Local Government Debt Risk Assessment And Early Warning System Based On Machine Learning

Shuai Zhang  
*School of Economic Information Engineering, Southwestern University of Finance and Economics, China*

Haichao Zheng  
*School of Economic Information Engineering, Southwestern University of Finance and Economics, China*

Follow this and additional works at: [https://aisel.aisnet.org/whiceb2022](https://aisel.aisnet.org/whiceb2022)

**Recommended Citation**  
[https://aisel.aisnet.org/whiceb2022/6](https://aisel.aisnet.org/whiceb2022/6)

This material is brought to you by the Wuhan International Conference on e-Business at AIS Electronic Library (AISeL). It has been accepted for inclusion in WHICEB 2022 Proceedings by an authorized administrator of AIS Electronic Library (AISeL). For more information, please contact elibrary@aisnet.org.
Local Government Debt Risk Assessment And Early Warning
System Based On Machine Learning

Shuai Zhang, Haichao Zheng

1School of Economic Information Engineering, Southwestern University of Finance and Economics, China

Abstract: The management and prevention of government debt risk is a global topic. In China, due to problems such as implicit debt and uneven regional fiscal performance, it is particularly necessary to explore how to effectively measure and prevent local government debt risks. In this article, we comprehensively consider the debt status and fiscal performance to design a local government debt risk assessment system. According to the debt risk index (DRI), we define the debt risk levels of local governments and find that debt risk has a rapidly increasing pattern and distinct regional characteristics. In addition, we further design a machine learning-based early warning system to predict the risk level of local government debt in the future. We extensively collect explanatory variables based on the previous literature and illustrate variables with high feature importance. Finally, our local government debt risk early warning system achieves an overall accuracy of 92% on the testing set and has a better performance by comparing it to the general debt risk indicator.

Keywords: local government debt risk, debt risk assessment, early warning system, machine learning.

1. INTRODUCTION

From the Latin American sovereign debt crisis caused by the rapid expansion of short-term debt, the Russian financial crisis caused by fiscal deterioration, and the sovereign debt crises in Iceland and Greece triggered by the 2008 financial crisis, it is obvious that effectively managing and preventing government debt risk is the foundation of national stability and the guarantee of government credibility. In 2014, with the revision of the Budget Law of China, the State Council began to allow local governments to raise funds by issuing bonds independently. The State Council delegated the power of fundraising to local governments, hoping that local governments could raise and use funds more flexibly, but this decision may also lead local governments to unfavorably control debt scale.

As we all know, people have the attribute of "voting with their feet", so people tend to gather in cities with a good economy and well-developed infrastructure. Therefore, local governments have to raise funds for urban construction and provide better welfare for citizens, which will undoubtedly increase their financial burden. Some Chinese government officials lack awareness of debt risks or even blindly pursue political achievements, which can easily lead to debt crises[1]. Besides, local governments often issue bonds through local financing institutions. These bonds appear to be sold by financing institutions, but in reality, they have credit endorsements from local governments. Once there is an economic downturn or capital operation failure, the implicit debt risk will be exposed[2]. Furthermore, The land transfer fee is the main source for Chinese local governments to repay debts. However, due to the scarcity of land and the depression of the real estate industry in China, selling land to alleviate the debt burden is no longer healthy and applicable[3].

To sum up, Chinese local government debt risk management has various complicated problems that need to be solved urgently, and it is necessary to turn "soft constraints" into "hard constraints". Therefore, we are looking forward to designing an effective and comprehensive local government debt risk assessment and early warning system, thus providing a reference for local governments to conduct debt risk management.
2. LITERATURE REVIEW

2.1 Controversy over government financing.

The debate about government financing has been around for a long time. Adam Smith expressed his opposition to government financing in *The Wealth of Nations* since he believed it would occupy private capital and hinder the natural development of the national economy. Coincidentally, Ricardo’s "equivalence theory" also expressed opposition to government financing. He suggested that government debt was merely a delay in raising the tax rate in the future, but it would cause the government to profligate and waste money. Kumar also pointed out that excessive government debt will lead to an increase in the country's long-term interest rate and at the same time cause inflation[4]. Furthermore, Westphal put forwards that government debt will harm economic growth when the government debt ratio is higher than 90%[5].

Conversely, a group of economists led by Keynes argued that government financing can not only stimulate economic growth when demand is insufficient but also provide more employment opportunities. According to the economic recovery plan of the US government during the financial crisis, Robert found that government financing can promote economic recovery in the short term since governments can raise funds to help small and mid-sized enterprises increase the prices of their financial assets[6]. Cai claimed that government financing can effectively balance the capital allocation in the regional economy and alleviate polarization effects[7]. Furthermore, Zhao stated that local government financing can promote urban development in three aspects: enhancing residents' welfare, improving infrastructure construction, and strengthening the ability to deal with risks[8].

Therefore, we find government financing is conducive to promoting economic growth, balancing resource allocation, and increasing residents' welfare, as long as it is on an appropriate scale. However, how to reasonably assess the government debt scale and how to prevent debt risk in advance are still remained obscure.

2.2 Previous research on analysis and prevention of local government debt risk.

In the research on local government debt analysis, Polackova proposed a fiscal risk matrix, which classifies government debt into direct debt, indirect debt, explicit debt, and implicit debt. This matrix has deepened our understanding of government debt more comprehensively[9]. Ruzzante claimed that government fixed assets should not be counted as debt-repayable assets, and officials should pay more attention to government liquid assets. Therefore, the GDP growth and fiscal surplus are the keys to effectively alleviating the pressure on the government debt burden[10]. Duca suggested that domestic systemic risks are often related to the global economic status, so he reconstructed the Financial Stress Index by introducing international macroeconomic variables[11].

Similarly, many scholars are concerned about the debt risks of local governments in China. Feng revealed that the disorderly development of local financing institutions is the main reason for the excessive expansion of local debt[2]. Besides, Zhong found that many local governments strongly rely on the central government’s fiscal transfer payments, and the phenomenon of issuing new bonds to repay old debts is becoming more and more serious[12]. By considering the competitive relationship, Wu stated that the local government will fully consider the financing strategies of neighboring governments when making financing decisions[13]. Furthermore, Zhao analyzed debt from the aspect of the debt stock and fiscal revenue, thus suggesting doing adjustments in the fiscal balance to prevent local government debt risks[14]. However, we can easily find that these previous researches only focused on one factor or aspect that causes debt risk, and different factors were lack of comparison. Also, the prediction of local government debt risk seems deficient. By combing through the literature on government debt risk analysis and prevention, we find the following three problems in previous studies:

- Most of the literature only used the simple indicator "debt ratio" as the explained variable. However, due to the large differences in debt status and fiscal performance between different governments, the "debt ratio" indicator cannot truly reflect the debt resolution capacity of a government itself. It also should not be used as a general indicator to delineate the risk threshold for all governments.
Previous studies only explored one factor that causes debt risk, so there was a lack of horizontal comparison between different factors. In this case, we do not know which factor is more significant. This deficiency creates difficulties for local governments to focus on key issues in debt management.

In prior researches, we find that most researchers used econometric methods to analyze the causes of debt risk, but there was little literature focused on debt risk prediction. However, forecasting future debt risk levels is of great significance to the local government debt management.

2.3 Machine learning.

Machine learning is a science (and art) of computer programming because they can learn from data. It is one of the fastest-growing technical fields today and is also the core of artificial intelligence. Atthey mentioned several advantages of machine learning in dealing with economic problems: (1) Machine learning can deal with unstructured data and capture non-linear features; (2) Machine learning can completely describe the model selection process; (3) Machine learning can better complete prediction and classification tasks, which introduces more possibilities for dealing with economic problems\textsuperscript{[15]}. In this research, we will not only use classic machine learning algorithms but also implement ensemble models, such as Random Forest and Gradient Boosting Decision Tree\textsuperscript{[16]}. These machine learning algorithms and ensemble models already have a wide range of applications in economic scenarios, thus providing a solid theoretical foundation and possible solutions for our study.

3. DESIGNING THE LOCAL GOVERNMENT DEBT RISK ASSESSMENT SYSTEM

3.1 Debt risk index.

The previous literature only used the simple indicator "debt ratio" as the explained variable, which is the ratio of the total balance of government bonds to the fiscal revenue. However, due to issuing new bonds to repay old debts and the existence of central transfer payments in fiscal revenue, the "debt ratio" can no longer truly evaluate the debt resolution capacity of local governments themselves. Also, because of the existence of implicit debts from local financing institutions, the total debt balance of local governments is opaque and underestimated. In this case, we need to take local financing institutions into assessment consideration. Furthermore, because of the discrepancies in local government fiscal performances, it is unreasonable to delineate a general risk threshold by "debt ratio" for all provinces. Thus, it is urgent to construct a better debt risk assessment system.

In this case, we establish a comprehensive and effective risk assessment system from two aspects: debt status and fiscal performance. In debt status, we introduce two secondary indicators of scale risk and repayment risk, aiming to characterize debt status from the debt scale and repayment pressure. It is worth noting that the debt scale and repayment amount include not only government bonds but also local financing institution bonds; In fiscal performance, we introduce two secondary indicators of expenditure risk and revenue risk, thus characterizing fiscal performance from two perspectives: fiscal self-sufficiency and local government revenue dependence on issuing new bonds. The indicators of the debt risk assessment system are shown in Table 1.

<table>
<thead>
<tr>
<th>Primary indicators</th>
<th>Secondary indicators</th>
<th>Indicator calculation</th>
<th>Indicator meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>Debt status ($P_{1}$)</td>
<td>Scale risk ($S_{11}$)</td>
<td>(Annual debt balance / annual GDP) *100%</td>
<td>Debt burden ratio: measure debt balance and debt scale relative to GDP</td>
</tr>
<tr>
<td>Repayment risk ($S_{12}$)</td>
<td>(Annual repayment amount / annual fiscal revenue) *100%</td>
<td>Debt service ratio: reflect the debt repayment ability of government revenue</td>
<td></td>
</tr>
<tr>
<td>Fiscal performance ($P_{2}$)</td>
<td>Expenditure risk ($S_{21}$)</td>
<td>(Annual fiscal expenditure / annual fiscal revenue) *100%</td>
<td>Fiscal self-sufficiency ratio: measure the excess of fiscal expenditure</td>
</tr>
<tr>
<td>Revenue risk ($S_{22}$)</td>
<td>(Annual debt revenue / annual fiscal expenditure) *100%</td>
<td>Debt dependence ratio: reflect the dependence of fiscal revenue on issuing new bonds</td>
<td></td>
</tr>
</tbody>
</table>
According to the methodology of constructing the financial stress index by Duca\cite{11}, we similarly assign the same weight to the secondary indicators (SI). We use DRI to represent the local government debt risk index. Besides, we introduce the international generally recognized risk threshold of secondary indicators. The risk threshold (RT) of the debt burden ratio is 60%, the debt service ratio is 20%, the fiscal self-sufficiency ratio is 100%, and the debt dependence ratio is 30%. The DRI is computed for province \( i \) at year \( t \) as follows:

\[
DRI_{it} = \frac{\sum_{j=1}^{4} SI_{ij}}{4}
\]  

Therefore, we can easily find that \( DRI = 1 \) is the risk threshold. However, the general risk threshold is usually based on the governments of developed countries. As a developing country with a unique political system, it is acceptable that Chinese local governments’ \( DRI \) is higher than 1. Therefore, we define that \( DRI \leq 1 \) is the low-level risk, \( 1 < DRI \leq 2 \) is the mid-level risk, and \( DRI > 2 \) is the high-level risk. In this case, we establish a better debt risk assessment system than the general debt risk indicator.

### 3.2 Descriptive statistics based on DRI.

We collect relevant data of all 31 provinces in China from 2001 to 2020 and calculate their DRI. Among the entire 620 samples, 140 cases are high-level debt risk. High-level risk cases account for 23% of the total samples, indicating that the debt risk of local government is an urgent problem that needs to be managed immediately. The numbers of cases on different debt risk levels in different regions are shown in Table 2.

<table>
<thead>
<tr>
<th>Regions</th>
<th>Cases on low-level risk</th>
<th>Cases on mid-level risk</th>
<th>Cases on high-level risk</th>
<th>Total number</th>
</tr>
</thead>
<tbody>
<tr>
<td>The eastern provinces and municipalities</td>
<td>166</td>
<td>36</td>
<td>38</td>
<td>240</td>
</tr>
<tr>
<td>Other inland provinces</td>
<td>206</td>
<td>72</td>
<td>102</td>
<td>380</td>
</tr>
<tr>
<td>All provinces</td>
<td>372</td>
<td>108</td>
<td>140</td>
<td>620</td>
</tr>
</tbody>
</table>

Besides, we find that the proportion of debt risk has distinct regional characteristics. The proportions of different debt risk levels are shown in Figure 1. Although the amount of debt in the eastern provinces and municipalities is large, the proportion of high-level risk cases is significantly lower than other inland provinces because of their developed economy, reasonable economic structure, and high fiscal revenue. In contrast, the inland provinces have more cases of mid-level risk and high-level risk.

![Figure 1. The proportions of debt risk levels in different regions](image-url)
The development of local government debt risk also has an increasing pattern in time. We average the risk levels of all provinces yearly from 2001 to 2020 and visualize the debt risk development pattern in Figure 2. We find that the average debt risk level of inland provinces is always higher than that of eastern provinces and municipalities. Before 2014, the average local government debt risk level was between low-level and mid-level. After 2014, the debt risk increases significantly and has risen to between mid-level and high-level. In 2020, the average local government debt risk level of all provinces is 2.93, which is very close to 3.

Figure 2. Local government debt risk development pattern

4. DESIGNING THE LOCAL GOVERNMENT DEBT RISK EARLY WARNING SYSTEM

4.1 Explained variables and explanatory variables.

In this section, we construct the explained variable and explanatory variable sets for the debt risk early warning system. We use the debt risk level calculated by DRI as the explained variable. By labeling Low-level Risk as 1, Mid-level Risk as 2, and High-level Risk as 3, we construct a three-category explained variable.

In terms of explanatory variables, we refer to the conclusions of the past literature on the causes of debt risk, thus constructing the explanatory variable set from four aspects: local government debt profile, local government fiscal profile, local economic development profile, and national macroeconomic profile. The following variables are the annual data of all 31 provinces from 2001 to 2020, collected from Wind and CSMAR databases.

4.1.1 Local government debt profile

Local government bond stock: the number of government bonds ($x_1$), the balance of government bonds ($x_2$), the proportion of government bonds ($x_3$).

Local financing institution bond stock: the number of local financing institution bonds ($x_4$), the balance of local financing institution bonds ($x_5$), the proportion of local financing institution bonds ($x_6$).

Total bond stock: the total number of bonds ($x_7$), the total balance of bonds ($x_8$), the growth rate of total bond balance ($x_9$).

Issuance and redemption of local government bonds: the issuing amount of government bonds ($x_{10}$), the issuing number of government bonds ($x_{11}$), government bond redemption amount ($x_{12}$), government bond redemption number ($x_{13}$), government bond net financing amount ($x_{14}$).

Issuance and redemption of local financing institution bonds: the issuing amount of local financing institution bonds ($x_{15}$), the issuing number of local financing institution bonds ($x_{16}$), local financing institution bond redemption amount ($x_{17}$), local financing institution bond redemption number ($x_{18}$), local financing institution bond net financing amount ($x_{19}$).
bonds \( (x_{15}) \), the issuing number of local financing institution bonds \( (x_{16}) \), local financing institution bond redemption amount \( (x_{17}) \), local financing institution bond redemption number \( (x_{18}) \), local financing institution bond net financing amount \( (x_{19}) \).

Issuance and redemption of total bonds: total issuance amount \( (x_{20}) \), total issuance number \( (x_{21}) \), total redemption amount \( (x_{22}) \), total redemption number \( (x_{23}) \), and total net financing amount \( (x_{24}) \).

### 4.1.2 Local government fiscal profile

Fiscal revenue status: local government fiscal revenue \( (x_{25}) \), fiscal revenue growth rate \( (x_{26}) \).

Fiscal expenditure status: local government fiscal expenditure \( (x_{27}) \), fiscal expenditure growth rate \( (x_{28}) \).

### 4.1.3 Local economic development profile

Economic scale and development: GDP \( (x_{29}) \), GDP growth rate \( (x_{30}) \).

Economic structure: primary industry GDP \( (x_{31}) \), secondary industry GDP \( (x_{32}) \), tertiary industry GDP \( (x_{33}) \), the proportion of primary industry GDP \( (x_{34}) \), the proportion of secondary industry GDP \( (x_{35}) \), the proportion of tertiary industry GDP \( (x_{36}) \).

### 4.1.4 National macroeconomic profile

Macroeconomic trends: Shanghai Composite Index \( (x_{37}) \), inflation rate \( (x_{38}) \).

Therefore, we constructed an explanatory variable set containing 38 variables from the above four aspects of all 31 provinces from 2001 to 2020 annually.

### 4.2 Constructing the training set and testing set.

In order to predict the future debt risk levels, we need to use the explanatory variables of the past years to predict the future explained variable. In this case, we concatenate the explanatory variable sets of the past three years \( (X_{t-3}, X_{t-2}, X_{t-1}) \) as new explanatory variable set \( X \) and the explained variable of the current year \( (Y_t) \) as the forecast target \( Y \).

Finally, we get 496 samples, we divided them into training and testing samples according to the proportion of 4: 1, thus getting our training and testing set for debt risk early warning system based on machine learning.

### 4.3 Modeling and results.

We implement machine learning to construct the local government debt risk early warning system. We not only use classic machine learning, such as Naive Bayes, the K-Nearest Neighbors, and Support Vector Machines but also use ensemble models, such as Random Forest and Gradient Boosting Decision Tree. The entire local government debt risk early warning system based on machine learning is shown in Figure 3.

![Figure 3. The entire local government debt risk early warning system](image-url)
We use the training set to train these machine learning models, and acquire their prediction performance on the testing set. The prediction performance of these machine learning models is shown in Table 3.

<table>
<thead>
<tr>
<th>Machine learning models</th>
<th>Accuracy on the testing set</th>
</tr>
</thead>
<tbody>
<tr>
<td>Support Vector Machines</td>
<td>59.4%</td>
</tr>
<tr>
<td>Naive Bayes</td>
<td>75.0%</td>
</tr>
<tr>
<td>K-Nearest Neighbors</td>
<td>80.2%</td>
</tr>
<tr>
<td>Decision Tree</td>
<td>81.3%</td>
</tr>
<tr>
<td>Gradient Boosting Decision Tree</td>
<td>89.6%</td>
</tr>
<tr>
<td>Random Forest</td>
<td>90.6%</td>
</tr>
</tbody>
</table>

According to the results, we can easily find that Random Forest has the highest overall accuracy. In this case, we further calculate its precision, recall, and F-measure, thus exploring its prediction performance on different risk levels. The prediction performance on different risk levels of Random Forest is shown in Table 4.

<table>
<thead>
<tr>
<th>Risk Level</th>
<th>Precision</th>
<th>Recall</th>
<th>F-Measure</th>
<th>Support</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low-level Risk</td>
<td>0.98</td>
<td>0.98</td>
<td>0.98</td>
<td>57</td>
</tr>
<tr>
<td>Mid-level Risk</td>
<td>0.71</td>
<td>0.71</td>
<td>0.71</td>
<td>17</td>
</tr>
<tr>
<td>High-level Risk</td>
<td>0.88</td>
<td>0.88</td>
<td>0.88</td>
<td>25</td>
</tr>
<tr>
<td>Overall accuracy</td>
<td>0.92</td>
<td></td>
<td></td>
<td>99</td>
</tr>
</tbody>
</table>

We find that Random Forest has better performances on low-level risk prediction and high-level risk prediction, and performs slightly worse on mid-level risk prediction. This may be due to the rapid increase of debt risk in many provinces after 2014, resulting in a small sample size of mid-level risk. However, the local government debt risk early warning system performs well in the prediction of high-level risk (F-measure = 0.88). In this case, we can accurately forecast high-level debt risk in advance.

### 4.4 Feature importance.

Exploring explanatory variables that have a greater impact on the future debt risk level and making the horizontal comparison of different explanatory variables is a focus of this study. In this way, we are looking forward to helping the local government focus on key issues in debt management.

We use the feature importance of Random Forest to help us achieve this goal. Feature importance is generally calculated by Gini Index. We use $G_{I}$ to represent Gini Index, $FI$ to represent feature importance.

Gini Index is calculated as follows:

$$GI_d = \sum_{k=1}^{K} \sum_{k'} \sum_{s} p_{d,k} p_{d,k'} = 1 - \sum_{k=1}^{K} p_{d,k}^2$$  \hspace{1cm} (2)

The number of sample categories is $K$. $p_{d,k}$ represents the proportion of category $k$ at node $d$.

The feature importance of the explanatory variable $x_j$ at node $d$ is the change of Gini Index before and after node $d$ branching:

$$FI_{j,d}^{(Gini)} = GI_d - GI_l - GI_r$$  \hspace{1cm} (3)

$GI_l$ and $GI_r$ is the Gini Index of the two new nodes $l$ and $r$ after branching at node $d$. Supposing $D$ is the node set of $x_j$ in decision tree $i$, then the feature importance of $x_j$ in tree $i$ is:

$$FI_{i,j}^{(Gini)} = \sum_{d \in D} FI_{i,d}^{(Gini)}$$  \hspace{1cm} (4)

Supposing that the number of decision trees in the Random Forest is $n$, then:
Finally, assuming that the number of explanatory variables is $M$, we normalize the feature importance of the explanatory variable $x_j$:

$$FI_j = \frac{FI_j^{(Gini)}}{\sum_{m=1}^{M} FI_m^{(Gini)}}$$ (6)

In this way, we can acquire the feature importance of each explanatory variable. We show the ten variables with the highest feature importance in Random Forest and their feature importance scores in Figure 4.

![Feature importance scores of the ten variables with the highest FI](image)

**Figure 4. Feature importance scores of the ten variables with the highest FI**

For predicting the risk level of year $t$, $x_{17,t-1}$ is the local financing institution bond redemption amount of year $t-1$, $x_{10,t-1}$ is the issuing amount of government bonds of year $t-1$, $x_{1,t-1}$ is the number of government bonds of year $t-1$, $x_{22,t-1}$ is the total redemption amount of year $t-1$, $x_{16,t-2}$ is the issuing number of local financing institution bonds of year $t-2$, and $x_{14,t-1}$ is the government bond net financing amount of year $t-1$. We find that bond redemption pressure ($x_{17,t-1}$) is the most important feature, and bond redemption pressure often comes from excessive debt issuance scale ($x_{10,t-1}$ and $x_{16,t-2}$). Therefore, the primary task of preventing debt risks is to control the scale of debt issuance, thereby alleviating the future debt redemption pressure. This also requires the joint efforts of local governments and financing institutions.

Besides, we find that local economic structure also has an important impact on future debt risk. $x_{36,t-3}$ is the share of tertiary industry GDP of year $t-3$, $x_{36,t-1}$ is the share of tertiary industry GDP of year $t-1$, and $x_{32,t-3}$ is the secondary industry GDP of year $t-3$. It is easy to find that adjusting the economic structure and accelerating the development of the tertiary industry is crucial to increasing fiscal revenue. $x_{28,t-1}$ is the local government fiscal expenditure growth rate of year $t-1$, which indicates that we also need to control fiscal expenditure within a reasonable scale and maintain the balance between fiscal revenue and expenditure.

### 4.5 Comparing our debt risk assessment and early warning system to the general indicator.

Different debt risk indicators are different in constructing concepts and modeling methods. In this research, we select several indicators from two perspectives of debt status and fiscal performance to design a comprehensive assessment system, thus ensuring the robustness of the assessment results. However, the general debt risk assessment indicator "debt ratio" is still worth to be considered and compared to our assessment system.

We calculate the yearly debt ratio ($DR$) for all 31 provinces and find that the risk threshold of the debt ratio
is 100%. By implementing the same methodology for defining the risk level in our assessment system, we define that $DR \leq 1$ is the low-level risk, $1 < DR \leq 2$ is the mid-level risk, and $DR > 2$ is the high-level risk.

Among the entire 620 samples, 82 cases are high-level risk under the assessment of $DR$, accounting for 13%. However, high-level risk cases account for 23% under our assessment system based on $DRI$. This comparison indicates that our comprehensive debt risk assessment system can effectively recognize more potential high-level risk cases than the general indicator. Furthermore, we find that our debt risk assessment system can detect the deterioration of debt status earlier than the general indicator. We compare the average risk level growth pattern of all provinces based on $DR$ and $DRI$ in Figure 5. Our assessment system detects the rapid risk growth trend in 2014. In contrast, the general assessment system based on $DR$ responses to the deterioration in 2016.

![Figure 5. The comparison of the average debt risk level based on DR and DRI](image)

We also use the debt risk level based on $DR$ as explained variable to construct an early warning system. Random Forest still has the best performance on the testing set with the overall accuracy of 91%. Similarly, we further explore the ten variables with the highest feature importance for the early warning system based on $DR$. However, these variables are mainly about the issuance and redemption of local government because the assessment indicator $DR$ is relatively simple and partial. These variables are $x_{10,t-1}$, $x_{14,t-1}$, $x_{2,t-1}$, $x_{1,t-1}$, $x_{11,t-1}$, $x_{20,t-2}$, $x_{15,t-2}$, $x_{21,t-1}$, $x_{20,t-1}$, $x_{10,t-2}$, illustrating that they are all belonged to local government debt profile. Apparently, these variables cannot indicate the key issues from the economic structure and the balance between fiscal revenue and expenditure like our early warning system does.

To sum up, our debt risk assessment system based on $DRI$ can effectively recognize more potential high-level risk cases and detect the deterioration of debt status earlier than the general indicator $DR$. Besides, our debt risk early warning system has a better overall prediction accuracy than the system based on $DR$, and our early warning system can indicate the key issues of debt risk management more comprehensively.

5. CONCLUSIONS AND LIMITATIONS

In this research, we design a local government debt risk assessment system from two perspectives of debt status and fiscal performance. According to the debt risk index ($DRI$), we define the debt risk levels of local governments. We find that the debt risk has distinct regional characteristics since the proportion of high-level risk cases in eastern provinces and municipalities is significantly lower than that in other inland provinces. Also, the development of local government debt risk has a rapidly increasing pattern especially after 2014, which indicates
that the debt risk of local government is an urgent problem that needs to be managed immediately.

Based on the debt risk index (DRI) and extensively collected explanatory variables, we implement machine learning to establish a local government debt risk early warning system. We use various machine learning algorithms and ensemble models, thus acquiring their prediction performance on the testing set. We find that Random Forest has the best prediction performance, with an overall accuracy of 92% and a high-level risk accuracy of 88%. Furthermore, we use feature importance to make the horizontal comparison of explanatory variables. We illustrate ten variables with the highest feature importance and explain how we can better arrange the debt management by focusing on key issues according to these variables. Finally, we verify that our debt risk assessment and early warning system has a better performance than the general debt risk indicator.

There are still some limitations of our research. Due to the limitation of data sources, we are unable to obtain some of the explanatory data we expect, such as land transfer fees of local governments. Besides, we only assess and forecast the local government debt risk from the perspective of a single province. However, there are also risk spillover effects between neighboring provinces. Therefore, how to comprehensively consider individual government risk and regional risk spillovers is an important topic of future work on debt risk management.

ACKNOWLEDGEMENT

This research was supported by the Natural Science Foundation of China (72071160), and the Chinese Ministry of Education Humanities and Social Sciences Fund (18YJAZH142).

REFERENCES


