Guan, H., Li, S., Wang, Q., Lyulyov, O. & Pimonenko, T. (2022). Financial Fraud Identification of the Companies Based on the Logistic Regression Model. *Journal of Competitiveness,* 14(4), 155–171. https://doi.org/10.7441/joc.2022.04.09

Financial Fraud Identification of the Companies Based on the Logistic Regression Model

▪ *Heshan Guan, Siying Li, Qian Wang, Oleksii Lyulyov, Tetyana Pimonenko*

Abstract

Companies' financial fraud provokes declining market asset allocation efficiency and significantly impacts trust and loyalty among all company's stakeholders. Most investigations focused on the prediction of accounting fraud; less research concentrated on financial restatements. In this case, the paper aims to develop a model for identifying the companies' financial fraud according to the developed index system construction based on financial statements and relationships between their items. The study applies the following stages: 1) analysis of the theoretical framework of the core determinants, impulses and factors of financial fraud and their identification; 2) development of the methodology for timely identification of financial fraud, which is based on index system construction using the Logistic regression model. The object of investigation is Chinese companies listed by China Stock Market & Accounting Research database, excluding J66 (remaining financial industry except for the monetary and financial services), J67 (capital market services), J68 (insurance industry) and J69 (other financial industry) enterprises. The period of investigation is 2017–2020. The data sample includes 53 fraudulent and 53 normal Chinese enterprises. The results show that the overall prediction accuracy of the developed model is 83% and robustness test results further verify the rationality and effectiveness of the method. The company's stakeholders could apply the proposed approach for fraud identification to improve the efficiency of financial fraud identification from the technical level.

Keywords: cryptocurrency fraud, financial statement, investment activities, logistic regression model
JEL Classification: M41, G20, G32, O16

Received: October, 2022
1st Revision: October, 2022
Accepted: November, 2022

1. INTRODUCTION

In recent years, a series of corporate financial fraud cases have appeared in the securities markets of various countries. Al-Hashedi and Magalingam (2021) classify financial fraud into four groups: bank fraud, insurance fraud, financial statement fraud, and cryptocurrency fraud. Besides, the studies by Miśkiewicz (2019), Kuzior and Kwilinski (2022), Bharadwaj and Deka (2021), Gajdzik et al. (2021), and Kuzior (2022) underline that development of digital technologies is conducive to the company's financial performance. At the same time, it opens new windows for fraud. The scholars Aggarwal et al. (2015), Wu et al. (2016), Liang et al. (2022), and Al-Hashedi and Magalingam (2021) underline that financial fraud is a crucial issue for the financial sector and has a significant impact on trust and loyalty among all company's stakeholders (investors, creditors, banks, government, consumers, society). Financial fraud is defined as financial abuse (unlawful or illegal behaviour that generates benefit for individuals or organisation in unethical and illegal ways) which provokes economic and reputation losses for the government, corporate sector, investors, and company's shareholders (Choi & Lee, 2018; Al-Hashedi & Magalingam, 2021; West & Bhattacharya, 2016). Ewelt-Knauer et al. (2015) confirm that shareholder values in German companies decline due to the declaration of information on financial fraud. On average, the overall value of losses is 81 million euros during 1998–2014. The financial fraud scandal with Luckin Coffee in 2020 provoked the exclusion of companies from the listing on the Chinese stock exchange; the scandal has declined the reputation of other listed companies (Qiu et al., 2021; Carleton, 2021). Such cases justify the relevant reaction from the government, particularly to increase the financial control of Chinese companies.

In 2020, China issued the Securities Law of the People's Republic of China (2020), which increased the punishment for companies' financial fraud. The new legislation increases the penalty for an enterprise's financial fraud, but as Wang and Wang (2022) and Lu (2021) remark, it has not brought the desired results. Mostly, it is because the fraud benefits for companies bring high earnings, which allows for overcoming the fraud cost and penalties. It should be noted that the Shenzhen Stock Exchange developed the Guide on Information Disclosure Evaluation Systems for Firms Listed in the SZSE in 2001 (Ho et al., 2022) to eliminate financial fraud among the listed companies. This document has been revised five times; the latest version was published in 2020. Considering this guide, fraud identification should consist of two stages: 1) to define the features and preconditions of financial fraud (index system construction); 2) to develop the identification model, which may affect the efficiency of fraud identification. The enormous negative consequences of financial fraud stimulate researchers and experts to develop approaches for the identification of financial fraud. However, most investigations (Ren et al., 2021; Su et al., 2021; Liang et al., 2022; Zhao et al., 2021; Xu et al., 2022) focused on corporate fraud; less research concentrated on financial fraud.

The paper fills the research gap by 1) the contributing to the theoretical basis by identifying the core determinants, impulses, and factors of financial fraud; 2) developing the methodology for timely identification of financial fraud based on index system construction. Thus, this study aims to develop a model for identifying the companies' financial fraud based on index system construction.

The paper has four upcoming sections. The first section is a literature review containing an analysis of the theoretical framework of financial fraud and the approaches to identify it in time. The second section identifies sources for data compilation and explanation of the methods and instruments to achieve the paper's aims. The results and discussion section contains the analysis of the empirical findings of the investigation. The last section is a conclusion summarising the core findings, highlighting the limitations, and outlining directions for future investigations

2. THEORETICAL BACKGROUND.

The analysis of the theoretical framework shows that scholars apply a vast range of methods for fraud identification: the Fraud Triangle Theory (Xiong & Zhang, 2022; Hu et al., 2020); GONE theory (Hong, 2012; He & Gao, 2020); Risk Factor Theory (Zheng & Tang, 2021; Wang & Zhang, 2020); or Iceberg theory (Huang, 2022). It should be noted that the United States and the Chinese Association of Certified Public Accountants have formulated relevant audit standards according to the Fraud Triangle Theory. According to the audit standards for Chinese certified public accountants (International Federation of Accountants, 2022), financial fraud includes false reporting of financial information and embezzling assets, such as false disclosure and related party transactions (Su & Zhong, 2021; Fang, 2020). The analysis of fraud also provides the basis for feature extraction. However, the characteristics of these documents are highly subjective, have different extraction results, and lack systematic internal logical analysis. For example, they evaluate the solvency, operation, and development abilities (Xiong & Zhang, 2022; Qian & Luo, 2015). It often appears in the relevant literature on financial fraud research, but there are still no uniform selection criteria.

Financial data are recorded in accordance with specific accounting standards and calculation formulas, and there is a particular logical relationship between the data (Liu, 2016). Therefore, the abnormal multi-relationship between the items related to the report can be used as an important basis for fraud identification (Lin, 2020). If the principle of correlation between the financial characteristics of the statements is violated, then there will be financial fraud in the statements (Shen et al., 2021). Multi-relationship of accounting statements include accurate and fuzzy kinds. An accurate accounting relationship refers to the relationship between the report items in an equation (Zhi, 2006). A precise relationship in the balance sheet means that assets equal the sum of liabilities and owner's equity. According to the relevant equation relationship, the enterprise financial department prepares the statement. Theoretically, they should have the above balance relationship (Lin, 2020). A fuzzy accounting relationship refers to the reasonably expected relationship between the report items or different report items (Wang, 2016). Under normal circumstances, when enterprises commit fraud, they will whitewash the statements to meet the accurate multi-relationship. However, data fraud will make it difficult for related projects to meet the normally expected relationship and be prone to contradictory relationships between related projects. Therefore, it is appropriate to study the fuzzy multi-relationship to judge whether the enterprise has financial fraud.

The existing research on the financial fraud-related literature in the accounting relationship remains at the level of qualitative analysis, lacking the necessary quantitative analysis and relevant empirical evidence. In this case, the paper aims to extract the characteristics of fraud from the multi-relationship between the report items and establish an index system for fraud identification.

3. RESEARCH OBJECTIVE, METHODOLOGY AND DATA

The results of companies' activities are mainly reflected in the balance sheet, income statement, and cash flow statement (Fedorko et al., 2021; Zadorozhnyi et al., 2021; Kwilinski et al., 2020).

The multi-relationship between the internal items and the enterprise's financial situation, operating results, and cash flow are presented in Fig. 1.

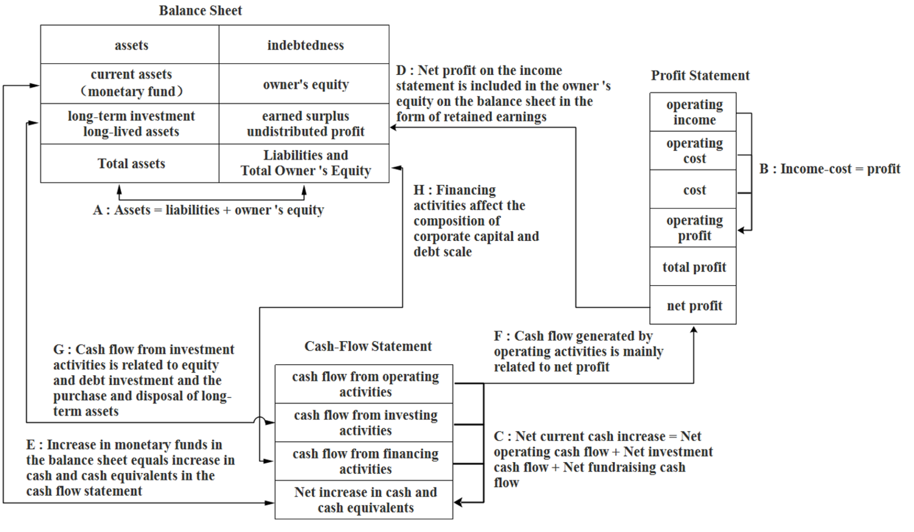

Fig. 1 – The diagram of the relationship between three major enterprise financial statements

The parameters A, B, and C reveal multi-relationships between the internal items of the report (Figure 1). D shows that the production and operation results of each period of the enterprise finally flow back to the owner's equity in the form of retained income while causing changes in the asset project. Relationship E indicates that the net increase in cash and cash equivalents after the three major activities equals the changes in monetary funds. Relationship F shows that business activities mainly bring the profits of an enterprise, and the quality of the enterprise's net profit can be judged according to the cash flow generated by business activities. Relationship G shows that the cash flow related to investment activities is mainly related to equity debt investment and the purchase and disposal of long-term assets. Relationship H indicates that the cash flow from financing activities will affect enterprises' capital and debt composition.

The abnormal linkage between the items in the balance sheet and the income statement from the major enterprise's activities (operation, investment, financing and synthetics judgment) indicates the fraud of the enterprise (Fig. 2).

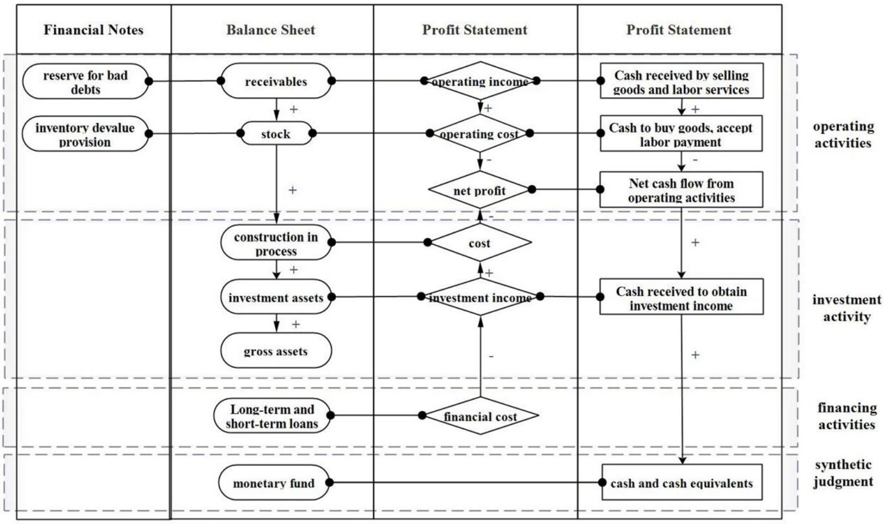

Fig. 2 – Schematic diagram of the report items and their checking relationships

Note: The in-table relationship is expressed in → "··", " + " and " - " and has a positive and negative impact on the final result (the three tables are total assets, net profit, cash and cash equivalents in turn); the inter-table relationship is expressed in "•—•".

Considering Fig. 1–2, the study uses the following indicators:

for assessment of operating activities:

X_1= (Accounts receivable growth rate)/(Operating revenue growth rate) (1)

where X_1 is used to judge the relationship between the growth rate of accounts receivable and operating income;

X_2=(Cash)/(Operating income received from selling goods and providing labour services) (2)

where X_2 is used to judge the proportion of cash received from selling goods and providing labour services in the operating income, which reflects the quality of enterprise sales income.

In order to generate false income for false transactions, it is difficult to generate real cash flow. Thus, the accounts receivable rely on commercial credit for transactions. There is no physical form, and it is difficult to distinguish between real and false. By forming a false asset ledger of accounts receivable, and after bad debt provisions to offset, it will realise the capital cycle from having to no. Suppose X_1 is unusually large and X_2 is abnormally small. In that case, most of the current period's income is accounted for accounts receivable. The real cash flow has not been formed, so we should be on guard against the fictional income behaviour of enterprises.

The enterprises use the proportion of accounts receivable bad debt provision to adjust profits. Suppose the current less bad debt provision will reduce asset impairment losses and inflated profits. By the way, the company uses bad debt provision to write off the accumulated receivables to cover up the anomalies of the asset project. From this, the indicator (X_3) is as follows:

X_3= (Increase in bad debt provision this year)/(Average accounts receivable this year) (3)

If X_3 becomes unusually large, enterprises will use bad debts to write off abnormal assets. The unusually small value of X_3 shows that the enterprises could adjust profits through bad debt provisions.

Enterprises purchase raw materials and other commodities processing to form enterprise inventory. Suppose the inventory growth rate is greater than the increased operating cost rate. In that case, it indicates that the enterprise may have unmarketable goods or less transfer cost, and a profit-inflated behaviour may exist. As a consequence, we construct the indicators X_4 as follows:

X_4= (Inventory growth rate)/(Operating cost growth rate) (4)

where X_4 – is the ratio of inventory growth rate to operating cost growth rate in business activities.

In addition, enterprises can prepare for the current withdrawal and turn back to adjust their profits through the decline in inventory prices. Therefore, the index X_5 is constructed.:

X_5= (Inventory decline increase this year)/(Average amount of inventory this year) (5)

where X_5 – is the proportion of inventory decline provision.

When there is an abnormal growth in X_4, it is suspected that enterprises have the behaviour of reducing transfer costs and falsely increasing profits. When X_5 is abnormally large or small, enterprises should pay attention to adjusting profits.

To some extent, the ratio of net cash flow to the net profit of an enterprise's operating activities can reflect the quality of its operating profit. The larger the proportion of cash flow of operating activities to net profit, the higher the quality of profits. On the contrary, when the value is low, it shows that the profit quality of the enterprise is poor, and the profit cash risk is higher, so the possibility of falsely increasing profits through a large number of receivables can not be ruled out. As a result, the indicator X_6 was constructed as follows:

X_6=(Cash flow)/(Net profit generated from operating activities) (6)

X_6 – the ratio of net cash flow to the net profit of an enterprise's operating activities.

The low value of X_6 shows that the quality of enterprise profit is poor, and the risk of false profit is higher.

for assessment of investment activities:

The fraud within investment activities could be realised in two ways: manipulation in the index of "project under construction;" 2) and manipulation in the investment assets.

The project under construction could be recorded into the asset cost to achieve a false asset increase. In this case, attention should be paid to the capitalised amount of enterprise interest. Compared with fixed assets, projects under construction do not need to provide for depreciation and will not affect profits and losses. Thus, enterprises could adjust profits and losses by delaying the conversion of projects under construction. As a result, the indicators are constructed like that:

X_7= (Accumulated amount of capitalisation of interest)/(Balance of work under construction) (7)

where X_7 – capitalisation amount of interest under construction in investment activities.

X_8 = (Construction Balance)/(Total Assets) (8)

where X_8 – is the ratio of the balance of the projects under construction to the total assets.

Shareholders could use a series of false foreign investments to transfer funds to encroach on the enterprise's interests. The values of indicators X_9 and X_{10} allow judge whether the enterprise has the behaviour of false investment and transferring funds:

By the ratio of investment assets to investment income and the ratio of cash received from investment income, we can understand the income of investment and the recovery of cash and judge whether the enterprise has the behaviour of false investment and transferring funds. As a result, the indicators are constructed:

X_9=(Investment income)/(Ending balance of investment assets) (9)

where X_9 – is the ratio of investment assets to investment income

X_{10}=(Cash)/(Investment income received from the investment income) (10)

where X_{10} – is the ratio of cash received from investment income to investment income.

When X_9 and X_{10} are small, the enterprise could have a false investment and not get the actual income and cash.

for assessment of financing activities:

If the financial expenses and loans are relatively high, the enterprises use high-cost funds, and there is a risk in their ability to repay debts. As a result, the indicator is constructed:

X_{11}=(Financial expenses)/(-Short-term loan + long-term loan) (11)

X_{11} – the ratio of financial expenses in short and long-term loans

After business activities, investment activities, and financing activities, the indicator X_{12} is equal to the number of monetary funds at the end of the year minus the beginning of last year.

If monetary funds are limited or occupied, the increase of cash and equivalents in the statement (Cash Flow Statement) will be less than monetary funds.

X_{12}=(Net increase of cash and cash equivalents)/(Number of monetary funds-beginning of the year of monetary funds) (12)

Based on studies (Itoo & Singh, 2021; Qin, 2021; Mishra & Pandey, 2021; Mehbodniya et al., 2021) the fraud is indicated by the logistic regression model:

$$\text{Fraud}=\ln p/(1-p)=\beta_0+\beta_1 x_1+\beta_2 x_2+\ldots+\beta_{12} x_{12}+u_1 C_1+u_2 C_2+\ldots+u_5 C_5+\varepsilon \quad (13)$$

where p is the probability of fraud, X_1, X_2,…, X_{12} is the explanatory variable of hook-check relationship, C_1, C_2,…, and C_5 is the explanatory variable of non-financial indicators. β_1,…,β_{12},u_1,…,u_5 is the coefficient, and ε is the random error term.

Based on Zhang and Xu (2021), and Sun and Shen (2022), this investigation selected the explanatory variables of non-financial indicators, which are indicated in Tab.1.

Tab. 1 – The description of the explanatory variables of non-financial indicators

Explanatory variables	Explanation of explanatory variables	Features of assessment
C1	Equity concentration degree	expressed by the largest shareholder shareholding ratio
C2	Whether the general manager is concurrently chairman	is recorded as 1, whether recorded as 0
C3	The proportion of independent directors	expressed as the number of independent directors (total number)
C4	Whether the internal control is valid	valid as 1, invalid as 0
C5	Type of audit opinion	standard audit opinion recorded as 1, otherwise recorded as 0

The dependent variable of fraudulent enterprises is set at 1, the dependent variable of normal enterprises is set at 0, and the number of fraudulent enterprises is the same as that of normal enterprises. Therefore, the judgment principle of the logistic regression model is as follows: if the probability P-value is higher than 0.5, it will be judged as a fraudulent enterprise. Otherwise, it will be judged as a normal enterprise.

The study selects the enterprises with financial violations in the China Stock Market and Accounting Research (2022) database (CSMAR) as the fraud sample for 2017–2020. There are two screening principles:

1. for enterprises that have violated the rules for many years in a row. The first violation year is selected as the fraud time to extract data samples to prevent the possibility of overestimation of fraud caused by repeated extraction of samples.
2. the financial industry data are very different from other industries. The financial enterprises whose main categories are J66 (remaining financial industry except for the monetary and financial services), J67 (capital market services), J68 (insurance industry) and J69 (other financial industry) are excluded according to the industry classification standard of CSRC (China Security Regulatory Commission) version 2012 to compile a sample of fraudulent enterprise which is in the same industry, and the total assets are closest.

Finally, the data sample includes 53 fraudulent and 53 normal enterprises. In the empirical study, SPSS26 is used for data analysis. The results of the analysis of descriptive statistics are shown in Tab. 2.

Tab. 2 – The findings of descriptive statistics

Fraud		N	Mean	St. Dev	Fraud		N	Mean	St. Dev	Fraud		N	Mean	St. Dev
X1	1	53	1.95	5.60	X7	1	53	12.02	11.19	non-financial indicators				
	0	53	5.06	23.29		0	53	6.91	3.19					
X2	1	53	1.02	0.21	X8	1	53	0.04	0.06	C1	1	53	33.37	14.74
	0	53	0.95	0.21		0	53	0.04	0.03		0	53	30.28	12.37
X3	1	53	8.96	5.89	X9	1	53	0.55	1.35	C2	1	53	0.45	0.50

X3	0	53	15.66	18.47		0	53	0.33	0.62		0	53	0.42	0.50
X4	1	53	2.48	9.88	X10	1	53	-5.31	29.34	C3	1	53	0.39	0.07
	0	53	-2.03	18.92		0	53	0.71	0.47		0	53	0.37	0.04
X5	1	53	0.05	0.13	X11	1	53	0.11	0.31	C4	1	53	0.96	0.19
	0	53	0.02	0.02		0	53	0.46	3.28		0	53	0.91	0.30
X6	1	53	0.57	3.31	X12	1	53	0.46	1.77	C5	1	53	0.92	0.27
	0	53	-0.42	7.33		0	53	1.42	3.54		0	53	0.92	0.27

Note: St. Dev – Standard Deviations; 1 – means company with fraud; 0 – means the company without fraud.

Considering the multi-collinear analysis results, the VIF of all variables is less than 10, which confirms no multi-collinearity in the regression model (13).

4. RESULTS AND DISCUSSION

Logistic regression results are shown in Tab. 2. Considering the findings, X3, X5, X7, X10, and C3 are significant at the 10% level, and the overall prediction accuracy of the model is 83%. The values of indicators of the model's fit goodness (-2Log likelihood, Cox&Snell R2 and Nagelkerke R2) are 82.506a, 0.456 and 0.607, respectively.

Tab. 3 – The outputs of logistic regression

Variable	Coeff.	St. Er.	Variable	Coeff.	St. Er.	Variable	Coeff.
X1	-0.01	0.06	0.93	X10	-1.06	0.41	0.01
X2	0.65	1.47	0.66	X11	0.00	0.13	1.00
X3	-0.10	0.05	0.05	X12	-0.44	0.34	0.20
X4	0.02	0.04	0.60	C1	0.03	0.02	0.17
X5	24.03	13.15	0.07	C2	-0.24	0.60	0.69
X6	0.04	0.06	0.54	C3	10.20	5.42	0.06
X7	0.13	0.08	0.10	C4	1.98	2.04	0.33
X8	-3.19	7.59	0.67	C5	0.41	1.77	0.82
X9	0.27	0.30	0.36	const.	-7.80	3.38	0.02
-2 Log likelihood							82.506a
Cox&Snell R2							0.46
Nagelkerke R2							0.61

Note: Coeff. – Regression Coefficient; St. Er. – Standard Error; Consp. – Conspicuousness

Further results show that the coefficient of X3 is negative. If other variables remain unchanged, the lower proportion of provision for bad debts will show the probability of fraud. This indicates that enterprises are more likely to falsely increase profits by reducing the provision for bad debts. The coefficient of X5 is also positive. It should be noted that the higher the inventory proportion decline provision values, the greater the probability of fraud. Enterprises are more likely to withdraw inventory price provisions on a large scale in the current period. The higher values of the capitalisation of interest under construction (X7 is positive), the greater the probability of fraud. It indicates that enterprises may falsely reduce costs through cost capitalisation. X10 is negative, which reflects the rapid decline of the cash received from investment income. It could also indicate fraud.

The findings of the recognition effect of the model (13) are shown in Tab. 4. The overall prediction accuracy of the model is 83% which is higher than the past studies (He & Gao, 2020; Wang, 2020; Li et al., 2015). The findings prove that the developed model of fraud identification based on multi-relationship has a comparative advantage in the accurate judgment of fraud.

Tab. 4 – The results of logistic accuracy statistics of regression (13)

Classify		Engage in Embezzlement		Precision %
Engage in Embezzlement	0	45	8	84.9
	1	10	43	81.1
Overall Percentage				83
Logistic accuracy statistics of regression in the existing literature				
The Literature Name	Number of Indicators	Precision %		
He & Gao (2020)	27	79.8		
Wang (2020)	25	80.46		
Li et al. (2015)	17	70.6		

Note: 1 – means the company with fraud; 0 – means the company without fraud.

To check the validity and reliability of the findings, the new control samples are generated based on the one-to-one matching principle. The new set includes 100 fraudulent and nonfraudulent enterprises.

The findings in Tab. 5 show that X7, X8, X10, X12, and C3 are significant at the 10% level, and the overall prediction accuracy of the model is 92%. The regression coefficients of X7 and X8 are positive, and X10 and X12 are negative.

Tab. 5 – The findings of the robustness test

Variable	Coeff.	St. Er.	Variable	Coeff.	St. Er.	Variable	Coeff.
X1	0.12	0.11	0.93	X10	-1.29	0.47	0.01
X2	0.46	2.98	0.66	X11	-2.25	2.15	0.30
X3	-0.08	0.07	0.05	X12	-1.88	1.04	0.07
X4	-0.01	0.05	0.60	C1	-0.03	0.04	0.37
X5	7.72	11.36	0.07	C2	1.19	0.98	0.22
X6	0.04	0.17	0.54	C3	21.35	8.53	0.01
X7	0.67	0.20	0.10	C4	-3.33	7.03	0.64
X8	22.93	13.40	0.67	C5	8.11	5.72	0.16
X9	-0.38	0.34	0.36	const.	-17.38	7.69	0.02
-2 Log likelihood							40.235a

Cox&Snell R2	0.63
Nagelkerke R2	0.84

Note: Coeff. – Regression Coefficient; St. Er. – Standard Error; Consp. – Conspicuousness

Thus, considering the findings in Tab. 5, within the other variables remaining unchanged, the following conditions prove the probability of fraud:

X7: the growth of the interest capitalisation amount of the project under construction provokes the increasing probability of fraud, indicating that the enterprise may falsely reduce the cost through cost capitalisation;

X8: the higher the proportion of the balance of projects under construction to total assets, the greater the possibility of fraud. It indicates that enterprises may falsely reduce costs by delaying the conversion of projects under construction to fixed assets;

X10: the lower the cash received from investment income, the greater the possibility of fraud. It means that the enterprise may have the behaviour of false foreign investment to transfer assets;

X12: the lower the ratio of the net increase of cash and cash equivalents to the change of monetary funds, the greater the possibility of fraud which indicates that fraudulent enterprises occupy more serious funds.

The empirical results in Tab. 6 show that the overall prediction accuracy of the model is 92%, which is comparable with Tab. 4.

Tab. 6 – The results of overall prediction accuracy for robust model (13)

Classify		Engage in Embezzlement		Precision %
		0	1	
Engage in Embezzlement	0	48	2	96
	1	6	44	88
Overall Percentage				92

With the comprehensive analysis of Tab. 3 and 5, most multi-relationship indicators do not enter the regression model, which does not indicate that these tick relationship indicators are redundant. As mentioned above, not all fraud enterprises will adopt the same fraud methods. Therefore, there is no significant difference between the empirical fraud samples and part of the normal samples. They do not enter the regression model, which does not affect the rationality of the index system extracted in this paper. On the contrary, another perspective shows the diversity of corporate fraud behaviour and the characteristics of identification difficulties.

5. CONCLUSION

This paper studies the feature extraction of enterprise financial fraud identification based on the relationship between statement items. According to the classification of enterprise activities and the relationship between the three major financial statements, twelve indicators are constructed, and five non-financial indicators are selected to identify financial fraud. The results show that the overall prediction accuracy of the developed model is 83% and robustness test results further verify the rationality and effectiveness of the method. The empirical results show that the multi-relationship index can improve the accuracy of fraud recognition. This study expands the relevant research on selecting an index system for financial fraud identification. The company's stakeholders could apply the proposed approach for fraud identification to improve the efficiency of financial fraud identification from the technical level.

Despite the valuable results, the investigation has a few limitations. Machine Learning algorithms in fraud identification using Logistic regression require the compilation of accurate samples for analysis. If the fraudulent enterprises are not found and regarded as normal enterprises for empirical analysis, it allows allocating the anomalies from the probability perspective. Besides, the analysed samples should be enlarged for further investigation to obtain more accurate findings. It should be noted that the asymmetry of information is one factor that provokes financial fraud. In this case, it should be considered in further investigations. In addition, financial fraud significantly impacts the companies' attractiveness to investors, which should be considered in future investigations.

Funding: A Project Supported by Scientific Research Fund of Hunan Provincial Education Department "Enterprise Financial Fraud Portrait and Early Warning Research" (21A0262).

References

1. Aggarwal, R., Hu, M., & Yang, J. (2015). Fraud, market reaction, and the role of institutional investors in Chinese listed firms. *Journal of Portfolio Management,* 41(5), 92–109.

2. Al-Hashedi, K. G., & Magalingam, P. (2021). Financial fraud detection applying data mining techniques: A comprehensive review from 2009 to 2019. *Computer Science Review*, 40, 100402. https://doi.org/10.1016/j.cosrev.2021.100402

3. Bharadwaj, S., & Deka, S. (2021). Behavioural intention towards investment in cryptocurrency: An integration of Rogers' diffusion of innovation theory and the technology acceptance model. *Forum Scientiae Oeconomia*, 9(4), 137–159. doi:10.23762/FSO_VOL9_NO4_7

4. Carleton, E. (2021, January 31). Luckin Coffee stock takes a dive as short sellers respond to fraud claim. Barron's. https://www.barrons.com/articles/luckin-coffee-stock-muddy-waters-citron-research-short-sellers-fraud-claim-51580495986

5. China Stock Market and Accounting Research (CSMAR) Database. (2022). Wharton Research Data Services. Retrieved October 1, 2022, from https://wrds-www.wharton.upenn.edu/pages/about/data-vendors/china-stock-market-accounting-research-csmar/

6. Choi, D., & Lee, K. (2018). An artificial intelligence approach to financial fraud detection under IoT environment: A survey and implementation. *Security and Communication Networks*, 2018, 5483472. https://doi.org/10.1155/2018/5483472

7. Ewelt-Knauer, C., Knauer, T., & Lachmann, M. (2015). Fraud characteristics and their effects on shareholder wealth. *Journal of Business Economics*, 85(9), 1011–1047. https://doi.org/10.1007/s11573-015-0773-5

8. Fang, Z. (2020). Literature review of financial statement whitewashing behavior identification.

International Journal of Education and Economics, 3(4), 75–79.

9. Fedorko, R., Skerhakova, V., Markova, S., Bacik, R., & Fejercak, L. (2021). Differences in the perceived financial risk factors between digital natives in the e-commerce. *Marketing and Management of Innovations*, 12(4), 137–152. https://doi.org/10.21272/mmi.2021.4-11

10. Gajdzik, B., Grabowska, S., & Saniuk, S. (2021). Key socio-economic megatrends and trends in the context of the industry 4.0 framework. *Forum Scientiae Oeconomia,* 9(3), 5–21. doi:10.23762/FSO_VOL9_NO3_1

11. He, Y., & Gao, J. (2020). Case study of St. Kangmei's financial fraud based on GONE theory. 2020 International Conference on Economics, Business and Management Innovation, 6–10. doi: 10.25236/icebmi.2020.002

12. Ho, K., Yang, L., & Luo, S. (2022). Information disclosure ratings and continuing overreaction: Evidence from the Chinese capital market. *Journal of Business Research,* 140, 638–656. https://doi.org/10.1016/j.jbusres.2021.11.030

13. Hong, H., Hu, H., & Guo, C. (2012). Research on fraud identification of listed companies' financial reports based on GONE theory. Accounting Research, 298, 84–97. https://kns.cnki.net/kcms/detail/detail.aspx?dbcode=CJFD&dbname=CJFD2012&filename=KJYJ201208014&uniplatform=NZKPT&v=rj4pMvnCAI4cNTrR8JqQpA4XL1F_CaaRWowX0ykztkvv6dxp0Mu5jcc3rm2bxDpl

14. Hu, M., Ma, Q., & Qian, Q. (2020). Analysis of the Kangdexin fraud case based on the Fraud Triangle Theory. Finance and Accounting, 6, 58–61. https://kns.cnki.net/kcms/detail/detail.aspx?dbcode=CJFD&dbname=CJFDLAST2020&filename=CWKJ202006028&uniplatform=NZKPT&v=YFP4sX2XH8N846PRw3UjfOYAvzO2T6ThgxWAK1PiAJT6KEvlq_nbVgp1taI-2iM3

15. Huang, T. (2022). Analysis of financial fraud cases in Shen Xianyuan Company based on Iceberg theory. *Academic Journal of Business & Management*, 4(2), 50–54.

16. International Federation of Accountants. (2022). Audit standards for Chinese certified public accountants. https://www.ifac.org/about-ifac/membership/profile/china

17. Itoo, F., & Singh, S. (2021). Comparison and analysis of logistic regression, Naïve Bayes and KNN machine learning algorithms for credit card fraud detection. *International Journal of Information Technology*, 13(4), 1503–1511. https://doi.org/10.1007/s41870-020-00430-y

18. Kuzior, A. (2022). Technological unemployment in the perspective of Industry 4.0. *Virtual Economics,* 5(1), 7–23. https://doi.org/10.34021/ve.2022.05.01(1)

19. Kuzior, A., & Kwilinski, A. (2022). Cognitive technologies and artificial intelligence in social perception. *Management Systems in Production Engineering,* 30(2), 109–115. doi:10.2478/mspe-2022-0014

20. Kwilinski, A., Slatvitskaya, I., Dugar, T., Khodakivska, L., & Derevyanko, B. (2020). Main effects of mergers and acquisitions in international enterprise activities. *International Journal of Entrepreneurship*, 24(1), 1–8.

21. Li, X., Kang, J., & Li, N. (2015). Application of factor analysis and logistic regression in research on accounting fraud. *Statistics and Decision-making*, 426, 89–92. https://kns.cnki.net/kcms/detail/detail.aspx?dbcode=CJFD&dbname=CJFDLAST2015&filename=TJ

JC201506028&uniplatform=NZKPT&v=XI8I0I-fdG39v9KANZ9WYVlVPD9WNb-bGmMGMBCh2VQHsCI3VH_l22G2uGRT5CtG

22. Liang, Q., Gao, W., & Xie, H. (2022). Do foreign investors deter corporate fraud? Evidence from China. *The Quarterly Review of Economics and Finance*, 84, 92–111. https://doi.org/10.1016/j.qref.2022.01.002

23. Lin, L. (2020). Identification of abnormalities in financial statements from the project links. *Accounting Monthly*, 873, 76–81. https://kns.cnki.net/kcms/detail/detail.aspx?dbcode=CJFD&dbname=CJFDLAST2020&filename=CKYK202005012&uniplatform=NZKPT&v=ANhpNK1Zt2T_pb2gyZUsa7C-gj6ch1cnYy8-kOVDU7981WsIW3nP1e49YCo0gaSa

24. Liu, Y. (2016). Financial fraud "Audit Trail Discovery" research. *Accounting and Communications,* 705, 93–95. https://kns.cnki.net/kcms/detail/detail.aspx?dbcode=CJFD&dbname=CJFDLAST2016&filename=CKTX201613029&uniplatform=NZKPT&v=IVypJwIFHP8Srq85WItoPqrOHTSU6OYwGYhyHfiNDzMc6FoXdal3rQ4D878zc7wP

25. Lu, L. (2021). PRC securities law 2020: Promoting more efficient and transparent capital markets in China. *Butterworths Journal of International Banking and Financial Law*, 36, 237–239.

26. Mehbodniya, A., Alam, I., Pande, S., Neware, R., Rane, K. P., Shabaz, M., & Madhavan, M. V. (2021). Financial fraud detection in healthcare using machine learning and deep learning techniques. *Security and Communication Networks,* 2021, 9293877. https://doi.org/10.1155/2021/9293877

27. Mishra, K. N., & Pandey, S. C. (2021). Fraud prediction in smart societies using logistic regression and k-fold machine learning techniques. *Wireless Personal Communications,* 119(2), 1341–1367. https://doi.org/10.1007/s11277-021-08283-9

28. Miśkiewicz, R. (2019). Challenges facing management practice in the light of Industry 4.0: The example of Poland. *Virtual Economics,* 2(2), 37–47. https://doi.org/10.34021/ve.2019.02.02(2)

29. Qian, P., & Luo, M. (2015). Prediction model of Chinese listed companies' financial fraud. Accounting Research, 96, 18–25. https://kns.cnki.net/kcms/detail/detail.aspx?dbcode=CJFD&dbname=CJFDLAST2015&filename=KJYJ201507003&uniplatform=NZKPT&v=iFbslh7vazS7PmVFy2HhGr9N87iFToxLSF3cf6HjtW9_IiXgMuCubfIA6G7hxDOk

30. Qin, R. (2021). Identification of accounting fraud based on support vector machine and logistic regression model. *Complexity*, 2021, 5597060 https://doi.org/10.1155/2021/5597060

31. Qiu, S., Luo, Y., & Guo, H. (2021). Multisource evidence theory-based fraud risk assessment of China's listed companies. *Journal of Forecasting,* 40(8), 1524–1539. https://doi.org/10.1002/for.2782

32. Ren, L., Zhong, X., & Wan, L. (2021). Missing analyst forecasts and corporate fraud: Evidence from China. *Journal of Business Ethics,* 181, 171–194. https://doi.org/10.1007/s10551-021-04837-w

33. Securities Law of the People's Republic of China. (2020). http://www.npc.gov.cn/englishnpc/c23934/202109/9886ca6f805e4663a9a725d6f72066dd.shtml

34. Shen, Y., Guo, C., Li, H., Chen, J., Guo, Y., & Qiu, X. (2021). Financial feature embedding with knowledge representation learning for financial statement fraud detection. *Procedia Computer Science,* 187, 420–425. doi: 10.1016/j.procs.2021.04.110

35. Su, B., & Zhong, L. (2021). Lessons learned for auditing: Luckin's financial fraud incident. *Friends of Accounting,* 652, 135–140. https://kns.cnki.net/kcms/detail/detail.aspx?dbcode=CJFD&dbname=CJFDLAST2021&filename=KJZY202104019&uniplatform=NZKPT&v=xRNy-8EHaDlPCeUEWdxmTcoEyFao3n5UzEiXUiZ_5KljZu5oJAcXcEhPLdtVHeNP

36. Su, F., Feng, X., & Tang, S. (2021). Do site visits mitigate corporate fraudulence? Evidence from China. *International Review of Financial Analysis*, 78, 101940. https://doi.org/10.1016/j.irfa.2021.101940

37. Sun, J., & Shen, Z. (2022). Financial fraud, audit quality and social audit supervision: An empirical analysis based on the data of listed companies with financial fraud in China. *Friends of Accounting,* 9, 108–115. https://kns.cnki.net/kcms/detail/detail.aspx?dbcode=CJFD&dbname=CJFDLAST2022&filename=KJZY202209014&uniplatform=NZKPT&v=YqlR2ZZr9VzXPKiP00tIIYh16HQ0MGt-TDGBshNU-_bBnUXMKNfs5Ts-kBW8xgP5

38. Wang, J., & Wang, D. (2022). Corporate fraud and accounting firm involvement: Evidence from China. *Journal of Risk and Financial Management,* 15(4), 180. https://doi.org/10.3390/jrfm15040180

39. Wang, W. (2020). Sparse group Lasso-logistic regression model in the identification of financial reporting fraud. *Practice and Understanding of Mathematics*, 50, 49–58. https://kns.cnki.net/kcms/detail/detail.aspx?dbcode=CJFD&dbname=CJFDLAST2020&filename=SSJS202009006&uniplatform=NZKPT&v=SCsiVw-Xwuj-SK1mYNPntWk8NDH3HTUVLe4IQUvMuvNS1iNk0jvL6YnYzZ_bFBMa

40. Wang, X. (2016). Risk identification of audit material misstatements based on financial statement checking relationship. *Accounting Monthly,* 749, 71–73. https://kns.cnki.net/kcms/detail/detail.aspx?dbcode=CJFD&dbname=CJFDLAST2016&filename=CKYK201601020&uniplatform=NZKPT&v=9PNWWc0VICtw_kyCQgavpbAaQMIi4HFLEleg9ECEk9WXDz0crS7llJi6xhFl5xNh

41. Wang, X., & Zhang, Y. (2020). The first illegal delisting case study under the new fraud risk factor theory. *Friends of the Accounting,* 646, 46–52. https://kns.cnki.net/kcms/detail/detail.aspx?dbcode=CJFD&dbname=CJFDLAST2020&filename=KJZY202022007&uniplatform=NZKPT&v=E1EsGj_htApG7qv2_g0X-lSgGqkmQ30kXPGpPh4sNRL8OnQQphkI6NrjT9IKS6bS

42. West, J., & Bhattacharya, M. (2016). Intelligent financial fraud detection: A comprehensive review. *Computers & Security,* 57, 47–66. https://doi.org/10.1016/j.cose.2015.09.005

43. Wu, W., Johan, S. A., & Rui, O. M. (2016). Institutional investors, political connections, and the incidence of regulatory enforcement against corporate fraud. *Journal of Business Ethics,* 134(4), 709–726. https://doi: 10.1007/s10551-014-2392-4

44. Xiong, F., & Zhang, L. (2022). Research on financial fraud risk identification and governance countermeasures of listed companies: Taking Luckin Coffee as an example. *Friends of Accounting*, 3, 55–61. https://kns.cnki.net/kcms/detail/detail.aspx?dbcode=CJFD&dbname=CJFDLAST2022&filename=KJZY202203009&uniplatform=NZKPT&v=YqlR2ZZr9Vwfoa8Lwr0mzcfh7ToXFcqQYYgkWvNzV8ICroHNzvnlg-TNzmfhTTUy

45. Xu, X., Xiong, F., & An, Z. (2022). Using machine learning to predict corporate fraud: Evidence based on the GONE framework. *Journal of Business Ethics*, 2022, 1–22. https://doi.org/10.1007/s10551-022-05120-2

46. Zadorozhnyi, Z.-M., Ometsinska, I., & Muravskiy, V. (2021). Determinants of firm's innovation: Increasing the transparency of financial statements. *Marketing and Management of Innovations,* 2, 74–86. https://doi.org/10.21272/mmi.2021.2-06

47. Zhang, D., & Zhang, X. (2018). Research on the relationship between accrual items and company performance. *Accounting and Communications*, 129, 26–29. https://kns.cnki.net/kcms/detail/detail.aspx?dbcode=CJFD&dbname=CJFDLAST2018&filename=CKTX201828007&uniplatform=NZKPT&v=YtJ1Inj8E5ToQQDbIjP3J85u9AjSAWqXYUg0UYLJeIaMje0et2xk-ZOObujihIDo

48. Zhang, Y., & Xu, C. (2021). Top management team characteristics, social responsibility awareness and financial fraud risk based on the threshold effect test of A-share listed companies. *Friends of Accounting,* 670, 58–65. https://kns.cnki.net/kcms/detail/detail.aspx?dbcode=CJFD&dbname=CJFDLAST2021&filename=KJZY202122009&uniplatform=NZKPT&v=xRNy-8EHaDnULAqLnB0MArUkB-oluiNxbZ1XoZoHUN0Jq0P9apsaFv0e3GLTobYV

49. Zhao, X., Yang, D., Li, Z., & Song, L. (2021). Multiple large shareholders and corporate fraud: Evidence from China. *Frontiers of Business Research in China,* 15(1), 1–21. https://doi.org/10.1186/s11782-021-00106-9

50. Zheng, J., & Tang, Y. (2021). Case study of Luckin coffee financial fraud based on theory of fraud risk factors. *Academic Journal of Business & Management*, 3(11), 37–41. doi: 10.25236/AJBM.2021.031107

51. Zhi, X. (2006). Analysis of financial statements based on the checking relationship of statements. *Friends of the Accounting,* 10, 27. https://kns.cnki.net/kcms/detail/detail.aspx?dbcode=CJFD&dbname=CJFD2006&filename=KJZY200610012&uniplatform=NZKPT&v=mKsq1Va9E-VKmN9WiVTN_Qzql4hVESgxL9UtQWGs27Fl9ZkSp_Ir4wxUbXYLgKof

Contact information

Heshan Guan
University of South China
School of Economics, Management and Law
China
E-mail: 2009001042@usc.edu.cn

Siying Li
University of South China
School of Economics, Management and Law
China
E-mail: 916920074@qq.com

Qian Wang
Fujian Normal University
School of Economics
China
E-mail: m15211476503@.163.com

Oleksii Lyulyov, Dr. Sc.
WSB University
Faculty of Applied Sciences
Department of Management
Poland
Sumy State University
Department of Marketing
Academic and Research Institute of Business, Economics and Management
Ukraine
https://orcid.org/0000-0002-4865-7306
E-mail: alex_lyulev@econ.sumdu.edu.ua

Tetyana Pimonenko, Dr. Sc.
WSB University
Faculty of Applied Sciences
Department of Management
Poland
Sumy State University
Department of Marketing
Academic and Research Institute of Business, Economics and Management
Ukraine
https://orcid.org/0000-0001-6442-3684
E-mail: tetyana_pimonenko@econ.sumdu.edu.ua
Corresponding Author: Tetyana Pimonenko, tetyana_pimonenko@econ.sumdu.edu.ua