Risk Screening of Tax Compliance of Manufacturing Listed Companies Based on Association Rules

Qu Ying¹*, Gao Lanlan¹ and Han Xiaoxin¹

¹School of Economics and Management, Hebei University of Science and Technology, Shijiazhuang City, Hebei Province, 050018, China

*Corresponding author’s e-mail: 732887983@qq.com

Abstract. With the increasing variety of taxes and corresponding tax laws and regulations, the complexity of tax collection and management is increasing, and listed manufacturing companies are one of the important sources of tax revenue in China. Therefore, it is of great significance to construct a scientific and effective tax risk identification and assessment system. In order to improve tax compliance under the premise of complex tax work, this paper proposes a method of tax compliance risk screening for listed manufacturing companies based on association rules. The financial index data of listed manufacturing companies that are crawled are processed, and the Apriori association rule analysis is used to realize the analysis and mining that have an impact on corporate tax compliance. By dividing the tax compliance risk levels of enterprises, it is concluded that the common characteristics of enterprises with low tax compliance are weak profitability, poor debt paying ability, and a large number of overdue repayment.

1. Introduction

The average contribution rate of listed manufacturing companies to tax revenue is 35.83%, and the largest contribution to tax revenue is in all categories of enterprises [1]. As far as listed companies are concerned, the staffs of the company, especially the financial staff, have broad knowledge and higher ability than those of other unlisted small and medium-sized enterprises. Therefore, the methods used in dealing with corporate accounts for tax evasion and tax evasion are more concealed and difficult for tax inspectors to find, which increases the difficulty of tax collection and management. At the same time, with the continuous improvement of China's tax law, tax laws and regulations are increasingly diverse, tax supervision is becoming more and more standardized, and much tax work is constantly detailed, which challenges the traditional work mode of tax collection and management. The current work mode of many tax collection and management has gradually not adapted to the requirements of the development of The Times.

At present, the research on enterprise tax compliance risk trajectory is analysed, some scholars from the enterprise tax behaviour xiao-qin han using Logistic regression model to analyze the taxpayer's tax compliance behavior motivation, it is concluded that the tax environment for taxpayers are following some influence, taxpayer due to the tax authority enforcement selection with "Passive compliance" [2]. Some scholars have also studied the characteristics and means of financial statement fraud of listed companies, and concluded that the financial statement fraud means of listed companies are characterized by diversity and concealment [3]. In the research on tax risks in the manufacturing industry, some scholars have used the COSO risk management framework to draw conclusions about the problems existing in the tax risk management in the equipment manufacturing industry, mainly
due to inadequate tax risk response methods, lack of risk information, communication and report, etc. [4]. After the launch of the third phase of the golden tax system, some scholars studied the application of big data technology in tax risk management and control. With the advent of the era of big data, the predictability of big data technology and the accurate analysis of historical data can be used to reduce the blindness in tax inspection [5]. Yang shanlin and Zhou kaile believe that big data is a strategic information resource [6]. In the era of big data, data mining is constantly applied in various fields, among which the association rule algorithm of data mining technology is also continuously optimized and developed. Lu xiaochen et al. used "Internet + big data algorithm" to introduce artificial intelligence into tax work, which can improve taxpayers' sense of gain and strengthen internal management of tax authorities [7]. To sum up, there are different studies on the tax risks of enterprises, and the relevant literature on the tax control of listed manufacturing companies in China is relatively scarce. Therefore, this paper aims at the current tax risks of listed manufacturing companies, and uses the existing rich information resources to establish a tax risk identification and evaluation system for listed manufacturing companies, so as to reduce the loss of tax revenue in China and improve the tax compliance of listed manufacturing companies.

2. Research design

2.1. Theoretical application
Data of manufacturing the financial indicators of listed companies by using samples of all the financial and tax risk index data is global item sets, using association rule algorithm to find frequent item sets, first in all candidate set C1 for 1 set that meet user pre-set minimum support threshold of item 1 - integration for frequent 1 - item sets L1, on the basis of L1, calculate all the 2 - item sets in the candidate set C2 support, support is greater than or equal to the user to specify the minimum support threshold of 2 L2 - two - integration as a frequent item sets. By using the concept of association rule algorithm, we can obtain the financial index of frequent item set which is related to the high tax risk of enterprises. Set the high tax risk of manufacturing listed companies as the latter item, so that every rule generated in the result of association rules has the item of high tax risk. Financial index data refer to the preceding paragraph. Frequent item sets with high correlation with tax risks are generated through association rules. Financial indicators with high correlation with tax risks are found according to confidence, and the main influencing factors of tax risks of enterprises are obtained.

2.2. Index selection
On the basis of the validity of the results, the financial index data of China's listed manufacturing companies published by domestic well-known financial websites were taken as the sample data. Because of manufacturing listed companies have different size, the same financial indicators data value is large, in order to ensure the data partition is relative, in this paper, the use of financial indicators is mostly through ratio calculated data, such as long-term debt ratio, asset-liability ratio, other payables, main business profit margins, main business cost rate, such as indicators, using these indicators also can response enterprise's profit ability, growth ability, operation ability. The corporate credit records were combined with the financial index data of the same period for association mining. The abnormal financial index values of enterprises with high tax risks were more obvious. In order to ensure the effectiveness of the results, the sample data used in this paper were all enterprises with high tax risks (Figure 1).
2.3. Data preprocessing
The initial data of listed manufacturing companies are continuous data, and there is no clear boundary between the data. The association rule algorithm used in data mining requires discrete values, so the association rule analysis cannot be carried out directly. Therefore, the data must be discretized. Considering the complexity of space and time, after several times of strict debugging, the data was divided into four parts. According to the size of data value, the data was divided into four parts from small to large. The index value in the first part was the minimum, and the index value in the fourth part was the maximum. The results show the Partial data after discretization are shown in Table 1.

| Indicators   | Type | Discrete Data |
|--------------|------|---------------|
| V1           | Numerical  | 1 2 2 2 2 ... |
| V2           | Numerical  | 4 2 3 1 ... |
| ...          | ... | ... |
| V11          | Numerical  | 3 3 2 2 ... |
| ...          | ... | ... |

3. Data processing and analysis

3.1. Association rules
According to the characteristics of the sample data collected, in order to ensure the smooth analysis of association rules, after several debugging settings, the support degree of association rules is more than 10%, and the confidence degree is more than 80%. Before the analysis of data association rules, the high tax risk level is set as the latter item, and the former item is set as all financial indicators. Due to space constraints, the results of some association rules are shown here in Table 2.

| Consequent            | Antecedent                                                                 | Support (%) | Confidence (%) |
|-----------------------|-----------------------------------------------------------------------------|-------------|----------------|
| Risk Level = High     | Long-term debt ratio=4                                                      | 14.385      | 94.62          |
| Risk Level = High     | Profit margin of main business=2; Asset-liability ratio=4                   | 13.273      | 93.333         |
| Risk Level = High     | Asset-liability ratio=4; Long-term debt ratio=4; Debt equity ratio=4        | 11.504      | 92.308         |
| Risk Level = High     | Asset-liability ratio=3; Long-term debt ratio=4; Debt equity ratio=4; Cost Rate of Main Business=4 | 11.504 | 92.308 |
| Risk Level = High     | Asset-liability ratio=3; Long-term debt ratio=4; Debt equity ratio=4; Cost Rate of Main Business=4 | 10.619 | 91.667 |
| ...                   | ...                                                                         | ...         | ...            |
3.2. Result analysis of association rules

Through the main association rules listed in Table 3, the rules are analyzed.

Rule 1: Long-term debt ratio \((V29) = 4\) → Risk Level = High, Support=14.385%, Confidence=94.62%. From this rule, we can get that the long-term debt ratio index \((V29) = 4\), in the fourth sub-point of the index data value, indicating that the index data value is large. When the debt-to-production ratio of enterprises is large, it shows that the debt-to-capital ratio of enterprises' total capital is relatively high. Enterprises should face a large number of debt-to-repayment businesses. Therefore, the degree of protection of debt-to-capital is weak and the ability to repay debt is poor, which leads to high tax risk.

Rule 2: main business profit margin \((V23) = 2\); asset liability ratio \((V30) = 4\) → risk level = high, Support=13.273%, Confidence=93.333%. From this rule, we can get that when the main business profit margin = 2 and asset-liability ratio = 4, the enterprise risk level is high. From this point of view, the main business profit margin reflects the profitability of the enterprise. When the profitability of the enterprise is relatively low, it shows that the market competitiveness of the main business is limited, and accompanied by large assets and liabilities; the tax risk level is high.

Rule 3: Asset-liability ratio \((V30) = 4\); Long-term debt ratio \((V29) = 4\); Debt equity ratio \((V34) = 4\) → Risk Level = High, Support=11.504%, Confidence = 92.308%. From this rule, it can be concluded that when enterprises have more liabilities, they will have higher tax risk.

Rule 4: Asset-liability ratio \((V30 ) = 4\); Long-term debt ratio \((V29 ) = 4\); Debt equity ratio \((V34) = 4\); Main Business Cost Rate \((V11) = 4\) → Risk Level = High, Support=11.504%, Confidence = 92.308%. From this rule, it can be concluded that when the enterprise has more liabilities and the cost of operating at the same time is high, the risk of tax payment will be high.

Rule 5: Asset-liability ratio \((V30) = 3\); V29Long-term debt ratio=4; Debt equity ratio \((V34) = 4\); Cost Rate of Main Business \((V11) = 4\) → Risk Level = High, Support=10.619%, Confidence=91.667%. From this rule, we can see that the tax risk is higher when the assets and liabilities of enterprises are on the high side, the ratio of other liabilities is high and the cost rate of main business is high, which is lower than the support and confidence of Rule 5 above. This shows that the tax risk is higher when the enterprises are in high liabilities.

The association rules with high confidence listed above is representative, and can also explain which financial indicators are closely related to high tax risk. The association rules with high confidence listed above is very representative. As a total of 11,053 rules are generated, they cannot be fully displayed due to limited space. The following is illustrated in figure. According to all rules mined by association rules, there is a high tax risk in each rule. The frequency statistics of each financial index in all rules generated are carried out, and the weighted statistics are obtained by multiplying the confidence of the corresponding rules (see figure 2).

![Figure 2](image)

**Figure 2** Histogram chart of financial indicators.

It can be intuitively obtained from the histogram of the proportion of each financial index to the high tax risk of the company in figure 2. The higher the range of the financial indexes in the histogram, the greater the correlation strength with the high tax risk of the company. It can be intuitively seen
from the figure that the ratio of debt to equity, long-term debt ratio, main business profit rate, asset-liability ratio, main business cost ratio and other indicators account for a relatively large ratio, indicating a large impact on the tax risk of enterprises. Objectively speaking, when the debt ratio of an enterprise increases, along with the increase in the cost of main business and the reduction in the profit margin of main business, it will lead to high tax risk. It can be obtained from financial indicators such as accounts receivable turnover rate, profit rate of costs and expenses, other payables, payable taxes, etc. Although these financial indicators and the financial indicators mentioned above have a low degree of correlation with high tax risk, they cannot ignore the impact on enterprise tax risk. Because these indicators also represent the financial status and development of enterprises, the changes of these indicators will also affect the decision-making of enterprises in some development strategies and whether enterprises choose to pay taxes according to law. Therefore, when discussing the impact of corporate tax risk, we should fully consider the impact of financial indicators on corporate tax payment.

4. Conclusion

Through the association rule analysis of the financial index data of listed manufacturing companies, it is concluded that the main financial indicators affecting tax compliance are asset-liability ratio, main business cost ratio, debt-equity ratio, long-term debt ratio, main business profit rate and other indicators. It can be found from the results of association rules that the main financial indicators affecting the tax risk of enterprises are weak in operation and profitability, poor in debt paying ability, and a large number of repayments in arrears. According to the results obtained from the analysis of association rules above, tax collection and management personnel should focus on enterprises with declining profitability and weak debt paying ability when carrying out tax supervision and administration, and such enterprises are at high risk of tax evasion in order to maintain their own development.

Acknowledgments

Thank you for your help and encouragement, and thank you for your guidance. This paper is supported by the Project of Hebei Education Department of China (ZD2017029) and Hebei Social Science Foundation Project (HB18GL045).

References

[1] Bai Jingming, He Ping. 2017. Research on Tax Contribution of China's Manufacturing Industry J. Price Theory and Practice, Res. 3, (May. 2017) 19-23.
[2] Han Xiaoqin. 2018. Empirical Analysis on the Influencing Factors of Enterprise Taxpayer's Tax Compliance: An Empirical Study Based on the Data of Tax Administration and Survey in a City of Guangdong Province. J. Taxation and Economy, Res. 4 (July 2018), 79-87.
[3] Liu Yuan, Lin Aimei, Shan Yadi. 2015. The characteristics and means of financial reporting fraud of Listed Companies in China - based on the punishment announcement of the SFC from 2008 to 2013. J. Monthly Financial Journal, Res. 28 (October 2018), 16-19.
[4] Guo Biao, Xie Mingzhu. Research on Tax Risk of Equipment Manufacturing Industry Based on COSO Risk Management Framework [J]. Journal of Jilin Business College, 2017, 33 (05): 64-67.
[5] Hou Huan. Control of Tax Inspection Risk in Big Data Era [J]. Tax and Economy, 2018 (03): 89-94.
[6] Yang Shanlin, Zhou Kaile. Management Problems in Big Data: Resource View Based on Big Data [J]. Journal of Management Science, 2015, 18 (05): 1-8.
[7] Lu Xiaochen, Qu Zhen, Ma Zefang, Zhang Fan. Application of Internet plus big data algorithm in tax work [J]. tax research, 2017 (02): 108-110.