

Research on Green Bond Default Risk Prediction Based on Deep Learning

Zhaoyang Wang

Jiangsu Normal University, Nantong, Jiangsu, 226600, China

ABSTRACT

In order to realize the "double carbon target", China's green bond market has been developing and expanding, and has become the largest green bond market in the world. Therefore, in order to effectively regulate green bonds, it is indispensable to effectively assess the default risk of green bonds. This study mainly combed the relevant literature on green bonds and bond default risk, and found that the existing green bond research mainly adopts linear econometric models, while there are many nonlinear factors in the field of green bonds, such as ESG rating indicators, etc. Combined with the results of the deep learning model in the default risk of ordinary bonds, it puts forward a new idea for the research of green bond default risk, i.e., to start from the deep learning technology instead of the traditional linear model.

KEYWORDS

Green bonds; Bond default; Deep learning

1 Introduction

The bond market, as an important part of China's financial system, optimizes resource allocation and promotes the development of the real economy, however, bond defaults are commonplace, according to the wind database between 2018 and 2023, more than 800 bonds defaulted, with an average annual default amount of more than 120 billion, far more than the more than 30 billion in 2017. More importantly, AAA-rated state-owned enterprises such as Yong coal and Ziguang defaulted beyond expectation (Jiang Fuwei, Chai Bailin et al., 2024) [1], reflecting the gradual expansion of bond risks and the need to strengthen the reliability of the bond rating model. In 2020, against the background of the increasingly prominent global environmental problems, Chinese leader proposed the "dual-carbon goal" at the 75th session of the United Nations General Assembly, the green bond market has been expanding, and China has become the largest country in the world's green bond market.

Green bonds are financial instruments to invest in environmentally friendly projects. Due to the existence of project specificity, the prediction of default risk of green bonds has higher requirements compared to ordinary bonds. In order to promote the green economy, since May 1, 2024, Shanghai, Shenzhen and North three exchanges jointly issued the "Guidelines on Sustainability Reporting for Listed Companies", marking the domestic ESG information disclosure into a unified and standardized stage; March 28, 2025, the China Securities Regulatory Commission (CSRC) issued the revised "Measures for the Administration of Disclosure by Listed Companies" and other documents, for the first time to clarify the requirements for ESG disclosure at the level of departmental regulations. In order to realize the goal of "Double Carbon" and respond to the call of "Green Water and Gold Mountain is Golden Mountain and Silver Mountain", in accordance with the call of "Green Water and Gold Mountain is Golden Mountain and Silver Mountain", the CSRC, through the "Green Water and Gold Mountain is Golden Mountain and Silver Mountain", has been working hard to achieve the goal of "Dual Carbon". In order to realize the goal of "double carbon" and respond to the call of "green water and gold mountain is golden mountain and silver mountain", we should solve the problem of sustainable development by practicing the concept of ESG and playing the market level. In this context, how to predict bond risk under ESG indicators has become a new problem to be solved. This study intends to use deep learning models, combined with the information reflected in ESG indicators, innovative credit modeling methods, more able to make credit risk under ESG indicators more accurate, so that the risk warning ability to improve, to help investors make better decisions. Combing through the existing literature, the research on the bond default risk of green bonds can be viewed from two perspectives: the first perspective is to study the impact of the attributes of the green bonds themselves on the bond default risk; the other perspective is to study the bond risk prediction model to predict the default risk.

2 Literature review

2.1 Green bonds

The 2024 Bund Green Finance Report points out that in order to achieve the "dual carbon" goal, the amount of capital required is very large, up to one trillion dollars, of which the amount of capital for long-term financing is larger and the investment risk is higher, so it is difficult for ordinary financial instruments to perform this function, which requires a new type of financial instruments --Green bonds. Green bonds are an environmentally friendly financial instrument that can

raise capital for corporate activities and at the same time fulfill environmental protection policies. With the continuous development of the green bond market, the types of green bonds have gradually become rich, and can be categorized into ordinary green bonds, carbon revenue green bonds, green project revenue bonds, green asset-backed bonds, and a number of innovative sub-products. With the continuous innovation and development of green bonds, the risk of bond default has gradually increased, and the requirements for financial supervision and risk prediction have been increasing, which will also have a certain impact on corporate financial risk.

On the relationship between the attributes of green bonds themselves and the risk of corporate debt default, there are now many scholars research. As a new type of financial instrument, the issuance of green bonds can increase public attention, thus reducing moral hazard and improving the financial situation of enterprises; Xia Juan and other scholars use the double-difference method, and through the test to discover the mechanism by which the green bonds themselves can reduce the risk of default: on the one hand, the enterprises can issue green bonds to improve the reputation of the enterprises, thus reducing the risk of default, on the other hand, the green bonds can optimize the allocation of resources between enterprises and banks, and the green bonds can be optimized for the enterprises and banks, thus reducing the risk of default. On the other hand, green bonds can optimize the resource allocation of enterprises and banks, banks can optimize their credit allocation, and enterprises can reduce the situation of short-term loans and long-term investment (Xia Juan, Yang Jianing et al., 2025) ^[2] These all support the view that green bonds may reduce the risk of corporate debt default.

However, Li Xuefeng and Yu Xinqi likewise present a different view through a double difference model. They divided the enterprise investment behavior from investment scale and investment efficiency, firstly, from the perspective of investment scale, due to the existence of policy support, the enterprise's capital term allocation is more flexible, and the investment scale is enlarged; from the perspective of investment efficiency, under the loose financing environment, the enterprise may produce irrational investment due to the issuance of green bonds, especially the overinvestment of the light pollution enterprises is more obvious, and the alleviation of financing constraints creates a complete intermediary effect between green bonds and overinvestment, which will make enterprises overestimate their investment capacity and then enhance the risk of debt default. (Xuefeng Li, Xinqi Yu, 2022) ^[3]

In summary, the relationship between green bonds and corporate debt default risk and the role of the mechanism under the traditional econometric modeling research has not appeared uniform conclusions, and still need to be further explored in depth, especially in the context of ESG factors are increasingly emphasized, the expansion of the relevant research will be a richer theoretical support for the practice of green finance and enterprise risk management.

2.2 Default risk research

2.2.1 Traditional research

Among the traditional credit models, linear regression models, such as the Altman Z-value model, are often used. However, traditional credit risk models and research ideas rely on linear assumptions and structured data, making it difficult to capture complex nonlinear relationships. Especially in the study of default risk of green bonds, green bonds are closely related to ESG concepts, but it is unrealistic to rely on traditional linear regression models to capture the nonlinear impact of ESG and other unstructured data on bond default risk. Among ESG indicators, there are many environmental factors, social factors, and corporate governance factors that cannot be summarized into a linear model, and naturally, precise conclusions cannot be obtained. For example, the above two studies on the attributes of green bonds on bond default risk, although both use double difference models, but due to the different perspectives of entry, inconsistent conclusions were reached, which further confirmed the lack of precision of the traditional linear regression model. Some scholars use the Logistic model to study the default risk of corporate bonds, linearize some information, and get the conclusion that the default risk of corporate bonds in China is increasing. (Pan Zeqing, 2018) ^[4] But its as a linear model, the same can not yet fully grasp the nonlinearized data, so the conclusion has the possibility of further improvement.

2.2.2 Machine learning and deep learning

Machine learning and deep learning are core technologies in the field of artificial intelligence, which can realize intelligent decision-making and prediction through data-driven methods, can capture nonlinear information more accurately, and have great potential in risk prediction. Combing through the existing literature, machine learning and deep learning can study bond default risk from the following aspects.

First, emotions can be extracted from textual information, and then the textual emotions can be analyzed, so as to control the market investment emotions and screen the bonds that have the possibility of default. Scholars such as Qi Xinlin and Lin Yihao started from two aspects of bond default factors and risk early warning models, and utilized machine learning models to conduct research. First, the study points out four levels of bond default factors by combing through the literature, which are divided into macro level, financial market level, corporate characteristics level, and bond characteristics level. Subsequently, it also starts with the judgments of classical models, nonlinear models and bond credit

rating models, pointing out the shortcomings of these models, thus leading to the study of bond risk under the perspective of machine learning. The specific perspectives are: firstly, the impact of different textual information on bond default will be different, so it is necessary to go to expand the textual analysis material; secondly, the direct use of sentiment dictionary to recognize textual sentiment is not accurate, and it is necessary to reconstruct a featured sentiment dictionary; lastly, it is possible to start from the aspect of the sentiment extraction technology, and adopt deep learning technology to extract textual sentiment more accurately. (Qi Xinlin, Lin Yihao, 2023) ^[5]

Secondly, the deep learning model is utilized to predict bond default risk. With the continuous deep research in the field of deep learning theory, there are many scholars who integrate deep learning models with the financial field and construct deep learning models that are more accurate than a single linear model and a single neural network. In this context, scholars such as Jiang Fuwei and Chai Bailin systematically studied the role of deep learning models in bond risk identification. Firstly, they refer to the GAN architecture and integrate the deep neural network to construct a deep learning-based corporate bond credit risk prediction model - CDL model. The CDL model they constructed can be mainly divided into two parts, the generative model and the discriminative model, and through the division of labor between the two parts of the model, it can efficiently learn the complex nonlinearity of high-dimensional data. It is found that compared with the classical machine learning model and common neural network model, the CDL deep learning model is able to predict the credit risk of corporate bonds more accurately, and the accuracy of risk prediction for the higher risk bonds is higher than that for the lower risk bonds, which demonstrates the great potential of the deep learning model in the prediction of bond default risk. (F. W. Jiang, B. L. Chai, 2024) ^[1]

2.3 Deep learning on green bonds

There are few studies in the existing literature that specialize with green bonds and deep learning on bond defaults, but the study of green bond defaults using deep learning is yet to be developed and is a highly innovative field. Scholar Tang Zhuoyuan utilizes deep learning techniques and knowledge pictures to study the green bond market. In the field of green bonds, the long and short-term memory network (LSTM) has advantages in time series data processing, while there are still some limitations in the long-term dependence of local features and long time series, so the study integrated the LSTM model, GRU model and knowledge graph technology for the prediction of the green bond index, and its experimental knot confirmed the effectiveness of deep learning models, and found that the deep learning model is not as effective as other deep learning models. The experimental results confirmed the effectiveness of the deep learning model and found some correlations between the green bond market and other markets such as the energy market. (Tang Zhuoyuan, 2024) This study verifies the reliability of the combination of deep learning models with new ideas in green bond research, while in terms of subject matter, this study mainly studies the green bond index rather than the default risk of green bonds, and there is still room for more in-depth research in this direction ^[6].

3 Conclusion

In summary, the default risk research focusing on green bonds is still mainly concentrated in the traditional measurement direction. In this paper, we mainly sort out the existing researches from the two aspects of green bonds and bond default risk prediction, and prove the feasibility of deep learning model in the study of green bond default risk that synthesizes the two directions from the existing deep learning model in green bonds and bond default risk prediction research. And, according to the existing deep learning research in the green bond market, the LSTM model has a great role, however, it also has certain limitations. Subsequent research can be based on the LSTM model with other deep learning models to construct a multimodal architecture to make up for the shortcomings of a single model.

References

- [1] F.W. Jiang, B.L. Chai, Y.H. Lin Deep. Learning and Corporate Bond Credit Risk.
- [2] Juan Xia, J.N. Yang, Shuangling Zhang. Can Green Bond Issuance Reduce the Risk of Corporate Debt Defaults? -- Based on the Perspective of Reputation Mechanism and Credit Allocation.
- [3] Xuefeng Li, Xinqi Yu, The Impact of Green Bond Issuance on the Investment Behavior of Listed Firms.
- [4] Zecheng Pan, Logistic Regression Early Warning Model of Corporate Debt Default Risk.
- [5] Xinlin Qi, Yihao Lin, Fuwei Jiang, Bond Default Risk and Machine Learning. Bond Default Risk and Machine Learning Methods: Research Progress and Future Directions.
- [6] Zhuoyuan Tang, "Green Bond Market Analysis and Index Prediction: An Investigation Based on Knowledge Graph and Deep Learning Techniques".