Decision Support Systems 170 (2023) 113964

Contents lists available at ScienceDirect

Decision Support Systems

journal homepage: www.elsevier.com/locate/dss

Financial risk prediction in supply chain finance based on buyer transaction behavior

Zelong Yi [a], Zhuomin Liang [a], Tongtong Xie [b], Fan Li [c,*]

[a] College of Economics, Shenzhen University, Shenzhen, China
[b] School of Foreign Languages, Shenzhen University, Shenzhen, China
[c] China Center for Special Economic Zone Research, Shenzhen University, Shenzhen, China

A R T I C L E I N F O

Keywords:
Supply chain stability
Delay payment scheme
Buyer transaction behavior
Financial risk prediction
XGBoost model

A B S T R A C T

Financial risk in supply chain finance (SCF) is defined as the possibility that suppliers fall into liquidity crisis due to delayed payment. Predicting financial risk is important for supply chain stability. In this paper, a financial risk prediction model is developed using XGBoost and then evaluated by applying buyer transaction behavior data. We further construct single and hybrid models, respectively, and compare their performance using receiver operating characteristic curve (ROC), area under the ROC curve (AUC), and F1-Score. Last, feature importance and partial dependence plots (PDPs) are employed for model interpretation. The results show that XGBoost model can effectively predict potential financial risks, and shed lights on managers' payment practice. This paper is one of the few studies that develop new models to examine financial risks in SCF empirically.

1. Introduction

Since the 2008 financial crisis, supply chain finance (SCF) has received considerable attention [35,41]. After the outbreak of COVID-19, SCF is increasingly recognized as a key instrument to mitigate the adverse financial impact of the pandemic [41]. Trade credit in delayed payment is the primary form of SCF [36]. 80% of UK companies adopt delayed payment in their transactions [48]. Credit payment has been the most typical form of payment [32].

Most literature on credit payment pays particular attention to its impact on investment decisions. For example, Taleizadeh et al. [50] analyze supply chain coordination by market segmentation and delay payment scheme (DPS) with incentive contracts to assess organizational and entrepreneurial profitability. Gupta et al. [20] study the ordering strategies with partially backlogged and time-varying deteriorating items that allow DPS, and construct an inventory model within a two-warehouse environment.

However, a report by Hilton-Baird Collection Services [22] shows an average delay of 25.2 days for firms with a turnover below £500,000, compared to an average delay period of 19.2 days for firms with a turnover over £3 million. This suggests that the degree of delayed payment varies from customer to customer, which tends to induce a liquidity crisis in suppliers [36]. The liquidity crisis of suppliers thus triggers SCF Financial risk. Interestingly, in the Internet era, SCF also features vulnerability and instability [3,62], and the risk can be transmitted among supply chain firms [24,66]. A UK review on the impact of buyers' delayed payments on suppliers' operation cost shows that excessively long delayed payment periods put enormous cash flow pressure on suppliers [17]. Thus, delayed payments is still a thorny issue for small and medium-sized enterprises (SMEs). Large enterprises often continuously fail to pay their suppliers on time [43], which causes significant financial losses to suppliers, especially SME suppliers. This suggests that large enterprises can be regarded as a major factor in delayed payment and should be taken seriously. And a full exploration of the payment behavior of large enterprises is beneficial for SMEs to improve their payment practices, thus better strengthening the stability of supply chain and promoting the development of SCF.

In our study, we define the possibility of a supplier falling into a liquidity crisis due to varying degrees of delayed payment by buyers as a financial risk to the customer in SCF. If buyers have a shorter delayed payment term, or if buyers use reverse factoring, as Wal-Mart initiated suppliers' SCF in 2009 [30], suppliers can receive payments faster and are less likely to fall into a liquidity crisis and suffer lower financial risks. Otherwise, there will be higher financial risks in SCF, leading to instability. However, the intra-firm information is often difficult to obtain [61]. Because of this, we retrieve the historical transactional data of

* Corresponding author.
E-mail addresses: yizl@szu.edu.cn (Z. Yi), liangzhuomin2019@email.szu.edu.cn (Z. Liang), xietongtong2018@email.szu.edu.cn (T. Xie), lifan@szu.edu.cn (F. Li).

https://doi.org/10.1016/j.dss.2023.113964
Received 24 October 2022; Received in revised form 23 March 2023; Accepted 27 March 2023
Available online 28 March 2023

large buyers published by the UK government, for the payment and transactional data are reliable, flexible, and show the operational conditions appropriately-timed [29]. Those with a higher percentage of late payments in total transactions within the required period are considered high risks, while others are low. In this way, we transfer our problem into a classification one. This paper is one of the few studies discussing financial risk in SCF. Therefore, our research seeks to tackle the following questions:

Q1. How to construct a model to predict risks in SCF?

Q2. How to evaluate this model's performance in prediction?

Q3. How to interpret this model and provide insights for supplier managers?

Hence, this paper constructs an SCF risk prediction model based on historical data of firms' transaction behavior using machine learning (ML) methods. Apart from the models, we use the Random Oversampling Technique algorithm to perform balanced sampling of the training data in case of imbalanced categorical distribution (*c.f.* [5]; in our study, the majority, i.e., low risks for non-default samples; the minority, high risks for default ones). We conduct a comprehensive experiment based on the XGBoost model on a large amount of data and compare it with different models. The results show that our XGBoost model outperforms others regarding ROC, AUC, and F1-Score values (*c.f.* section 4). Finally, this paper uses feature importance and partial dependence plots (PDPs) to perform an interpretability analysis of the model. Surprisingly, our result shows that changing the payment terms by large customers may lower the financial risk, which is against the idea of strengthening the risk (e.g., [35]). The reason could be the supplier may negotiate for more reasonable transaction conditions during term modification (e.g., ask to provide SCF), and strengthen the cooperation awareness, contributing to double wins. Hence, our study not only gives highlights for managers' data-mining in the transaction behavior of the customer but also enhances the understanding of cooperation awareness in the supply chain.

This paper advances the understanding of financial risk prediction in SCF with artificial intelligence (AI) and big data. First, this paper is among the few studies that evaluate financial risks in SCF, and we also focus on information about buyers' historical transaction behaviors. The results of this study provide practical inspiration to combine AI and big data in SCF financial risk prediction. Second, we use various methods to construct predictive models, including XGBoost tree models and Random Oversampling techniques, to deal with imbalanced data. Finally, we conduct comprehensive experiments to compare and analyze multiple models, including single and hybrid models, and generate the best model XGBoost using an interpretability framework. Within the limited literature, most of them adopt mathematical analysis (e.g., [50]). In contrast, we pioneer using empirical and data-driven approaches to study SCF risks, enlightening SCF participants in practice.

The rest of our study has been organized as follows. Section 2 reviews the relevant literature. In Section 3, we present the UK large-firm behavior data, the rationality of our variable definition, and the way we predict SCF financial risk. We also introduce the models and provide reasons in section 4. Section 5 presents and analyzes our results with robustness test and interpretability analysis. Section 6 provides insights to supplier managers. Section 7 concludes this paper, discusses the limitations and projects future research.

2. Literature review

Our research draws on three streams of literatures: (1) studies on the SCF risk analysis, (2) studies on payment scheme optimization, and (3) studies on SME credit risk prediction.

(1) SCF risk analysis

Regarding the study of risk in traditional supply chains, Blome and Schoenherr [6] consider supply disruptions and defaults as essential aspects of supply chain risk management. Assessing financial risk, Liu and Cruz [40] quantify financial risk based on a supply chain network with the capital asset pricing model (CAPM) and net present value (NPV) and explore how financial risk affects the value and profitability of supply chain companies, cash and credit transactions, and conclude that low-risk firms may be more profitable and valuable to the suppliers. Aqlan and Lam [2] argue that the supply chain risk factors are interrelated. They also propose a framework for assessing integrated supply chain risks based on expertise, historical data, and the structure of the supply chain, crystallizing supply chain agents, viz., supplier, customer, manufacturer, transportation, and commodity, with various corresponding risks. Ghadge et al. [18] reveal that financial risk affects the long-term relationships of supply chain stakeholders. Therefore, the financial risk in SCF is of great worth.

The financial assessment of SCF mainly includes the financing of SMEs, core-firm security, and the stability of the supply chain relationships and the industry [64]. Among them, credit risk prediction of SMEs' defaults or even bankruptcy is important [57,63,68]. Casanova [8] analyzes the data from 1003 companies and concludes that delayed payments lead to increased risks for firms and, unfavorably, defaults on their debts. Moretto et al. [42] include buyer-supplier relationships as a risk indicator, but they fail to focus on information about the historical transaction behavior of customers and suppliers. The research by van der Vliet et al. [54] reveals that longer delays negatively affect suppliers' ability to reduce financing costs.

Existing literature on the financial risk of SCF is very limited. Therefore, we consider the financial risk which arise from customers' DPS, i.e., the cash flow pressure on SMEs due to payment delays by large firms. Our study is among the few that focus on this kind of risks, which focus on the ability of large buyer-customers to pay on time, one significant indicator of which is risk transmission in the supply chain network [24]. Also, the adoption of SCF can be considered an essential factor. For example, reverse factoring has been increasingly popular recently in the industry as a way for companies to facilitate early payment of their trade credit obligations (e.g., [35,41]), helping many companies to be granted more generous payment terms [54], such as Wal-Mart starting using SCF in 2009 to enable suppliers to receive payments earlier (*c.f.* [30]). Liu et al. [37]'s study also elucidates the strategic significance of maintaining cash flow for SMEs' higher competitiveness.

(2) The payment scheme optimization

In a supply chain, a payment scheme is primarily a way of paying for transactions [1], and it correlates with the policies and profits of supply chain members as part of the supply chain contracts [19]. The study of payment schemes is also a hot topic in operations management [13]. More recent attention has focused on three payment schemes, advance payment scheme (APS), normal payment scheme (NPS), and delay payment scheme (DPS), respectively [7,60]. In supply chains, suppliers want buyers to pay earlier to manage cash flow better; however, customers prefer DPS [51]. DPS has become the most commonly employed payment scheme [32]. Seifert et al. [48] also demonstrate that over 80% of UK and US companies use DPS in their business, and approximately 60% of international transactions are conducted by DPS. DPS can be seen as a short-term financing approach within the supply chain [38]; van der Vliet et al. [54] propose that appropriate DPS may benefit suppliers in SCF and mitigate the payment stress and capital constraints, thus promoting more orders [60].

Numerous studies have attempted to explain the influence of payment schemes on operational decisions. For instance, Chang et al. [9] consider a combination of APS, NPS, and DPS to extend pricing and lot-sizing decisions for perishable goods and show the impact of combination on optimal strategies; the study by Li et al. [31] examines lot-sizing and backordering decisions of APS, NPS, and DPS to figure out the influence of mixed schemes on the optimal decisions; Cao et al. [7] analyze the optimal pricing and profit of suppliers and customers under different payment schemes based on Stackelberg model, identify the factors of different policies, and explore the selection of payment schemes in decentralized supply chains; Liu and Liu [38] discuss the green supply chain under various payment schemes. Wu et al. [60] investigate how APS, DPS, and reverse factoring affect financial performance of suppliers and customers, and analyze different indicators (e.g., capital cost). Wu et al. (2022) [59] develop a blockchain-based smart contract (BBSC) system for the late payment problem, and then implement it as a shadow system in a modular construction project with great certainty and efficiency in practice.

However, previous studies have not dealt with supplier selection for different customers in real business.

As mentioned earlier, these large buyers are the main factor in the delayed payment. Suppliers can make a better financial decision under efficient recognition for buyers' delayed payment degree, and we also argue that to strengthen the comprehensive management of the capital flow in the supply chain, especially the cooperation consciousness of participants in transaction, is the central issue for the optimal payment scheme. Furthermore, current studies lack a data-driven and empirical approach, so we model the delayed payment degree of buyers by a large dataset of customer released by the UK government, and provide advice to help suppliers improve their payment practice based on the model's result.

(3) SME credit risk prediction Model

A large and growing body of literature has investigated to predict SME credit risks, which can be briefly classified into 3 categories (*c.f.* [11,57,64,65,68]): ML-based, deep learning-based (DL-based), and hybrid approaches.

For machine learning methods, especially based on classification algorithms, Ciampi [12] predicts the credit risk of SMEs through logistic regression (LR) using financial ratios and corporate governance data. Similarly, Tian et al. [53] evaluate SCF credit risks by factor analysis and LR. Liu and Zeng [39] conduct the risk assessment by the comparison between gradient boosting decision tree (GBDT), support vector machine (SVM), LR and backpropagation (BP, a deep learning approach) which evidences that GBDT is superior in accuracy, compared with other models. Qian et al. [45] propose a PIMP-XGBoost model to predict the corporate financial distress, and prove its higher prediction accuracy and clearer interpretation, making it suitable for commercial use.

Much of the current literature on SCF risk prediction pays attention to deep learning methods, especially neural network. For instance, Sang [47] develops a backpropagation neural network (BPNN) to predict credit risks of Chinese automotive SMEs, concluding that the BPNN method outperforms the SVM. Li and Pan [33] adopt a neural network to study credit evaluation of SMEs in P2P, with surprisingly high accuracy. Wang [58] combines fuzzy neural network and block chain in the credit-risk study.

For the hybrid mechanisms, Zhu et al. [67] put forward integrated ensemble ML models in their prediction, showing a better accuracy than single model (i.e., decision tree). Zhu et al. [68] combine random subspaces (RS) and multiple enhancement techniques to propose an enhanced hybrid ensemble model (i.e., RS-MultiBoosting model) with a dataset of 46 SMEs and 7 core firms. Their findings suggest that their hybrid model outperforms traditional augmented learning methods in small-sized data. Kou et al. [29] construct a model by two-stage multi-objective feature selection, using transaction and payment data of Chinese SMEs. The results have proven the excellent model performance.

Our study draws on the research of credit risk predictive models for SMEs, with the difference that, to consider better model interpretability, we construct both a single tree-based model and a hybrid model and select the optimal one to improve the ability to predict financial risk.

In summary, the present study has made significant progress in three independent areas of SCF risk management, investment decision of payment schemes, and SCF risk control models. However, their integration has yet to be studied. Very little has been found in the literature on the financial risk of SCF. Even though DPS has been widely adopted, its extent varies among different customers, and effective identification of different customers is a vital issue in SCF operations. Thus, it is valuable to combine customer transaction behavior data and ML models to study the degree of DPS for different customers.

3. Data

In UK, large firms may have to report on their payment terms and practices. They should satisfy at least two of the following criteria, i) £36 million in turnover, ii) £18 million on its balance sheet, and iii) 250 employees. The companies are from various industries, e.g., health and medicine (Medtronic), oil and gas (Greenergy Fuels), etc. Hence, we collect data on large buyers' transactions with suppliers released by the UK government. This dataset involves payment policies, practices, and performance of large UK companies from year 2017 to 2020, with a total number of 21,820 valid observations.

For example, a report (2017.4–2017.10) by Medtronic Limited, reveals that the average time to pay invoices is 25 days, 77% of payments are paid in 30 days or less, 20% between 31 and 60 days, and 3% beyond 60 days. However, they report their invoices due but not paid within agreed terms is relatively higher at 89%. For its payment terms, the shortest and longest standard payment period is 30 days and 60 days, respectively, and the maximum contractual payment period is also 60 days. Medtronic does not provide payment code, SCF, any policy covering charges for remaining on supplier list, or any charges having been made for remaining on supplier list. This company either does not report whether payments were made in the reporting period or whether their suppliers were notified of changes.

Howorth and Reber (2003) [23] state that the late payment can be defined as the proportion of overdue invoices, and Bahrami et al. [4] code the customer's payment behavior to distinguish between risk categories in their study of customer invoice payment behavior as 0 and 1 (1 for customer having invoices not paid within the date, otherwise 0). Therefore, based on previous research, we construct the dependent variable "target" based on the variable "% Invoices not paid within agreed terms" in the data, and similarly 1 denotes the high risks, 0 as the low risks. Pareto 80:20 rule has been widely used in economic studies, and recent research for late payment also applies this rule and proves its rationality (e.g., [44]). Hence, we employ this method and define the last 20% as high-risk samples. The 80th quantile of the invoice data is given in Fig. 1.

Company payment behaviors have been used as variables in research, for example, Kou et al. [29] study the cash flow within 30 days. However, according to the UK act in 2017.4 (c.f. the report by the Department for Business, Energy & Industrial Strategy, UK in 2019), all large companies are required to publish their payment behavior data twice a year. Therefore, very few studies focus on SCF financial risk with this kind of information, and our study is one of them, including some interaction information by the large companies and their suppliers. Based on that, other variables in the dataset is employed as our independent variables, which mainly reflects customers' transaction behavior, such as historical payment behavior and the interaction behavior of transactions (c.f. Table 1).

The payment behavior of customers can effectively reflect their cash flow levels (e.g., [10,29]). However, the irregular payment behaviors can greatly influence the financial stability of the supply chain [49]. At this time, the effective mining of customer payment behavior is beneficial to participants' financial decisions in transaction. For example, Tangsucheeva and Prabhu [52] construct a corporate cash flow prediction model based on the historical payment time and other behavioral information of the company, which in turn can effectively predict the business situation of the customer; Bahrami et al. [4] examine the information of customer payment behavior (e.g., invoice payment,

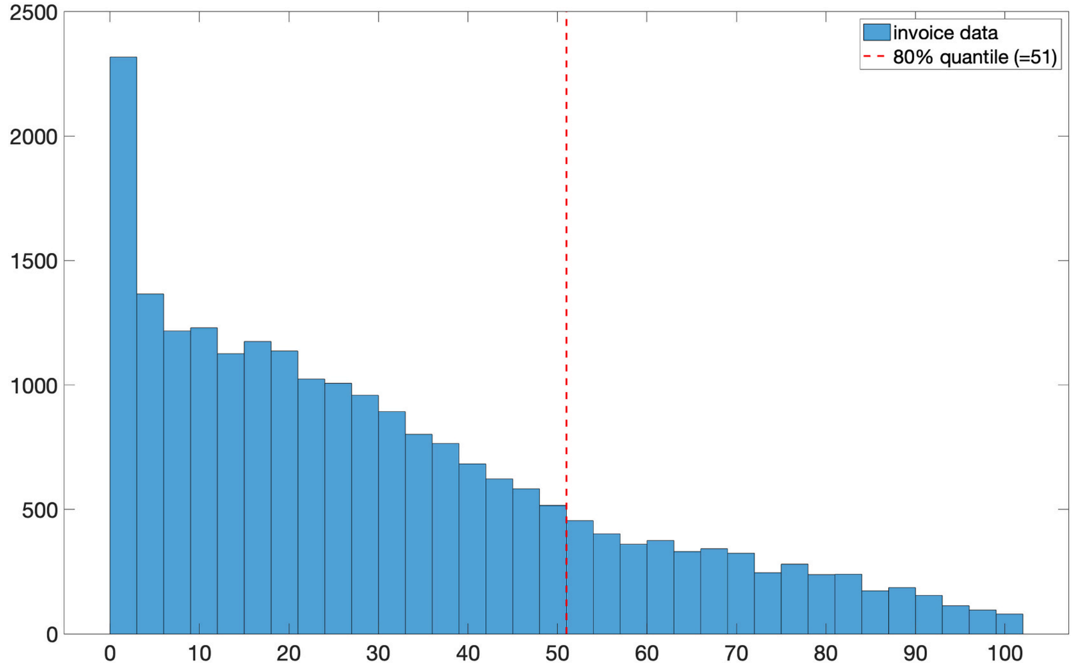

Fig. 1. Bar chart of the variable "% Invoices not paid within agreed terms", with 51 as the 80th quantile.

number of unpaid invoices, etc.) to predict whether customers will pay their invoices on time, and demonstrate that the model can help decision makers in cash flow management to promote the financial stability of the firm. In this paper, in addition to the customer's payment behavior, we also consider the interactive information of the transaction (e.g. whether the customer offers SCF, charges a fee for retaining supplier status, or provides electronic invoices, etc.), which allows us to better predict the extent of their late payment and manage SCF financial risk.

We first process the missing values of the data. For continuous variables, we choose to replace them with the mean for a better representation of the sample, while for discrete variables, we use the model for the missing ones.

Then a correlation analysis is performed, as shown in Fig. 2, in which the darker blue indicates a strong positive correlation between the variables. It is apparent that "% Invoices paid within 30 days", "% Invoices paid later than 60 days", "Suppliers notified of changes", and "Policy covers charges for remaining on supplier list" have significant correlations with other indicators, so we remove these indicators.

Finally, as shown in Fig. 3 (left panel), there is an unbalanced distribution between high-risk and low-risk samples. That is, the high-risk sample only accounts for about 20% of the entire data set, and the distribution of high-risk and low-risk samples is highly unbalanced, which can lead to a lower accuracy of prediction for high-risk samples and a higher one for low-risk samples, thus affecting the performance of the model prediction [65]. For this reason, the commonly used Random Oversampling algorithm is adopted to deal with the unbalanced distribution of samples in financial risk prediction (e.g., [29]).

4. Methodology

We first briefly summarize the models used in our study. Li et al. [34] stress the importance of interpreting in financial data mining. Hence, given great interpretability of tree-structured models [15], we selected some frequently used tree-structed models for comparison, i.e., XGBoost, Random Forest, and GBDT, LightGBM as single models. Inspired by He et al. [21], we also construct hybrid models. Since our dependent variable is binary classified, and linear regression (LR) also shares great interpretablity (ibid.), we also construct hybrid models (XGBoost + LR, Random Forest + LR, and GBDT + LR) for comparison.

For the single model, we choose the four commonly used tree model, i.e. XGBoost, Random Forest (RF), Gradient Boost Decision Tree (GBDT) and LightGBM. XGBoost is an ensemble learning algorithm, the latter of which refers to the construction and integration of multiple learners to complete the learning task. A recent study by Qin [46] adopts XGBoost as a major device to perform a risk model. RF is a Bagging strategy of the ensemble learning algorithm. The use of RF is a well-established approach in SCF-related research, e.g., Kalsyte and Verikas [25]'s study of entrepreneurial financial soundness with RF. Given its popularity, GBDT has been applied to many studies in finance (e.g., [37]). Similar to XGBoost, it also belongs to ensemble learning and uses the gradient boosting framework, but GBDT only supports CART base tree in learning. LightGBM is an advanced ML algorithm, which employs histogram algorithm and grows tree by leaf-wise strategy. Its histogram algorithm can merge mutually exclusive features, lower cost in time, and leaf-wise algorithm is used to avoid overfitting (see [56]). Since the dependent variable is binary, we construct 3 hybrid models, i.e., XGBoost + LR, RF + LR, and GBDT + LR, respectively.

4.1. Model applications

We choose scikit-learn (scikit-learn.org.cn), an open-source Python-based machine learning library as the framework. Also, for XGBoost model, we adopt XGBoost library. We then randomly select 70% of the data as the training set and 30% as the test set, and we also use the grid search method to determine the primary parameters.

Three indicators have been applied to evaluate model performance, viz., receiver operating characteristic curve (ROC), area under the ROC curve (AUC), and F1-Score, respectively (e.g., [28,55]).

FN: False Negative, the predictive result is a negative sample, but it is actually a positive one.

FP: False Positive, the predictive result is positive, but actually negative.

TN: True Negative, the predictive result is a negative sample, and in fact also a negative one.

TP: True Positive, the prediction is positive, the same as the actual observation.

First, the ROC curve is obtained by the values of true positive rate (TPR, or correct rate) and false positive rate (FPR, or error rate) based on

Table 1
Data description.

Variable	Definition	Range
Average time to pay	Average time to pay for all transactions during the reporting period	[0,1000]
% Invoices paid within 30 days % Invoices paid between 31 and 60 days % Invoices paid later than 60 days	The percentage of payments made within the reporting period which were paid: between day 1 and day 30 (including day 30), between day 31 and day 60 (including days 31 and 60), on or after day 61, receiving the invoices or any other form of notice of payment	[0,100]
Shortest (or only) standard payment period	Standard minimum payment cycle for the reporting period of the industry in which the customer operates	[0,1000]
Longest standard payment period	Standard minimum payment cycle for the reporting period of the industry by the customer	[0,1264]
Maximum contractual payment period	Maximum payment cycle for contractual agreements	[0,5475]
Payment terms have changed	Whether standard payment terms have been modified in transactions in the reporting period, e.g., additional extensions of terms	{0,1} 1 for modification
Suppliers notified of changes	Whether the suppliers have been required to modify standard payment terms	{0,1} 1 for requirement
Participates in payment codes	Whether the customer participates in standard payment guidelines at the time of the transaction, such as paying no later than the agreed-upon time	{0,1} 1 for participation
E-Invoicing offered	Whether the parties in the transaction adopt e-invoice	{0,1} 1 for adoption
Supply-chain financing offered	Whether the customer adopts SCF	{0,1} 1 for adoption
Policy covers charges for remaining on supplier list	Whether the customer may deduct a portion of the amount from the transaction as a fee to maintain the supplier relationship based on the payment terms or practices	{0,1} 1 for deduction
Charges have been made for remaining on supplier list	Whether the costs of maintaining supplier partnerships are deducted	{0,1} 1 for maintenance

the results of the predictive category combination of the samples. Then TPR and FPR will be plotted as horizontal and vertical coordinates, respectively. They can be computed as follows:

$$\mathrm{TPR} = \frac{TP}{TP + FN} \tag{1}$$

$$\mathrm{FPR} = \frac{FP}{TN + FP} \tag{2}$$

$$\mathrm{F1-Score} = \frac{2 \times TP}{2 \times TP + FP + FN} \tag{3}$$

If the ROC curve reaches to the upper left corner, our model will feature higher accuracy. AUC is the area under the ROC curve. Hence, the area of AUC and model performance are positively correlated.

F1-Score is another commonly used evaluation metric in classification (see Formula 3). It reflects the ability to identify the positive samples, i.e., the enterprises marked as high-risk. For the financial risk prediction studied in this paper, it is more costly to classify high-risk enterprises as low-risk ones than to predict low-risk as high-risk, when the cooperation with high-risk enterprises, managers are more likely to suffer from the delayed payment behavior of customers and puts greater pressure on their cash flow operations, especially for the SMEs prone to liquidity crises. The higher F1-Score indicates a lower possibility of predicting errors for the high-risk samples and a better performance. This further reflects the practical value of our model to identify enterprises with a higher degree of delayed payment and thus provide support for managers' financial decision making.

5. Results

5.1. Model comparison

The KS and AUC values are used to compare our six models, which are shown in Table 2. We find the following results.

First, the single models show better performance than the hybrid ones. Among the single models, the predictive performance of RF is the lowest. However, RF performance is better than the best hybrid model RF + LR, which suggests the financial risk predictive model based on the single model is more powerful than the hybrid model, and disagrees with Zhu et al. [67]'s study. An implication of this might be that some variables in our study (e.g., customers' trading behavior) are discrete ones, leading to a lower data diversity. Therefore, it is difficult for the hybrid models to fully adopt the linear combination of features to improve model performance.

Second, XGBoost shows the best performance among all models. In the single model, the F1-Score of LightGBM is improved by 1.1% and the AUC by 0.1%, compared to GBDT. Nevertheless, XGBoost supports parallel computation by the second-order Taylor expansion, which allows the loss function to use both first-order and second-order derivatives and improves its efficiency. Furthermore, it can also identify the appropriate classification information on financial risk assessment with customers' historical transaction behavior. Hence, the F1-Score and AUC values of XGBoost are improved by 3.7% and 1.2% over those of LightGBM, showing that XGBoost model has the most reasonable prediction performance.

In Fig. 4, it can be seen that the ROC curve of XGBoost is always located at the top left of the graph, enclosing the other models, indicating that the performance of the XGBoost model is better than the performance of the others.

Also, in the upper right corner (*c.f.* Fig. 5), the ROC curves of other models are almost wrapped by that of XGBoost, except for the deviations of LightGBM and GBDT models. Although deviations do not yield our result, i.e., XGBoost as the optimal model, we will still bear in mind the possible bias.

5.2. Robust and sensitivity analysis

5.2.1. Robust analysis for sampling methods

In order to test the stability for different sampling methods, we choose 3 classical methods, i.e., random under-sampling, random oversampling, and SMOTE oversampling (c.f. [5,29]). Random oversampling refers to randomly copying the minority samples to increase their size, while random under-sampling randomly samples the majority. Compared to random oversampling, SMOTE oversampling implements the analysis and simulation of the minority samples and adds the simulated new samples to the training data. Specifically, we compare the performance for these methods by AUC and F1-Score, and the results are shown in Table 3. We can see that the random over-sampling method has better accuracy, which also reflects better performance of our model. (See Table 4.)

5.2.2. Sensitivity analysis for risk proportion

In our study, the last 20% samples are selected as high-risk. In order to evaluate the sensitivity of our financial risk prediction model, the degree of customers' delayed payment is considered by different percentages of high-risk samples (i.e., 10%, 30%, 40%, 50%), the higher the risks, the more sensible the model classification of customer. Based on our XGBoost model, the performance of different risk division is assessed by ROC, AUC and F1-score, the results of which are as follows.

In Fig. 6, the model outperforms when the high-risk proportion is

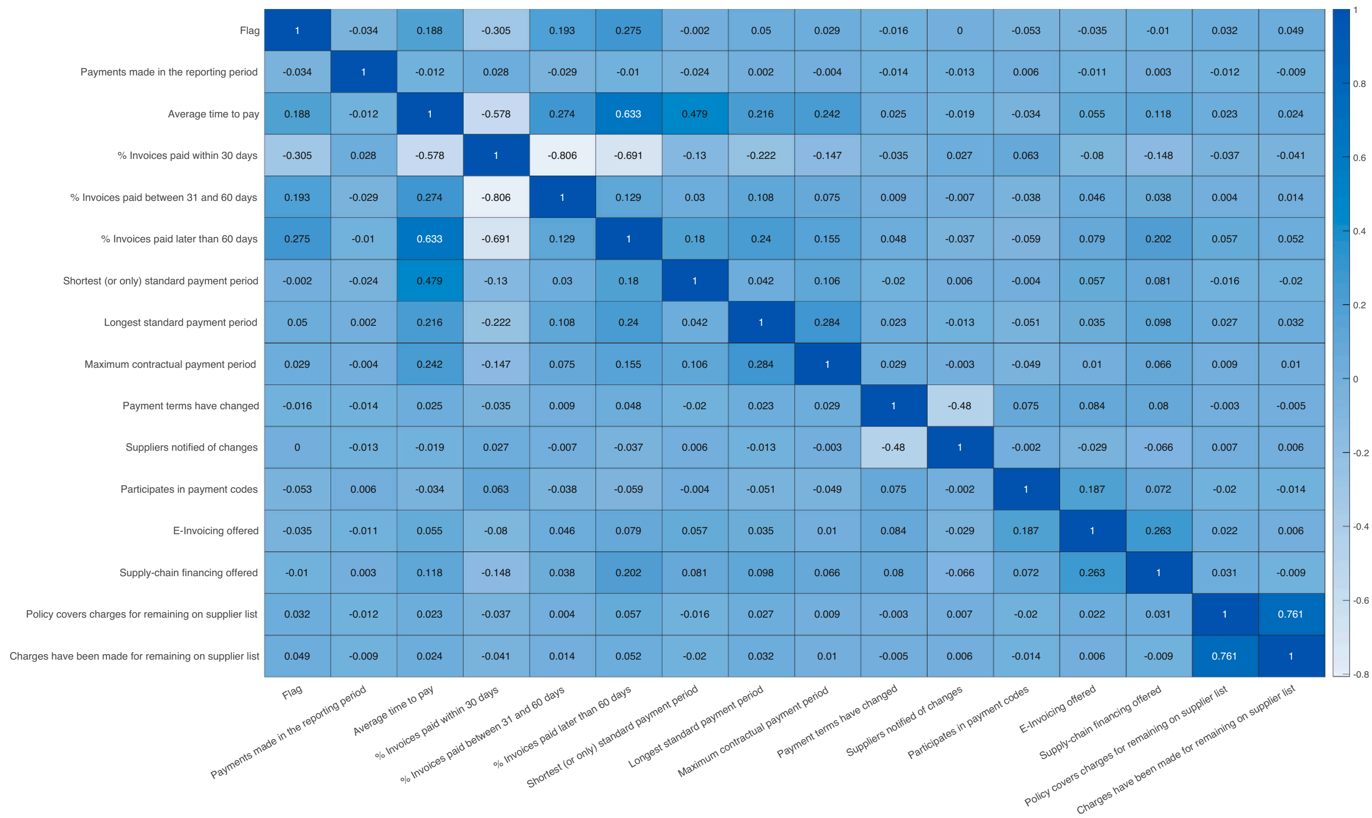

Fig. 2. The correlation of indicators.

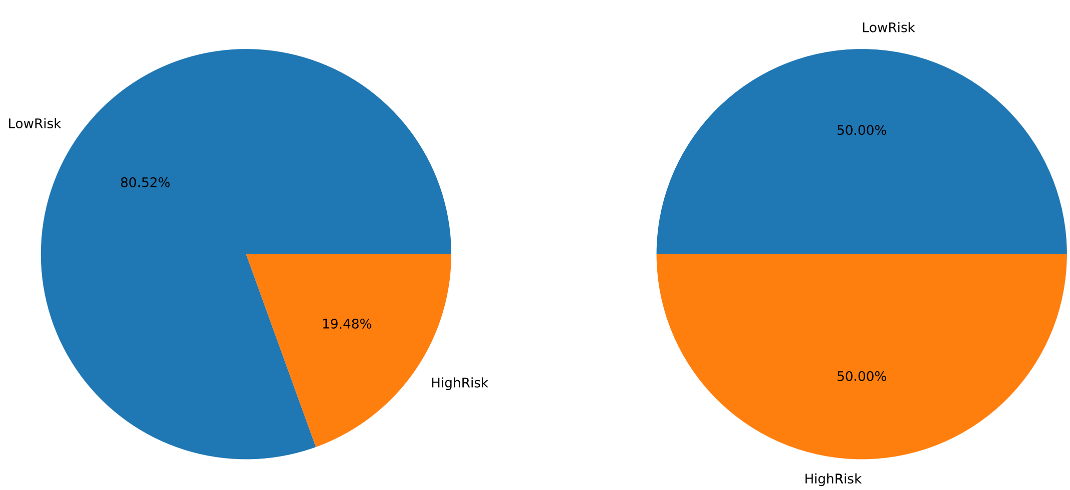

Fig. 3. Pie chart of low-risk and high-risk samples (left panel: original definition, i.e., 20% as high-risk; right panel: processed by Random Oversampling algorithm, See [29]).

Table 2
F1-Score and AUC Values.

Model		F1-Score	AUC
Single models	XGBoost	0.548	0.820
	RF	0.517	0.801
	GBDT	0.522	0.809
	LightGBM	0.528	0.810
Hybrid models	XGBoost + LR	0.484	0.776
	RF + LR	0.506	0.778
	GBDT + LR	0.499	0.767

marked as 20%, i.e., the ROC curve wraps around the other models, and also its fluctuations are smoother. In addition, the AUC and F1-Score values for the 20% are also the highest, which highlights the best performance of the model at this time.

5.2.3. *10-fold cross-validation*

To further examine the performance of the model, we choose 10-fold cross-validation method. It refers to splitting the initial training data into 10 parts, one for validating the model, and the others for training. This process is repeated 10 times, with each part for once-validation, and finally 10-times performance scores are obtained by AUC and F1-Score, it can be seen that our model features high prediction accuracy,

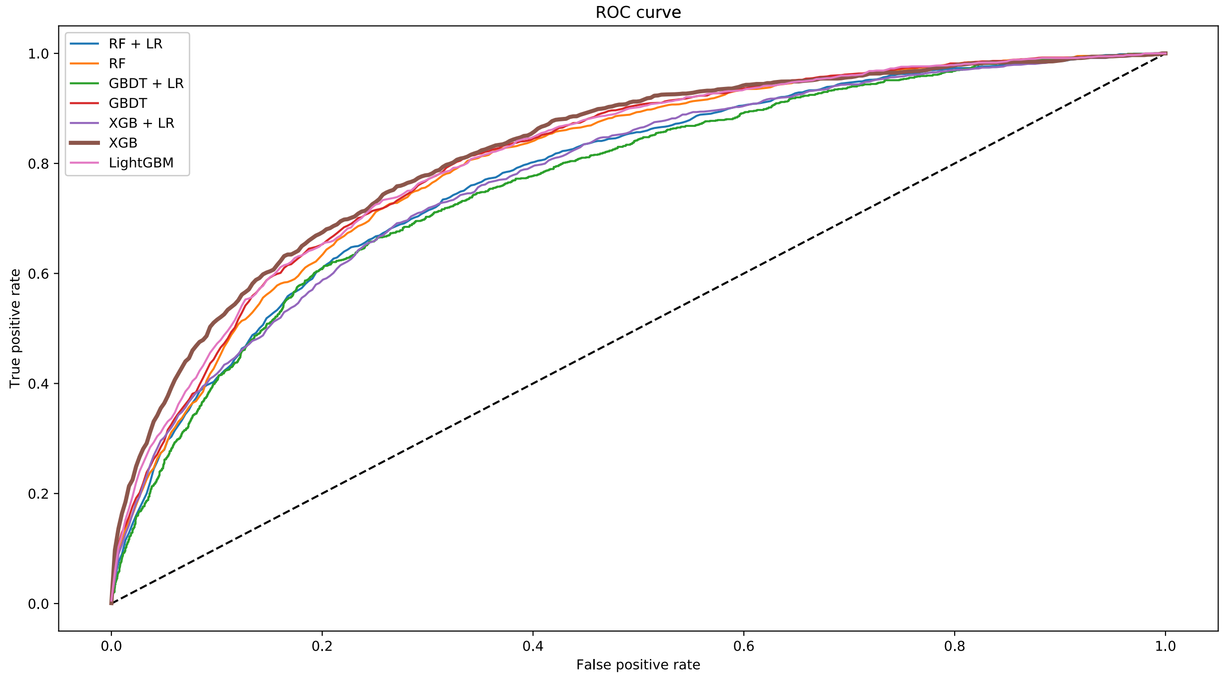

Fig. 4. ROC curve.

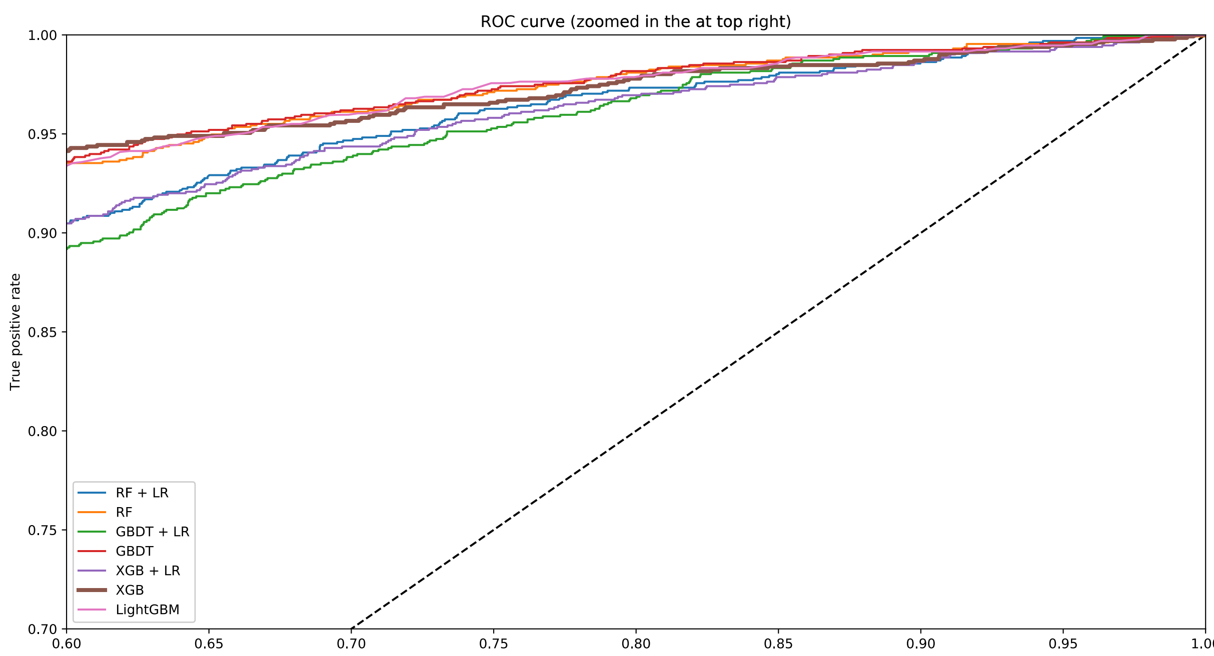

Fig. 5. The right corner of ROC.

Table 3
Results for F1-score and AUC by sampling methods.

	F1-Score	AUC
SMOTE oversampling	0.508	0.826
random oversampling	0.538	0.827
random undersampling	0.519	0.805

Table 4
Results for sensitivity of risk proportion.

Proportion	F1-Score	AUC
10%	0.484	0.787
20%	0.548	0.820
30%	0.532	0.810
40%	0.495	0.781
50%	0.464	0.747

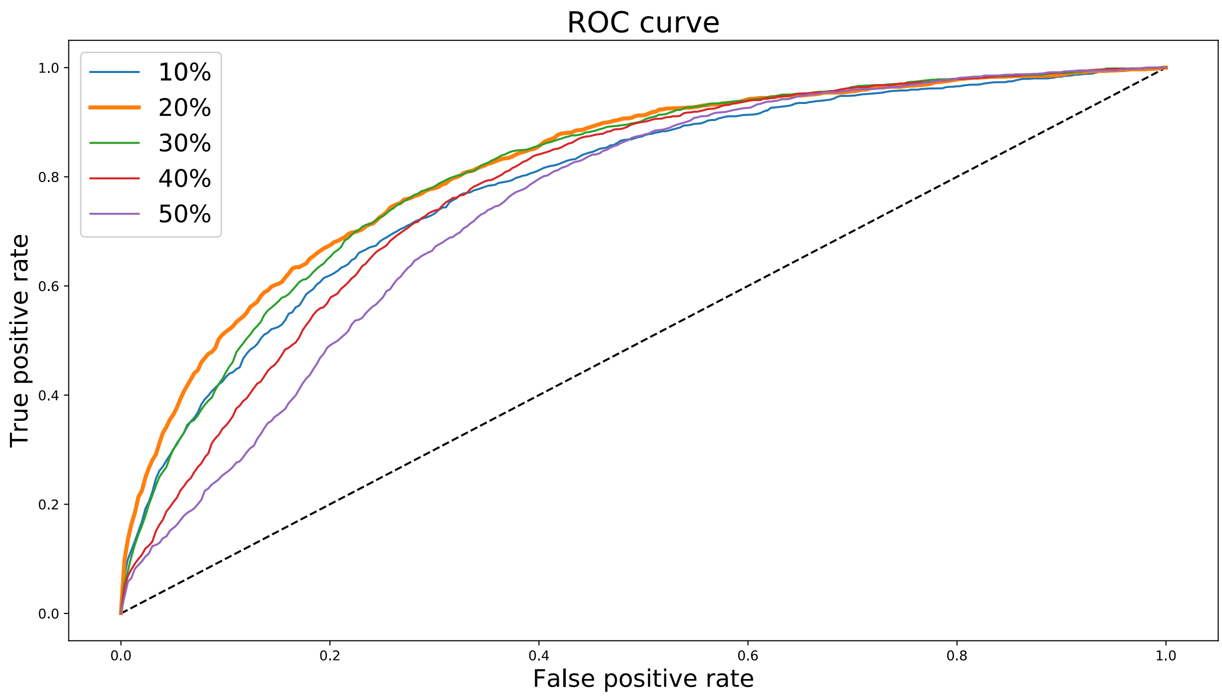

Fig. 6. The ROC of different Risk Proportion.

low fluctuation, and stable performance (For AUC, *mean* = 83.71%, *SD* = 1.53%; For F1-Score, *mean* = 44.69%, *SD* = 2.91%).

5.3. Model interpretability

With our XGBoost model, we show the ranking of feature importance by the IVs (c.f. Fig. 7, produced by Python library scikit-learn) and select the top 4 (i.e, "Average time to pay", "Supply-chain financing offered", "Payment terms have changed", "Charges have been made for remaining on supplier list") for further analysis by PDP (also constructed with with scikit-learn), a powerful tool to visualize the impact of selected variables on the results [16]. It should be noted that these variables are discrete ones, except for "Average time to pay" (details c.f. section 2).

From Fig. 8, as the average time to pay increases, the likelihood of the sample being classified as high-risk raises, but the marginal effect also decreases.

Two variables ("Supply-chain financing offered", "Payment terms have changed", see Figs. 9 and 10) show a lower degree of being considered high-risk when the variable takes a value of 1, and contrast to "Charges have been made for remaining on supplier list" (0 for high risks). It is also evident from the vertical axis of the PDP plots of these variables that they are all essentially more homogeneous for risk category prediction, except for the average time to pay, a continuous variable that contains more categorical information. All the above supports

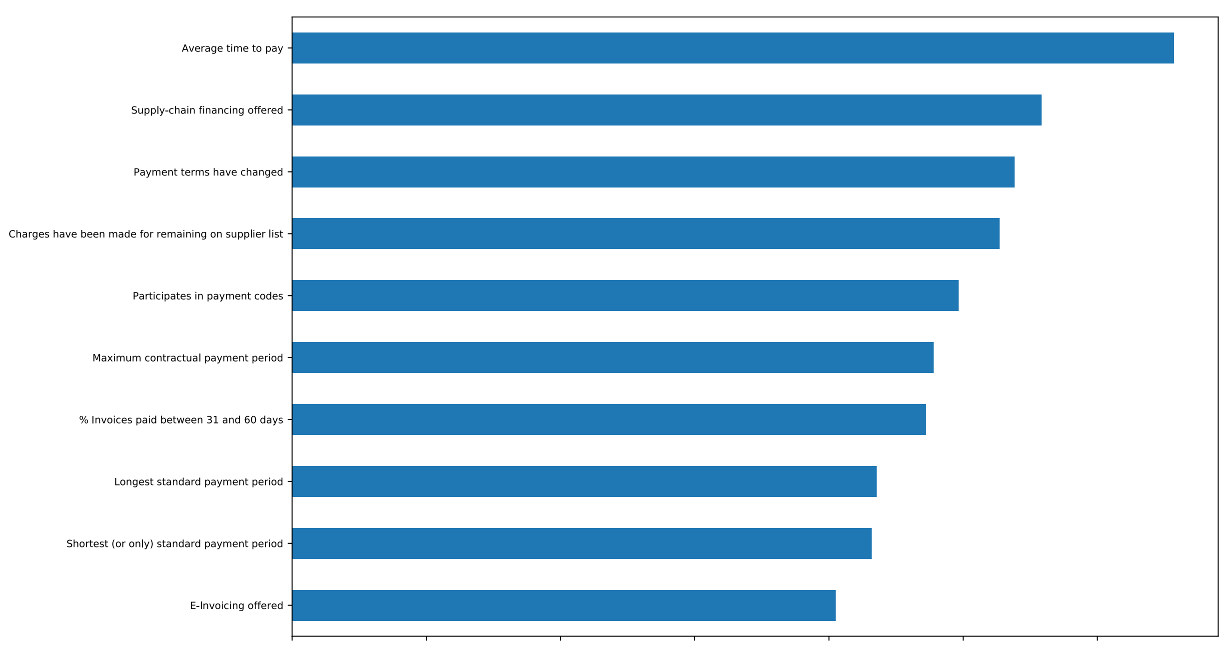

Fig. 7. Feature importance.

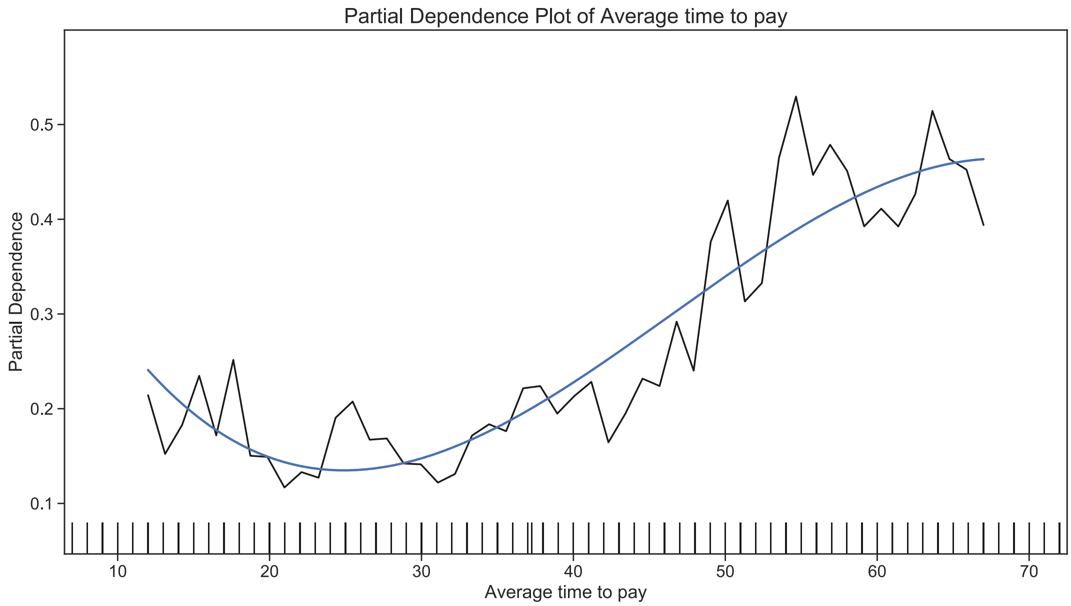

Fig. 8. PDP of average time to pay.

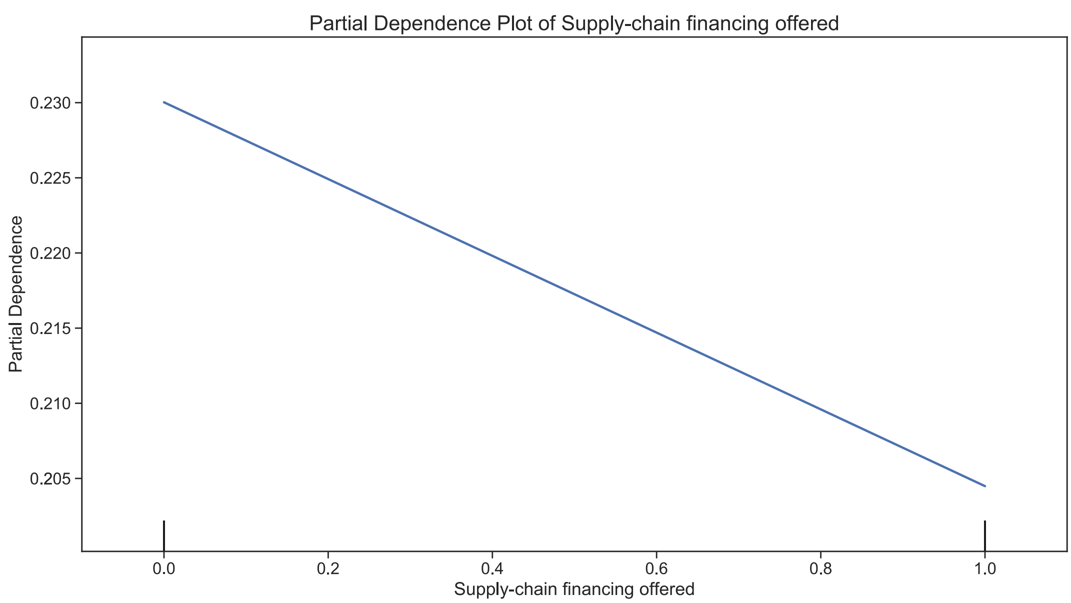

Fig. 9. PDP of supply-chain offered.

the idea that risk elements within the supply chain are interrelated [2].

First, on the interpretations of the figures above, average payment time reflects the level of cash management of the company; the higher the level, the faster the company can pay for the goods. However, suppose the average payment is more time-consuming. In that case, the cash flow cycle will be significantly extended, which makes its partner suppliers vulnerable to liquidity crises due to the inability to receive payments timely and quickly, leading to suppliers' higher financial risks from customers.

Second, for the discrete variable "Supply-chain financing offered", the financial risk faced by suppliers decreases when the customer provides SCF, which also indicates that reverse factoring is important to control the SCF risks. As previously mentioned, Wal-Mart has been using SCF since 2009 to enable suppliers to receive payments earlier [30]. Our PDP analysis on this enhances the understanding of SCF and highlights its role to maintain supply chain stability.

Third, the "Payment terms have changed" and "Charges have been made for remaining on supplier list", are categorized as transactional interaction behaviors (viz. the content of contract negotiations).

The PDP of the "Payment terms have changed" crystallizes that the modification the standard payment contract terms during the transaction can reduce the financial risk of SCF operations. The standard payment terms fixes customers' payment time, while it may be varied in different situations (e.g., business location, credit term). Klapper et al. [27] study a dataset of nearly 3000 trade credit contracts, and show that the stronger bargaining power of large customers can lead to more favorable trade credit contract terms. Liu et al. [35] point out that large customers are prone to exploit their stronger bargaining power for benefits, which in turn increases the likelihood of their counterparties falling into liquidity crises, i.e., higher SCF risks. However, Fig. 10 also shows that modifications to the standard payment contract terms lower the risks. The reason could be the standard payment terms can only represent the trading practices of a certain industry; but for these large firms with stronger bargaining power, they generally do not choose to

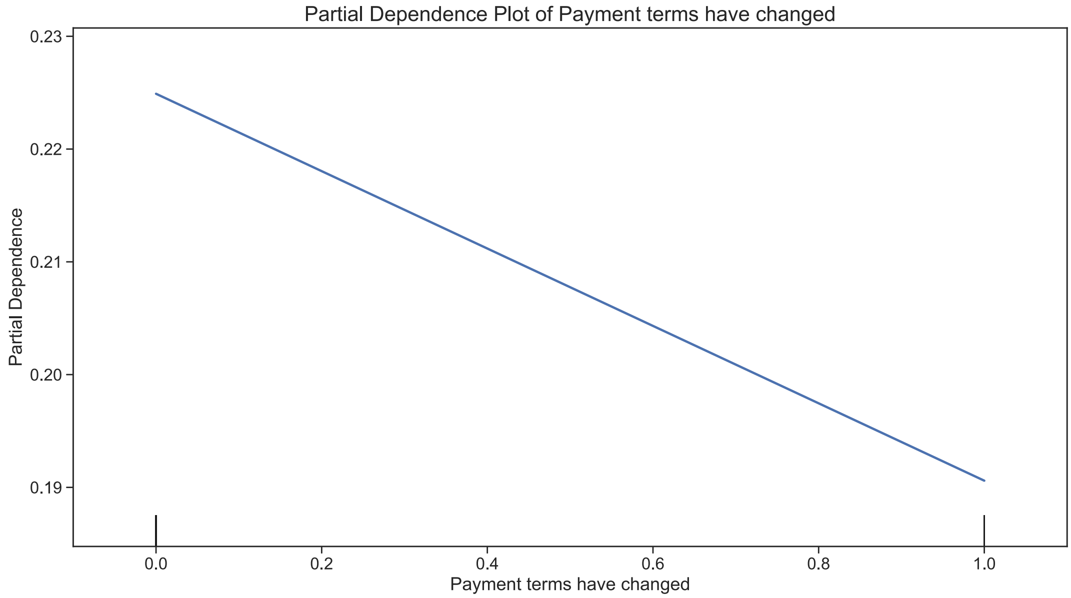

Fig. 10. PDP of payment terms have changed.

deal on standard payment terms directly, such as longer payment terms instead of the standard ones for flexible cash flow. If the customer will fully and carefully communicate with these suppliers during the transaction to determine reasonable payment terms, credit conditions, etc., and reach close cooperation with the customer, the supply chain capital flow can be greatly managed, thus improving the overall stability of operations and reducing financial risks. (See Fig. 11.)

For "Charges have been made for remaining on supplier list", it is shown that charges for retaining supplier status increase the financial risk, which is also in line with the reality. For example, the Chinese government released the regulations to protect the payment of SMEs in 2020, which clearly states no charges for so-called qualification maintenance. Our results enhance the understanding that such charges to suppliers can be seen as exploitation of suppliers by customers, which is not conducive to a healthy and stable supply chain.

6. Management inspiration

This study provides some valuable operational and practical insights for managers.

First, the results of the PDP analysis suggest that managers should carefully analyze the type of customer payment behavior. For example, we can assume that customers with shorter average payment times feature a lower risk of default. Based on our model, managers can predict their customers' risk types with their payment behavior, enabling managers to prioritize the right customers and payment terms in their transactions and thus focus more on the most value-added customers, which will benefit their business in the long run.

Second, managers should be more collaborative and work more closely with their clients. Our research shows that customers who have modified payment terms in a transaction have less delayed payment, meaning that the negotiation of payment terms in a transaction can be a means for suppliers to control financial risk [27], and although such

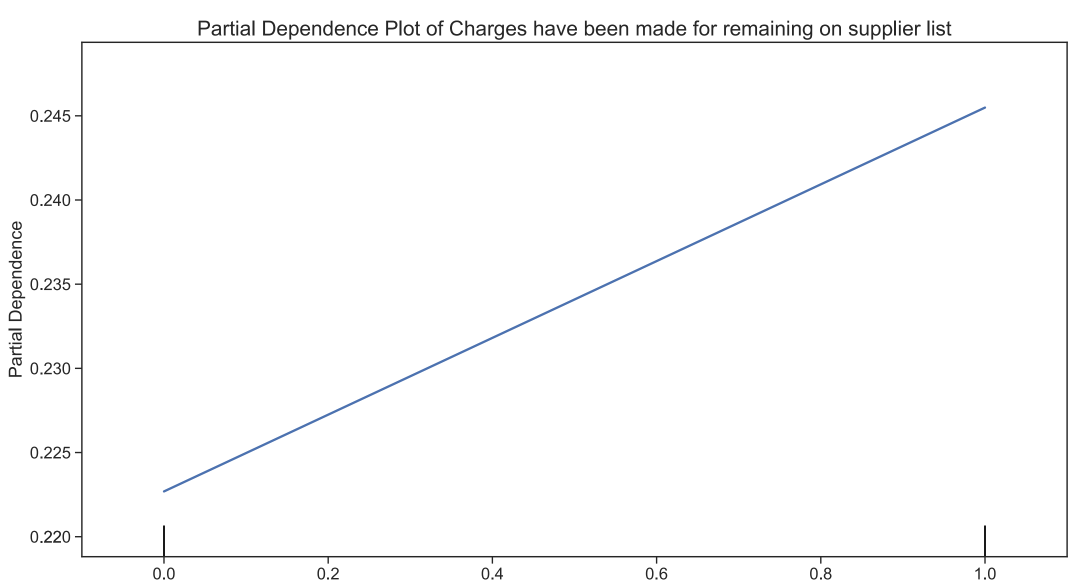

Fig. 11. PDP of charges have been made for remaining on supplier list.

modification can be considered as a large customer exercising power for profit, managers can negotiate with them for better payment terms, for example, to provide SCF.

Our results also show that when the customer provides SCF, the supplier can recover the payment faster through the third-party factoring agent, which may reduce the increase of financial burden, and the payment pressure of the customer is also transferred to the agent at this time. The customer can also relieve the payment pressure by asking the agent to set a corresponding grace period, simultaneously improving the stability of supply chain. Also, in addition to finding and developing large customers, managers of suppliers need to increase their market power as much as possible to promote more cooperative customers. Therefore, managers should strengthen the cooperation with customers in transactions, enhance communication before providing trade credit, and determine the most reasonable payment terms for both parties in the negotiation with customers.

Finally, this paper sheds light on managers' use of AI and data mining. Titman (2021) [69] argues that the SCF risk depends on the information available throughout the supply chain. Based on information about the historical transaction behavior of customers, our study analyzes the SCF risk arising from varying degrees of delayed payments by customers, which is one of the few studies that focus on this issue and adopts a data-driven approach in SCF. Firstly, we clarify the sources of the data used and provide ideas for managers to collect data, and secondly, our modeling process can provide effective guidance for managers to construct predictive models and provide useful insights from big data analysis, which supports managers' decision making in selecting and evaluating customers in SCF. In conclusion, this study applies artificial intelligence (AI) and big data in finance, which provides a feasible solution for managers to predict the financial risk of customers in practice. The powerful data mining also improves companies' competitivity, especially in the era of Industry 4.0 (e.g., [14]).

7. Conclusion

In this paper, we develop models to predict SCF financial risk triggered by different degrees of delayed payments from customers. First, we retrieve data on customers' transaction behavior, and then we construct different machine learning models (single and hybrid ones) for comparative analysis. Since a minority of samples are high-risk ones and a majority of low-risk, we use the Random oversampling algorithm to deal with unbalanced data distribution. The results show that the XGBoost outperforms the others in terms of ROC, AUC and F1-Score values. We also interpret the model using feature importance and PDP methods. The results reveal that our model can effectively predict the financial risk in SCF and provide application-worthy inspiration for managers to adequately evaluate their customers.

Though this paper is among the few studies that discuss the issue of financial risk prediction in SCF, there are a few limitations. Our study is limited to data of large buyers in the UK from 2017 to 2020, which may limit the generalizability to other countries. Future research will contribute to SCF financial risk in other countries (e.g., China, the US) and retrieve more recent data. Furthermore, the transmission effect of SCF financial risk on credit risk needs more analysis. Third, given the relatively monotonous data structure (e.g., as a "Supply-chain financing offered" binary variable), the hybrid models may fail to realize their full potential. In the future, we can collect more types of data, such as customer financial information (accounts receivable turnover, profitability, operating income, etc.), and combine more methods (e.g., deep learning) to improve the model performance of SCF financial risk prediction.

Credit author statement

Zelong Yi initiates the research idea and develops the estimation models.

Zhuomin Liang collects data, predicts the financial risk, conducts robustness check and writes the manuscript.

Tongtong Xie analyzes the research background and writes the manuscript.

Fan Li develops the theoretical framework, illustrates the practical application and revises the manuscript.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could be construed as a potential conflict of interest.

Data availability

Please visit the UK government website (https://check-payment-practices.service.gov.uk/export) for further information.

Acknowledgements

This work was partially supported by the National Natural Science Foundation of China (No. 72031004, 72171154, 72222019). It was also supported by Key Project of Key Research Institute for Humanities and Social Sciences of the Chinese Ministry of Education (No. 22JJD790054) and The Project of Longgang Innovation Research Institute of Shenzhen University (No. SZJR006).

Appendix A. Parameters for XGBoost

Parameter	Description	Value
n_estimators	Number of generated maximum trees	900
learning_rate	Step size of each iteration	0.01
max_depth	Maximum depth of the tree	10
min_child_weight	Minimum sum of instance weight in leaf nodes	1
subsample	Proportion of random samples per tree	0.85

References

[1] E. Adida, H. Mamani, S. Nassiri, Bundled payment vs. fee-for-service: impact of payment scheme on performance, Manag. Sci. 63 (5) (2016) 1–19, https://doi.org/10.1287/mnsc.2016.2445.

[2] F. Aqlan, S.S. Lam, A fuzzy-based integrated framework for supply chain risk assessment, Int. J. Prod. Econ. 161 (2015) 54–63, https://doi.org/10.1016/j.ijpe.2014.11.013.

[3] M. Bai, Y. Qin, Short-sales constraints and liquidity change: cross-sectional evidence from the Hong Kong market, Pac. Basin Financ. J. 26 (2014) 98–122, https://doi.org/10.1016/j.pacfin.2013.11.004.

[4] M. Bahrami, B. Bozkaya, S. Balcisoy, Using behavioral analytics to predict customer invoice payment, Big Data 8 (1) (2020) 25–37, https://doi.org/10.1089/big.2018.0116.
[5] G.E.A.P.A. Batista, R.C. Prati, M.C. Monard, A study of the behavior of several methods for balancing machine learning training data, ACM SIGKDD Explor. Newsl. 6 (1) (2004) 20–29, https://doi.org/10.1145/1007730.1007735.
[6] C. Blome, T. Schoenherr, Supply chain risk management in financial crises—a multiple case-study approach, Int. J. Prod. Econ. 134 (1) (2011) 43–57, https://doi.org/10.1016/j.ijpe.2011.01.002.
[7] B.-B. Cao, T.-H. You, C.X.J. Ou, H. Zhu, C.-Y. Liu, Optimizing payment schemes in a decentralized supply chain: a Stackelberg game with quality investment and bank credit, Comput. Ind. Eng. 168 (2022), 108077, https://doi.org/10.1016/j.cie.2022.108077.
[8] C. Casanova, China Corporate Payment Survey 2018: Payment Delays Increase despite Rapid and Robust Growth, Paris, France. Retrieved from, www.coface.com/Economic-Studies-and-Country-Risks, 2018.
[9] C.-T. Chang, L.-Y. Ouyang, J.-T. Teng, K.-K. Lai, L.E. Cárdenas-Barrón, Manufacturer's pricing and lot-sizing decisions for perishable goods under various payment terms by a discounted cash flow analysis, Int. J. Prod. Econ. 218 (2019) 83–95, https://doi.org/10.1016/j.ijpe.2019.04.039.
[10] N.S. Cheng, R. Pike, The trade credit decision: evidence of UK firms, Manag. Decis. Econ. 24 (6–7) (2003) 419–438, https://doi.org/10.1002/mde.1049.
[11] X. Chen, X. Wang, D.D. Wu, Credit risk measurement and early warning of SMEs: an empirical study of listed SMEs in China, Decis. Support. Syst. 49 (3) (2010) 301–310, https://doi.org/10.1016/j.dss.2010.03.005.
[12] F. Ciampi, Corporate governance characteristics and default prediction modeling for small enterprises. An empirical analysis of Italian firms, J. Bus. Res. 68 (5) (2015) 1012–1025, https://doi.org/10.1016/j.jbusres.2014.10.003.
[13] S. Deng, K. Fu, J. Xu, K. Zhu, The supply chain effects of trade credit under uncertain demands, Omega 98 (102113) (2021) 1–22, https://doi.org/10.1016/j.omega.2019.102113.
[14] B. Dorneanu, S. Zhang, H. Ruan, M. Heshmat, R. Chen, V.S. Verykios, H. Arellano-Garcia, Big data and machine learning: a roadmap towards smart plants, Front. Eng. Manage. 9 (4) (2022) 623–639, https://doi.org/10.1007/s42524-022-0218-0.
[15] R. Elshawi, M.H. Al-Mallah, S. Sakr, On the interpretability of machine learning-based model for predicting hypertension, BMC Med. Inform. Decis. Mak. 19 (1) (2019) 146, https://doi.org/10.1186/s12911-019-0874-0.
[16] J.H. Friedman, Greedy function approximation: a gradient boosting machine, Ann. Stat. 29 (5) (2001), https://doi.org/10.1214/aos/1013203451.
[17] J. Gbaf, The real cost of late payments for SME's in the UK, Global Banking & Finance Review | Global Banking, 2021. https://www.globalbankingandfinance.com/the-real-cost-of-late-payments-for-smes-in-the-uk/.
[18] A. Ghadge, S.K. Jena, S. Kamble, D. Misra, M.K. Tiwari, Impact of financial risk on supply chains: a manufacturer-supplier relational perspective, Int. J. Prod. Res. 59 (23) (2021) 7090–7105, https://doi.org/10.1080/00207543.2020.1834638.
[19] X. Guan, G. Li, Z. Yin, The implication of time-based payment contract in the decentralized assembly system, Ann. Oper. Res. 240 (2016) 641–659, https://doi.org/10.1007/s10479-014-1579-5.
[20] M. Gupta, S. Tiwari, C.K. Jaggi, Retailer's ordering policies for time-varying deteriorating items with partial backlogging and permissible delay in payments in a two-warehouse environment, Ann. Oper. Res. 295 (1) (2020) 139–161, https://doi.org/10.1007/s10479-020-03673-x.
[21] X. He, J. Pan, O. Jin, T. Xu, B. Liu, T. Xu, Y. Shi, A. Atallah, R. Herbrich, S. Bowers, J.Q. Candela, Practical lessons from predicting clicks on ads at Facebook, in: Proceedings of the Eighth International Workshop on Data Mining for Online Advertising, 2014, https://doi.org/10.1145/2648584.2648589.
[22] Hilton-Baird Collection Services, Hilton-Baird's late payment survey (pp. 1–29), Retrieved from, www.hiltonbaird.co.uk/cs, 2015.
[23] C. Howorth, B. Reber, Habitual late payment of trade credit: an empirical examination of UK small firms, Manag. Decis. Econ. 24 (6–7) (2003) 471–482, https://doi.org/10.1002/mde.1128.
[24] D. Ivanov, A. Dolgui, Low-certainty-need (LCN) supply chains: a new perspective in managing disruption risks and resilience, Int. J. Prod. Res. 57 (15–16) (2019) 5119–5136, https://doi.org/10.1080/00207543.2018.1521025.
[25] Z. Kalsyte, A. Verikas, A novel approach to exploring company's financial soundness: Investor's perspective, Expert Syst. Appl. 40 (13) (2013) 5085–5092, https://doi.org/10.1016/j.eswa.2013.03.031.
[27] L. Klapper, L. Laeven, R. Rajan, Trade credit contracts, Rev. Financ. Stud. 25 (3) (2012) 838–867, https://doi.org/10.1093/rfs/hhr122.
[28] G. Kou, Y. Peng, G. Wang, Evaluation of clustering algorithms for financial risk analysis using MCDM methods, Inf. Sci. 275 (2014) 1–12, https://doi.org/10.1016/j.ins.2014.02.137.
[29] G. Kou, Y. Xu, Y. Peng, F. Shen, Y. Chen, K. Chang, S. Kou, Bankruptcy prediction for SMEs using transactional data and two-stage multiobjective feature selection, Decis. Support. Syst. 140 (2021), 113429, https://doi.org/10.1016/j.dss.2020.113429.
[30] H.-H. Lee, J. Zhou, J. Wang, Trade credit financing under competition and its impact on firm performance in supply chains, Manuf. Serv. Oper. Manag. 20 (1) (2018) 36–52, https://doi.org/10.1287/msom.2017.0640.
[31] R. Li, Y. Liu, J.-T. Teng, Y.-C. Tsao, Optimal pricing, lot-sizing and backordering decisions when a seller demands an advance-cash-credit payment scheme, Eur. J. Oper. Res. 278 (1) (2019) 283–295, https://doi.org/10.1016/j.ejor.2019.04.033.
[32] R. Li, H.-L. Yang, Y. Shi, J.-T. Teng, K.-K. Lai, EOQ-based pricing and customer credit decisions under general supplier payments, Eur. J. Oper. Res. 289 (2) (2021) 652–665, https://doi.org/10.1016/j.ejor.2020.07.035.
[33] S. Li, Y. Pan, Application of neural network in evaluation of credit evaluation by SMEs in P2P, Product. Res. 5 (2019) 14–22, https://doi.org/10.3969/j.issn.1004-2768.2019.05.003.
[34] T. Li, G. Kou, Y. Peng, P.S. Yu, An integrated cluster detection, optimization, and interpretation approach for financial data, IEEE Trans. Cybernet. 52 (12) (2022) 13848–13861, https://doi.org/10.1109/TCYB.2021.3109066.
[35] B. Liu, T. Ju, H.K. Chan, The diverse impact of heterogeneous customer characteristics on supply chain finance: empirical evidence from Chinese factoring, Int. J. Prod. Econ. 243 (108321) (2022) 1–13, https://doi.org/10.1016/j.ijpe.2021.108321.
[36] B. Liu, Y. Wang, Y. Shou, Trade credit in emerging economies: an interorganizational power perspective, Ind. Manag. Data Syst. 120 (4) (2020) 768–783, https://doi.org/10.1108/IMDS-05-2019-0292.
[37] F. Liu, M. Fang, K. Park, X. Chen, Supply chain finance, performance and risk: how do SMEs adjust their buyer-supplier relationship for competitiveness? J. Compet. 13 (4) (2021) 78–95, https://doi.org/10.7441/joc.2021.04.05.
[38] Y. Liu, C. Liu, Research on product pricing and financing decision of green supply chain—analysis based on capital constraint, Price 10 (2020) 114–118, https://doi.org/10.19851/j.cnki.CN11-1010/F.2020.10.447.
[39] P. Liu, Z. Zeng, Medical supply chain risk evaluation based on GBDT, Friends Account. 24–31 (2021), https://doi.org/10.3969/j.issn.1004-5937.2021.16.005.
[40] Z. Liu, J.M. Cruz, Supply chain networks with corporate financial risks and trade credits under economic uncertainty, Int. J. Prod. Econ. 137 (1) (2012) 55–67, https://doi.org/10.1016/j.ijpe.2012.01.012.
[41] A. Moretto, F. Caniato, Can supply chain finance help mitigate the financial disruption brought by Covid-19? J. Purch. Supply Manag. 27 (4) (2021), 100713 https://doi.org/10.1016/j.pursup.2021.100713.
[42] A. Moretto, L. Grassi, F. Caniato, M. Giorgino, S. Ronchi, Supply chain finance: from traditional to supply chain credit rating, J. Purch. Supply Manag. 25 (2) (2019) 197–217, https://doi.org/10.1016/j.pursup.2018.06.004.
[43] J. Owen, Majority of large firms fail to pay suppliers promptly, Supply Manage (2019). https://www.cips.org/supply-management/news/2019/january/majority-of-large-firms-fail-to-pay-suppliers-promptly/.
[44] S.Y. Paul, S.S. Devi, C.G. Teh, Impact of late payment on Firms' profitability: empirical evidence from Malaysia, Pac. Basin Financ. J. 20 (5) (2012) 777–792, https://doi.org/10.1016/j.pacfin.2012.03.004.
[45] H. Qian, B. Wang, M. Yuan, S. Gao, Y. Song, Financial distress prediction using a corrected feature selection measure and gradient boosted decision tree, Expert Syst. Appl. 190 (2022), 116202, https://doi.org/10.1016/j.eswa.2021.116202.
[46] R. Qin, The construction of corporate financial management risk model based on XGBoost algorithm, J. Math. 2022 (2022) 1–8, https://doi.org/10.1155/2022/2043369.
[47] B. Sang, Application of genetic algorithm and BP neural network in supply chain finance under information sharing, J. Comput. Appl. Math. 384 (113170) (2021) 1–11, https://doi.org/10.1016/j.cam.2020.113170.
[48] D. Seifert, R.W. Seifert, M. Protopappa-Sieke, A review of trade credit literature: opportunities for research in operations, Eur. J. Oper. Res. 231 (2) (2013) 245–256, https://doi.org/10.1016/j.ejor.2013.03.016.
[49] A. Serrano, R. Oliva, S. Kraiselburd, Risk propagation through payment distortion in supply chains, J. Oper. Manag. 58–59 (1) (2018) 1–14, https://doi.org/10.1016/j.jom.2018.01.003.
[50] A.A. Taleizadeh, N. Rabiei, M. Noori-Daryan, Coordination of a two-echelon supply chain in presence of market segmentation, credit payment, and quantity discount policies, Int. Trans. Oper. Res. 26 (2018) 1576–1605, https://doi.org/10.1111/itor.12618.
[51] A. Tananbaum, Beyond the Zombies: Financing SMEs to Strengthen Trade, Retrieved from, https://risnews.com/beyond-zombies-financing-smes-strengthen-trade, 2011.
[52] R. Tangsucheeva, V. Prabhu, Stochastic financial analytics for cash flow forecasting, Int. J. Prod. Econ. 158 (2014) 65–76, https://doi.org/10.1016/j.ijpe.2014.07.019.
[53] K. Tian, X. Zhuang, W. Zhao, Credit risk evaluation of SMEs under the mode of supply chain finance — analysis based on sample data of automobile manufacturing industry, J. Indus. Technol. Econ. 331 (5) (2021) 15–20, https://doi.org/10.3969/j.issn.1004-910X.2021.05.002.
[54] K. van der Vliet, M.J. Reindorp, J.C. Fransoo, The price of reverse factoring: financing rates vs. payment delays, Eur. J. Oper. Res. 242 (3) (2015) 842–853, https://doi.org/10.1016/j.ejor.2014.10.052.
[55] C. Wang, D. Han, Q. Liu, S. Luo, A deep learning approach for credit scoring of peer-to-peer lending using attention mechanism LSTM, IEEE Access 7 (2019) 2161–2168, https://doi.org/10.1109/ACCESS.2018.2887138.
[56] D. Wang, L. Li, D. Zhao, Corporate finance risk prediction based on LightGBM, Inf. Sci. 602 (2022) 259–268, https://doi.org/10.1016/j.ins.2022.04.058.
[57] W. Wang, L. Feng, Y. Li, F. Xu, Q. Deng, Role of financial leasing in a capital-constrained service supply chain, Transp. Res. Part E 143 (102097) (2020) 1–21, https://doi.org/10.1016/j.tre.2020.102097.
[58] Y. Wang, Research on supply chain financial risk assessment based on Blockchain and fuzzy neural networks, Wirel. Commun. Mob. Comput. 7 (2021) 1–8, https://doi.org/10.1155/2021/5565980.
[59] L. Wu, W. Lu, J. Xu, Blockchain-based smart contract for smart payment in construction: a focus on the payment freezing and disbursement cycle, Front. Eng. Manage. 9 (2) (2022) 177–195, https://doi.org/10.1007/s42524-021-0184-y.
[60] Y. Wu, Y. Wang, X. Xu, X. Chen, Collect payment early, late, or through a third party's reverse factoring in a supply chain, Int. J. Prod. Econ. 218 (2019) 245–259, https://doi.org/10.1016/j.ijpe.2019.04.040.

[61] D.A. Wuttke, E.D. Rosenzweig, H.S. Heese, An empirical analysis of supply chain finance adoption, J. Oper. Manag. 65 (3) (2019) 242–261, https://doi.org/10.1002/joom.1023.
[62] Q. Yang, Y. Wang, Y. Ren, Research on financial risk management model of internet supply chain based on data science, Cogn. Syst. Res. 56 (2019) 50–55, https://doi.org/10.1016/j.cogsys.2019.02.001.
[63] C. Yin, C. Jiang, H.K. Jain, Z. Wang, Evaluating the credit risk of SMEs using legal judgments, Decis. Support. Syst. 136 (113364) (2020) 1–13, https://doi.org/10.1016/j.dss.2020.113364.
[64] H. Zhang, Y. Shi, X. Yang, R. Zhou, A firefly algorithm modified support vector machine for the credit risk assessment of supply chain finance, Res. Int. Bus. Financ. 58 (101482) (2021) 1–21, https://doi.org/10.1016/j.ribaf.2021.101482.
[65] W. Zhang, S. Yan, J. Li, X. Tian, T. Yoshida, Credit risk prediction of SMEs in supply chain finance by fusing demographic and behavioral data, Transp. Res. Part E 158 (102611) (2022) 1–25, https://doi.org/10.1016/j.tre.2022.102611.
[66] X. Zheng, Z. Zheng, X. Xu, Research review on risk Management of Supply Chain: risk evaluation, Technol. Econ. 32 (7) (2013) 123–130, https://doi.org/10.3969/j.issn.1002-980X.2013.07.021.
[67] Y. Zhu, C. Xie, G.-J. Wang, X.-G. Yan, Comparison of individual, ensemble and integrated ensemble machine learning methods to predict China's SME credit risk in supply chain finance, Neural Comput. & Applic. 28 (2017) 41–50, https://doi.org/10.1007/s00521-016-2304-x.
[68] Y. Zhu, L. Zhou, C. Xie, G.-J. Wang, T.V. Nguyen, Forecasting SMEs' credit risk in supply chain finance with an enhanced hybrid ensemble machine learning approach, Int. J. Prod. Econ. 211 (2019) 22–33, https://doi.org/10.1016/j.ijpe.2019.01.032.
[69] S. Titman, Risk Transmission Across Supply Chains, Prod Oper Manag 30 (2021) 4579–4587. https://doi.org/10.1111/poms.13542.

Zelong Yi is an Associate Professor in College of Economics at Shenzhen University. He received his Ph.D. in Operations Management from the Hong Kong University of Science and Technology. His research interests include supply chain management and supply chain finance. His papers have been published in *Manufacturing & Service Operations Management, Production and Operations Management, Decision Sciences, Transportation Research Part E,* etc.

Zhuomin Liang is a PhD student in College of Economics at Shenzhen University. His research interest is supply chain finance.

Tongtong Xie is a Master student in School of Foreign Languages, Shenzhen University. His research interest is international relationship and international economics.

Fan Li is a professor in China Center for Special Economic Zone Research at Shenzhen University. She received her Ph.D. in Economics from University of Florida. Her research interests include the economics of information systems, electronic commerce, and regulatory economics. Her papers have been published in *Information & Management, Decision Sciences, Omega: The International Journal of Management Science, Electronic Commerce Research and Applications,* etc.