Exploring the Credit Rating System Inefficiencies by Using Machine Learning*

Lei Ge† , Pedro Gete‡ and Ziyan Yu§

September 2025

Abstract

Corporate credit rating systems, whose opacity and inefficiencies were starkly exposed during the 2008 financial crisis, function as a "black box", obscuring the rationale behind critical financial assessments. This paper employs machine learning techniques to decode this opaque process and demonstrates that such methods are effective in identifying systematic inefficiencies within the credit rating system. Utilizing Chinese corporate credit rating data, we identify a systematic rating bias against private firms that lack implicit government backing, which exacerbates financing disparities between state-owned and private enterprises. Furthermore, this biased rating system generated unintended consequences under government credit policies. Specifically, we show that this bias amplified the funding difficulties encountered by private firms during the 2016 macroprudential "Deleveraging" campaign.

Keywords: Machine Learning, Credit Rating, Credit Policy

JEL Classification: C45, C53, G17, E44, G28, G30, G32, G34

*We appreciate the comments of Francois Schmitt, Lan Wu, Dacheng Xiu, Jingping Yang, Steven Zhu and seminar participants at The 9th PKU-NUS Annual International Conference on Quantitative Finance and Economics.

†Renmin University of China, School of Economics. Email: gelei@ruc.edu.cn

‡IE University, IE Business School. Email: pedro.gete@ie.edu.

§Tsinghua University, School of Economics and Management. Email: yuziyan25@mails.tsinghua.edu.cn

1 Introduction

A prevailing narrative regarding the onset of the 2008 financial crisis in the United States posits that credit rating agencies, including Moody's, Standard & Poor's (S&P), and Fitch, significantly underestimated the risks associated with large companies and financial institutions before the crisis (Gorton and Metrick 2012). Moreover, these agencies inaccurately assessed the corporate ratings, fostering an illusion of credit prosperity that precipitated the rapid expansion of debts and ultimately culminated in the subprime mortgage crisis (Dell'Ariccia et al. 2012). This pattern of corporate rating misjudgment appears to have recurred in China in recent years. The misjudgment of credit ratings has exerted substantial adverse effects on China's financial markets (Hossain et al. 2023). Notably, when the heavily indebted Chinese real estate giant Evergrande Group faced a severe liquidity crisis and was ultimately ordered to liquidate by a Hong Kong court in early 2024, most domestic credit rating agencies in China had persistently maintained their highest AAA rating for the company.

In this way, the fairness and efficiency of corporate credit rating systems are crucial for financial system stability, as they guide capital allocation toward productive economic activities (Almeida et al. 2017). However, assessing the fairness and efficacy of credit rating agencies is challenging due to their often opaque ("black box") rating methodologies (Sangiorgi and Spatt 2017). While some agencies disclose general frameworks, stakeholders frequently lack clarity on the specific factors prioritized, their relative weights in assessments, and, most critically, whether these factors and their applications promote equitable and efficient outcomes. This opacity can obscure potential biases, undermine accountability, and erode trust in the ratings that shape financial decisions and market dynamics. In the second part of our results, our research further demonstrates that systemic biases and inefficiencies within the rating system can distort public credit policies, leading to unintended negative consequences (Sy 2009).

To decipher the inefficiency in the corporate credit rating system, we employ machine learning models. Fuster et al. (2022) utilizes such methods to investigate inefficiencies in the mortgage application market, while Gu et al. (2020) applies various machine learning algorithms to study asset pricing, capturing non-linearities and interactions among financial factors in the stock market.

Our research employs machine learning algorithms over traditional models like logistic regression due to the latter's limitations in capturing the complexity inherent in corporate credit rating systems (Galil et al. 2023). Specifically, linear models tend to systematically underestimate the inefficiencies in the credit rating system. In practice, inefficient factors

within these systems do not operate in isolation; they interact with others, often amplifying the overall inefficiency through complex, non-linear pathways. Furthermore, the marginal effects of such factors are rarely linear and can exhibit escalating side effects when coupled with other variables. To address these intricacies, our methodologies, including artificial neural networks and gradient boosting trees, leverage non-parametric machine learning approaches, which are capable of modeling both the interactions and the non-linear dynamics of these inefficient factors in the credit rating system. The critical importance of modeling such interaction effects is also a key result highlighted in Fuster et al. (2022) and Gu et al. (2020), although their studies focus on different economic domains.

We utilized the extensive Chinese corporate rating data spanning from 2010 to 2020, which included a rich set of financial indicators, to train our machine learning models for the purpose of reverse-engineering the credit rating mechanism. By systematically analyzing 61 features within the dataset, we identified a systematic rating bias against private firms that lack implicit government support. This bias exacerbates existing financing disparities between state-owned and private enterprises. The consequences of bias within the credit rating system are both significant and far-reaching. Subsequent to the identification of this factor with inefficiency, implicit government support, within the corporate credit rating system, our analysis further demonstrates that its presence significantly undermines the efficacy of government credit policy transmission through the credit rating channel. To empirically test the impact of this distorted rating system, we leverage China's 2016 "deleveraging" macro-prudential policy as a quasi-natural experiment and employ a difference-in-difference-in-differences (DDD) model. Our results show that implicit government support, which artificially lowers the credit ratings of private firms, distorts the allocation of credit under the government macro prudential "deleveraging" credit policy. This bias renders government policy less effective and creates a significant barrier to borrowing for private companies that rely on the corporate credit rating channel. The estimated negative impact is both statistically significant and economically substantial.

In this context, our findings suggest that machine learning offers a robust and effective methodological framework for evaluating the transparency, fairness, and efficiency of corporate credit rating systems, with significant implications for policy design and oversight.

Our paper is organized as follows. Section 1 provides an introduction. Section 2 reviews the relevant literature. Section 3 describes China's credit rating framework and presents our corporate rating dataset. Section 4 applies machine learning models to uncover inefficiencies in the rating system. This section is the core contribution of our work and includes the pivotal Figures 2 and 3. Section 5 examines the ramifications of rating bias under a government credit policy scenario. Finally, Section 6 offers our concluding remarks.

2 Literature Review

The purpose of corporate ratings is precisely to address the problem where loan institutions often lean towards companies with government guarantees under information asymmetry (Diamond 1984). If corporate ratings objectively reflect all information of listed companies, they should act as a lubricant for the financing market, helping financing resources tilt towards companies with higher marginal productivity and benefiting borrowers with greater returns (Houston et al. 2010). Conversely, if corporate ratings fail to break the problem of information asymmetry and accurately reflect a company's financing value, they can distort the lending distribution in the financing market and reduce market efficiency (Demyanyk and Van Hemert 2011).

Unfortunately, real markets are often imperfect. The objectivity of corporate ratings is challenged by various factors (Cornaggia et al. 2018). Firstly, there is a positive correlation between corporate ratings and issuer size, number of employees, and tax payment, demonstrating a "too big to fail" effect (Dong et al. 2021). Secondly, there is a positive association between corporate ratings and state ownership of issuers, revealing a "government connection" effect (Dong et al. 2021). Our research further proves that asset size and ownership background are the main factors considered in China's credit rating industry. Corporate rating agencies also smooth ratings upwards for rated companies under the drive for higher profits (Becker and Milbourn 2011).

In the case of unfair corporate ratings, they not only fail to act as a lubricant for the financing market by improving financing efficiency, but also distort the lending distribution in the financing market, allocating more credit loans to large companies and companies with government background, placing small and private companies at a disadvantage in the competitive market (Cong et al. 2019). During the period of credit expansion in China, a large amount of funds flowed into the market, enabling almost all listed companies to obtain relatively sufficient sources of financing (Song et al. 2011). The distortion effect of corporate ratings during the credit boom period is not significant because although small companies and private enterprises cannot become the preferred entities in the financing market, there are still many funding sources available for them when large companies and state-owned enterprises are fully financed (Almeida et al. 2017).

However, as China's market debt problem continues to worsen, the Chinese government chose to implement a deleveraging policy in 2016, significantly reducing the influx of funds into the market, aiming to eliminate highly leveraged and inefficient enterprises (Qiu and Cheng 2022). In the case of insufficient financing resources, the distortion effect of unfair corporate ratings is intensified. When the financing market is in short supply, the market

tends to allocate funds to companies with higher corporate ratings (Ivashina et al. 2022). Due to the influence of company size and background, larger companies and those with a government background often receive higher corporate ratings (Dong et al. 2021). Therefore, in times of market financing tightening, resources are preferentially allocated to large companies and those with government backgrounds, leading to asset differentiation between large and small companies, and between state-owned and private enterprises(Dong et al. 2021). However, under the backdrop where large companies and state-owned enterprises are not always the most productive, such allocation can lead to a decrease in market efficiency (Jin et al. 2023).

Previous studies have revealed the unfairness of corporate credit ratings, which are influenced by factors such as company size, background, and the interests of rating agencies (Dong et al. 2021, Becker and Milbourn 2011). However, these studies mainly rely on econometric methods, typically based on linear models, and are limited by data dimensions and computational power, making it difficult to simultaneously consider a large number of related factors and complex nonlinear relationships between them (Varian 2014). Therefore, quantitative research often only focuses on the impact of individual variables on credit rating, while in reality, credit rating is a comprehensive evaluation system that involves the balance of multidimensional factors (Altman and Rijken 2004). In contrast, machine learning methods can break through the limitations of traditional econometric methods and reveal the operating mechanism of the "black box" of credit rating. By constructing predictive models, machine learning can not only effectively identify the core factors that affect credit ratings, but also quantify the relative contributions of each factor (Fuster et al. 2022). In recent years, the application of machine learning in revealing implicit mechanisms in financial markets has received increasing attention and has shown high reliability in economic research (Zhang et al. 2022, Gu et al. 2020).

3 China Credit Rating and Data

3.1 Credit Rating in China

Since the 1990s, professional credit rating agencies have emerged in China. Currently, some of the well-known corporate rating agencies in the Chinese market include China Chengxin International Corporate Rating Co., Ltd. (CCXI), United Corporate Rating Co., Ltd., Dagong Global Corporate Rating Co., Ltd., among others. These rating agencies are mainly domestically developed and conduct credit ratings independently within China. Even in the case of CCXI, which has a stake in Moody's, Moody's only holds 30% of the shares and has

never been able to fully control CCXI. The credit rating market in China has maintained its unique operating trajectory and has not yet fully integrated with the international credit rating market. Over the past twenty years, these major rating agencies have completed issuer corporate ratings and bond corporate ratings for over 50% of listed companies in China. There are certain differences in the credit rating markets between China and the United States, reflected in the persistent inflation of credit ratings in China.

Since 2006, China's rating market has adopted a standardized universal evaluation system with a three-level, nine-grade scale, ranging from AAA to C, with adjustments permitted using "+" or "-" signs, where AAA+ represents the highest corporate rating. However, China's rating market has been prone to inflate corporate ratings for a long time. The vast majority of credit rating agencies in China classify the risk assessment of Chinese borrowers as AAA or AA. In contrast, top global rating agencies headquartered outside of China typically rate bonds issued by Chinese companies overseas at much lower levels - an average of 6 to 7 levels lower (Jiang and Packer 2019). In 2015, companies with AAA ratings in China accounted for 16.87% of issuers' credit ratings., and by December 2020, this proportion had risen to 21.84%. In 2020, 85.98% of companies received an AA or higher rating from their issuers, with 46.33% receiving an AA+ or higher rating. Table 1 shows the changes in the number of companies with different credit rating categories in the Chinese market from 2010 to 2020. The number of companies with AAA ratings skyrocketed from 58 in 2010 to 1018 in 2020.

[Insert Table 1 about here]

Credit rating criteria in the Chinese and American markets also differ. Previous studies have shown that asset size is considered a positive factor and has a higher weight in the Chinese institutional evaluation system, while international rating agencies place greater emphasis on the positive impact of profitability (Jiang and Packer 2017). In addition, the long-standing credit rating defects in China have also been widely criticized. Usually, the ratings provided by rating agencies should have an inverse relationship with bond default rates. However, in the Chinese market, there has long been an abnormal inversion between credit ratings and bond default rates, where bonds with lower credit ratings have lower default rates than bonds with higher credit ratings. For example, in 2021, multiple rating agencies such as Dagong Global and CCXI have experienced inverted default rates between different credit ratings. In 2023 and 2024, this inversion still persists.

3.2 Credit Rating's Impact in China

The Chinese government imposes many regulations on the credit rating market. In the bond market, only companies with an issuer credit rating of AA or above can issue ultra-short-term bonds. According to the regulations of the exchange, if the credit rating of the issuer of corporate bonds is adjusted to below AA during the term, the trading method cannot be selected as competitive bidding. According to the regulations of the National Development and Reform Commission, for bonds with issuer ratings not lower than AA and bond ratings not lower than AA+, companies are allowed to use up to 40% of the raised funds to repay bank loans and supplement working capital. For urban investment enterprises and general production and operation enterprises, the National Development and Reform Commission has relaxed the leverage requirements for providing guarantee measures to 65% and 75%, respectively. If the subject rating reaches AA+, the corresponding leverage requirements can be relaxed to 70% and 80%; If the subject rating reaches AAA, the leverage requirement for urban investment enterprises will be further relaxed to 75%, and for general production and operation enterprises, it will be relaxed to 85%. These provisions show that corporate ratings play an important role in the Chinese financial market. The financing market favors companies with higher corporate ratings.

As the AAA rating is the highest level in credit ratings, it represents a company's strong debt-paying ability and outstanding credit performance. In the Chinese market, companies with AAA ratings usually have significant advantages in financing, policy support, and other aspects. Therefore, this study designates companies with AAA ratings as high-credit- rating companies, while other companies are classified as low-credit-rating companies.

3.3 Data

The credit rating data of listed companies in this study are sourced from the Wind database, covering almost all publicly traded firms that received third-party credit ratings from 2010 to 2020. Financial data, including leverage, log of total assets, ROE, quick ratio, cash ratio, and net profit, etc., is obtained from the Wind database. The Wind database is a leading financial data service platform in China, widely used for financial market research and investment analysis (Jiang and Packer 2019, Cong et al. 2019, Jin et al. 2023).

To ensure the robustness of the results and eliminate outliers, we exclude firms with a leverage ratio greater than 1, as well as those whose key performance indicators, such as asset turnover and total asset turnover, fall outside the normal range of 0 to 1 for the given year. Furthermore, given the unique characteristics of the financial sector-such as monetary intermediation, high leverage, and high liquidity (Ali et al. 2008)-financial firms are excluded

from the sample to mitigate industry-specific effects on the results.

To maintain data balance, we classify firms into a high-credit-rating group (AAA) and a low-credit-rating group, allowing us to assess the impact of credit ratings on corporate financial conditions and financing ability. If a company receives different ratings from multiple third-party agencies in that year, we uniformly take the highest rating. Based on existing literature, we further control for several key factors that may influence firm growth and financing capacity: (1) firm size-measured by the logarithm of total assets, as larger firms tend to have greater access to financing and higher leverage ratios (Beck et al. 2008); (2) firm performance-proxied by ROE, quick ratio, cash ratio, and net profit, etc., since firms with stronger financial performance often inspire greater investor confidence and thus maintain higher leverage levels (Beck et al. 2006); (3) corporate governance and ownership structure-considering the presence of implicit government guarantees, we include a control variable indicating whether the firm's ultimate controlling shareholder has a government background (Jin et al. 2023); (4) industry effects-as leverage levels vary across industries, with sectors such as construction and manufacturing typically exhibiting higher leverage (Korteweg 2010); (5) credit rating agency effects-different rating agencies may adopt varying methodologies and criteria, potentially influencing firms' assigned ratings. To account for this, we include rating agencies as a control variable in our empirical analysis, mitigating potential selection biases arising from rating agency differences. We used 61 features from these five dimensions in our machine learning model. Table 2 describes the variables used in machine learning. Table 3 shows the summary statistics of machine learning variables.

[Insert Table 2 about here]

[Insert Table 3 about here]

By performing a statistical analysis of the data, we find that the proportion of implicit guarantees in high credit ratings shows a relatively stable trend. Starting from 2011, the proportion of implicit guarantees in high credit ratings was 88.79%, and this proportion remained above 88% in the following years, demonstrating a consistent pattern. This trend indicates that, despite some fluctuations, implicit guarantees consistently account for a large proportion of high credit ratings. From these data, we can roughly infer that the ultimate controlling shareholders may have some influence on a company's ability to obtain a high credit rating, especially among those companies with implicit guarantees. However, based on this alone, we still cannot fully deduce the specific credit rating mechanism, especially considering that the ultimate controlling shareholders may be involved in other factors that affect credit ratings. Therefore, to further explore this relationship, we decided to build a

machine learning model aimed at accurately predicting whether a listed company belongs to the high credit rating group through the analysis of historical data. This model can not only reveal the potential impact of the ultimate controlling shareholders on a company's credit rating but also identify other factors-such as financial condition, industry background, and debt structure-also play a significant role in credit ratings.

[Insert Table 4 about here]

On the basis of building a machine learning model for predicting credit rating determinants, we further developed a triple difference (DDD) model to systematically evaluate whether private enterprises with low credit ratings experienced a decline in financing ability compared to other types of companies during the implementation of the deleveraging policy in 2016. The treatment variables of the model are whether it is a private enterprise and whether it has a low credit rating. We use the leverage ratio as the interpreted variable to measure a firm's financing capacity, and introduce a series of control variables in the model to account for other factors that might influence leverage. Specifically, we controlled the company's size, measured by total assets (Beck et al. 2008), and also control the company's performance indicators: ROE, quick ratio, cash ratio, and net profit (Beck et al. 2006). In addition, we include industry dummy variables to control for structural differences across industries (Korteweg 2010). We also control for credit rating agency fixed effects to avoid systematic biases in results due to methodological differences across rating agencies. Table 5 presents the statistics of the variables used in the empirical regression.

[Insert Table 5 about here]

4 Deciphering the Black Box of Corporate Rating

4.1 Corporate Credit Rating Data

Is the credit rating in the Chinese market fair? This study takes a high credit rating (whether it is AAA or not) as the target variable and uses machine learning methods to reveal the black box of credit rating and the key factors that affect credit rating. To ensure the accuracy and robustness of the model evaluation, we use a ten-fold cross-validation method to train and validate the model. Ten-fold cross-validation involves randomly dividing the dataset into ten equally sized subsets, selecting one subset as the validation set, and the remaining nine subsets as the training set. This process is repeated ten times, with each subset selected as the validation set. Finally, a reliable estimate of the model performance is obtained by averaging the evaluation results ten times. The advantage of this method is that it can

fully utilize data, reduce the variance and bias of evaluation results, and thus improve the robustness of the results.

4.2 Machine Learning Algorithm for Credit Rating

Linear Model In this study, we initially employed logistic regression, a linear probabilistic model, for credit rating prediction to establish a baseline performance benchmark for subsequent comparison with non-parametric machine learning models discussed in later sections. The logistic regression transforms the output of linear regression into a probability value within the interval $[0, 1]$ via the sigmoid function to determine the likelihood of a company belonging to the high-credit-rating category based on its characteristics. Let Y denote the target variable representing whether the credit rating is high or low, $X = \{x_1, x_2, \ldots, x_n\}$ represent the feature variables, β_0 be the intercept term, β_i be the regression coefficients of the features, and σ be the Sigmoid activation function that converts the linear combination result into a probability value. The formula for logistic regression is as follows:

$$P(Y = 1|X) = \sigma(\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_n x_n) = \frac{1}{1 + e^{-(\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_n x_n)}} \quad (1)$$

Tree-Based Boosting Model We use a tree-based machine learning model with gradient boosting to predict the credit rating of listed companies. Gradient boosting is an ensemble learning method that improves predictive ability by constructing multiple decision trees. Another tree-based model is random forest, but it is weaker than gradient boosting in capturing complex nonlinear relationships. Therefore, we choose gradient boosting as the representative tree-based model. The advantage of gradient boosting lies in its boosting strategy, which iteratively trains a series of decision trees, each optimized based on the residuals of the previous tree, gradually reducing prediction errors. Let $F_{t-1}(x)$ be the prediction at the $(t-1)$-th iteration, and $r_t = y - F_{t-1}(x)$ be the residual fitted by the t-th new tree. The model's iterative formula is as follows:

$$F_t(x) = F_{t-1}(x) + \nu \cdot h_t(x) \quad (2)$$

Here, ν represents the learning rate, controlling the contribution of each tree to the final prediction. The loss function for gradient boosting is the same as Equation (2), with the core being the use of the negative gradient direction to update the model. Take, for example, the task of predicting whether a listed company belongs to the high credit rating category. The initial prediction of the model may be simple, for example, assuming that the probabilities of

all categories are equal. Subsequently, in each iteration, the model uses a regression tree to fit the current loss function to adjust the assumptions and classifications for the next round of predictions. For example, companies with asset sizes exceeding a certain threshold typically have higher credit ratings, while companies below the threshold often have lower ratings. Based on this partitioning, the model constructs one or more decision trees to reduce the error between predicted ratings and actual ratings, thus better adapting to actual ratings. As the iteration progresses, the model continues to build new decision trees. Each new decision tree attempts to reduce residuals that the previous tree cannot explain, gradually improving overall prediction performance. Ultimately, the gradient boosting model aggregates the predictions of all decision trees to obtain the final credit rating prediction. Compared with traditional linear prediction methods, tree-based models are more flexible, able to better adapt to the complexity of credit rating prediction, and capture nonlinear relationships between variables. For example, factors such as company asset size and profitability often have complex interactions. These interactions can be effectively identified through tree-based models, particularly gradient enhanced models. This is crucial for predicting the credit ratings of financial markets and the banking industry.

Artificial Neural Networks To enhance robustness, we also constructed a three-layer artificial neural network (ANN) for credit rating prediction. This neural network consists of three layers: an input layer with 32 neurons, a hidden layer with 16 neurons using the ReLU (rectified linear unit) activation function, and an output layer with 1 neuron using the sigmoid activation function to map the output value to the interval $[0, 1]$. Let $a^{(L-1)}$ be the input to the L-th layer, $W^{(L)}$ be the weight matrix, and $b^{(L)}$ be the bias term. The calculations for each layer are as follows:

$$z^{(L)} = W^{(L)} a^{(L-1)} + b^{(L)} \tag{3}$$

$$a^{(L)} = \text{ReLU}(z^{(L)}) \quad \text{(Hidden Layer)} \tag{4}$$

$$a^{(L)} = \sigma(z^{(L)}) \quad \text{(Output Layer)} \tag{5}$$

The construction of a three-layer network enhances the model's nonlinear expression capability. During training, we computed the gradients using backpropagation and updated the parameters using Adam (adaptive moment estimation) to minimize the prediction error, as shown in Equations (8) and (9):

$$\text{Loss} = -\frac{1}{N}\sum_{i=1}^{N}[y_i \log(\hat{y}_i) + (1 - y_i)\log(1 - \hat{y}_i)] \tag{6}$$

$$\theta = \theta - \alpha \cdot \nabla_\theta \text{Loss} \tag{7}$$

α is the learning rate, and ∇_θLoss is the gradient of the loss with respect to the weights. Compared to logistic regression and tree-based models, the advantage of neural networks lies in their ability to automatically learn complex feature representations and demonstrate better predictive performance on large-scale datasets. However, ANN are also prone to overfitting. Therefore, we incorporated dropout mechanisms during training to improve the model's generalization.

Outcome: This study employed ten-fold cross-validation to assess the predictive performance of logistic regression, gradient boosting, and artificial neural networks (ANN). The evaluation metrics for the models included accuracy, F1 score, and AUC. To understand accuracy and F1 score, it is essential to introduce the core concepts used in classification model evaluation: TP (true positives), TN (true negatives), FP (false positives), and FN (false negatives). TP represents the number of samples correctly predicted as positive. In credit rating prediction, if a company has a high actual rating and the model correctly predicts a high rating, it is considered a TP. TN represents the number of samples correctly predicted as negative. If a company has a low actual rating and the model correctly predicts a low rating, it is considered a TN. FP represents the number of samples incorrectly predicted as positive when they are actually negative. If a company has a low actual rating but the model mistakenly predicts a high rating, it is considered an FP. FN represents the number of samples incorrectly predicted as negative when they are actually positive. If a company has a high actual rating but the model mistakenly predicts a low rating, it is considered an FN. Table 6 presents the confusion matrix, showing TP, TN, FP, and FN for the three models. The gradient boosting model performs the best in correctly predicting TP and TN samples, with the lowest number of false positives and false negatives, indicating that the classification performance of the gradient boosting model is superior to other models on this dataset.

[Insert Table 6 about here]

Starting from the confusion matrix, many metrics can be derived to measure the performance of machine learning, and accuracy is one of them. Accuracy measures the proportion

of correctly predicted samples out of the total samples, calculated as follows:

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \tag{8}$$

Similar to accuracy, precision and recall are also important concepts. Precision measures the proportion of true positives among the samples predicted as positive, while recall measures the proportion of true positives among all actual positives. Given their inverse relationship, the F1 score, which is the harmonic mean of precision and recall, is introduced:

$$\text{Precision} = \frac{TP}{TP + FP} \tag{9}$$

$$\text{Recall} = \frac{TP}{TP + FN} \tag{10}$$

$$\text{F1 Score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \tag{11}$$

In addition to classification-based metrics, this study also reports the pseudo-R^2 statistic to evaluate model fit from a likelihood perspective. Unlike the traditional R^2 used in linear regression, pseudo-R^2 measures the relative improvement of the fitted model over a null model (a model with only an intercept). One of the most commonly used forms is **McFadden's** R^2, defined as:

$$R^2_{\text{McFadden}} = 1 - \frac{\ln L_{\text{full}}}{\ln L_{\text{null}}} \tag{12}$$

where $\ln L_{\text{full}}$ is the log-likelihood of the fitted model, and $\ln L_{\text{null}}$ is the log-likelihood of the null model. A higher pseudo-R^2 indicates better model fit. Although pseudo-R^2 values are typically much lower than the R^2 in linear regression, values between 0.2 and 0.4 are generally considered indicative of a good fit in classification problems.

In order to strengthen the robustness test of the model and comprehensively evaluate its discriminative ability, this study further adopted the receiver operating characteristic (ROC) curve. The ROC curve depicts the trade-off relationship between True Positive Rate (TPR) and False Positive Rate (FPR) at different thresholds. Here, TPR (also known as recall) measures the proportion of high credit–rated companies correctly identified, while FPR measures the proportion of low credit–rated companies misclassified as high credit–rated. The closer the ROC curve is to the upper left corner, the stronger the discriminative ability of the model. To summarize the performance of the ROC curve using a single number, AUC

(area under the rOC curve) is commonly used.

$$\text{TPR} = \frac{TP}{TP + FN} \tag{13}$$

$$\text{FPR} = \frac{FP}{FP + TN} \tag{14}$$

The AUC value ranges from 0.5 to 1: 0.5 indicates that the model has no discriminative power like random guessing, while 1 represents perfect classification. The higher the AUC, the more effectively the model can distinguish between two types of samples. We plotted the ROC curves of out-of-sample predictions for three models-gradient boosting, ANN, and logistic regression-and computed their AUC values. As shown in Figure 1, the gradient boosting model achieves the highest AUC of 0.85, indicating the best discriminatory power. The ANN model follows with an AUC of 0.83, while logistic regression lags slightly behind at 0.81. These results suggest that machine learning models exhibit stronger predictive capabilities in predicting high credit ratings.

[Insert Figure 1 about here]

Table 7 presents the performance of each model across different metrics after ten-fold validation. The results show that gradient boosting performed the best across all metrics, with an accuracy of 0.914, an F1 score of 0.784, and an AUC of 0.845, indicating its superior performance in overall prediction accuracy, balance in classifying positive and negative samples, and distinction capability. ANN followed closely, with an accuracy of 0.906, an F1 score of 0.765, and an AUC of 0.832, still demonstrating good performance and the ability of neural networks to effectively capture complex data features. In contrast, logistic regression showed similar performance, suggesting fewer redundant features in the data. Its accuracies (0.892), F1 scores (0.724), and AUC (0.806) were slightly lower, possibly due to its linear nature being insufficient to fully capture the nonlinear relationships in the data, which are often present in credit rating prediction involving complex interactions among company financial conditions and market information.

[Insert Table 7 about here]

4.3 Feature Importance

Credit rating is crucial for financial markets and corporate financing. Therefore, revealing the key factors that affect credit ratings and understanding the decision-making logic of the

model can help us understand the credit allocation mechanism in the real world. On the basis of successfully constructing a high-precision credit rating model for listed companies, this study further explores the key factors that affect credit rating prediction. In order to measure the importance of factors, this study used the feature permutation method to explore the impact of each feature on the model's prediction accuracy, and converted to percentages for easier interpretation. Specifically, we sum the importance of all factors and calculate the percentage contribution of each factor.

Feature permutation measures the importance of a factor by randomly shuffling its values in the test dataset and then observing the change in model performance (e.g., accuracy). If shuffling a particular feature significantly degrades model performance, it indicates that the feature is crucial for credit rating prediction. Let P_{original} denote the original model performance, and P_{permuted} denote the performance after shuffling feature f. The importance score of feature f can be calculated as:

$$\text{Importance}(f) = P_{\text{original}} - P_{\text{permuted}} \tag{15}$$

In the context of classification problems, model performance is often measured using the log-loss (also known as logistic loss or binary cross-entropy). Log-loss evaluates the difference between the predicted probability distribution and the actual class labels, providing a way to quantify how well the model's predicted probabilities match the true outcomes. Specifically, the log-loss function for binary classification is defined as:

$$\text{log-loss} = -\frac{1}{N}\sum_{i=1}^{N}[y_i \log(p_i) + (1 - y_i)\log(1 - p_i)] \tag{16}$$

where y_i is the true label of the i-th sample (either 0 or 1), p_i is the predicted probability that the i-th sample belongs to class 1, and N is the total number of samples. A lower log-loss indicates better model performance, as the model's predicted probabilities are closer to the true labels.Thus, in the feature permutation process, a significant drop in model performance (log-loss increase) after shuffling a feature implies that the feature is important for making accurate predictions.

According to the results of the permutation importance method (Table 8), firm size, as measured by total assets, is consistently the most important feature across all models, with importance values of 45.591%, 31.757%, and 28.618% in the logistic regression, gradient boosting, and ANN models, respectively. This suggests that larger firms are more likely to receive higher credit ratings, confirming the existence of the "too big to fail" effect. Another highly important feature is the ownership status of the firm, particularly whether it is a

state-owned enterprise (SOE), with importance values of 9.964%, 11.696%, and 9.531% in the respective models. This indicates that firms with government affiliations enjoy a significant advantage in credit ratings, lending support to the "government connection" effect. Other important financial features include operating revenue, the ratio of interest-bearing debt to invested Capital, and the quick Ratio. operating revenue and the quick ratio reflect the firm's operating performance and short-term debt-paying ability, respectively. Interest-bearing debt often stems from bank loans or bond financing; a higher level of such debt may imply a broader financing network, reflecting the firm's financial stability and investor trust-factors that can contribute to better credit ratings. These financial characteristics collectively indicate the company's growth potential, which also improves its credit status.

Credit rating agencies also play an important role in determining the level of ratings. The choice of agency often reflects the strength of the company. Generally, companies with larger assets and better financial indicators are more willing to choose highly ranked domestic or international credit rating agencies. For example, "Agency_Fitch" is a reputable international rating agency, which is also an important factor affecting the rating level. The characteristics of the industry also have a significant impact on the rating level. Whether the manufacturing industry has a greater impact on the rating is also a factor that leads to a higher concentration of high/low ratings due to the characteristics of the industry.

[Insert Table 8 about here]

The results from feature permutation confirm that total assets and SOE are core determinants of credit ratings. These findings provide empirical support for the "too big to fail" and "government connection" effects. Nevertheless, a firm's operational performance-such as revenue generation-and financial health, indicated by debt and liquidity ratios, also play vital roles. This suggests that while credit ratings in China prioritize firm size and implicit government backing, actual business performance is not ignored. Overall, the permutation method yield consistent results, revealing unfair biases in the credit rating system. Although machine learning models can capture rating determinants more accurately, they are still heavily influenced by firm scale and state ownership. These findings offer empirical evidence for advancing the market-oriented reform of China's credit rating system and provide important implications for policymakers seeking to enhance credit assessment fairness and market efficiency.

In the field of machine learning, Partial Dependency Plot (PDP) can provide interpretability of a model by visualizing the impact of specific features on its prediction results. By fixing other features in the model and only changing the value of the target feature, the changes in the model's prediction results when the target feature changes can be displayed, which helps

to reveal the dependency relationship between the features and the prediction results. On this basis, a two-dimensional PDP can also be constructed to simultaneously examine the changes in two features, thereby revealing their interaction effects. Based on this method, this study focuses on analyzing the interaction between the identity of state-owned enterprises and total assets, in order to explore potential unfairness in credit rating. Figure 2 shows the partial dependency plot (PDP) under the interaction effect between total assets and SOE in gradient boosting, ANN, and logistic regression models. Analyzing the interaction effect between SOE and total assets is crucial for understanding the fairness of credit ratings. The results indicate that in all three models, state-owned enterprise identity significantly enhances the positive impact of total assets on the probability of high credit ratings. At the same asset level, companies with SOE status are more likely to receive high credit ratings. This discovery reveals potential unfairness in the credit rating market. Considering that total assets are the most important factor in all models, the interaction effect between total assets and SOE amplifies the implicit guarantee effect of the government, further expanding the advantages of state-owned enterprises in the credit rating market and even the financing market. Figure 3 shows a partial dependency plot (PDP) of the interaction effect between the important financial factor Debt to Invested Capital and state-owned enterprises. In gradient boosting and ANN models, as the debt to investment capital ratio increases, the probability of both state-owned and non-state-owned enterprises obtaining high credit ratings decreases, but the probability of non-state-owned enterprises decreases even faster. The advantage of state-owned enterprises in logistic regression models is not as significant as the other two algorithms, which may be due to the lower predictive precision of logistic regression. Overall, the identity of state-owned enterprises has shown a positive impact on the probability of credit ratings, revealing unfair phenomena in the credit rating market.

[Insert Figure 2 about here]

[Insert Figure 3 about here]

5 How Biased Credit Ratings Distort Credit Policy

5.1 China's deleveraging policy in 2016

[Insert Figure 4 about here]

In 2016, China implemented a series of deleveraging policies aimed at resolving high leverage risks and promoting sustainable economic development. These policies are mainly implemented at the enterprise and macro levels. At the enterprise level, the focus is on promoting

the restructuring and integration of state-owned enterprises to restructure or eliminate state-owned enterprises with high asset-liability ratios in heavy-asset industries such as steel and coal, and reduce the occupation of funds by ineffective assets. At the macro level, policies focus on financial regulation and monetary policy control. One is to include off-balance sheet wealth management in the macro prudential assessment, regulate off-balance sheet wealth management business to compress the scale of shadow banking, and reduce the leverage level of the financial market. The second is to stop illegal interbank business, prohibit local governments from using public-private partnership (PPP) projects to indirectly borrow money, regulate local government financing behavior, and curb the disorderly expansion of debt. Thirdly, the central bank has raised reverse repurchase and MLF interest rates, guiding market interest rates to rise, increasing corporate financing costs, curbing excessive borrowing, and prompting financial institutions to tighten credit allocation.

During the policy implementation period, due to the credit crunch, financing resources were more inclined towards high-credit-rated companies, resulting in a significant differentiation in the leverage levels of high and low-credit-rated enterprises. Our machine learning research reveals that the main influencing factors for high credit ratings are asset size and government background. Specifically, companies with larger asset sizes often have stronger risk resistance and financing capabilities, making them more likely to gain favor from financial institutions; state-owned enterprises are also considered to have higher credit ratings due to the implicit support of the government behind them. Therefore, under the 2016 deleveraging policy, state-owned enterprises with large asset sizes are more likely to obtain financing resources and maintain or even expand their leverage levels; private enterprises with small asset sizes face greater financing pressure and relatively lower leverage levels. This phenomenon of differentiation reflects that although policies are formulated with the expectation of deleveraging the entire society, they may also have different impacts on different types of enterprises, thereby exacerbating structural differences between enterprises.

5.2 Consequence of Biased Crediting Rating Under Credit Policy

Our identification assumption is that the 2016 deleveraging policy served as an exogenous shock that intensified the financing disparities among firms with different credit ratings and ownership structures. These differences are expected to manifest in firms' leverage ratios. To identify the heterogeneous impact of the deleveraging policy-particularly whether low-rated private firms were more adversely affected-we construct the following difference-in-differences-in-differences (DDD) model. We employ two specifications to enhance robustness. The baseline model, shown in Equation (17), includes firm fixed effects and year fixed effects,

controlling for time-invariant firm characteristics and common macroeconomic shocks across years:

$$\begin{aligned} Leverage_{it} = {} & \beta_0 + \beta_1 \text{Low-rating}_i + \beta_2 \text{Policy}_t + \beta_3 \text{Private}_i \\ & + \beta_4 (\text{Low-rating}_i \times \text{Policy}_t) + \beta_5 (\text{Low-rating}_i \times \text{Private}_i) \\ & + \beta_6 (\text{Policy}_t \times \text{Private}_i) + \beta_7 (\text{Low-rating}_i \times \text{Policy}_t \times \text{Private}_i) \\ & + \gamma \text{Controls}_{i,t-l} + \gamma_i + \delta_t + \varepsilon_{i,t} \end{aligned} \tag{17}$$

Machine learning models have demonstrated a strong correlation between private ownership backgrounds and low ratings. We intentionally chose low ratings and private enterprises as indicators to investigate whether the financing ability of low-rated private companies would be further weakened compared to other types of companies under the dual disadvantage of low ratings and private enterprise status under the deleveraging policy in 2016. In both specifications, $Leverage_{it}$ denotes the firm's leverage ratio; Low-rating_i is a binary indicator for low credit rating (=0 if high, =1 if low); Private_i equals 1 if the firm is privately owned; and Policy_t equals 1 for the post-policy, $\varepsilon_{i,t}$ is the error term. In Equation (16), γ_i represents firm fixed effects, capturing time-invariant differences between firms that could affect leverage, such as industry or size. δ_t represents year fixed effects, controlling for common shocks or macroeconomic factors that affect all firms in a given year. As a robustness check, we estimate an alternative model in Equation (18), replacing firm fixed effects with rating agency fixed effects and adding industry and year fixed effects, in order to account for unobserved heterogeneity across credit rating agencies and industry-level characteristics:

$$\begin{aligned} Leverage_{it} = {} & \beta_0 + \beta_1 \text{Low-rating}_i + \beta_2 \text{Policy}_t + \beta_3 \text{Private}_i \\ & + \beta_4 (\text{Low-rating}_i \times \text{Policy}_t) + \beta_5 (\text{Low-rating}_i \times \text{Private}_i) \\ & + \beta_6 (\text{Policy}_t \times \text{Private}_i) + \beta_7 (\text{Low-rating}_i \times \text{Policy}_t \times \text{Private}_i) \\ & + \gamma \text{Controls}_{i,t-l} + \delta_{\text{Agency}_i} + \delta_j + \delta_t + \varepsilon_{i,j,t} \end{aligned} \tag{18}$$

In Equation (17), δ_{Agency_i} represents rating agency fixed effects, which control for any unobserved heterogeneity across different rating agencies that could affect the firms' ratings. δ_j represents industry fixed effects, which account for unobserved factors at the industry level that may influence leverage ratios. Control variables include financial indicators such as total assets, ROE, quick ratio, cash ratio, and net profit. We focus on the triple interaction term β_7, which captures whether low-rated private firms experienced a sharper decline in leverage-

interpreted as a more pronounced tightening of financing capacity-after the deleveraging policy, relative to SOE or highly rated firms.

The regression results support our hypothesis. Across all model specifications, the coefficient on the triple interaction term is significantly negative, with values of -7.014, -7.774, -6.153, and -4.138 in columns (1) through (4), respectively-all statistically significant at the 5% level or better. This indicates that, after the policy implementation, low-rated private firms significantly reduced their leverage levels, suggesting their financing capacity was more severely constrained. This is consistent with our expectations: under tightened credit conditions, the deleveraging policy disproportionately affected private firms with already limited financing channels, especially those with lower credit ratings.

The inclusion of control variables and fixed effects substantially improves model fit (with R^2 increasing to 0.455 and 0.319 in columns (2) and (4), respectively), highlighting the importance of firm-level heterogeneity in explaining financing behavior. Overall, the results are robust and confirm our hypothesis: the 2016 deleveraging policy amplified financing inequality along the intersecting dimensions of credit rating and ownership structure, with low-rated private firms being the most adversely affected group.

[Insert Table 9 here]

6 Robustness Test

6.1 Dynamic Regression

We examined the dynamic impact of deleveraging policies in 2016 to further validate the robustness of our conclusions, with a focus on analyzing the interaction between credit ratings, policy implementation, and different years. This test aims to verify the parallel trend hypothesis and evaluate whether there are significant differences in the financing ability of different types of companies in the capital market before the implementation of deleveraging policies. Through the interaction term β_7, we will test the differential impact of deleveraging policies on different types of companies in the years after policy implementation. These tests help address potential reverse causality in the relationship between policy implementation and financing outcomes. Although we believe that the deleveraging policy issued by the government is a reasonable exogenous shock to the financing market of high credit rating companies with a government background, the faster growth of financing levels of high credit rating companies with a government background compared to other types of companies may be caused by the potential decline and investment decline of other types of companies.

$$
\begin{aligned}
Leverage_{it} = & \beta_0 + \beta_1 \text{Low-Rating}_i + \beta_2 \text{Policy} + \beta_3 \text{Year}_t \\
& + \beta_4 (\text{Low-Rating}_i \times \text{Policy}_t) + \beta_5 (\text{Low-Rating}_i \times \text{Year}_t) \\
& + \beta_6 (\text{Policy}_t \times \text{Year}_t) + \beta_7 (\text{Low-Rating}_i \times \text{Policy}_t \times \text{Year}_t) \\
& + \gamma \text{Controls}_{i,t-l} + \gamma_i + \delta_t + \varepsilon_{i,t}
\end{aligned} \tag{19}
$$

In this equation, the variable $Leverage_{it}$ represents the leverage ratio of firm i at time t. The dummy variables for different years ($Year_t$) indicate the periods before and after the 2016 deleveraging policy, with $Year_{-1}$ serving as the benchmark. β_7 represents the three-way interaction between firm rating, the policy, and time. Our primary focus is on β_7, the coefficient for the three-way interaction term ($\text{Rating}_i \times \text{Policy}_t \times \text{Year}_t$). The reason for emphasizing β_7 is that it provides the most direct evidence of how the policy's impact on leverage differs across firm types and over time.

The plot results of β_7 under the parallel trend assumption are shown in Figure 5, where the positive significance coefficient during the pre-policy period indicates that before policy implementation, private enterprises with low credit ratings had higher leverage ratios compared to other companies. This may be because private enterprises with low credit ratings may have received more government support prior to policy implementation, or it may be due to their poor asset structure, resulting in lower leverage ratios. However, after the implementation of the deleveraging policy in 2016, it can be seen that the coefficient is significantly negative, which confirms that low credit-rated private enterprises are in a weak position in the financing market following the policy implementation in 2016, with a significant decrease in leverage ratio and an immediate reduction in financing capacity.

[Insert Figure 5 about here]

In the years following the implementation of the policy, private enterprises with low credit ratings faced a more severe financing environment. With the deepening of deleveraging policies, banks and other financial institutions have become more stringent in their loan approvals for these enterprises, resulting in a significant contraction of credit resources. Although some companies have attempted to address challenges by optimizing their asset structure or adjusting their business strategies, overall, their financing difficulties continue to intensify, limiting their development space. Meanwhile, with the gradual manifestation of policy effects, the market's risk vigilance towards high-leverage enterprises has increased, further exacerbating the financing pressure on low credit-rated enterprises.

6.2 Nonlinear Policy

We previously set Policy_t as a dummy variable to represent the implementation of the deleveraging policy in 2016, but in fact, the strength of the Chinese government's policy implementation is likely to be non-linear and can be flexibly adjusted according to the economic situation. Therefore, we use the ratio of China's non-financial credit loans (credit) to GDP to represent the policy strength. We subtract the ratio of non-financial credit loans to GDP from the previous year to construct a nonlinear $\text{Policy Intensity}_t$ variable instead of a linear Policy_t for regression, as shown in Figure 4. It can be seen that the ratio value reached its peak in 2016 and rapidly decreased, stabilizing in 2018, which is in line with the trend of "deleveraging" policy. The construction of the variable is reasonable.

$$\begin{aligned} Leverage_{it} = {} & \beta_0 + \beta_1 \text{Low-Rating}_i + \beta_2 \text{Policy Intensity}_t + \beta_3 \text{Private}_i \\ & + \beta_4 (\text{Low-Rating}_i \times \text{Policy Intensity}_t) + \beta_5 (\text{Low-Rating}_i \times \text{Private}_i) \\ & + \beta_6 (\text{Policy Intensity}_t \times \text{Private}_i) + \beta_7 (\text{Low-Rating}_i \times \text{Policy Intensity}_t \times \text{Private}_i) \\ & + \gamma \text{Controls}_{i,t-l} + \gamma_i + \delta_t + \varepsilon_{i,t} \end{aligned} \tag{20}$$

The triple interaction term ($\text{Low-Rating}_i \times \text{Policy Intensity}_t \times \text{Private}_i$) is negative and highly significant in all regressions ($p<0.01$), which is consistent with the regression results of the benchmark model. This means that during the deleveraging period in 2016, for every 1-unit reduction in the credit-to-GDP ratio, the leverage ratio of low credit-rated private enterprises decreased by 0.268-0.416 percentage points. These results indicate that nonlinear deleveraging policies still have a stronger inhibitory effect on low credit-rated private enterprises, which is consistent with the regression results of the benchmark model and enhances the robustness of the results.

[Insert Table 10 about here]

7 Conclusion

In our research, we find that non-parametric machine learning algorithms such as artificial neural networks and gradient boosting trees serve as a powerful tool for conducting a comprehensive health scan of our corporate credit rating system. By leveraging these advanced techniques, we can effectively identify and quantify the impact of inefficient factors within the system. The significant value of employing machine learning for credit rating estimation lies not only in its superior predictive accuracy but also in its capacity for enhanced model

explainability. This enhanced transparency enables stakeholders to gain a clearer understanding of the credit rating system's deficiencies, thereby facilitating the development of more transparent and evidence-driven policies for systemic improvement.

In this research, we utilize Chinese corporate credit rating data to evaluate the performance of our machine learning models in identifying the influence of inefficient elements within the corporate rating system. Our findings indicate that implicit government guarantees for state-owned enterprises (SOEs) contribute approximately 10% to their high credit ratings. This factor ranks as the second most important determinant of high credit ratings, immediately following the total asset size of companies. Given that most SOEs are also very large, private firms without implicit government guarantees face significant disadvantages within this credit rating framework.

Furthermore, by employing partial dependence plots (Figure 2 and Figure 3), we visually demonstrate that implicit government guarantees interact with other variables, such as total assets and debt to invested capital leverage, to amplify the likelihood of achieving a high credit rating. In contrast, traditional linear models tend to underestimate these interaction effects, particularly in relation to how implicit government guarantees influence the credit ratings of private companies.

To demonstrate the importance of conducting a "body check" of credit rating systems using machine learning algorithms, we utilize the 2016 macro-prudential "deleveraging" policy as an empirical example. Within a triple difference-in-difference (DDD) framework, our analysis reveals that the government's implicit guarantees favoring state-owned enterprises (SOEs) distort credit allocation mechanisms. These distortions unintentionally amplify the adverse impact of deleveraging policies on private firms through biased corporate credit rating channels.

Our research demonstrates that machine learning models can serve as a powerful tool for auditing the fairness, efficiency, and transparency of opaque credit rating systems. Using a triple difference-in-differences (DDD) design, our analysis further underscores, from the perspectives of both policymakers and rating agencies, the critical importance of maintaining these principles in corporate credit ratings to ensure both market stability and equitable treatment across different types of firms. Moreover, machine learning algorithms provide academic researchers with a robust methodological framework to evaluate whether credit rating agencies adhere to these guidelines when assessing companies.

References

Ali, A., Klasa, S. and Yeung, E.: 2008, The limitations of industry concentration measures constructed with compustat data: Implications for finance research, *The Review of Financial Studies* **22**(10), 3839–3871.

Almeida, H., Cunha, I., Ferreira, M. A. and Restrepo, F.: 2017, The real effects of credit ratings: The sovereign ceiling channel, *The Journal of Finance* **72**(1), 249–290.

Altman, E. I. and Rijken, H. A.: 2004, How rating agencies achieve rating stability, *Journal of Banking & Finance* **28**(11), 2679–2714.

Beck, T., Demirgüç-Kunt, A., Laeven, L. and Maksimovic, V.: 2006, The determinants of financing obstacles, *Journal of international money and finance* **25**(6), 932–952.

Beck, T., Demirgüç-Kunt, A. and Maksimovic, V.: 2008, Financing patterns around the world: Are small firms different?, *Journal of financial economics* **89**(3), 467–487.

Becker, B. and Milbourn, T.: 2011, How did increased competition affect credit ratings?, *Journal of financial economics* **101**(3), 493–514.

Cong, L. W., Gao, H., Ponticelli, J. and Yang, X.: 2019, Credit allocation under economic stimulus: Evidence from china, *The Review of Financial Studies* **32**(9), 3412–3460.

Cornaggia, J., Cornaggia, K. J. and Israelsen, R. D.: 2018, Credit ratings and the cost of municipal financing, *The Review of Financial Studies* **31**(6), 2038–2079.

Dell'Ariccia, G., Igan, D. and Laeven, L. U.: 2012, Credit booms and lending standards: Evidence from the subprime mortgage market, *Journal of Money, Credit and Banking* **44**(2-3), 367–384.

Demyanyk, Y. and Van Hemert, O.: 2011, Understanding the subprime mortgage crisis, *The review of financial studies* **24**(6), 1848–1880.

Diamond, D. W.: 1984, Financial intermediation and delegated monitoring, *The review of economic studies* **51**(3), 393–414.

Dong, Y., Hou, Q. and Ni, C.: 2021, Implicit government guarantees and credit ratings, *Journal of Corporate Finance* **69**, 102046.

Fuster, A., Goldsmith-Pinkham, P., Ramadorai, T. and Walther, A.: 2022, Predictably unequal? the effects of machine learning on credit markets, *The Journal of Finance* **77**(1), 5–47.

Galil, K., Hauptman, A. and Rosenboim, R. L.: 2023, Prediction of corporate credit ratings with machine learning: Simple interpretative models, *Finance Research Letters* **58**, 104648.

Gorton, G. and Metrick, A.: 2012, Securitized banking and the run on repo, *Journal of Financial economics* **104**(3), 425–451.

Gu, S., Kelly, B. and Xiu, D.: 2020, Empirical asset pricing via machine learning, *The Review of Financial Studies* **33**(5), 2223–2273.

Hossain, A., Hossain, T., Jha, A. and Mougoué, M.: 2023, Credit ratings and social capital, *Journal of Corporate Finance* **78**, 102338.

Houston, J. F., Lin, C., Lin, P. and Ma, Y.: 2010, Creditor rights, information sharing, and bank risk taking, *Journal of financial Economics* **96**(3), 485–512.

Ivashina, V., Laeven, L. and Moral-Benito, E.: 2022, Loan types and the bank lending channel, *Journal of Monetary Economics* **126**, 171–187.

Jiang, X. and Packer, F.: 2017, Credit ratings of domestic and global agencies: What drives the differences in china and how are they priced?

Jiang, X. and Packer, F.: 2019, Credit ratings of chinese firms by domestic and global agencies: Assessing the determinants and impact, *Journal of Banking & Finance* **105**, 178–193.

Jin, S., Wang, W. and Zhang, Z.: 2023, The real effects of implicit government guarantee: evidence from chinese state-owned enterprise defaults, *Management Science* **69**(6), 3650–3674.

Korteweg, A.: 2010, The net benefits to leverage, *The journal of finance* **65**(6), 2137–2170.

Qiu, B. and Cheng, B.: 2022, Is a deleveraging policy effective? evidence from china, *International Review of Economics & Finance* **77**, 471–480.

Sangiorgi, F. and Spatt, C.: 2017, Opacity, credit rating shopping, and bias, *Management Science* **63**(12), 4016–4036.

Song, Z., Storesletten, K. and Zilibotti, F.: 2011, Growing like china, *American economic review* **101**(1), 196–233.

Sy, A. N. R.: 2009, The systemic regulation of credit rating agencies and rated markets.

Varian, H. R.: 2014, Big data: New tricks for econometrics, *Journal of economic perspectives* **28**(2), 3–28.

Zhang, S., Lee, D., Singh, P. V. and Srinivasan, K.: 2022, What makes a good image? airbnb demand analytics leveraging interpretable image features, *Management Science* **68**(8), 5644–5666.

Figures

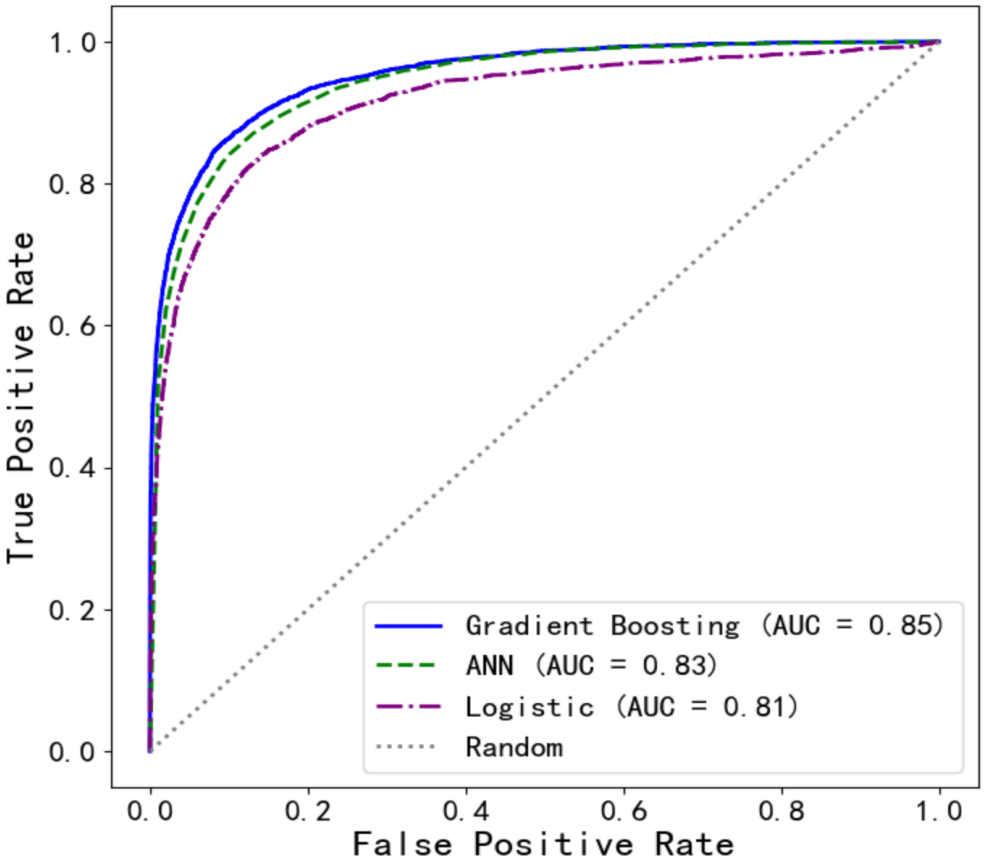
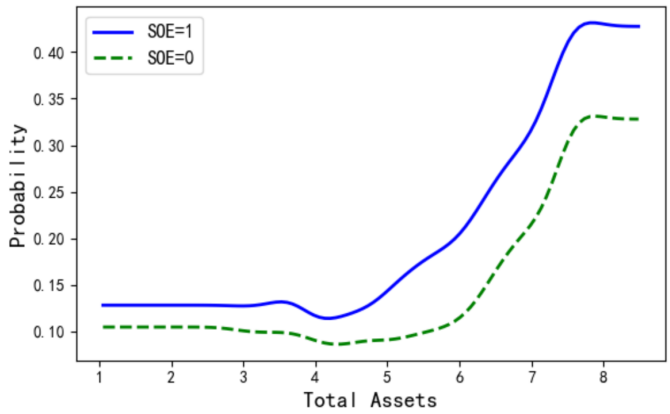
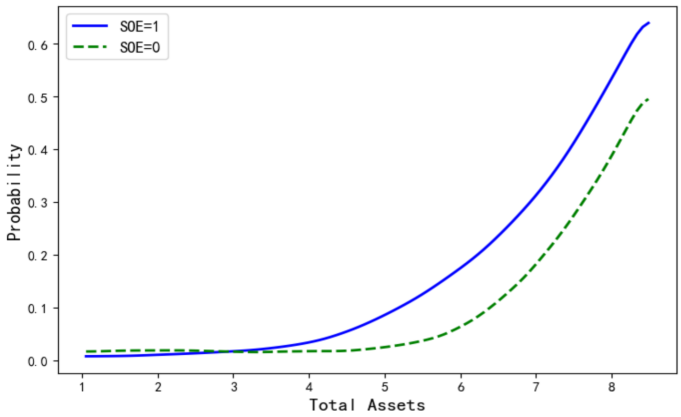


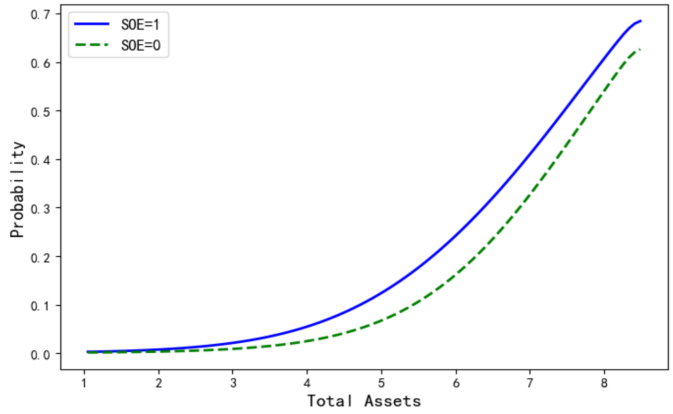
Figure 1: **Performance of Different Algorithms. ROC Curves of Models.**
This figure presents the ROC curves of gradient boosting, ANN, and logistic regression models in a single plot. A higher curve indicates better model performance in predicting credit rating.



Panel A: PDP of Total Assets (gradient boosting)

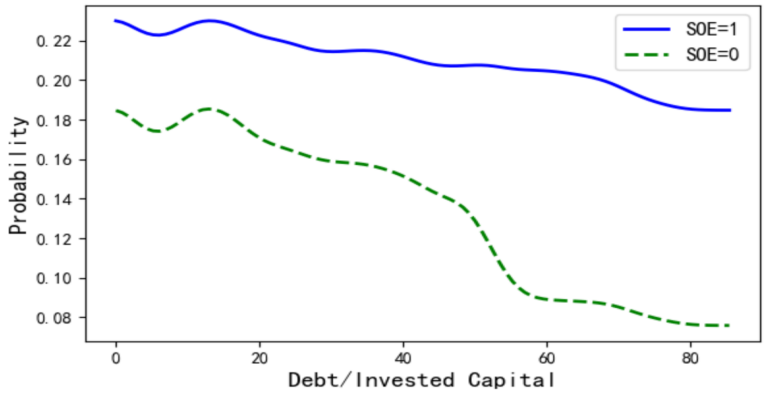


Panel B: PDP of Total Assets (ANN)

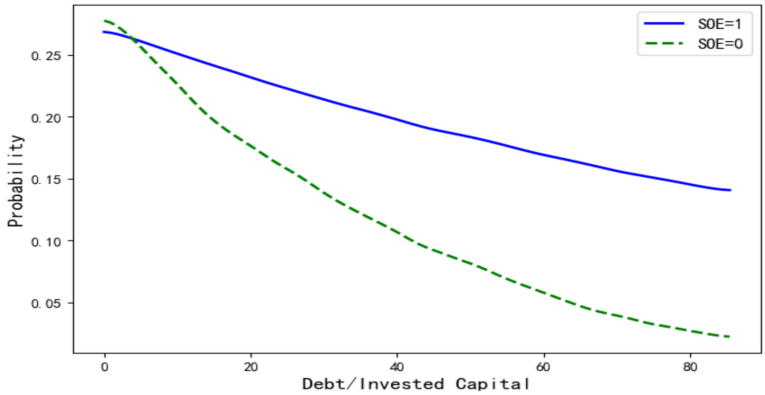


Panel C: PDP of Total Assets (logistic regression)

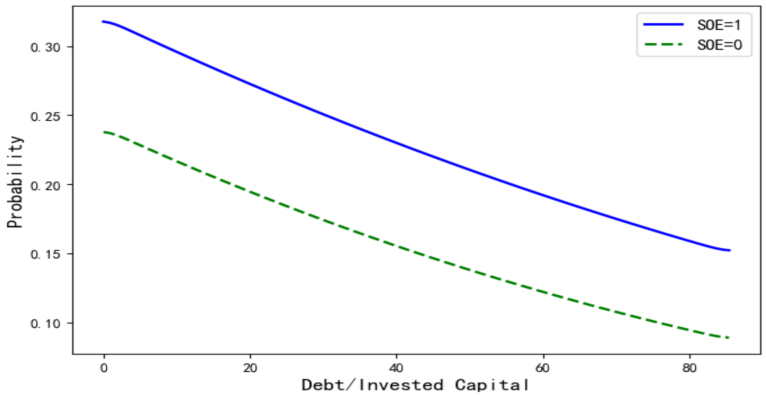
Figure 2: **Non-linear Marginal Effects of Total Assets for High Ratings. SOE vs Private Companies.**
This figure shows the non-linear marginal effects of total assets (log-transformed) for high ratings by using the partial dependence plots (PDP) method. The solid line is for SOEs, and the dashed line is for the private companies. The Panel A, Panel B and Panel C are for gradient boosting trees, ANN, and logistic regression, respectively.



Panel A: PDP of Debt to Invested Capital (gradient boosting)



Panel B: PDP of Debt to Invested Capital (ANN)



Panel C: PDP of Debt to Invested Capital (logistic regression)

Figure 3: **Non-linear Marginal Effects of *Debt to Invested Capital* for High Ratings. SOE vs Private Companies.**
This figure shows the non-linear marginal effects of Debt to Invested Capital for high ratings by using the partial dependence plots (PDP) method. The solid line is for SOEs, and the dashed line is for the private companies. The Panel A, Panel B and Panel C are for gradient boosting, ANN, and logistic regression, respectively.

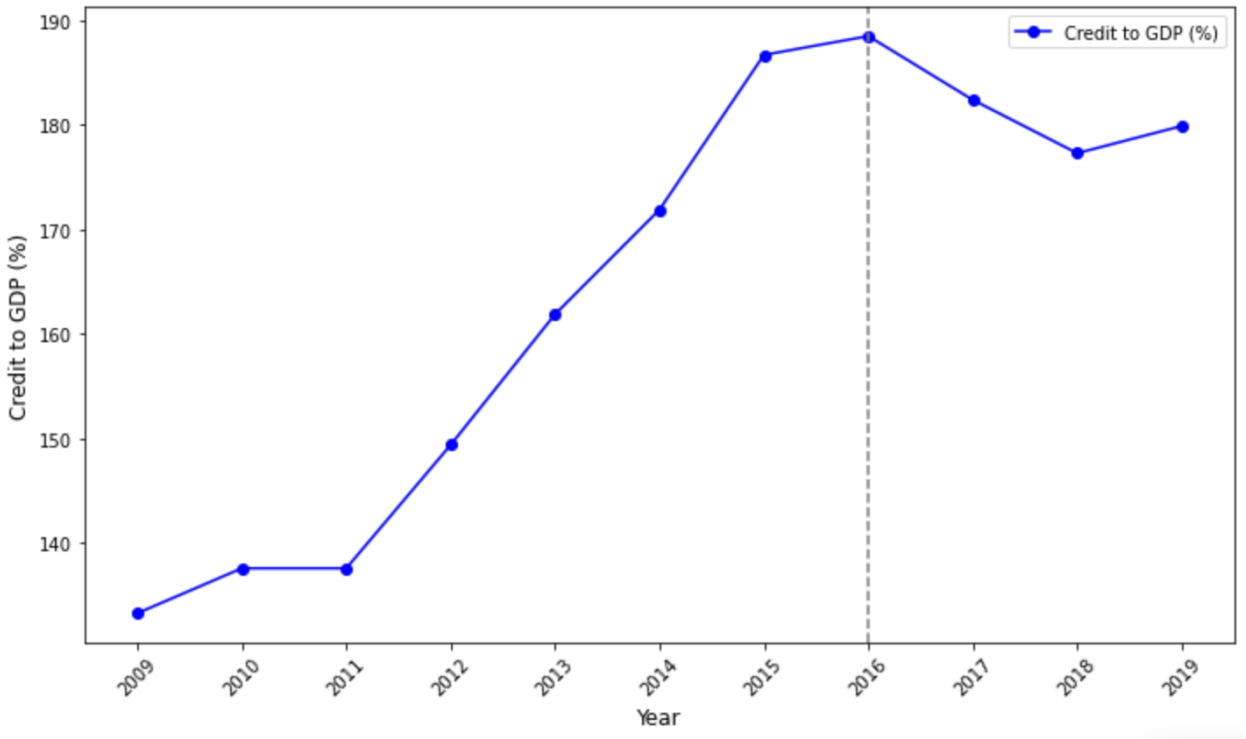


Figure 4: **Debt Leverage Decrease Sharply After 2016 Macro-Prudential "Deleveraging" Policy.**

This chart illustrates the trajectory of China's macroeconomic level debt leverage from 2010 to 2021, as represented by the annual credit debt to GDP ratio. And this data comes from the Bank for International Settlements. The dashed vertical line indicates the "Deleveraging" policy which was initially announced in December of 2015 and applied in 2016.

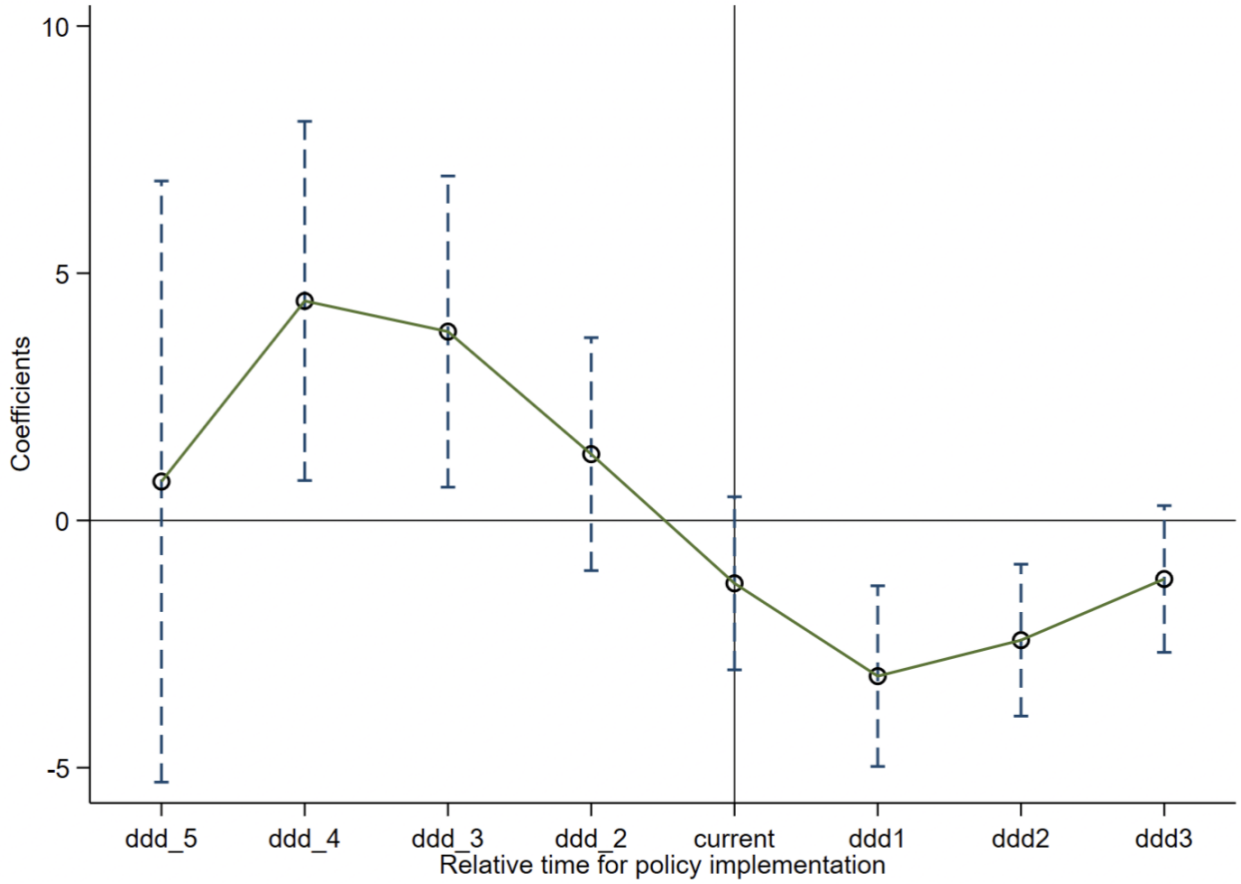


Figure 5: **Results of Dynamic Regression**

This figure shows the coefficients of dynamic regression. After the 2016 "deleveraging" policy, the leverage ratios of low-credit-rated private enterprises decreased significantly, confirming their weakened position in the financing market.

Tables

Table 1: **Distribution of Credit Ratings by Year**

Year	AAA	AA+	AA	AA-	A	BBB	Below BB
2010	58	47	46	13	8	3	6
2011	119	136	145	33	31	6	13
2012	180	216	390	81	52	7	24
2013	244	287	597	132	67	12	36
2014	419	497	1159	296	93	16	61
2015	504	662	1269	328	98	16	68
2016	623	780	1397	325	105	20	93
2017	704	797	1438	329	110	22	109
2018	784	857	1570	356	116	26	120
2019	872	1005	1843	395	116	31	127
2020	1018	1105	1804	400	118	28	135

Note: The table shows the changes in the number of companies in the Chinese market across different credit rating categories from 2010 to 2020. If a company receives different ratings from multiple third-party agencies in that year, we take the highest rating.

Table 2: **Variable Definitions**

Variable Name	**Content Explanation**
Total Assets	Total assets, originally measured in billion yuan, have been log-transformed.
SOE	State-owned enterprise(SOE) indicator (1 if SOE, 0 otherwise)
EBITDA	Earnings Before Interest, Taxes, Depreciation, and Amortization (in billion yuan)
Operating Revenue	Revenue from the main business (in billion yuan)
Operating Profit	Operating profit from the main business (in billion yuan)
Net Profit	Net profit (in billion yuan), have been log-transformed.
Profit Margin	Operating profit margin (percentage)
Revenue Growth	Growth rate of operating revenue (percentage)
ROA	Return on total assets (percentage)
ROE	Return on equity (percentage)
Quick Ratio	Quick Ratio measures a company's ability to meet short-term liabilities with its liquid assets.
Inventory Turnover	Inventory turnover rate
Leverage	Debt-to-asset ratio
Short Debt/Total Debt	Ratio of short-term debt to total debt
Debt/Invested Capital	Ratio of interest-bearing debt to total invested capital
Cash Ratio	Cash Ratio is the ratio of cash and securities to current liabilities, typically expressed as a percentage.
Cash Flow from Ops	Cash flow from operating activities (in billion yuan)
Cash Flow from Invest	Cash flow from investing activities (in billion yuan)
Cash Flow from Financing	Cash flow from financing activities (in billion yuan)
Interest Debt	Interest-bearing debt (in billion yuan)
Industry Dummies	Dummies for industries (e.g., Transport, IT, Public Utilities, Utilities, Healthcare, Materials, Manufacturing, Construction, Financial, Real Estate, Services, Wholesale, Science, Energy, Culture, Mining, Agriculture, Renting Services)
Year Dummies	Dummies for different years (e.g., 2010–2019)
Agency Dummies	Dummies for credit rating agencies (e.g., Fitch, New Century, China Chengxin, United Credit, United Ratings, Pengyuan, China Chengxin Data, United Ratings Intl., Dongfang Jincheng, Guangzhou Pucai, S&P China, Other)

Note: The table shows our main variables used in the machine learning model.

Table 3: **Summary Statistics of Machine Learning Data**

Indicator	Count	Mean	Std	Min	Max
EBITDA	19,454	10.33	17.18	-0.08	65.62
Operating Revenue	19,454	3.16	1.71	-0.44	7.94
Operating Profit	19,454	6.99	19.67	-18.50	160.07
Net Profit	19,454	0.90	1.34	-8.52	4.85
Profit Margin	19,454	11.18	10.70	-2.09	30.71
Revenue Growth	19,454	16.27	27.34	-18.19	71.68
ROA	19,454	3.08	3.16	-8.83	26.59
ROE	19,454	4.09	6.62	-38.24	47.40
Quick Ratio	19,454	1.50	1.04	0.34	4.21
Inventory Turnover	19,454	8.46	24.46	0.05	115.62
Leverage	19,454	54.56	16.16	9.25	90.81
Short Debt/Total Debt	19,454	50.34	22.51	10.21	100.00
Interest Debt/Invested Capital	19,454	42.26	17.44	0.00	85.38
Cash Ratio	19,454	0.46	0.40	0.07	1.75
Cash Flow from Ops	19,454	4.57	18.09	-21.09	55.08
Cash Flow from Invest	19,454	-15.44	24.22	-84.85	3.31
Cash Flow from Financing	19,454	14.07	25.43	-21.23	79.13
Interest Debt	19,454	4.09	1.38	-1.59	6.36
Total Assets	19,454	5.37	1.10	1.05	8.49
Rating	19,454	0.20	0.40	0.00	1.00
SOE	19,454	0.84	0.37	0.00	1.00

Note: The table summarizes the numerical variables used in machine learning models, where Rating is the learning objective in the machine learning model. The description of variables excludes dummy factors such as financial factors, agency indicators, and industrial indicators. Due to the median imputation of all missing values in the machine learning model, there are slight differences in the mean values of some variables compared to the DDD model.

Table 4: **Counts of High and Low Credit Ratings for the State-Owned and Private Enterprises**

Year	Private		SOE	
	High Rating	Low Rating	High Rating	Low Rating
2011	13	86	103	271
2012	18	175	136	513
2013	14	175	176	713
2014	26	305	357	1613
2015	51	354	407	1683
2016	77	417	495	1779
2017	59	287	408	1586
2018	72	293	545	1809
2019	81	410	644	2356
2020	92	398	818	2285

Note: Table 2 reports the number of firms by rating status for the state-owned and private enterprises.

Table 5: **Summary Statistics for DDD Regression**

Variable	Count	Mean	Std	Min	Max
Dependent variable					
Leverage	19,454	54.56	16.16	9.25	90.81
Core explanatory variables					
Low-Rating	19,454	0.80	0.40	0.00	1.00
Private	19,454	0.16	0.37	0.00	1.00
Policy	19,454	0.61	0.49	0.00	1.00
Control variables					
Total Assets	18,721	5.36	1.09	1.05	8.49
ROE	18,721	4.66	5.58	-38.24	47.40
Quick Ratio	18,721	1.52	1.05	0.34	4.21
Cash Ratio	18,721	0.47	0.40	0.07	1.75
Net Profit	18,721	0.90	1.36	-8.52	4.85

Note: Table reports summary statistics of the main variables. Control variables are statistically analyzed after removing missing values.

Table 6: **Confusion Matrices of Machine Learning Models**

Model	TP	TN	FP	FN
Gradient boosting	2847	15058	477	1072
ANN	2784	14930	605	1135
Logistic	2548	14928	607	1371

Note: This table reports the aggregated confusion matrices of four machine learning classification models based on cross-validation. TP = True Positive, TN = True Negative, FP = False Positive, FN = False Negative.

Table 7: **Machine Learning Credit Rating Models' Performance**

Model	Accuracy	F1 Score	AUC Score	R^2 Score
Gradient boosting	0.920	0.786	0.848	0.548
ANN	0.909	0.759	0.832	0.553
Logistic	0.898	0.720	0.806	0.521

Note: This table reports the out-of-sample prediction performance of four machine learning classification models - gradient Boosting, artificial neural networks (ANN), and logistic regression - evaluated using Accuracy, F1 Score, AUC Score, and pseudo-R^2 Score. All metrics are averaged over 10-fold cross-validation to assess out-of-sample performance.

Table 8: **Feature Importance for the Machine Learning Credit Rating Models (%)**

Feature Name	Gradient boosting	ANN	Logistic
Total Assets	31.757	28.618	46.591
State-Owned-Enterprise (SOE)	11.696	9.531	9.964
Operating Revenue	5.064	6.990	5.387
Debt/Invested Capital	4.589	5.697	4.609
Interest Debt	3.833	3.531	1.158
Quick Ratio	3.917	3.517	5.650
Agency_Fitch	1.343	2.771	4.594
Leverage	2.658	1.828	1.453
Cash Ratio	1.427	1.748	2.495
Net Profit	6.771	1.666	0.269
Operating Profit	1.735	1.586	0.082
EBITDA	1.483	1.189	0.046
Inventory Turnover	6.883	1.125	0.496
ROA	1.847	0.879	1.094
Short Debt/Total Debt	0.755	0.855	0.000
ROE	1.567	0.566	0.001
Agency_China Chengxin	0.811	0.498	1.334
Agency_United Credit	0.951	0.314	0.282
Cash Flow from Ops	1.203	0.291	0.023

Note: Table 7 presents the top 20 feature importance scores based on the permutation method, with results expressed as percentages. The importance of each feature is determined by how much shuffling that feature increases the log loss in its loss function. For readability, feature importance values are normalized to percentages.

Table 9: **Firm Leverage for the Low-rated Private Companies Under the Government "Deleveraging" Policy**

VARIABLES	(1)	(2)	(3)	(4)
Low-Rating×Private×Policy	-7.014***	-7.774***	-6.153***	-4.138***
	(-3.58)	(-4.51)	(-5.32)	(-4.31)
Low-Rating×Policy	6.267***	3.928***	8.230***	5.918***
	(9.72)	(6.78)	(18.85)	(13.70)
Private×Policy	4.058**	4.330***	1.150	-0.482
	(2.23)	(2.68)	(1.14)	(-0.59)
Low-Rating×Private	8.474***	6.908***	1.535	0.992
	(3.82)	(3.69)	(1.04)	(0.83)
Private	-5.406***	-4.434**		
	(-2.58)	(-2.49)		
Low-Rating	-8.980***	-0.594	-3.875***	-2.716***
	(-11.63)	(-0.86)	(-6.36)	(-4.90)
Controls	NO	YES	NO	YES
Industry FE	YES	YES	NO	NO
Firm FE	NO	NO	YES	YES
Year FE	YES	YES	YES	YES
Agency FE	YES	YES	YES	YES
Observations	19,454	18,721	19,454	18,721
R-squared	0.217	0.455	0.184	0.319

Note: This table presents the regression results analyzing the impact of the government's 2016 "deleveraging" policy on corporate debt leverage. The dependent variable is the firm's leverage ratio. The Policy is an indicator, which takes a value of 1 for the year 2016 and all subsequent years, and 0 otherwise. Robust t-statistics are reported in parentheses, with standard errors clustered at the firm level. Significance levels are denoted by: *** p<0.01, ** p<0.05, * p<0.1.

Table 10: **Firm Leverage for the Low-rated Private Companies Under the Government "Deleveraging" Policy (Nonlinear Policy Intensity Version)**

VARIABLES	(1)	(2)	(3)	(4)
Low-Rating×Private×Policy Intensity	-0.324***	-0.416***	-0.387***	-0.268***
	(-2.95)	(-4.55)	(-6.50)	(-5.33)
Low-Rating×Policy Intensity	0.372***	0.234***	0.419***	0.302***
	(10.47)	(7.17)	(17.81)	(13.46)
Private×Policy Intensity	0.111	0.189**	0.114**	0.005
	(1.11)	(2.26)	(2.30)	(0.12)
Low-Rating×Private	2.863**	0.542	-3.179***	-2.121**
	(2.07)	(0.51)	(-2.69)	(-2.02)
Low-Rating	-3.794***	2.661***	2.734***	2.000***
	(-6.90)	(5.04)	(5.04)	(4.07)
Policy Intensity	1.392**	-0.269	3.809***	-3.430***
	(2.07)	(-0.49)	(7.79)	(-5.34)
Private	-2.367*	-0.981		
	(-1.85)	(-1.00)		
Controls	NO	YES	NO	YES
Industry FE	YES	YES	NO	NO
Firm FE	NO	NO	YES	YES
Year FE	YES	YES	YES	YES
Agency FE	YES	YES	YES	YES
Observations	19,454	18,721	19,454	18,721
R-squared	0.216	0.454	0.167	0.311

Note: This table presents the regression results analyzing the impact of the government's 2016 "deleveraging" policy on corporate debt leverage. The dependent variable is the firm's leverage ratio. Policy Intensity is a continuous variable indicating the intensity of the government's 2016 "deleveraging" policy. Robust t-statistics are reported in parentheses, with standard errors clustered at the firm level. Significance levels are denoted by: *** p<0.01, ** p<0.05, * p<0.1.

A Appendix: Additional Data Exploration

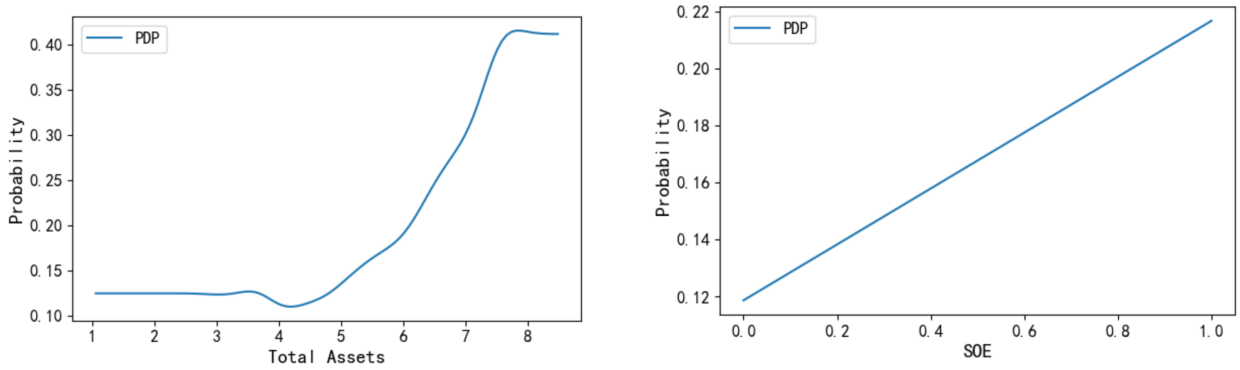
A.1 PDP with other important financial features

Figure A.1 shows the PDP of two important variables, total assets and Soe, in gradient boosting, ANN, and logistic regression models. In the gradient boosting model (Panel A), the total assets did not fluctuate much before 5 billion yuan, and the probability of obtaining a high credit rating slightly decreased as the assets increased. However, as the total assets increased to 5 billion yuan, the probability of obtaining a high credit rating significantly increased, especially when the total assets reached around 7, the probability rapidly increased. In addition, the status of state-owned enterprises significantly increases the probability of obtaining high credit ratings. In the ANN model (Panel B), the increase in total assets also significantly increases the probability of obtaining a high credit rating, but its growth trend is smoother. And the impact of state-owned enterprise identity is similar to that of the gradient boosting model. In the logistic regression model (Panel C), the impact of total assets on obtaining a high credit rating is the smoothest, and the probability increases significantly with the increase of total assets from the lowest asset to the highest asset. The impact of state-owned enterprise identity is similar to the first two models. The results indicate that the increase in total assets significantly increases the probability of obtaining a high credit rating, and the identity of state-owned enterprises also shows a positive impact in all models. The functional forms of these features vary slightly in different models, but the overall trend is consistent.

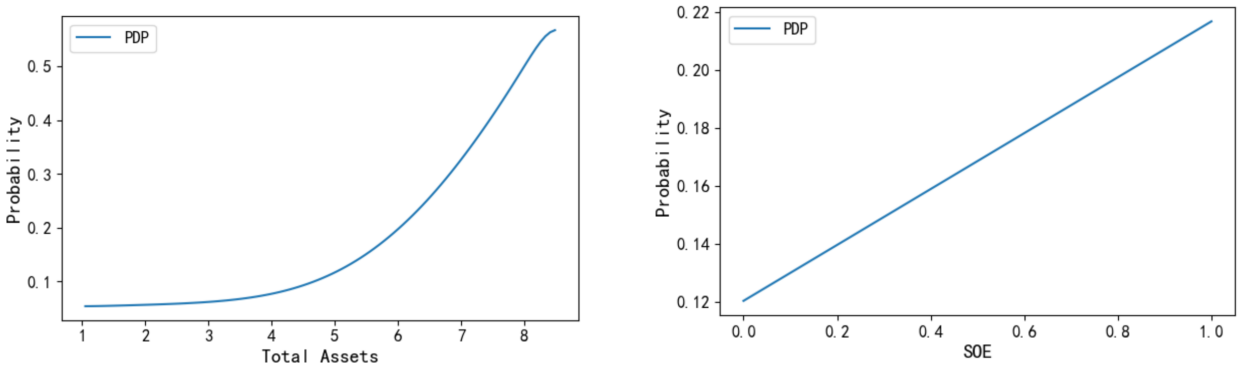
In addition to total assets and SoE as two significant features, the study ranks the importance of four other key financial variables (Operating Revenue, Debt to Invested Capital, Interest-bearing Debt, and Quick Ratio) and visualizes their impact on the target variable using partial dependence plot (PDP) across three different machine learning models. The results indicate that Operating Revenue exhibits a significant positive influence in all models, with the probability of the target variable notably increasing as Operating Revenue rises. However, the impact of Debt to Invested Capital varies depending on the model: in the gradient boosting model, its influence shows a nonlinear trend that first rises and then falls; whereas in the ANN and logistic regression models, its impact decreases as the ratio increases, demonstrating a negative relationship. The influence of Interest-bearing Debt displays a U-shaped trend in the gradient boosting and ANN models, while in the logistic regression model, it shows a linear upward trend. The Quick Ratio consistently exhibits a negative impact across all models, with the probability of the target variable significantly decreasing as the ratio increases. By comparing the partial dependence plots of important financial characteristics under different models, the study further reveals the relationship

between financial variables and corporate performance.

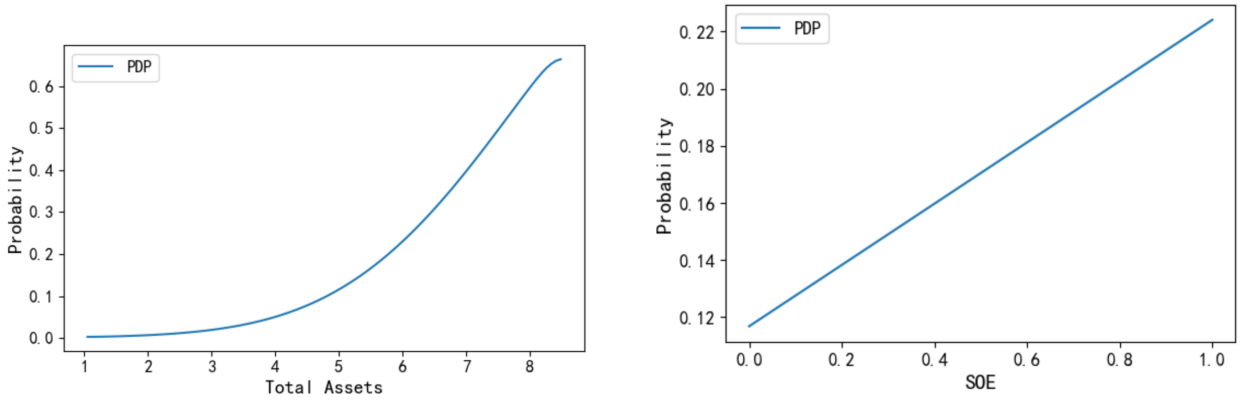
To further comprehend the equity within the credit rating market, this study investigates the moderating role of state-owned enterprise (SOE) status on the impact of four key financial variables-Operating Revenue,Interest-bearing Debt, and Quick Ratio-on the target variable. The findings reveal that SOE status confers significant advantages across multiple financial dimensions. Specifically, concerning Operating Revenue, SOEs can more effectively utilize their operating income to enhance their likelihood of securing a high credit rating. Furthermore, with respect to Debt to Invested Capital, SOEs demonstrate a greater capacity to counteract the adverse effects of high debt levels on credit ratings compared to non-SOEs. The relative advantages of SOEs are also apparent in the indicators of Interest-bearing Debt and Quick Ratio, indicating that SOEs possess certain strengths in liquidity management. The positive influence of SOE status on corporate credit ratings across several key financial indicators suggests that, under similar performance conditions, SOEs are more likely to attain higher credit ratings.



Panel A: PDP of Total Asset and SOE (Gradient boosting)



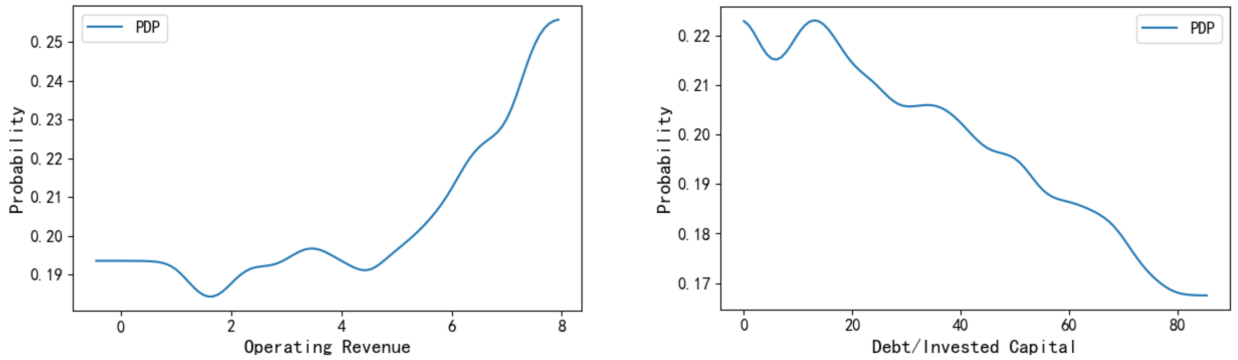
Panel B: PDP of Total Asset and SOE (ANN)



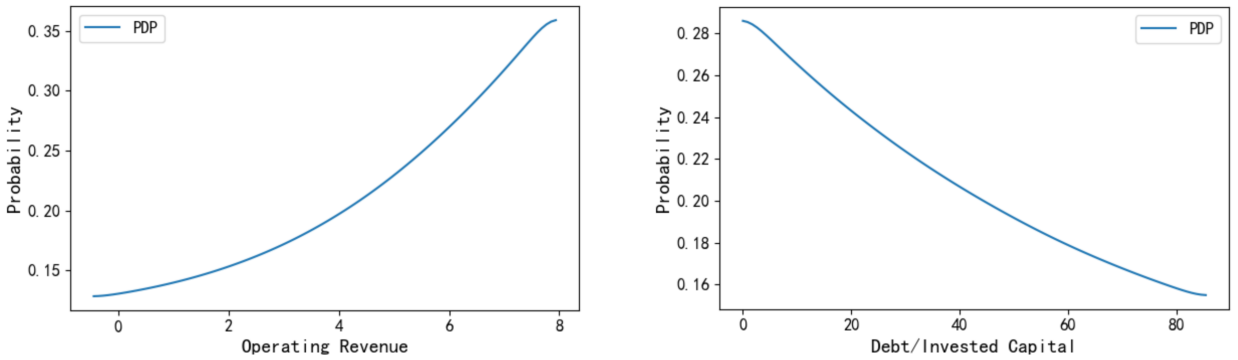
Panel C: PDP of Total Asset and SOE (logistic regression)

Figure A.1: **Partial Dependence Plots for Total Asset and SOE across Models.**

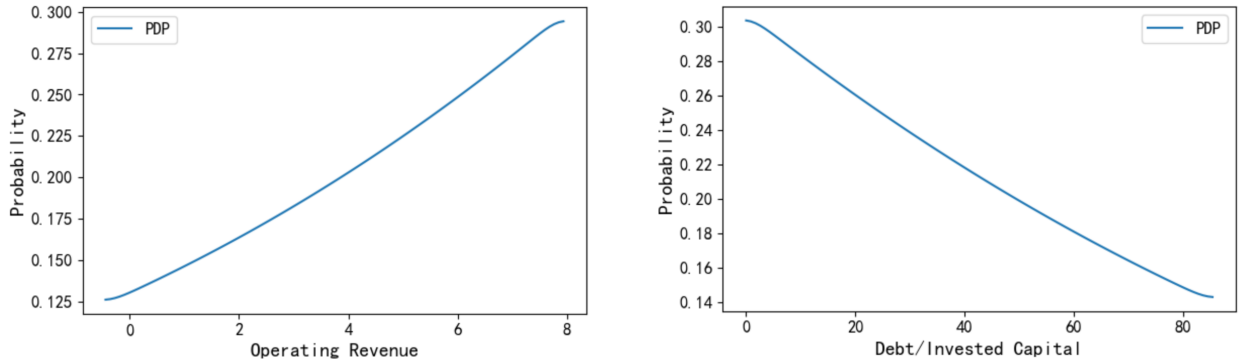
This figure presents the partial dependence plots of SOE and Total Asset (log-transformed) for three models: gradient boosting (Panel A), ANN (Panel B), and logistic regression (Panel C).



Panel A: PDP of Operating Revenue and Debt to Invested Capital (gradient boosting)



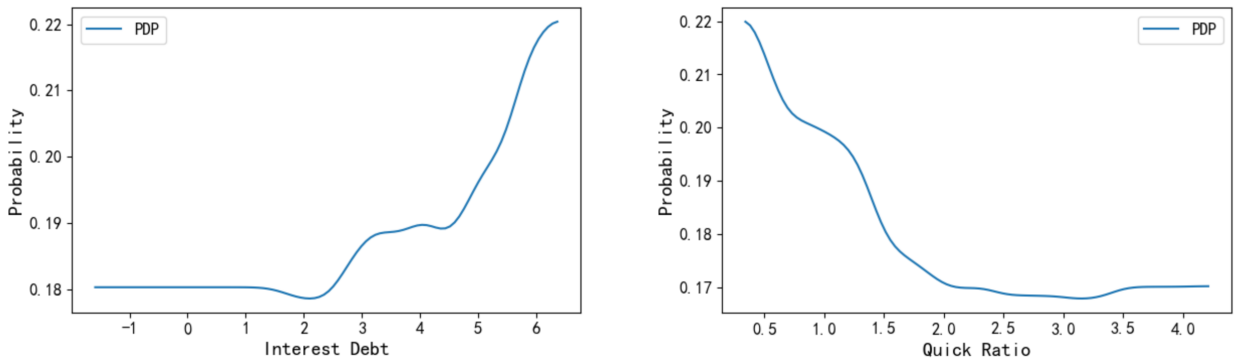
Panel B: PDP of Operating Revenue and Debt to Invested Capital (ANN)



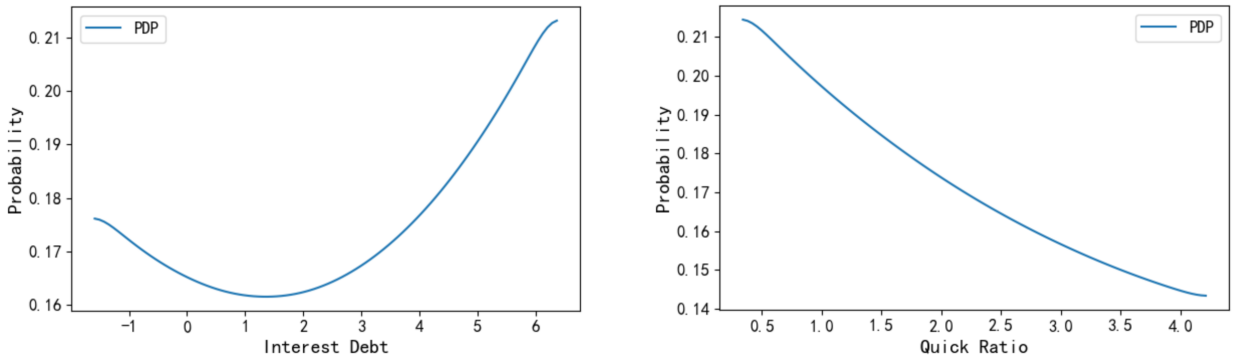
Panel C: PDP of Operating Revenue and Debt to Invested Capital (logistic regression)

Figure A.2: **Partial Dependence Plots (PDPs) for Operating Revenue and Debt to Invested Capital across Models.**

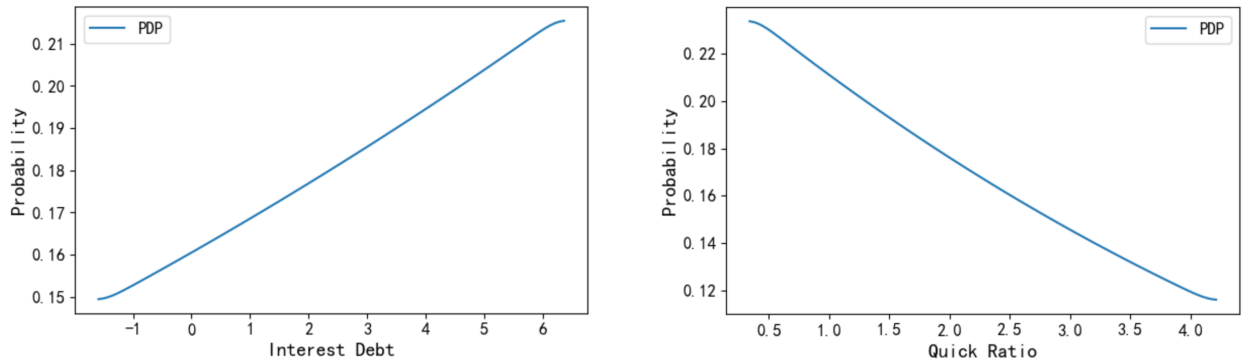
This figure presents the partial dependence plots of two financial features for three models: gradient boosting (Panel A), ANN (Panel B), and logistic regression (Panel C).



Panel A: PDP of *Interest-bearing Debt* and *Quick Ratio* (gradient boosting)



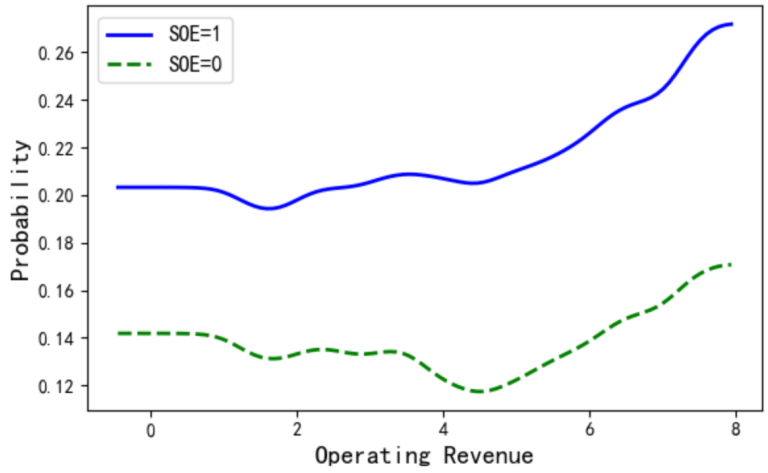
Panel B: PDP of *Interest-bearing Debt* and *Quick Ratio* (ANN)



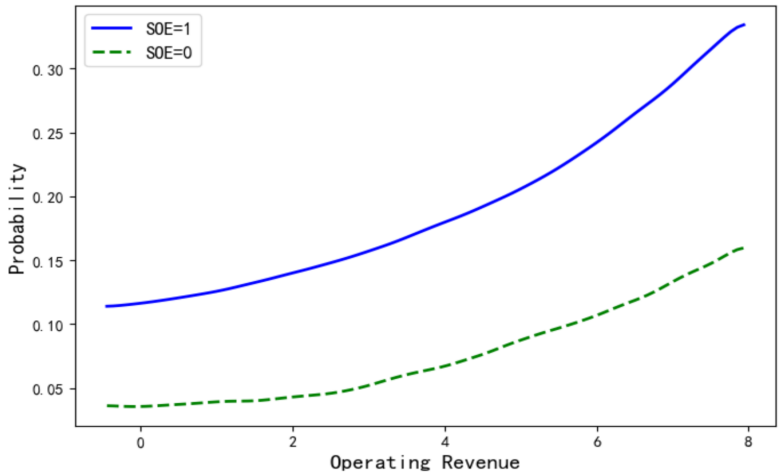
Panel C: PDP of *Interest-bearing Debt* and *Quick Ratio* (logistic regression)

Figure A.3: **Partial Dependence Plots (PDPs) for *Interest-bearing Debt* and *Quick Ratio* across Models.**

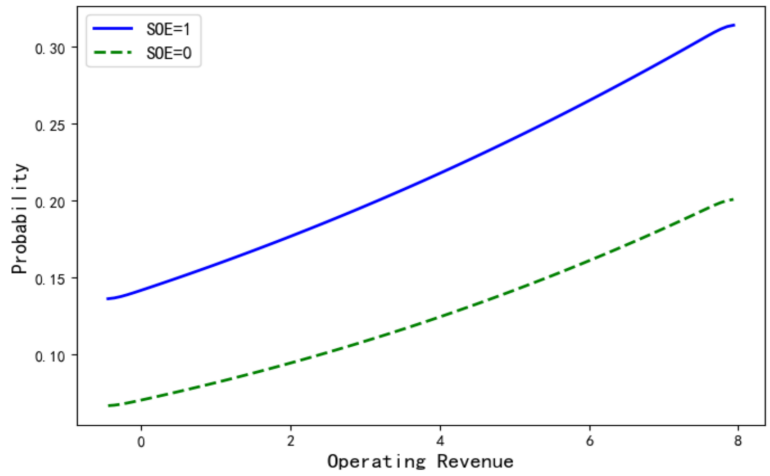
This figure presents the PDPs of two key financial features across three models: gradient boosting (Panel A), ANN (Panel B), and logistic regression (Panel C).



Panel A: PDP of Operating Revenue (gradient Boosting)

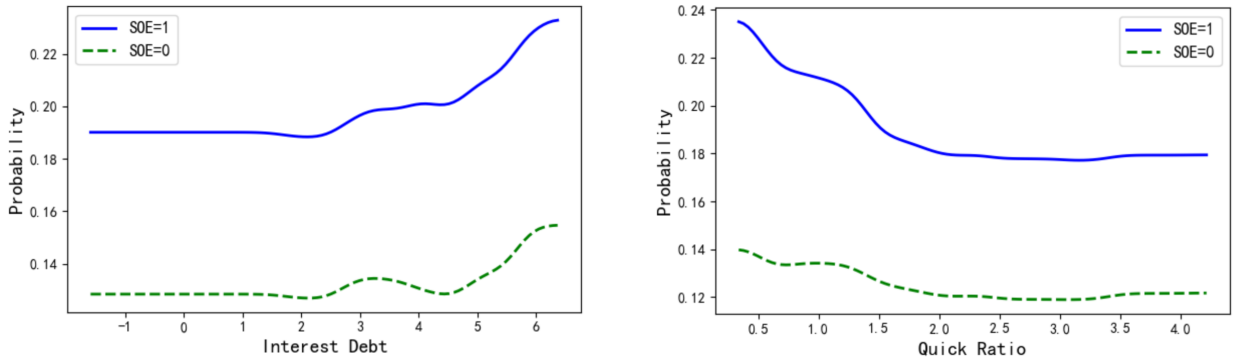


Panel B: PDP of Operating Revenue (ANN)

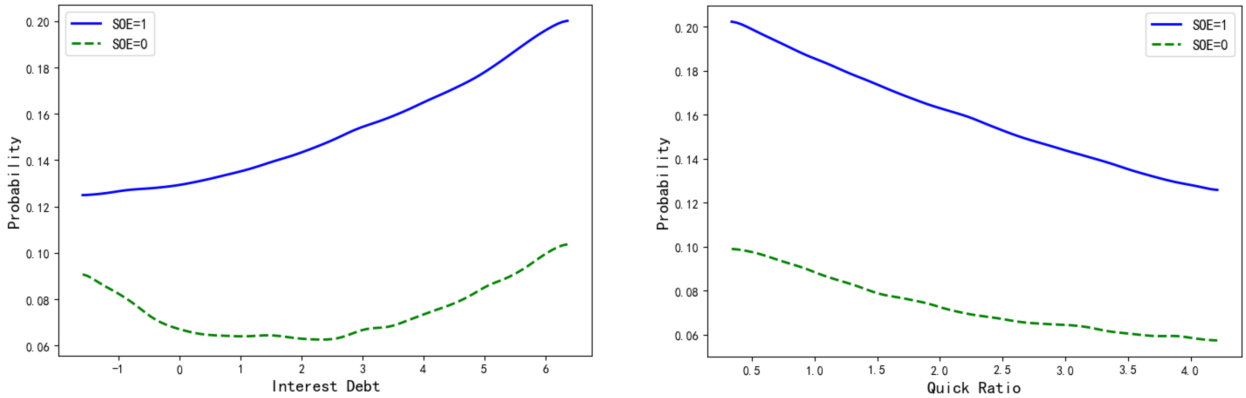


Panel C: PDP of Operating Revenue (logistic regression)

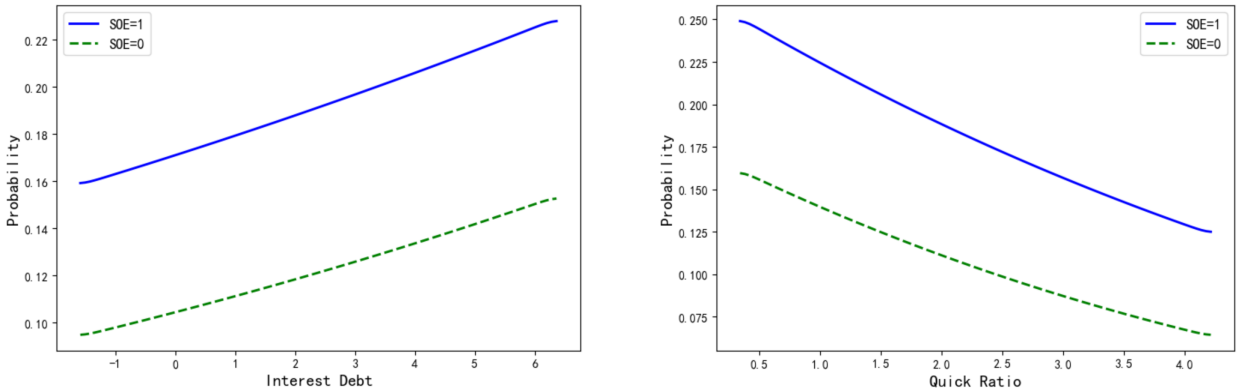
Figure A.4: **Partial Dependence Plots of *Operating Revenue* by SOE Status.** This figure displays the PDP of Operating Revenue (log-transformed) for SOE and non-SOE firms, separately modeled using gradient boosting (Panel A), ANN (Panel B), and logistic regression (Panel C).



Panel A: Conditional PDP of *Interest-bearing Debt* and *Quick Ratio* (gradient boosting)



Panel B: Conditional PDP of *Interest-bearing Debt* and *Quick Ratio* (ANN)



Panel C: Conditional PDP of *Interest-bearing Debt* and *Quick Ratio* (logistic regression)

Figure A.5: **Conditional Partial Dependence Plots (PDPs) for *Interest-bearing Debt* and *Quick Ratio* by SOE Status across Models.**

This figure shows the conditional PDPs of two financial features across three models: gradient boosting (Panel A), ANN (Panel B), and logistic regression (Panel C), with separate lines for SOE and non-SOE firms.