

Liu, Chang; Shi, Haoming; Cai, Yujun; Shen, Shu; Lin, Dongtao

## Article

# A new pricing approach for SME loans issued by commercial banks based on credit score mapping and Archimedean Copula simulation

Journal of Business Economics and Management (JBEM)

## Provided in Cooperation with:

Vilnius Gediminas Technical University

*Suggested Citation:* Liu, Chang; Shi, Haoming; Cai, Yujun; Shen, Shu; Lin, Dongtao (2019) : A new pricing approach for SME loans issued by commercial banks based on credit score mapping and Archimedean Copula simulation, Journal of Business Economics and Management (JBEM), ISSN 2029-4433, Vilnius Gediminas Technical University, Vilnius, Vol. 20, Iss. 4, pp. 618-632, <https://doi.org/10.3846/jbem.2019.9854>

This Version is available at:

<https://hdl.handle.net/10419/317344>

## Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

## Terms of use:

*Documents in EconStor may be saved and copied for your personal and scholarly purposes.*

*You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.*

*If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.*



<https://creativecommons.org/licenses/by/4.0/>

## A NEW PRICING APPROACH FOR SME LOANS ISSUED BY COMMERCIAL BANKS BASED ON CREDIT SCORE MAPPING AND ARCHIMEDEAN COPULA SIMULATION

Chang LIU<sup>1\*</sup>, Haoming SHI<sup>1</sup>, Yujun CAI<sup>1</sup>, Shu SHEN<sup>1</sup>, Dongtao LIN<sup>2</sup>

<sup>1</sup> Southwestern University of Finance and Economics, Chengdu, China

<sup>2</sup> Sichuan University, Chengdu, China

Received 29 June 2018; accepted 27 March 2019

**Abstract.** The traditional loans pricing methods are usually based on risk measures of individual loan's characteristics without considering the correlation between the defaults of different loans and the contribution of individual loans to the entire loan portfolio. In this study, using account-level loans data of 2010–2016 abstracted from 2 databases kindly provided by a Chinese commercial bank, the authors choose Archimedean Copula to fit the default relationship between loans, combined with the loss distribution function constructed to measure the economic capital of the loan portfolio, to propose a loan pricing method that is more suitable for measuring the unique risk characteristic of SMEs loans. Empirical evidence shows that compared with the traditional loan pricing model, this new proposed one, requiring lower loan interest rates from customers with higher credit rating, while higher loan interest rates from customers with lower credit rating, could thus be able to provide higher risk-adjusted returns, higher economic capital adequacy ratios, and ultimately stronger banks' capabilities to tolerate risk events. Although there might still be some issues and limitations in the study, the method proposed in this study could be of interest not only to the banks' management, but also to banking regulators as well.

**Keywords:** loan pricing, economic capital, Archimedean Copula, SMEs loans, Internal Rating Model, RAROC, capital adequacy, risk tolerance.

**JEL Classification:** C63, G21, G12.

### Introduction

The approach for pricing loans of commercial banks basically adopts the way of adjusting certain premium/discount points based on the benchmark interest rates, with premiums/discounts often reflects the credit ratings of the firms applying the loans. On the other hand, although the client rating system has a rating module specific to small and medium-sized enterprises (SMEs), the specific rating indicators tend to follow the metrics of large-scale

---

\*Corresponding author. E-mail: [changliu0826@qq.com](mailto:changliu0826@qq.com)

enterprises (e.g., the relevant financial indicators in the reports) only with a few adjustments in the weight of indicators, which means no specific module is developed for the actual situation of SMEs. Consequently, there is no essential difference between SMEs' credit ratings and those of large-scale enterprises, which often lead to catastrophic such as a large simultaneous delinquencies of banks' SME loans since the correlation between SMEs loans are extremely high and the risk characteristics of SMEs loans are quite different from those of large enterprises. Using pricing methods developed for large firms' loans would definitely miss some essential risk profiles if the methods are applied to SME loans.

On the other hand, the Economic Capital has become one of the most important parameters of risk management and resource allocation in commercial banks, playing an increasingly crucial role in loan decision-making, pricing, and risk management process. The measure of economic capital itself takes into account the issue of the correlation between the default assets, while the traditional measure does not. Therefore, in this paper, a deep analysis is conducted on loan decision-making and pricing of SMEs loans, based on the measure of economic capital. The idea of this study can be summarized as follows: First, based on the analysis of SMEs loans' risk characteristics, this paper constructs a scoring system suitable for the credit risk measurement of SME applicants for loans of one Chinese commercial bank. Second, the scores are calibrated and mapped to the default probabilities table released by Basel Accord III to build an internal rating system that complies with the risk profiles of SME loans' applicants. The mapping system classifies the SME loans portfolio into 19 different credit risk levels and the Archimedean Copulas are used to construct the loss distribution at the loans portfolio level to determine the unanticipated losses and their corresponding economic capital required. Finally, according to the marginal risk contribution of each incremental individual loan to the entire portfolio, economic capital is allocated, followed by a RAROC model constructed to make loan decision and loan pricing.

This paper consists of five parts. The Introduction section presents the basic ground for the research, followed by detail literature reviews in section 1. While section 2 presents the models, the empirical results are discussed by section 3. The final section concludes the paper.

## **1. Literature review**

### **1.1. Economic capital**

Aas and Puccetti (2014) assessed the model risk related to the choice of a specific copula function and computed bounds on the total economic capital for the aggregate loss distribution. Economic capital was again studied by Alie, and Jean (2016)'s introduced Business Risk Economic Capital Model and Earnings Volatility Model to assess economic capital frameworks from the global and local industry viewpoint. In the same year, Scandizzo (2016) extended their framework by caught up with the ICAAP, which prescribed that institutions should have a process for assessing the institutes' overall capital adequacy in relation to their risk profile and risk preferences. Moreover, Gubareva and Borges (2018) addressed the economic capital estimation based on an integrated derivative-based treatment of Internal Rating Models applied to the EM sovereign debt portfolios.

## **1.2. Loan pricing**

Wang and Libaers (2015) discussed loan pricing from the cost-price model and customer-profit-analysis model, reaching the conclusion that modern loan pricing theory based on RAROC was more suitable for SMEs. Petropoulos, Chatzis, and Xanthopoulos (2016) presented a novel corporate credit rating system based on Student's *t* hidden Markov models (SHMMs). Allen, Chapman, Echenique, and Shum (2016) also came up with an improved estimated approach for risk modelling processes of loan pricing, stress testing and capital allocation, which could obtain more robust credit loss estimates in the risk modeling. Besides, Gorelaya (2016) examined the impact of selected loan characteristics (terms, purpose, amount and borrowers' creditworthiness) on the loan price, applying regression analysis based on a lender provided dataset covering the period from the beginning of 2007 until the mid-2015. Hauptmann (2017) found that borrowers with strong sustainability performance paid lower loan prices than borrowers with weak sustainability performance. Additionally, Walke, Fullerton, and Tokle (2018) researched on the risk-based loan pricing consequences for credit unions, suggesting potential overcharges for low-risk borrowers and lower charges for high-risk borrowers, which is also confirmed by this study.

## **1.3. Copula and its application in risk management**

Ghosh, and Ray (2016) constructed risk models by implementing a series of Copulas, such as bivariate normal copula, AMH (Ali-Mikhail-Haq), and Marshall-Olkin copula. Khan et al. (2016) also created a Copula Bayesian Network (CBN) model for dynamic probability updating which can be integrated with the copula-based multivariate loss modeling to provide dynamics structure in risk modeling. And Pfeifer et al. (2016) constructed new multivariate copulas on the basis of a approach allowing for asymmetry in tail-dependence. Apart from these, Oh and Patton (2017) used the expounded copula models to study a collection of daily credit default swap (CDS), finding the joint probability of distress substantially higher now than before 2008. Furthermore, Abdelkafi, Ghorbel, and Khoufi (2018) considered a mixture of the dynamic factor copulas of Duffee in risk modeling. Karmakar and Paul (2018) forecast intraday portfolio risk by using CGARCH-EVT-Copula model. On the other hand, Archimedean Copula described by Kim et al. (2017) would be adapted in this study to generate the probability loss distribution of the loan portfolio.

## **1.4. Comments**

The method suggested in this study is related to the previous literatures in the following aspects. This paper first selects the factors that can identify the risk profiles of SMEs to construct the scoring model as proposed by Liu, Guo, Mo, and Zhang (2012). Then, the internal rating model of Liu, Zhang, and Mo (2013) has been revised in complying with the requirements of the Basel Accord III to map the estimated scoring to risk categories divided into 19 classes according to the output of the internal rating model. The default correlations between different credit ratings are simulated and measured by the Archimedean Copula function proposed by Karmakar and Paul (2018) and Tang (2015). As a result, the empirical

result presented in this study could provide a more applicable framework than it in Aas and Puccetti (2014).

## **2. Research methodology**

This paper builds a scoring system firstly, making it more suitable to measure the credit risk of SME loans. Secondly, used Basel Accord III, an internal rating system is constructed, which categorizes the SME loans portfolio into nineteen different levels. Eventually, based on the established loss distribution and the marginal risk contribution, this study constructs a RAROC model for loan pricing.

### **2.1. Basel III and the proposed method**

The core content of Basel III related to this research are:

1. The new regulation introduced a “capital retention buffer” that must be greater than 2.5%, which is a core parameter in the pricing model;
2. The Basel Committee also recommended the establishment of a counter-cyclical capital buffer between 0% and 2.5%. The purpose of the regulation is to correct the procyclicality of Basel II, especially during the period of economic expansion;
3. The proposed regulations introduce a leverage ratio of 3.0% to stabilize the system: in any case, the ratio of capital to total assets must exceed this threshold.

The core contents of 2 and 3 are considered in our stress testing framework.

As a permitted by Basel III, this paper selects the Archimedean Copula method to fit the default relationship between loans, together with the loss distribution function constructed to measure the economic capital of the loan portfolio and the marginal contribution of a single loan considered when computing the marginal economic capital consumption and the price of that loan. Thus it effectively avoided two disadvantages of referring to Libor as the benchmark interest rate:

1. The benchmark interest rate has not yet established a market formation mechanism;
2. SME loan pricing mechanism has not been formed due to imperfections in related computer systems and historical data based on internal ratings.

### **2.2. Sample selection and descriptive statistics**

We obtained a data sample from a Chinese commercial bank from 2010 to 2016, including 2 databases, an application database and a database of loan repaying behavior tracking database. The first database contains applicants' data for loan application. The second one has the loan repaying data, including over 300 of defaults and 16773 transaction records. The data is pre-processed according to the following rules:

1. Any missing data field will let the entire record being removed;
2. Outliers, identified as the value passed 3 standard deviation (SD) spaces, would be set as the value equal to the mean value plus 3 SDs.

The following risk factors and their descriptive statistics are extract from the commercial banks' application database shown in Table 1. As indicated by the Table 1, there are total 16

risk factors of 5 dimensions, consisting of dimension of Enterprise production status, Basic ability of employees, External, Quantitative Financial Indicators, and Innovation. Table 1 presents each indicator's Mean, Variance, Skewness, and Kurtosis.

Table 1. Risk factors descriptive statistics (source: own processing)

Dimension	Variables	Mean	Variance	Skewness (Standardized)	Kurtosis (Standardized)
Enterprise production status	Process Capability Index (Mcap)	256	76.66	1.9	-3.3
Basic ability of employees	Age of Employee (ECT)	32	10	-4.6	-10.2
	Employee Compensation (Ework2)	70K	3.2K	-5.6	-4.63
	Number of Employees (Ework3)	58	20	0.36	-5.87
	Sales Growth Rate (Msi)	5%	2.43%	0.97	-4.65
External	Industry Prosperity Index (Mpbj)	351	36	-3.3	-6.76
Quantitative Financial Indicators	Assets and Liabilities (Fin1)	2,808K	568K	-2.6	-4.65
	Current ratio (Fin2)	56%	24.87%	-1.02	-5.65
	Operating Income Coverage of Debt (Fin3)	67.21%	52.54%	-2.01	-3.54
	Net Profit Margin (Fin4)	10.99%	14.65%	0.03	-5.65
	Total Asset Turnover (Flsc)	3	8.32	-1.32	-6.56
	Working Capital/Total Assets (Fin6)	30.65%	94.76%	-3.36	-6.78
	Net Asset Growth Rate (Fin7)	23.98%	24.65%	2.31	-6.23
	Accounts Receivable Turnover (Mpol)	6.9	6.15	1.24	-8.37
Innovation	R&D staff in the proportion of employees (MW1)	13	6.36	0.06	-5.87%
	R&D expenditure as a percentage of sales revenue (DIsl)	23.14%	29.93%	-3.32	-4.98

Table 1 shows that all variables are distributed with bigger (fatter) tails than normally distributed tails (Kurtosis < 0 with Standardized), which means that their values are distributed wider than normally distributed values, the phenomenon common found on financial variables, Baltussen, van Bekkum, and van der Grient (2018). On the other hand, a variable with a skewness value smaller than 0 is called tilted to right (having a fatter left tail than those normally distributed ones) which are the cases for ECT, Ework2, Mpbj, etc. On the contrary, skewness values, larger than 0, are deemed as titled to the left, indicating that they have thinner tails than those of normally distributed ones, which are the case for Mcap, Ework3, Msi, etc.

## 2.3. Modelling procedures

### 2.3.1. Logit regression model

The data of SMEs loan default and single lagged indicator ( $x$ ) are used to select single indicator as an independent variable  $x_i$ , loan default as dependent variable  $y$ , together to build a Logit regression model:

$$Y = f(x) = e^{(\alpha + \beta x_i)} / (1 + e^{(\alpha + \beta x_i)}), \quad i = 1, \dots, 39. \quad (1)$$

The Logit regression model is chosen over nonparametric model based on the following reasons: According to Giacomini and Rossi (2016), the risk metrics of non-parametric models may not be comparable across banks, and the stability of the calculations is at question. Therefore, there are problems in the cross-comparison of banking institutions. Due to the instability of the non-parametric model, it is not the primary recommended model of BASEL III. Since this research requires that banks can develop risk metrics for BASEL in practice, the Logit model, rather than non-parametric models, is selected to perform further analysis.

According to the  $P$  value t-test of and discriminatory power of the variable tested, the factors with statistical significance and discrimination power are selected. The discrimination power is often measured by Accurate Ratio (AR) value which is equal to area  $A / (A + B)$  as presented by Figure 1. The variables that meet one of the following two criteria will be included in the multivariate analysis.  $AR > 60\%$  or  $P < 0.05$  for the t-test.

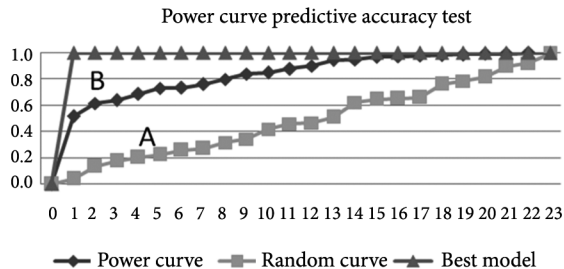


Figure 1. Accurate ratio (source: own processing)

In Figure 1, the middle curve visualizes the established logistic model's estimation accurate ratio, the top curve shows the results of the ideal prediction model, and the lower curve represents the result of the random prediction model. Large area of  $A$  indicates the superiority of the model in classifying accuracy.

### 2.3.2. Model validation

Through the correlation analysis and multivariate analysis, the best candidate models are chosen. The next step is to use out-of-sample  $k$ -fold Man-Whitney test each candidate's stability in prediction for out-of-sample data. The  $k$ -fold method is commonly known as cross-validation, in which the original data is separated into  $k$  parts randomly. The process of cross-validation is actually to repeat the experiment  $k$  times. Each experiment selects one

of the  $k$  parts data (in-sample) to build the multivariate candidate model and the remaining  $k - 1$  parts data (out-of-sample) are used to test the model's out-of-sample performance. The best model with stable AR both for the in-sample data and the  $k - 1$  pieces of out-of-sample data is identified as the final model for further analysis.

### 2.3.3. Scores saturation transformation and weighting

The final candidate multivariate model produces a score with value distributed between 0 and 1. The score is a theoretical number and not a real default probability. It requires mapping and calibration before the score could be used in loan decision and loan pricing. Since the scores estimated from the model denoted as  $Score(x)$  has extreme values and outliers, they need to be further processed by the saturation transformation and weighting. The saturation transformation function is given as:

$$score(x) = \phi(x) = 100 \times e^{\frac{x-\mu}{\sigma}} / (e^{\frac{x-\mu}{\sigma}} + 1), \quad (2)$$

where  $x$  denotes the direct model output,  $\mu$  the median of  $x$ ,  $\sigma$  the variance of  $x$ .

### 2.3.4. Model calibration and mapping

The process of calibration and mapping refers to the procedure in which the SMEs scoring got mapped to the true default rate to obtain the customer rating grade by the primary mapping scale published by Basel Accord III. The probability of default and the rating result must satisfy the following conditions:

1. The average PD value obtained after calibration is 3.09% of the long-term default rate (this is required by the Basel Accord III);
2. The average probability of default at each level of the main scale after calibration should be consistent with the actual probability;
3. The distribution of risk levels should be consistent with the actual situation of SMEs and risk management requirements;
4. The proportion of debtors under each risk level should not exceed a certain percentage.

The calibration model proposed by Cao Jin (2011), who thought that the calibration method is actually a nonlinear optimization problem with multiple goal constraints as showed in Eq. (3):

$$\left\{ \begin{array}{l} PD = f(score) \\ \text{subject to:} \quad (1) \ E(PD) = 3.09\% \\ \quad \quad \quad (2) \ Min_i \leq Spd_i \leq Max_i, \\ \quad \quad \quad (3) \ Peri \leq point_i \\ \quad \quad \quad (4) \ i = Aaa, \dots, Ccc \end{array} \right. \quad (3)$$

where  $Max_i$  denotes the upper limit of the  $i$ th main scale level default rate,  $Mini$  the lower limit,  $Spd_i$  the actual sample default rate of the  $i$ th level,  $Per_i$  the proportion of debtors under the  $i$ th main scale level, and  $Point_i$  as a set value, which is determined by the rating distribution of SMEs. The following steps are used to solve the nonlinear optimization in Eq. (3): First, group the applicants according to their rating scores. Based on the historical



data of the debtor’s rating and the repayment behaviors for each grouped applicants granted the loans, the ratings and the defaults probabilities could be lined up for each credit groups, or categories  $i = Aaa, \dots, Ccc$ . Second, calibration mapping curve and a fitting calibration function could be used to sort all debtors’ scores from low to high (the lower the score, the lower the risk and vice versa) and put them in X-axis from *Aaa*, ..., *Ccc*. The probability of default (PD) is set as the Y-axis and the rating level is descending, from *Maxi* to *Mini* (*Aaa* is the highest level and *Ccc* is the lowest level), then the X-axis scoring results and the Y-axis default rate has a mapping relationship. Continue adjusting the distribution and calculating the corresponding default rate, so that the constraints are satisfied, and the model calibration mapping curve is obtained, and ultimately, the functional relationship between the model output score and default probability is established through the fitting mapping curve  $f(\cdot)$ , as our final default rate measurement model. Finally, Binomial Test and chi-square test are used to verify the average probability of default at each level of the main scale, so that the main scale can be verified by the actual data of the banks. After the calibration and mapping, the ratings scale and the default rate model  $PD = f(score(X))$  are finally determined, and the accuracy of the default rate estimation is tested by Binomial Test and chi-square test.

2.4. Build loan pricing model

2.4.1. The economic capital measure of default risk

Assuming there is a portfolio of SME loans issued by commercial banks, its economic capital is the estimated risk compensation for the unexpected losses from defaults in the next one year with 99% confidence level. The economic capital of SMEs loans is estimated according to the procedure presented in Figure 2 as follows:

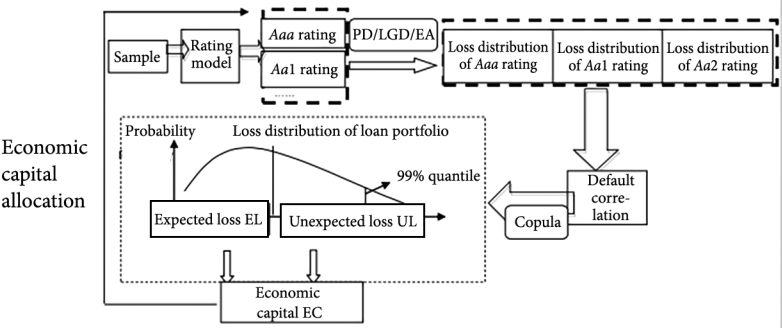


Figure 2. The estimation procedure of economic capital (source: own processing)

According to the mapping procedure, SME loans applicants can be divided into 19 different credit ratings levels as the risk exposures of the economic capital,  $X_p, i = Aaa, \dots, Ccc$ . by inputting the SMEs loan data of a bank to the default rate measurement model  $PD = f(score(X))$ . Then use Archimedean Copula to generate the probability loss distribution of the loan portfolio. The marginal distribution  $u_i = F_p, i = Aaa, \dots, Ccc$  of single exposure loss  $X_i$  is fitted by a non-parametric kernel function. The Archimedean copula of 19 variables is used to carve the correlation between the 19 risk levels,

$$C(u_{Aaa}, \dots, u_{Ccc}) = \hat{G}[\hat{G}^{-1}(u_{Aaa}) + \dots + \hat{G}^{-1}(u_{Ccc})], \quad (4)$$

where  $\hat{G}(\cdot)$  is the Laplace transformation of the distribution function,  $\hat{G}^{-1}(\cdot)$  is the inverse function of  $\hat{G}(\cdot)$ ,  $\hat{G}^{-1}(\cdot) = \varphi(\bullet)$  is the generator of the Archimedean Copula function. Taking the marginal loss distribution  $u_i = F_p$ ,  $i = Aaa, \dots, Ccc$  of individual risk exposure  $X_p$ ,  $i = Aaa, \dots, Ccc$  into (5), Maximum Likelihood (ML) estimation is used to estimate the marginal contribution of different rating-grade loan portfolios to overall economic capital.

### 2.4.2. Loan pricing

As the final step, the basic formula for loan pricing is derived based on RAROC as given in Eq. (5),

$$r_i = \frac{EL_i + OC_i + DC_i + ECC_i \times RAROC_i}{I_i};$$

$$RAROC_i = \frac{I_i - EL_i - OC_i - DC_i}{ECC_i}, \quad (i = Aaa, \dots, Ccc), \quad (5)$$

where  $I_i = (L \times r \times T + F) \times (1 - t)$  represents the loan income corresponding to each risk exposure of the bank loans for SMEs,  $L$  represents the loan amount,  $r$  the aggregated loan interest rates,  $T$  the agreed loan term,  $F$  the loan commitment fee,  $t$  the sales tax and the additional tax rate;  $EL$  represents the expected loss of single risk exposure at 99% confidence level, and  $OC$  the funding transfer price which represents the costs of financing resources of the same kind in financial markets. At present, the major banks generally transfer funds to and within the system of the banks, and the funds raised by the business units are fully transferred to the capital center through the internal transfer price (*IFTP*). The funds required for each loan business are also purchased in full from the capital center at the full internal transfer price. Therefore, the capital cost of using *IFTP* is applicable to the actual situation of commercial banking industry. The cost of the capital is  $OC = L \times IFTP \times T$ , operating costs  $DC$  include the fixed costs and the variable costs of daily operations and management. Fixed costs mainly refer to the credit investigation and analysis costs. The variable costs include the cost of workbooks, management fees, salaries of credit personnel. The allocation and collection of expenses can be carried out by using the method of job-based costing. As a result, operating costs  $DC = DC_f + DC_v$  proposed by Deng, Ao, Hu, and Wang (2010), where  $DC_f$  denotes the fixed cost; the variable cost. For each exposure,  $I_i, EL_i, OC_i, DC_i$ , and then  $RAROC_i$  ( $i = Aaa, \dots, Caa3$ ) can be calculated for each risk exposure. Finally, compare  $RAROC_i$  ( $i = Aaa, \dots, Caa3$ ) with the minimum required rate of return  $R$  of banks, which can be used as the basis for loan decision.

## 3. Empirical results

### 3.1. The estimation of internal rating model

Account level SMEs loans data of a Chinese commercial bank from 2010 to 2016 has been granted to access. The data has been abstracted from 2 databases. The application database consists of all the SMEs applicant for loans with some granted the loans and some denied the

loans later on. The other database consisting of all the granted loans accounts with remaining balances. The Eq. (6) is the final form of the rating model:

$$\begin{aligned} \text{Score}(X) &= 0.485 \times \phi(f_1(x_1, x_2 \cdots x_m)) + 0.514 \times \phi(f_2(x_1, x_2 \cdots x_m)); \\ f_1(x_1, x_2 \cdots x_m) &= 71.718 - 9.523 \times Ecp - 0.503 \times Mpbi - 2.656 \times Fin2 - \\ &\quad 7.65459 \times Fin4 - 15.1667 \times Fin6 - 20.0099 \times DIsl; \\ f_2(x_1, x_2 \cdots x_m) &= 5.597 - 2.627 \times ECT - 2.303 \times Mcap - 1.871 \times Mpol + \\ &\quad 1.917 \times FIsc + 1.307 \times MW1, \end{aligned} \quad (6)$$

where  $f_1(x_1, x_2 \cdots x_m)$  represents the optimal quantitatively model,  $f_2(x_1, x_2 \cdots x_m)$  the optimal qualitatively model. After model calibration, mapping and accuracy test of probability of default, the main scale and the function of model calibration are gained. Figure 3 is a model calibration mapping curve, and the curve fitted to  $f(\bullet)$ , the final default rate model, is given by Eq. (7):

$$f(\text{Score}) = 8.95 \times 10^{-5} \times \exp(0.058 \text{Score}) + 9.94 \times 10^{-5} \times \exp(0.144 \text{Score}). \quad (7)$$

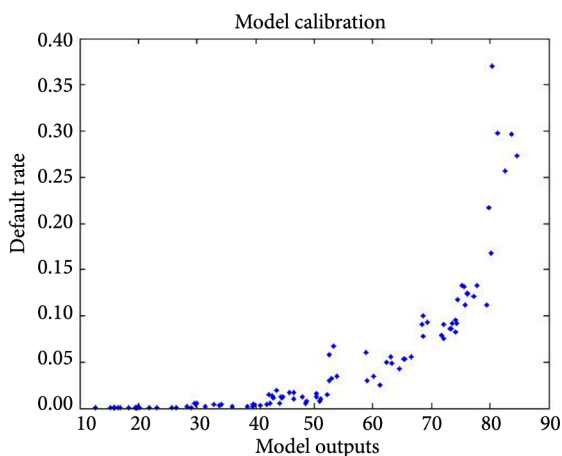


Figure 3. Model calibration mapping curve (source: own processing)

Figure 3 presents the mapping curve of the model calibration by plotting the scattered scores of Eq. (6) as the X-axis and their default rates from the database as Y-axis. As shown in Figure 3, from 0–40 of the model output, the default rate changes at a small ratio and the values are close to 0. From 40–50, while default rate starts increase, it tops at 0.03. After output = 50, they grow exponentially, which means the higher scores would produce higher default rate estimated at accelerated paces. The topology of the Figure 3 fits the real status of the SME loans portfolio of this commercial bank.

### 3.2. Risk categories mapping

The procedure to estimate economic capital is as follows: First, the variables of SMEs loans applicant are substituted into Eq. (7) and ultimately Eq. (8) to get rating results based on the main scale (Figure 4).

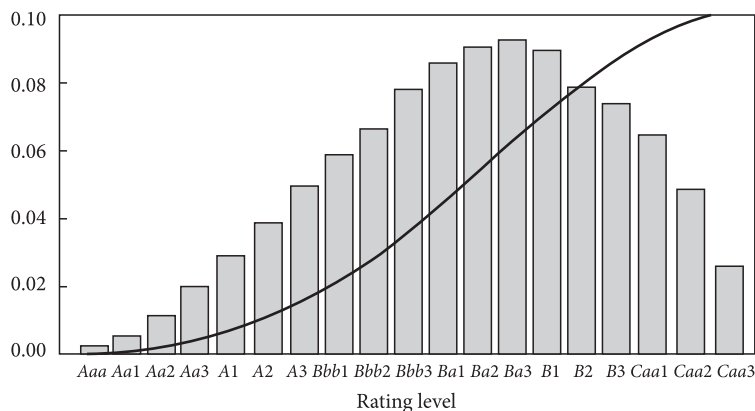


Figure 4. Rating results (source: own processing)

Figure 4 shows the distribution of the SMEs loans for this particular bank. As one might found that the rating level and scale are in inverted-U shaped, with slight right skewed, which indicates that the bank has issued way too many loans with low ratings and higher risk profiles.

Therefore the SMEs loans applicants would be classified is divided into one of the 19 categories based on the risk exposures. The loans applicants belonging to the same categories  $X_i$ ,  $i = Aaa, Aa1, \dots, Ccc$  have homogenous defaults probabilities. Second, Based on the simulation of 1000 kinds of loss proportions with regard to  $PD_i, LGD_i, EAD_i$ ,  $i = Aaa, Aa1, \dots, Ccc$  of each risk exposure and the non-parametric kernel estimation being used to fit the loss distribution, the column vector  $L = (l_{Aaa,n}, \dots, l_{Caa3,n})$  is generated, where  $n = (1, 2, \dots, N)$  is a random vector that obeys the joint distribution  $F(l_{Aaa}, \dots, l_{Caa3})$  and is also the entire vector of loan loss proportions, and the distribution of which is the loss distribution for the entire loan portfolio. The three distribution functions of Archimedean Copula family, such as *Clayton*, *Gumbel*, *Frank* function, are chosen to characterize the loss distribution of SMEs loans. Finally, the *fsearch* function of MATLAB® (version 12) is used to select the optimal Copula function and obtain the final loss distribution, as shown in Figure 5:

Figure 5 present again a inverted -U relationship between PDF and loss ratio. The figure is left skewed with large fat tail, which indicates an active risk pursuit strategies in loan approval.

According to the requirements of the Basel Accord III, a 99% quantile loss distribution is taken as  $EC = F(l_{Aaa}, \dots, l_{Ccc})_{99\%} - EL$ ,  $EL = PD \times EAD \times LGD$ . As  $ECC_i = EC \times ULC_i / UL$ , ( $i = Aaa, \dots, Ccc$ ), the marginal contribution of each risk exposure to the overall economic capital can be calculated.

### 3.3. Loan decision and pricing

The parameters of each risk categories and  $I_i, EL_i, OC_i, DC_i$  and  $RAROC_i$ ,  $i = Aaa, Aa1, \dots, Ccc$  of the customers under each rating level can be calculated, as shown in Table 2. According to Eq. (7), the prices of loans belonging to each risk category and the average interest rates

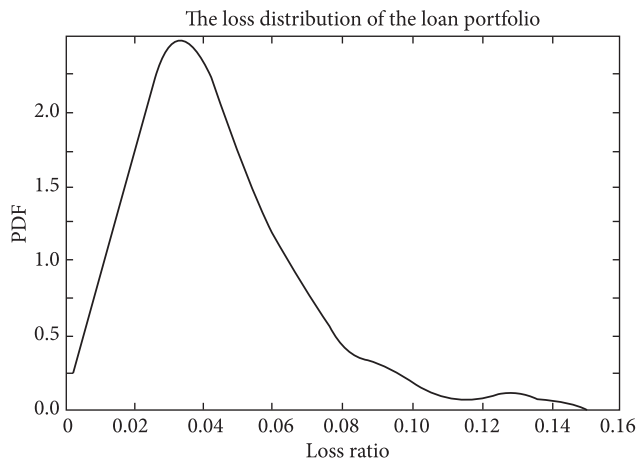


Figure 5. The loss distribution of the loan portfolio (source: own processing)

for that category can be calculated, as shown in Table 2. Since the management of this bank is risk averse, the bank refuses to grant loans to customers classified as Ba2 or above in risk categories.

Table 2. RAROC and return rate of each rating category (source: own processing)

Risk Categories	The average interest rate actually paid by banks	The estimated prices of the loans	Differences
<i>Aaa</i>	6.68%	6.14%	−0.54%
<i>Aa1</i>	6.74%	6.21%	−0.53%
<i>Aa2</i>	6.81%	6.28%	−0.53%
<i>Aa3</i>	7.11%	6.39%	−0.72%
<i>A1</i>	7.42%	6.88%	−0.54%
<i>A2</i>	7.43%	7.18%	−0.25%
<i>A3</i>	7.55%	7.93%	0.38%
<i>Baa1</i>	7.91%	8.09%	0.18%
<i>Baa2</i>	8.13%	8.37%	0.24%
<i>Baa3</i>	8.32%	8.82%	0.50%
<i>Ba1</i>	8.55%	9.12%	0.57%
<i>Ba2</i>	8.92%	9.79%	0.87%

As indicated by Table 2, the interest rates charged by the bank increase as the risk categories of the customers increase since the bank requires more risk-adjusted return from customers of high risks, as interpreted by Demarta and Mcneil (2010), Liu et al. (2012), and Flood and Korenko (2013). For risk category A2 or lower the differences between the average interest rate actually paid by banks and the estimated prices of the loans are nega-

tive, which indicates that the bank is charging too much for customers with less risk, which might lose some potential businesses over competition with other financial institutions. On the other hand, the bank is charging too little to its customers with riskier categories, as the differences between the average interest rate actually paid by banks and the estimated prices of the loans are positive, which means the banks would not have sufficient compensation by bearing potential credit risk and the RAROCs of loans belonging to these risk categories are actually overestimated.

## Conclusions

In this paper, the empirical data of a Chinese commercial bank from 2010 to 2016 would be used in order to remedy the existing problems in SMEs loan pricing for commercial banks, an internal rating model that measures the risk characteristics of SMEs loans is established and Archimedean Copula functions are used to describe the default correlations between different risk exposures and the economic capital (unexpected loss) at the loan portfolio level through the simulation of loss distribution. Ultimately, a loan decision-making and loan pricing system based on RAROC is constructed. Using the method introduced in this study, we found the bank providing commercial loans' data at account level is charging wrong interest rates regarding of the corresponding risk categories. As a result, the approaches could thus be able to provide higher risk-adjusted returns, higher economic capital adequacy ratios, and ultimately stronger banks' capabilities to tolerate risk events.

There are, however, several issues that should be mentioned. First, the dimensions of the risk factors should be extended furthermore to cover as comprehensive as possible aspects of the SME loans. Second, the risk modeling and mapping methods are based on historical data. Forward-looking models could provide more accurate results for out-of-sample observations. Finally, the *fsearch* function provided by MATLAB® (version 12) might be stuck at the local maximums if the data spaces are heterogeneous and the global optimization could be reached by more complicated algorithms. Because of the above mentioned issues, there are some limitations that should be elaborated before the method proposed in this study could be used in practice: The model would provide good out-of-sample estimation based on large volume of SME loans data since the solution to the *fsearch* function's local maximum problem would requires booth-trapping method which works well only with enough data. On the other hand, since the implementation of the simulation method requires the teamwork of personnel with fair amount of mathematical and computational backgrounds, the method might be too demanding for banks just in business with insufficient risk analysts. Although more actions could be easily added to make the study more practical, the simplified infrastructure of this setup is still useful in gaining a general picture about the dynamic decision making process.

## Acknowledgements

The authors thank two anonymous reviewers for their crucial comments. This study is funded by China National Natural Science Foundation, Grant No.:71473204, by Fundamental Re-

search Funds for the Central Universities, Grant No.: JBK1902054, and by Fundamental Research Funds for the Central Universities, Grant No.: JBK18FG16.

### Disclosure statement

The authors declare that they did not have any competing financial, professional, or personal interests from other parties.

### Funding

This work was supported by the <National Natural Science Foundation of China> under Grant [71473204]; < Fundamental Research Funds for the Central Universities > under Grant [JBK1902054 and JBK18FG16].

### Author contributions

CL, SS, and HS conceived the study and were responsible for the design and development of the data analysis. YC and SS were responsible for data collection and analysis. HS and SS were responsible for data interpretation. CL and DL wrote the first draft of the article.

### References

- Aas, K., & Puccetti, G. M. (2014). Bounds on total economic capital: the DNB case study. *Extremes*, 17(4), 693-715. Retrieved from <https://doi.org/10.1007/s10687-014-0202-0>
- Abdelkafi, S. Z., Ghorbel, A., & Khoufi, W. M. (2018). Energy portfolio risk management using time-varying copula methods: application to bonds, interest rate and VIX. *American Journal of Finance and Accounting*, 5(4), 371-393. <https://doi.org/10.1504/AJFA.2018.093633>
- Alie, & Jean, K. (2016). *Assessment of business risk economic capital for South Africa banks: a response to pillar 2 of Basel ii* (Doctoral dissertation). M.M. Thesis, University of the Witwatersrand, Johannesburg, The Republic of South Africa.
- Allen, J., Chapman, J., Echenique, F., & Shum, M. M. (2016). Efficiency and bargaining power in the interbank loan market. *International Economic Review*, 57(2), 691-716. <https://doi.org/10.1111/iere.12173>
- Baltussen, G., van Bakkum, S., & van der Grient, B. (2018). Unknown unknowns: uncertainty about risk and stock returns. *Journal of Financial & Quantitative Analysis*, 53(2018), 1615-1651. <https://doi.org/10.1017/S0022109018000480>
- Cao, J. M. (2011). Calibration and main scale development for credit risk model of non-retail risk exposure. *International Finance Research*, 7, 81-103.
- Demarta, S., & Mcneil, A. J. (2010). The t copula and related copulas. *International Statistical Review*, 73(1), 111-129. <https://doi.org/10.1111/j.1751-5823.2005.tb00254.x>
- Deng, C., Ao, H., Hu, W., & Wang, X. M. (2010). Research on loan pricing for small business from big bank based on relationship loans. *Economic Research*, 2, 83-96.
- Flood, M. D., & Korenko, G. (2013). *Systematic scenario selection: stress testing and the nature of uncertainty* (SSRN Scholarly Paper No. ID 1262896). Rochester, NY: Social Science Research Network. Retrieved from <https://papers.ssrn.com/abstract=1262896>

- Ghosh, I., & Ray, S. M. (2016). Some alternative bivariate Kumaraswamy-type distributions via copula with application in risk management. *Journal of Statistical Theory and Practice*, 10(4), 693-706. <https://doi.org/10.1080/15598608.2016.1215943>
- Giacomini, R., & Rossi, B. M. (2016). Model comparisons in unstable environments. *International Economic Review*, 57(2), 369-392. <https://doi.org/10.1111/iere.12161>
- Gorelaya, N. M. (2016). Otsenka vliyaniya faktorov na formirovanie tseny kredita [Evaluation of the Impact of Factors on Loan Pricing]. *Journal of Corporate Finance Research*, 10(1), 59-76. Retrieved from <https://ssrn.com/abstract=3073043>
- Gubareva, M., & Borges, M. R. M. (2018). Rethinking economic capital management through the integrated derivative-based treatment of interest rate and credit risk. *Annals of Operations Research*, 266(1-2), 71-100. Retrieved from <http://hdl.handle.net/10400.21/6962>
- Hauptmann, C. M. (2017). Corporate sustainability performance and bank loan pricing: it pays to be good, but only when banks are too. *Saïd Business School WP*, 2017(20). Retrieved from <https://ssrn.com/abstract=3067422>
- Karmakar, M., & Paul, S. M. (2018). Intraday portfolio risk management using VaR and CVaR: A CGARCH-EVT-Copula approach. *International Journal of Forecasting*, 35(2), 699-709. <https://doi.org/10.1016/j.ijforecast.2018.01.010>
- Khan, F., Hashemi, S. J., Paltrinieri, N., Amyotte, P., Cozzani, V., & Reniers, G. M. (2016). Dynamic risk management: a contemporary approach to process safety management. *Current opinion in chemical engineering*, 14, 9-17. <https://doi.org/10.1016/j.coche.2016.07.006>
- Kim, J., Shim, K., Cao, L., Lee, J., Lin, X., & Moon, Y. (2017, 23-26 May). *Advances in knowledge discovery and data mining: 21st Pacific-Asia Conference, PAKDD 2017, Proceedings, Part I*. Jeju, South Korea.
- Liu, C., Guo, M., Mo, N., & Zhang, X. M. (2012). Building an effective measurement and early warning index system for credit risk of SMES loans in China's commercial banks based on internal ratings approach of Basel Accord. *Financial Regulatory Research*, 7, 26-39.
- Liu, C., Zhang, X., & Mo, N. M. (2013). Construction and empirical study on internal ratings model of SMEs loans in China's commercial banks. *Investment Research*, 5, 3-16.
- Oh, D. H., & Patton, A. J. M. (2018). Time-varying systemic risk: evidence from a dynamic copula model of CDS spreads. *Journal of Business & Economic Statistics*, 36(2), 181-195. <https://doi.org/10.1080/07350015.2016.1177535>
- Petropoulos, A., Chatzis, S. P., & Xanthopoulos, S. M. (2016). A novel corporate credit rating system based on Student co hidden Markov models. *Expert Systems with Applications*, 53, 87-105. <https://doi.org/10.1016/j.eswa.2016.01.015>
- Pfeifer, D., Tsatedem, H. A., Mändle, A., & Girschig, C. M. (2016). New copulas based on general partitions-of-unity and their applications to risk management. *Dependence Modeling*, 4(1), 123-140. <https://doi.org/10.1515/demo-2016-0006>
- Scandizzo, S. B. (2016). Economic capital models. In S. Scandizzo (Ed.), *The validation of risk models* (pp. 193-204). London, UK: Palgrave Macmillan. [https://doi.org/10.1057/9781137436962\\_13](https://doi.org/10.1057/9781137436962_13)
- Shim, F., & Lee, B. (2017). An estimation of VaR based on Archimedean Copula. *Journal of Chongqing Institute of Technology*, 22(8), 326-336.
- Tang, X. (2015). *Credit loss distribution and copula in risk management* (MSc Thesis). Mathematical Institute Leiden University, Leiden, The Netherlands.
- Walke, A. G., Fullerton Jr, T. M., & Togle, R. J. M. (2018). Risk-based loan pricing consequences for credit unions. *Journal of Empirical Finance*, 47, 105-119. <https://doi.org/10.1016/j.jempfin.2018.02.006>
- Wang, T., & Libaers, D. M. (2015). Nonmimetic knowledge and innovation performance: empirical evidence from developing countries. *Journal of Product Innovation Management*, 33(5), 570-588. <https://doi.org/10.1111/jpim.12306>