between internal control quality effectiveness and Z-value is established at the significance level of 0.1, and the regression coefficient is 4.286, which is positively correlated, i.e., Z-value positively affects internal control quality effectiveness.

4) The regression relationship between internal control quality effectiveness and the nature of the ultimate controller is established at the significance level of 0.05, with a regression coefficient of 2.893, a positive correlation, and an OR value of 17.39, i.e., the nature variable positively affects the internal control quality effectiveness.

5) The regression relationship between the effectiveness of internal control quality and the proportion of independent directors is established at the significance level of 0.1, and the regression coefficient is 13.609, which is positively correlated, i.e., the proportion of independent directors positively affects the effectiveness of internal control quality.

6) The regression relationship between the effectiveness of internal control quality and the degree of separation of the two powers is only established at the significance level of 0.01, with a regression

coefficient of -0.275, so the influence of whether the chairman of the board of directors and the general manager are the same person on the effectiveness of the quality of internal control cannot be established.

7) The regression relationship between internal control quality effectiveness and profitability is established at 0.05 level of significance, with a regression coefficient of 0.611, showing a positive correlation, i.e. profitability positively affects internal control quality effectiveness.

Table 5. Logistic regression

	B	S.E	Wald	df	Sig.	Exp(B)	95% C.I.for EXP(B)	
							Lower	Upper
CC	3.745	1.356	7.783	1	0.047	40.099	2.963	542.674
CS	0.584	0.304	3.77	1	0.056	1.808	1.009	3.243
Z	4.286	1.668	6.734	1	0.01	68.157	2.764	1680.738
OP	2.893	1.35	4.678	1	0.033	17.39	1.297	233.123
IP	13.609	7.54	3.323	1	0.072	635326.995	0.324	1.24E+12
PS	-0.275	0.088	9.947	1	0.002	0.778	0.657	0.922
SS	0.611	0.305	4.098	1	0.046	1.855	1.033	3.332
Constant	-25.756	8.057	10.422	1	0.001	0		

3.2. Factors Influencing the Selection of Accounting Strategies in New Quality Productivity Enabling Internal Audit

There are no clear criteria for distinguishing between capitalization and expensing of R&D expenditures, and the judgment of the conditions for capitalization still relies on the subjective judgments of the internal and external auditors of New Quality Productivity Empowerment (NQPE). Having too much autonomy has led to differences in accounting methods for R&D expenditures by enterprises at present, and the quality of internal control audits is mixed. Studying the key factors influencing enterprises to make accounting policy choices for R&D expenditures can not only improve the quality of internal auditing, but also provide an effective basis for investment decisions.

3.2.1. Data calibration and truth table construction. Data calibration. After systematic research and in-depth analysis, this paper proposes a set of comprehensive and practical data processing procedures based on the operation manual of the fuzzy set qualitative comparative analysis method, combing in detail the domestic and international research results about the method. After collecting the raw data, it should go through the steps of carefully calibrating the data, preparing the truth table, executing the necessity analysis and grouping analysis in order to realize the effective results.

The five research variables selected for this paper are continuous variables and must be transformed for data. This is done by characterizing the data in terms of three anchors: full affiliation, cross fuzzy point and full non-affiliation. According to the relevant research results, this paper sets the critical values of full affiliation and full non-affiliation as 95% and 5% of each variable, and takes the average of the two critical values as the crossover point as shown in Table 6.

Table 6. Calibration anchor

Variable type	Variable name	Anchor point		
		Full membership	Crossing point	Completely unaffiliated
Result variable	RDCAP	63.944	12.76	0
Conditional variable	PROFIT	0.377	0.095	0.002
	RDINT	28.835	8.517	0.765
	LEV	59.864	31.283	8.894

	TAX	1.53	0.5	0.051
	AUDIT	4.58	1.153	0.367

3.2.2. Truth table construction. By calibrating the data using the fsqca 3.0 software, we can combine multiple conditional variables to better demonstrate the case scenarios, constructing the truth table shown in Table 7. The truth table summarizes all possible logical combinations. Typically, there may be 32 (24) different combinations of the 5 condition variables, but only 31 are actually observed, leaving 1 as a logical remainder. In this paper, we set the consistency threshold to 0.8 and the case frequency threshold to 1 to ensure the accuracy and reliability of the analysis results.

Table 7. Case variable truth table

PROFIT	RDINT	LEV	TAX	AUDIT	RDCAP	N
0	1	1	1	0	1	5
0	1	0	1	0	0	5
1	1	1	0	1	0	7
0	1	0	0	0	0	2
1	1	0	1	0	1	4
0	0	1	1	0	0	6
0	1	0	0	1	0	3
0	0	1	0	0	1	2
1	1	1	1	0	1	2
0	1	0	0	1	0	4
0	1	1	0	0	0	1
1	0	0	1	1	0	6
0	1	1	0	1	0	1
1	1	1	1	0	1	4
0	0	0	1	0	0	5
1	1	0	0	1	0	3
0	0	1	0	0	0	5
1	1	0	0	0	0	3
0	0	0	1	0	0	1
1	0	0	0	1	1	4
0	0	1	0	1	0	4
0	0	0	1	0	0	1
1	0	0	0	1	1	8
0	0	1	0	0	0	9
0	0	0	0	0	0	1
1	0	0	1	0	1	3
0	0	0	0	1	0	2
0	0	1	0	1	1	4
1	0	0	1	0	0	3
0	0	0	0	0	1	3
1	0	0	0	0	0	4

3.2.3. Analysis of results. After calibrating the study variables, necessity analysis of a single variable is first required. The necessity analysis is divided into two perspectives: consistency and coverage.

In the study of qualitative comparative analysis of fuzzy sets, the first step is to determine that the consistency of the condition grouping pattern meets the requirements, and subsequently the coverage is judged. In this paper, fsQCA3.0 software is used to test whether five antecedent conditions are necessary to constitute the choice between capitalization and expensing of corporate R&D expenditures. If an antecedent condition constitutes a necessary condition for the outcome, it means that the condition is essential for the outcome to occur. The results of the necessity test are shown in Table 8.

As can be seen from the table, when the outcome variable is capitalized R&D expenditures, the variables with higher necessity are ~profitability level (~PROFIT, 0.854) and R&D intensity (RDINT, 0.832). When the outcome variable is expensed R&D expenditures, the variables with higher necessity are: profit level (PROFIT, 0.832), ~R&D intensity (~RDINT, 0.836). That is to say, among the factors affecting the accounting treatment of corporate R&D expenditures, the factors that have a higher necessity effect on capitalized R&D expenditures are: lower level of profitability, higher R&D intensity; and the factors that have a higher necessity effect on expensed R&D expenditures are: higher level of profitability, lower R&D intensity.

In summary, the consistency and coverage of the single condition variables are lower than 0.8, indicating that profitability level, R&D intensity, debt level, R&D plus deduction tax incentives, and the quality of external auditing all have insufficient explanatory power for the outcome variables, and cannot be the sufficient and necessary conditions to influence the enterprises to carry out the accounting treatment of R&D expenditures. Therefore, the accounting policy choice of R&D expenditures will be affected by the linkage of many factors, and it is necessary to explore this complex mechanism of action from a global perspective to improve the existing research on the accounting choice of R&D expenditures.

Table 8. Necessity analysis of a single variable

	RDCAP		~RDCAP	
	Consistency	coverage	Consistency	coverage
PROFIT	0.435	0.494	0.832	0.786
~PROFIT	0.854	0.603	0.482	0.622
RDINT	0.832	0.679	0.389	0.627
~RDINT	0.396	0.456	0.836	0.775
LEV	0.687	0.586	0.587	0.659
~LEV	0.602	0.521	0.727	0.731
TAX	0.686	0.643	0.509	0.644
~TAX	0.632	0.479	0.712	0.752
AUDIT	0.613	0.566	0.547	0.704
~AUDIT	0.701	0.538	0.655	0.69

3.2.4. Grouping Analysis of Multiple Variables. The study above shows that the choice of enterprises in capitalizing and expensing R&D expenditures is influenced by a variety of factors. In order to better study this influence, this paper uses fsQCA3.0 software, sets the consistency threshold to 0.8 and the case threshold to 1, and conducts an in-depth study on the linkage matching mechanism between variables, and analyzes the results of running the software.

1) Group analysis of capitalized R&D expenditure. There are three groupings of capitalized R&D expenditures derived from the analysis, as shown in Table 9. The total consistency is 0.778, indicating that 77.8% of the cases of pharmaceutical manufacturing companies that satisfy the grouping state H1, grouping state H2 and grouping state H3 choose capitalized R&D expenditures. The total coverage is 0.536, indicating that grouping state H1, grouping state H2 and grouping state H3 explain about 53.6% of the cases of capitalized R&D expenditures. Therefore, these three groupings are the main reasons that lead firms to choose capitalized R&D expenditures.

Table 9. The configuration of the result of the capitalization of the r&d expenditure

Conditional variable	Capitalization development expenditure configuration		
	H1	H2	H3
PROFIT	No	No	
RDINT	Yes	Yes	Yes
LEV	No	Yes	Yes
TAX	Yes	No	Yes
AUDIT			No
Consistency	0.821	0.803	0.84
Raw coverage	0.375	0.335	0.352
Unique coverage	0.101	0.069	0.063
Solution consistency	0.778		
Solution coverage	0.536		

2) Analysis of the grouping of expensed R&D expenditures. The analysis resulted in four grouping patterns that generate expensed R&D expenditures, as shown in Table 10. The total consistency is 0.877, indicating that 87.7% of the cases of pharmaceutical manufacturing companies that satisfy the grouping states NH1, NH2, NH3 and NH4 choose to expense R&D expenditures. The total coverage is 0.616, indicating that group states NH1, NH2, NH3 and NH4 explain about 61.6% of the cases of expensed R&D expenditures. Therefore, these four group states are the main reasons that lead companies to choose expensed R&D expenditures.

Table 10. The configuration of the cost of r&d expenditure

Conditional variable	Capitalization development expenditure configuration			
	NH1	NH2	NH3	NH4
PROFIT		Yes	No	Yes
RDINT	No	No	No	No
LEV	No			
TAX		No	No	Yes
AUDIT		No	Yes	Yes
Consistency	0.896	0.912	0.911	0.91
Raw coverage	0.492	0.41	0.312	0.293
Unique coverage	0.083	0.053	0.034	0.014
Solution consistency	0.877			
Solution coverage	0.616			

4. Conclusion

This paper constructs a logistic regression model to explore the correlation between factors such as the proportion of independent directors and the degree of separation of powers and the effectiveness of internal audit quality of new quality productivity empowerment, and analyzes the factors influencing the choice of capitalization and expensing of R&D expenditures - profitability level, R&D intensity, debt level, R&D plus deduction tax incentives, and external audit quality The Role of. It is found that the regression relationship between internal control quality effectiveness and company size holds at a significance level of 0.1, and the regression coefficient is 0.584, which is positively correlated, i.e., company size positively affects internal control quality effectiveness. And whether the chairman and general manager are the same person does not have a significant effect on the internal control quality effectiveness. Lower profitability level and higher R&D intensity create a stronger necessity for capitalized R&D expenditures, and higher profitability level and lower R&D intensity create a stronger necessity for expensed R&D expenditures.

Funding

1. The 2023 Guangxi Universities Young and Middle-aged Teachers' Basic Scientific Research Capacity Enhancement Project: Research on the Mechanism of the Impact of ESG Performance on Corporate Performance (Project Number: 2023KY1610).

2. Guangxi Education Science '14th Five-Year Plan' 2023 Special Financial Literacy Project: Research on the Mechanism of Financial Literacy Education Based on Social Capital Theory in Promoting Rural Revitalization in Guangxi (2023ZJY936).

3. 2024 Guangxi Higher Education Undergraduate Teaching Reform Project: One Main Line, Two Combinations, Three Stages, Four Levels, and N Abilities: Research on the Innovation of Financial Management Practice Teaching Model Driven by Digital Intelligence (2024JGB459).

4. 2023 Guilin University Course Ideology Reform Project: Research on the Innovation of Financial Management Practice Teaching Model Driven by Artificial Intelligence and the Practice of Ideological and Political Education.

References

[1] Xu, T., Yang, G., & Chen, T. (2024). The role of green finance and digital inclusive finance in promoting economic sustainable development: A perspective from new quality productivity. Journal of Environmental Management, 370, 122892.

[2] Zhang, L. (2023, July). Intelligent Internal Audit Platform Architecture Based on Machine Learning. In 2023 2nd International Conference on Artificial Intelligence and Autonomous Robot Systems (AIARS) (pp. 531-534). IEEE.

[3] Kryvoruchko, O., Hnatchenko, D., Desiatko, A., Karpunin, I., Shestak, Y., & Kostiuk, Y. (2024, May). Methodology for Developing an Information System for Internal Audit Support. In 2024 IEEE 4th International Conference on Smart Information Systems and Technologies (SIST) (pp. 106-110). IEEE.

[4] Zhao, P., & Deng, L. (2024, April). Enterprise Financial Fraud Detection and Audit Optimization Based on Deep Learning. In International Conference on Computational Finance and Business Analytics (pp. 3-14). Cham: Springer Nature Switzerland.

[5] Ghafar, I., Perwitasari, W., & Kurnia, R. (2024). The Role of Artificial Intelligence in Enhancing Global Internal Audit Efficiency: An Analysis. Asian Journal of Logistics Management, 3(2), 64-89.

[6] Adelakun, B. O., Fatogun, D. T., Majekodunmi, T. G., & Adediran, G. A. (2024). Integrating machine learning algorithms into audit processes: Benefits and challenges. Finance & Accounting Research Journal, 6(6), 1000-1016.

[7] Kang, H. (2024). Optimization of Enterprise Accounting Audit Risk Identification and Prevention Strategy Based on Machine Learning. Journal of Electrical Systems, 20(9s), 79-84.

[8] Popara, J., Savkovic, M., Lalic, D. C., & Lalic, B. (2023, September). Application of digital tools, data analytics and machine learning in internal audit. In IFIP International Conference on Advances in Production Management Systems (pp. 357-371). Cham: Springer Nature Switzerland.

[9] Du, W. (2021, November). Enterprise Internal Audit System in the Context of Big Data. In 2021 IEEE International Conference on Emergency Science and Information Technology (ICESIT) (pp. 215-218). IEEE.

[10] Bondarenko, T., Dutchak, R., Kondratiuk, O., Rudenko, O., Shaikan, A., & Shubenko, E. (2020, March). Internal Audit of Conflicts in Enterprise's Accounting with the Help of Artificial Intellect's Instruments. In III International Scientific Congress Society of Ambient Intelligence 2020 (ISC-SAI 2020) (pp. 39-46). Atlantis Press.

[11] Thomas, M. S., & Mathew, J. (2022, April). Supervised machine learning model for automating continuous internal audit workflow. In 2022 6th International Conference on Trends in Electronics and Informatics (ICOEI) (pp. 1200-1206). IEEE.

[12] Liu, X., Ren, Y., Qi, G., Li, Y., & Fan, R. (2024, July). Artificial intelligence digital audit system under machine learning technology. In 2024 3rd International Conference on Artificial Intelligence and Autonomous Robot Systems (AIARS) (pp. 739-743). IEEE.

[13] Shivram, V. (2024). AUDITING WITH AI: A THEORETICAL FRAMEWORK FOR APPLYING MACHINE LEARNING ACROSS THE INTERNAL AUDIT LIFECYCLE. EDPACS, 69(1), 22-40.

[14] Zhang, H. (2024, May). Application of Hybrid Intelligent Optimization Algorithm in Internal Audit Informatization Construction. In 2024 International Conference on Telecommunications and Power Electronics (TELEPE) (pp. 170-175). IEEE.

[15] Duan, H. K., Vasarhelyi, M. A., & Codesso, M. (2024). Integrating Process Mining and Machine Learning for Advanced Internal Control Evaluation in Auditing. Journal of Information Systems, 1-21.

[16] Khanam Zobaida. (2024). Effectiveness of internal auditing from the lens of internal audit factors: empirical findings from the banking sector of Bangladesh. Journal of Financial Crime(6),1467-1484.

[17] Isac de Freitas Brandão & Vicente Lima Crisóstomo. (2024). Shareholding control, ownership concentration, and the value of the Brazilian firm. Borsa Istanbul Review(5),984-995.

[18] Mousaab El Khair Ghoujdam,Rachid Chaabita,Oussama Elkhalfi,Kamal Zehraoui,Hicham Elalaoui & Salwa Idamia. (2024). Consumer credit risk analysis through artificial intelligence: a comparative study between the classical approach of logistic regression and advanced machine learning techniques. Cogent Economics & Finance(1).