International Journal of Mathematics Trends and Technology
ISSN: 2231-5373/ https://doi.org/10.14445/22315373/IJMTT-V68I9P505

Volume 68 Issue 9, 21-33, September 2022

Original Article

Predicting and Optimizing Future Bank Performance

Qiaopeng Ma[1], Zhongqiang Liu[2]

[1,2]*School of Mathematics and Information Science, Henan Polytechnic University, Henan, China.*

Received: 23 July 2022 Revised: 29 August 2022 Accepted: 14 September 2022 Published: 27 September 2022

Abstract *- This article selects 25 listed banks in China, and collects their annual financial data from 2016 to2021. According to existing research, we identify 5 influencing variables to explain weighted average return on equity (ROE) of banks. Noticing good fitting effect of nonlinear regression, we use nonlinear regression model to fit the annual financial data. Then taking fitted regression models as objective function, we solve nonlinear programming problems and obtain corresponding optimized results, which are compared with actual weighted average ROE of the 25 banks. By comparison, we observe that optimized values are significantly greater than corresponding actual values, which indicates that there is room for improvement in weighted average ROE. Based on regression models obtained, one-year predicted performance and three-year mean predicted performance are calculated to predict future bank performance. Combining optimized and predicted values, banks can optimize their performance in the future by adjusting their own internal influencing variables.*

Keywords *- Mean predicted performance, Nonlinear regression model, Nonlinear programming, Optimized performance, Weighted average ROE.*

1. Introduction

With the development of economy, the number of banks in China is increasing, by the end of 2021, there are 57 A+H-share listed banks all over the country. In May 2021, Forbes magazine released the “Forbes Global 2000 List”, China’s big four banks all ranked in the top 15, among which Industrial and Commercial Bank of China (ICBC) topped the list for nine consecutive years. Nowadays, banks are no longer simply handling deposit and loan business, there are various types of financial services to choose from, and the banking business is becoming more and more diversified. However, in recent years, continuous downward pressure on the economy has made all industries face challenges, especially the banking industry in the financial system, whose performance has been significantly affected. Therefore, it is of great significance to control risks and improve performance by studying various factors affecting bank performance.

To clarify the relationship between performance of banks and corresponding influencing factors, scholars carried out a large number of theoretical and empirical studies from different perspectives. In the following, we classify and introduce research results about the factors that affect the performance of banks.

McConnell and Servaes investigated the relation between Tobin’s Q and the structure of equity ownership for a sample of 1173 firms for 1976 and 1093 firms for 1986. They found a strong curvilinear relation between Tobin’s Q and the fraction of shares owned by corporate insiders [18]. By using a sample of 117 Indonesian banks, and 28 public banks, Hanafi et al. found that ownership concentration and governance by larger numbers of commissioners improves a bank’s profitability and it’s handling of risk [11]. Massis et al. found that there is a U-shaped relationship between the degree of family ownership dispersion and firm performance when conducting empirical research on family firms [17]. Busta et al. investigated the relationship between ownership concentration and market value of European banks. By using GMM dynamic estimator on a sample of European banks over a 13-year period (1993−2005), they found on average a negative effect of ownership concentration on bank value, measured by Tobin’s Q [6].

Ducassy and Guyot studied the link between a firm’s ownership structure and its performance in the French context, their findings showed that a majority shareholder had a positive influence on a firm’s value, as suggested by agency theory[8]. Liu et al. conducted an empirical analysis using 5 years (2011−2015) of operational data for 71 Chinese commercial banks. Two two-stage meta-frontier data envelopment analysis network models and multiple regression models were used to estimate and analyze impacts of variations in bank ownership structure. The deposit efficiency of a bank was found to be mainly influenced by the nature of ownership and ownership concentration [16].

Berger and Patti were the first to employ a simultaneous-equations model that accounts for reverse causality from performance to capital structure. Their empirical results were consistent with the agency costs hypothesis that higher leverage or a lower equity capital ratio in banking is associated with higher profit efficiency [4]. Trujillo-Ponce empirically analyzed the main factors behind the high bank profitability in Spain for the period of 1999 through 2009 using an unbalanced panel data of 697 observations, and found that an increase in the equity-to-total-assets ratio reduces the ROE of the banks because of the fall in leverage [23]. By using small-business data for Japan, Tsuruta empirically investigated the relationship between firm performance and leverage and found that highly leveraged firms enjoy stronger performance (measured as firm sales growth or profitability) compared with low-leveraged firms [24]. Abdullah and Tursoy empirically examined the relationship between firm performance and capital structure. Their results confirmed a positive relationship between firm performance and capital structure [1].

Larger banks are usually safer and more efficient as compared to those with smaller asset size. Banz was the first economist to find the size effect. He found that both the total returns and the risk adjusted returns in the United States were negatively related to the size of the firms [3]. Daly and Frikha used Data Envelopment Analysis to examine the determinants of banking performance in Bahrain between 2005 and 2009. They found that increase in size of Islamic banks and the rapid growth in the customers' deposits are the important factors of performance [7].

Francis et al. investigated whether diversification/focus across assets, industries and borrowers affects bank performance. In particular, size, bank capital and solvency differences of banks are shown to influence bank performance, especially regarding diversification/focus strategies [10]. Adesina also found a significant positive relationship between size and bank performance [2]. Kendo and Tchakounte applied a panel quantile approach with nonadditive fixed effects to study the impact of asset size on financial performance and outreach of microfinance institutions. The results revealed that an increase in asset size leads to increased profitability, with a greater impact for microfinance institutions that have poor or low-end profitability levels than for those with satisfactory levels [14].

When a bank has too many non-performing loans in its balance sheet, it poses cash flow problems for the bank since it is no longer earning income from its credit business. Obviously, non-performing loan ratio has negative effect on overall bank profitability. Kjosevski and Petkovski examined selected macroeconomic and bank-specific determinants of nonperforming loans for a panel of 21 commercial banks from the Baltic States (Estonia, Latvia and Lithuania), using annual data for the period 2005−2016. They found that equity to total assets ratio, return on assets, return on equity and growth of gross loans have an impact on the amount of non-performing loans [15]. By employing U.S. commercial banks data from 1994 to 2018, Phung et al. adopted an advance Data Envelopment Analysis method to estimate bank efficiency and found a negative relationship between non-performing loans and bank efficiency [19].

Stickney and Brown decomposed return on assets into asset turnover and profit margin, and suggested that calculating the relative contributions of asset turnover and profit margin to current profitability is useful in providing insights into the firm's strategy [21]. Fairfield and Yohn provided evidence that disaggregating the change in return on assets into the change in asset turnover and the change in profit margin is useful in forecasting the change in return on assets one year ahead. Analysts and investors should focus on changes in asset turnover to improve forecasts of future profitability [9]. The asset turnover ratio is an indicator of the efficiency with which a company is deploying its assets. Different arrangements of asset structure will lead to different times of asset turnover.

Most scholars establish linear regression model or nonlinear regression model through correlation analysis, and then conduct descriptive analysis or hypothesis tests on the results to reveal the relationship between bank performance and corresponding influencing factors. On the basis of previous studies, combined with the characteristics of the banking industry, this paper identifies five influencing factors: equity concentration, debt to asset ratio, asset size, non-performing loan ratio and asset turnover ratio. By comparing the fitting effects of linear and nonlinear regression models, we choose to establish nonlinear regression model to explain bank performance. Finally, we can predict and optimize future bank performance by using annual nonlinear regression models obtained.

The remainder of the paper is organized as follows. Section 2 provides an overview of 25 listed banks in China. To characterize the relationship between bank performance and influencing factors, we build nonlinear regression models year by year in Section 3. With the established regression models as objective function, we solve nonlinear programming problems and obtain optimized annual performance values in Section 4. Combining fitted nonlinear regression equations and optimized bank performance, we provide ways to predict and optimize bank performance in Section 5. Section 6 concludes with some remarks and prospective research topics.

2. An overview of Research Object

With the development of China's economy, more and more banks have entered the A-share market. By the end of 2021, a total of 41 banks have been listed on the A-share market, including 15 A+H-share listed banks and 26 pure A-share listed banks. Among them 33 are listed on the Shanghai Stock Exchange and 8 are listed on the Shenzhen Stock Exchange. The first listed bank in China was PingAn Bank, whose predecessor was Shenzhen Development Bank, which was listed on the Shenzhen Stock Exchange on April 3, 1991. 8 years later Shanghai PuDong Development Bank was listed, and after that, commercial banks went public one after another. The first peak period was from 2000 to 2010, with 14 banks listed, of which 7 banks landed in the A-share market in 2007. After that, 2016 and 2019 ushered in the peak of bank listing, respectively, and 8 banks went public in each of these two years. In 2021, under the influence of the COVID-19, although the banking industry was affected to some extent, there were still 4 banks listed.

According to the financial data disclosed by commercial banks on the Shanghai Stock Exchange and Shenzhen Stock Exchange, and considering the completeness of the data from 2016 to 2020, this paper selects 25 Chinese listed banks for research, which are listed in Table 1 according to their listing time.

Table 1. Profiles of 25 A-share listed banks

Bank name	Listing code	Listing time	Bank name	Listing code	Listing time
Ping An Bank	000001.SZ	April 3, 1991	China Construction Bank	601939.SH	September 25, 2007
SPD Bank	600000.SH	November 10, 1999	Agricultural Bank of China	601288.SH	July 15, 2010
China Minsheng Bank	600016.SH	December 19, 2000	China Everbright Bank	601818.SH	August 18, 2010
ChinaMerchants Bank	600036.SH	April 9, 2002	Bank of Jiangsu	600919.SH	August 2, 2016
Hua Xia Bank	600015.SH	September 12, 2003	Bank of Guiyang	601997.SH	August 16, 2016
Bank of China	601988.SH	July 5, 2006	Jiangyin Bank	002807.SZ	September 2, 2016
Industrial and Commercial Bank of China	601398.SH	October 27, 2006	Wuxi Bank	600908.SH	September 23, 2016
Industrial Bank	601166.SH	February 8, 2007	Changshu Bank	601128.SH	September 30, 2016
China Citic Bank	601998.SH	April 27, 2007	Bank of Hangzhou	600926.SH	October 27, 2016
Bank of Communications	601328.SH	May 15, 2007	Bank of Shanghai	601229.SH	November 16, 2016
Bank of Ningbo	002142.SZ	July 19, 2007	Sunong Bank	603323.SH	November 29, 2016
Bank of Nanjing	601009.SH	July 19, 2007	Zhangjiagang Bank	002839.SZ	January 24, 2017
Bank of Beijing	601169.SH	September 19, 2007			

In order to have an intuitive and clear understanding of operating conditions of the 25 listed banks, we plotted weighted average ROE data of the 25 listed banks from 2016 to 2020 (Fig. 1). From Figure 1, we can clearly observe that on the whole, the weighted average ROE of 25 listed banks from 2016 to 2020 showed a decreasing trend year by year. Specifically, the decline rate of each bank is different, and the weighted average ROE of most banks has dropped significantly. ROE is an important indicator reflecting bank performance. The purpose of this paper is to clarify the relationship between influencing factors and weighted average ROE of banks, so as to help banks reduce operating risks and improve performance.

In 2019, the average return on total assets of China's banking industry was 0.87%, 0.03 percentage points lower than last year, and the average return on net assets was 10.96%, 0.77 percentage points lower than last year. Judging from a relatively long time period, indicators such as profit growth rate, net interest margin, and ROE of China's banking industry are all at low points in the past 10 years. In general, 2021 is the first year of China's "14th Five-Year Plan". Facing a new pattern of "dual circulation", China's banking industry will enter a new stage of development, and also face many challenges.

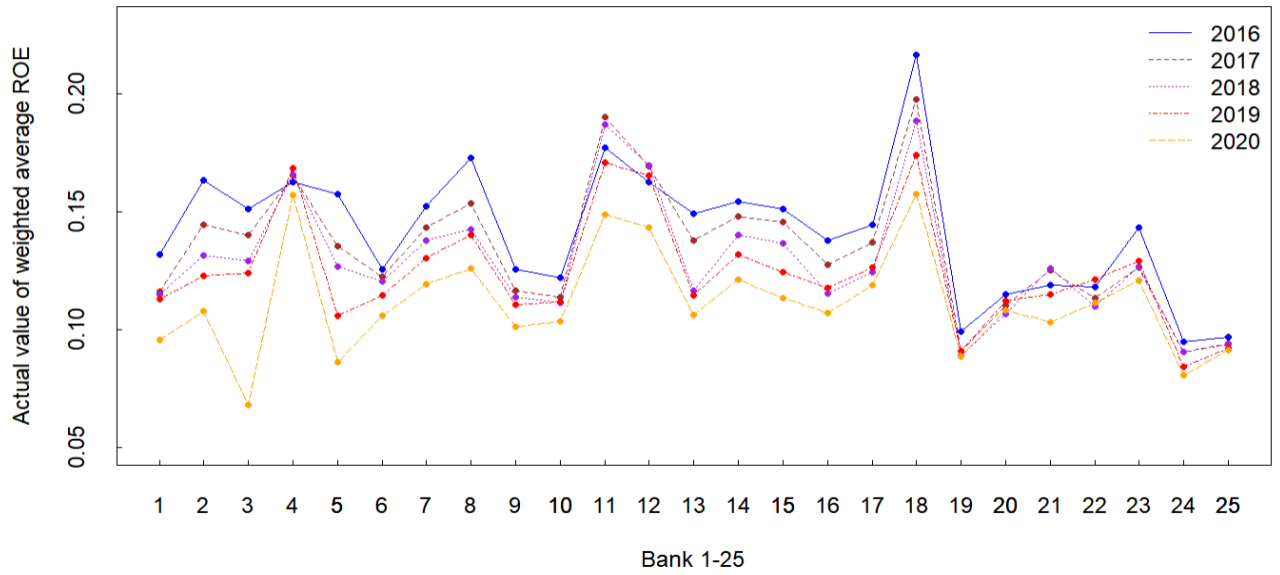

Fig. 1 Trend chart of weighted average ROE of 25 listed banks in the five years from 2016 to 2020.

3. Discovering Relationship between Bank Performance and Influencing Factors

The banking sector in China plays an important role in the development of the country's economy. The performance of Chinese banks has attracted great attention from the government, banking regulatory authorities and academic researchers [22]. Many scholars described the impact of influencing factors on bank performance. Of course, we are also interested in how influencing factors affect bank performance and how this effect changes over time. In this section we will build regression models based on actual financial data to explain bank performance year by year.

3.1. Determination of Research Variables

Regression models are useful tools applied in many research fields dealing with empirical data. They connect an outcome variable to one or several independent variables and quantify the strength of association between them. From the introduction in Section 1, we can see some influencing variables used by scholars in the study of bank performance. Following research results of scholars, we select and interpret influencing variables as follows:

3.1.1. ROE

ROE for the banking industry, factors such as government regulation, legal system, and social environment will all have an impact on performance. ROE is a relatively comprehensive indicator, which is the ratio of a company's net profit to the average shareholders' equity in a certain period of time. This indicator reflects the income level of shareholders' equity and measures the profitability of self owned capital. The higher the value of this indicator, the better the operation of a company. This article selects weighted average ROE as the dependent variable to measure bank's operating performance.

3.1.2. Equity Concentration

Equity concentration moderate equity concentration is conducive to the improvement of performance. If the equity concentration is too high, major shareholders may damage the rights and interests of other minority shareholders for their own interests, which is not conducive to the improvement of performance. If the equity concentration is too dispersed, it will make it impossible for the equity to reach an agreement, and it is not conducive to the improvement of performance. Therefore, moderate equity concentration will promote performance. Bouvatier et al. considered three ownership measures in the construction of clusters of banks with "similar" ownership characteristics: the percentage held by the largest shareholder, the percentage held by the second-largest shareholder and a Herfindahl index computed for a bank's ownership distribution. The first two measures give meaningful information on the shape of the ownership concentration, whereas the Herfindahl index captures the distribution of ownership for all shareholders [5]. In this article we select the percentage held by the largest shareholder to represent the equity concentration.

3.1.3. Debt to Asset Ratio

It is also called debt operation ratio, which is the ratio of debt to total assets, that is, debt to asset ratio = debt/total assets, which reflects how much of the total assets of a company is obtained through debt, representing the comprehensive ability to repay debts. Banks mainly operate by absorbing deposits, so a high debt to asset ratio is conducive to the improvement of bank performance to a certain extent.

3.1.4. Asset Size

Asset size usually, asset size and performance promote each other. The expansion of asset size may absorb more customer resources, and will not lose investment opportunities due to asset constraints. Since the value of asset size is generally large, the logarithm of total assets is used to represent the asset size, that is, asset size = ln(total assets).

3.1.5. Non-Performing Loan Ratio

The non-performing loan ratio is one of the important indicators to evaluate the safety of credit assets of financial institutions. The higher the non-performing loan ratio, the larger the proportion of loans that may not be recovered in the total loans; the lower the non-performing loan ratio, the smaller the proportion of loans that may not be recovered in the total loans by financial institutions. The main business of banks includes absorbing public deposits, issuing loans, and handling bill discounting. Compared with other industries, controlling credit risk and ensuring asset safety is key to the operation of the banking industry. The non-performing loan ratio reflects the quality of bank assets. China Banking Regulatory Commission has clear requirements on the non-performing loan ratio, and stipulates that the non-performing loan ratio of banks should not be higher than 5%.

3.1.6. Asset Turnover Ratio

Asset turnover ratio refers to the ratio of the company's total sales to the average assets in a certain period of time, which reflects the operating capacity of the company's overall assets. A high asset turnover ratio means that the operating efficiency of the asset is relatively high, which is conducive to the improvement of performance.

Next we will try to find the causal relationship among the above variables. For clarity, they are summarized in Table 2.

Table 2. Summary of research variables

Variable name	Variable symbol	Variable interpretation
Return on equity	ROE	Net profit/Average shareholders' equity
Equity concentration	EC	The percentage held by the largest shareholder
Debt to asset ratio	DA	Debt/Total assets
Asset size	AS	ln(Total assets)
Non-performing loan ratio	NPL	Loan provision ratio/Provision coverage ratio
Asset turnover ratio	AT	Total sales/Average assets

3.2. Establishment of Nonlinear Regression Models

To discover the relationship between dependent variable and independent variables, we first tried to establish multiple linear regression models based on annual financial data, however, regression results showed that the coefficient of determination R-square was small, indicating that the proportion of variation explained by the regressors was small and goodness of fit for linear models was not high. There may not be a simple linear relationship between weighted average ROE and influencing variables.

In practical application, the relationship between the dependent variable and independent variables is usually not strictly linear. We often encounter some curvilinear relationship between the dependent variable and independent variables of many economic phenomena. Therefore, it is necessary to investigate nonlinear relationship between research variables. Based on the bank's annual financial data, by exploring the relationship between weighted average ROE and each influencing variable, we find there is a nonlinear relationship between weighted average ROE and influencing variables. This is why nonlinear regression model is built below.

3.2.1. Nonlinear Regression Model for 2016 Data

In order to establish the most suitable regression model, we first study the relationship between weighted average ROE and each influencing variable. By observing scatterplots, we tried to establish linear model, quadratic curve model and exponential model for comparison, and finally came to the conclusion that the weighted average ROE in 2016 could be explained by equity concentration and asset size via quadratic function, by debt to asset ratio via exponential function, by non-performing loan ratio and asset turnover ratio via linear function. Therefore, nonlinear regression model is established as follows:

$$ROE_{16} = \beta_0 + \beta_1 EC_{16}^2 + \beta_2 exp(\beta_3 DA_{16}) + \beta_4 AS_{16}^2 + \beta_5 NPL_{16} + \beta_6 AT_{16} \quad (1)$$

Where β_0 is the regression constant, β_1, β_2, β_3, β_4, β_5, β_6 is corresponding regression coefficient, "16" represents year 2016. By using R software we can fit nonlinear regression model (1) based on 2016 data and obtain regression equation as follows:

$$ROE_{16} = -0.4975 - 0.0649EC_{16}^2 + 0.0642exp(2.3400DA_{16}) + 0.0002AS_{16}^2 - 1.7410NPL_{16} + 2.1916AT_{16} \quad (2)$$

Some other regression results like residual sum-of-squares, number of iterations to convergence, achieved convergence tolerance are displayed in Table 3.

Table 3. Nonlinear regression results of 2016 data

Coefficient	$\boldsymbol{\beta_0}$	$\boldsymbol{\beta_1}$	$\boldsymbol{\beta_2}$	$\boldsymbol{\beta_3}$	$\boldsymbol{\beta_4}$	$\boldsymbol{\beta_5}$	$\boldsymbol{\beta_6}$
Estimate	-0.4975	-0.0649	0.0642	2.3400	0.0002	-1.7410	2.1916
Residual sum-of-squares				0.0071			
Number of iterations to convergence				208			
Achieved convergence tolerance				8.42e-06			

3.2.2. Nonlinear Regression Model for 2017 Data

For 2017 data, by observing scatterplots between weighted average ROE and each influencing variable, we tried to establish linear model, quadratic curve model and exponential model for comparison, and finally came to the conclusion that the weighted average ROE in 2017 could be explained by equity concentration via quadratic function, by asset size via exponential function, by debt to asset ratio, non-performing loan ratio and asset turnover ratio via linear function. Therefore, nonlinear regression model is established as follows:

$$ROE_{17} = \beta_0 + \beta_1 EC_{17}^2 + \beta_2 DA_{17} + \beta_3 exp(\beta_4 AS_{17}) + \beta_5 NPL_{17} + \beta_6 AT_{17} \quad (3)$$

where β_0 is the regression constant, β_1, β_2, β_3, β_4, β_5, β_6 is corresponding regression coefficient, "17" represents year 2017. By using R software we can fit nonlinear regression model (3) based on 2017 data and obtain regression equation as follows:

$$ROE_{17} = -1.5940 - 0.0716EC_{17}^2 + 1.8360DA_{17} + 0.0001exp(0.3741AS_{17}) - 2.7070NPL_{17} + 1.9220AT_{17} \quad (4)$$

Some other regression results like residual sum-of-squares, number of iterations to convergence, achieved convergence tolerance are displayed in Table 4.

Table 4. Nonlinear regression results of 2017 data

Coefficient	$\boldsymbol{\beta_0}$	$\boldsymbol{\beta_1}$	$\boldsymbol{\beta_2}$	$\boldsymbol{\beta_3}$	$\boldsymbol{\beta_4}$	$\boldsymbol{\beta_5}$	$\boldsymbol{\beta_6}$
Estimate	-1.5940	-0.0716	1.8360	0.0001	0.3741	-2.7070	1.9220
Residual sum-of-squares				0.0045			
Number of iterations to convergence				7			
Achieved convergence tolerance				2.707e-06			

3.2.3. Nonlinear Regression Model for 2018 Data

For 2018 data, by observing scatterplots between weighted average ROE and each influencing variable, we tried to establish linear model, quadratic curve model and exponential model for comparison, and finally came to the conclusion that the weighted average ROE in 2018 could be explained by equity concentration and asset size via quadratic function, by non-performing loan ratio via exponential function, by debt to asset ratio and asset turnover ratio via linear function. Therefore, nonlinear regression model is established as follows:

$$ROE_{18} = \beta_0 + \beta_1 EC_{18}^2 + \beta_2 DA_{18} + \beta_3 AS_{18}^2 + exp(\beta_4 NPL_{18}) + \beta_5 AT_{18} \quad (5)$$

where β_0 is the regression constant, β_1, β_2, β_3, β_4, β_5 is corresponding regression coefficient, "18" represents year 2018. By using R software we can fit nonlinear regression model (5) based on 2018 data and obtain regression equation as follows:

$$ROE_{18} = -1.7760 - 0.0507EC_{18}^2 + 0.9397DA_{18} + 0.0003AS_{18}^2 + exp(-4.6263NPL_{18}) + 1.9896AT_{18} \quad (6)$$

Some other regression results like residual sum-of-squares, number of iterations to convergence, achieved convergence tolerance are displayed in Table 5.

Table 5. Nonlinear regression results of 2018 data

Coefficient	β_0	β_1	β_2	β_3	β_4	β_5
Estimate	-1.7760	-0.0507	0.9397	0.0003	-4.6263	1.9896
Residual sum-of-squares				0.0082		
Number of iterations to convergence				5		
Achieved convergence tolerance				2.406e-06		

3.2.4. Nonlinear Regression Model for 2019 Data

For 2019 data, by observing scatterplots between weighted average ROE and each influencing variable, we tried to establish linear model, quadratic curve model and exponential model for comparison, and finally came to the conclusion that the weighted average ROE in 2019 could be explained by equity concentration and asset size via quadratic function, by non-performing loan ratio via exponential function, by debt to asset ratio and asset turnover ratio via linear function. Therefore, nonlinear regression model is established as follows:

$$ROE_{19} = \beta_0 + \beta_1 EC_{19}^2 + \beta_2 DA_{19} + \beta_3 AS_{19}^2 + exp(\beta_4 NPL_{19}) + \beta_5 AT_{19} \quad (7)$$

where β_0 is the regression constant, β_1, β_2, β_3, β_4, β_5 is corresponding regression coefficient, "19" represents year 2019. By using R software we can fit nonlinear regression model (7) based on 2019 data and obtain regression equation as follows:

$$ROE_{19} = -2.3004 - 0.0468 EC_{19}^2 + 1.4901 DA_{19} + 0.0002 AS_{19}^2 + exp(-3.8676 NPL_{19}) + 2.2595 AT_{19} \quad (8)$$

Some other regression results like residual sum-of-squares, number of iterations to convergence, achieved convergence tolerance are displayed in Table 6.

Table 6. Nonlinear regression results of 2019 data

Coefficient	β_0	β_1	β_2	β_3	β_4	β_5
Estimate	-2.3004	-0.0468	1.4901	0.0002	-3.8676	2.2595
Residual sum-of-squares				0.0053		
Number of iterations to convergence				5		
Achieved convergence tolerance				1.729e-06		

3.2.5. Nonlinear Regression Model for 2020 Data

For 2020 data, by observing scatterplots between weighted average ROE and each influencing variable, we tried to establish linear model, quadratic curve model and exponential model for comparison, and finally came to the conclusion that the weighted average ROE in 2020 could be explained by equity concentration via quadratic function, by asset size via exponential function, by debt to asset ratio, non-performing loan ratio and asset turnover ratio via linear function. Therefore, nonlinear regression model is established as follows:

$$ROE_{20} = \beta_0 + \beta_1 EC_{20}^2 + \beta_2 DA_{20} + \beta_3 exp(\beta_4 AS_{20}) + \beta_5 NPL_{20} + \beta_6 AT_{20} \quad (9)$$

where β_0 is the regression constant, β_1, β_2, β_3, β_4, β_5, β_6 is corresponding regression coefficient, "20" represents year 2020. By using R software we can fit nonlinear regression model (9) based on 2020 data and obtain regression equation as follows:

$$ROE_{20} = -0.8046 - 0.0446 EC_{20}^2 + 1.0220 DA_{20} + 0.0003 exp(0.2892 AS_{20}) - 4.3086 NPL_{20} + 0.4721 AT_{20} \quad (10)$$

Some other regression results like residual sum-of-squares, number of iterations to convergence, achieved convergence tolerance are displayed in Table 7.

Table 7. Nonlinear regression results of 2020 data

Coefficient	β_0	β_1	β_2	β_3	β_4	β_5	β_6
Estimate	-0.8046	-0.0446	1.0220	0.0003	0.2892	-4.3086	0.4721
Residual sum-of-squares				0.0069			
Number of iterations to convergence				44			
Achieved convergence tolerance				7.904e-06			

From residual sum-of-squares and achieved convergence tolerance of each year, the fitting effect of nonlinear regression models is good. Obviously, nonlinear regression models corresponding to each year's data are different, which reflects the regression models change dynamically over time. However, it is relatively certain that bank performance has an inverted "U"-

shaped relationship with equity concentration, a positive relationship with debt to asset ratio, asset size and asset turnover ratio, and a negative relationship with non-performing loan ratio.

Specifically, observing the influence of each independent variable from annual nonlinear regression equations, we can draw the following conclusions:

1. The inverted "U"-shaped relationship between weighted average ROE and equity concentration indicates that too concentrated or too dispersed equity is not conducive to the improvement of performance. Only when the equity concentration is within a certain range will banks improve their performance.
2. The weighted average ROE has a positive relationship with the debt to asset ratio. This is because banks operate by absorbing deposits and belong to highly indebted industry. The larger the debt to asset ratio, the better the operation. At present, the debt to asset ratio of banks is generally more than 90%. Although according to the Basel Accord, the debt to asset ratio of banks should not exceed 92%, many banks exceed this upper limit.
3. The relationship between the weighted average ROE and the asset size is quadratic or exponential. This reflects when the asset size is too small, the ability to resist risks is poor and the operating value of the bank is small. When the asset size increases or reaches a certain height, there will be enough capital investment to improve performance.
4. The weighted average ROE has a negative correlation with the non-performing loan ratio, which is one of the indicators to evaluate the safety of bank assets. The more nonperforming loans, the more uncollectible loans and the poorer asset quality. Therefore, the lower the non-performing loan ratio of a bank, the better.
5. The weighted average ROE is positively correlated with the asset turnover ratio. It reflects the operating capacity of enterprise assets. The higher the asset turnover ratio, the stronger the sales capacity, and the assets have been fully utilized, which is conducive to the improvement of performance.

4. Optimizing Bank Performance based on Nonlinear Programming

Optimization (Mathematical Programming) method is applied in many fields such as mathematics, economy, engineering, geology, management and so on. It refers to a method of selecting a research programme to maximize or minimize the objective function under certain conditions [12]. The purpose of this section is to determine the range of values for each independent variable and optimize bank performance by nonlinear programming.

Specifically, after obtaining the annual nonlinear regression equations, we can determine the expressions of objective functions, then look for constraints on independent variables based on their actual context, which forms an explicit nonlinear programming problem. By solving the nonlinear programming problem we can obtain the optimized ROE of banks.

4.1. Determining Value Range of Influencing Variables

Huang and Wang studied the relationship between the percentage held by the largest shareholder and the company's performance and found that with the increase of the percentage held by the largest shareholder, the company's performance takes on four phrases: increasing, decreasing, increasing and decreasing again. The first increasing phrase is when the percentage held by the largest shareholder is less than 15%. The second increasing phrase is when the percentage held by the largest shareholder is 30% ~75% [13]. Combined with actual situation of the banking industry, the range of equity concentration is set at 30%~75%.

Generally, too high debt to asset ratio will cause insolvency, but the banking industry is mainly operated by absorbing deposits and providing loans, with a high debt to asset ratio level. In view of the particularity of the banking industry, according to the Basel Accord, the capital adequacy ratio of commercial banks cannot be lower than 8%, which means that the debt to asset ratio of banks should be below 92% [25].

There is no clear limit on asset size, so the value range can be determined according to the actual situation of the banking industry. In the 2016-2020 data of the 25 listed banks in this article, the minimum value of asset size is 11.3065 and the maximum value is 17.3224. Therefore, the value range of asset size is set at 11−18.

According to the Commercial Bank Risk Supervision Core Indicators issued by China Banking Regulatory Commission, the non-performing loan ratio shall not be higher than 5%.

Asset turnover ratio reflects the operating capacity of the enterprise's overall assets. There are obvious differences in turnover ratio for different industries, and there is no clear standard. In the banking industry, there may be many long-term loans and non-performing assets, resulting in slow capital turnover. Therefore, asset turnover ratio of the banking industry is generally low. In the 2016-2020 data of the 25 listed banks in this article, the minimum asset turnover ratio is 0.017 and the maximum is 0.0368. Therefore, the value range of asset turnover ratio is set at 0.015−0.040.

Through the discussion of influencing variables above, we come to the following constraints on them:

$$s.t.\begin{cases} 30\% \le EC \le 75\% \\ 0 < DA \le 92\% \\ 11 \le AS \le 18 \\ 0 < NPL \le 5\% \\ 1.5\% \le AT \le 4\% \end{cases} \quad (11)$$

4.2. Solving Nonlinear Programming Problems

After determining the value range of influencing variables, we now are ready to optimize bank performance by taking the nonlinear regression equation of each year as the objective function. Next is a flow of solving nonlinear programming problems for 2016−2020 objective functions, optimal solutions obtained after running R software are shown in corresponding table.

According to (2), objective function of bank performance in 2016 can be formulated as follows:

$$maxROE_{16} = -0.4975 - 0.0649EC_{16}^2 + 0.0642exp(2.3400DA_{16}) + 0.0002AS_{16}^2 - 1.7410NPL_{16} + 2.1916AT_{16} \quad (12)$$

Based on constraint condition (11), we solve nonlinear programming problem and show optimal value of the objective function and corresponding optimal solution in Table 8.

Table 8. Optimal results of nonlinear programming for 2016

Maximum	$\boldsymbol{ROE_{16}}$	**0.2018**
Optimal solution	EC_{16}	0.3000
	DA_{16}	0.9200
	AS_{16}	18
	NPL_{16}	4.0360e-09
	AT_{16}	0.0400

According to (4), objective function of bank performance in 2017 can be formulated as follows:

$$maxROE_{17} = -1.5940 - 0.0716EC_{17}^2 + 1.8360DA_{17} + 0.0001exp(0.3741AS_{17}) - 2.7070NPL_{17} + 1.9220AT_{17} \quad (13)$$

Based on constraint condition (11), we solve nonlinear programming problem and show optimal value of the objective function and corresponding optimal solution in Table 9.

Table 9. Optimal results of nonlinear programming for 2017

Maximum	$\boldsymbol{ROE_{17}}$	**0.2496**
Optimal solution	EC_{17}	0.3000
	DA_{17}	0.9200
	AS_{17}	18
	NPL_{17}	1.6406e-13
	AT_{17}	0.0400

According to (6), objective function of bank performance in 2018 can be formulated as follows:

$$maxROE_{18} = -1.7760 - 0.0507EC_{18}^2 + 0.9397DA_{18} + 0.0003AS_{18}^2 + exp(-4.6263NPL_{18}) + 1.9896AT_{18} \quad (14)$$

Based on constraint condition (11), we solve nonlinear programming problem and show optimal value of the objective function and corresponding optimal solution in Table 10.

Table 10. Optimal results of nonlinear programming for 2018

Maximum	$\boldsymbol{ROE_{18}}$	**0.2607**
Optimal solution	EC_{18}	0.3000
	DA_{18}	0.9200
	AS_{18}	18
	NPL_{18}	9.9885e-11
	AT_{18}	0.0400

According to (8), objective function of bank performance in 2019 can be formulated as follows:

$$maxROE_{19} = -2.3004 - 0.0468EC_{19}^2 + 1.4901DA_{19} + 0.0002AS_{19}^2 + exp(-3.8676NPL_{19}) + 2.2595AT_{19} \quad (15)$$

Based on constraint condition (11), we solve nonlinear programming problem and show optimal value of the objective function and corresponding optimal solution in Table 11.

Table 11. Optimal results of nonlinear programming for 2019

Maximum	$\boldsymbol{ROE_{19}}$	**0.2215**
Optimal solution	EC_{19}	0.3000
	DA_{19}	0.9200
	AS_{19}	17.9992
	NPL_{19}	1.0152e-16
	AT_{19}	0.0400

According to (10), objective function of bank performance in 2020 can be formulated as follows:

$$maxROE_{20} = -0.8046 - 0.0446EC_{20}^2 + 1.0220DA_{20} + 0.0003exp(0.2892AS_{20}) - 4.3086NPL_{20} + 0.4721AT_{20} \quad (16)$$

Based on constraint condition (11), we solve nonlinear programming problem and show optimal value of the objective function and corresponding optimal solution in Table 12.

Table 12. Optimal results of nonlinear programming for 2020

Maximum	$\boldsymbol{ROE_{20}}$	**0.2052**
Optimal solution	EC_{20}	0.3000
	DA_{20}	0.9200
	AS_{20}	17.9984
	NPL_{20}	5.1338e-15
	AT_{20}	0.0400

From optimal results in Tables 8−12, we observe that although there is significant difference in the maximum ROE from 2016 to 2020, the difference in corresponding optimal solutions is very small, or the same, which means that there is no essential difference in the influence of independent variables on objective function. In other words, the difference among nonlinear regression equations is only in form, not in essence, which also reflects the rationality of the regression equations obtained in Section 3.

In order to visually describe the difference between optimized and actual annual weighted average ROE of 25 listed banks, here we draw a graph (Fig. 2) for comparison. In Figure 2, actual weighted average ROE of 25 listed banks in each year are connected by solid lines, and horizontal dashed lines represent the optimized weighted average ROE from 2016 to 2020. It is easy to see that, except for the weighted average ROE of Guiyang Bank in 2016, optimized annual weighted average ROE are all above corresponding actual annual weighted average ROE, indicating that the bank's annual weighted average ROE still has room for improvement.

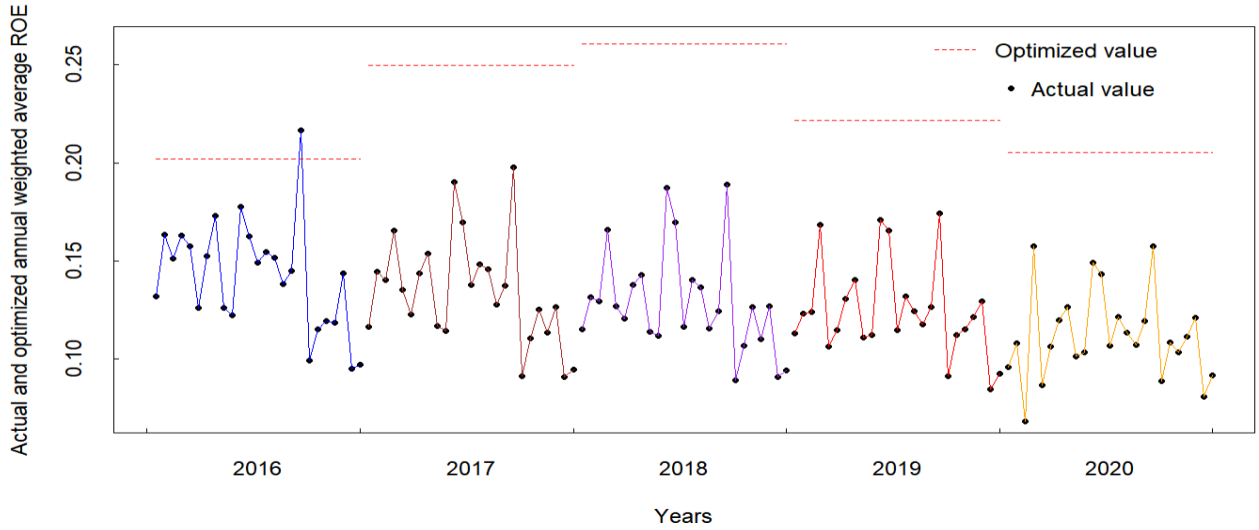

Fig. 2 Comparison of actual and optimized annual weighted average ROE of banks in the five years from 2016 to 2020.

According to the optimized annual weighted average ROE shown in Figure 2, combined with corresponding optimal solution in tables 8−12, banks can make their own performance tend to be optimal by adjusting the values of internal influencing variables. This article builds nonlinear models for the bank's financial data from 2016 to 2020, and obtains the optimized weighted average ROE for five consecutive years through the optimization method, forming a time series, from which the changing direction of the market environment over time can be found. In the future, banks can dynamically adjust internal variables according to the market environment and their own conditions to achieve expected performance.

5. Predicting and Optimizing Future Bank Performance

Predicting bank performance is important for both bank management and investors. After obtaining nonlinear regression models of bank performance from 2016 to 2020, we could predict the future bank performance by applying them. Notice that we also calculate optimized annual weighted average ROE and corresponding independent variable values from 2016 to 2020, which gives a reference for optimizing future bank performance. To show how to predict and optimize future bank performance, we will take four commercial banks in China as examples to illustrate, they are Bank of Ningbo, China Merchants Bank (CMB), China Construction Bank (CCB) and Industrial and Commercial Bank of China (ICBC).

Notice that Chinese commercial banks have released their 2021 annual financial reports, to predict the future bank performance and compare them with the true performance, we first collect key 2021 annual financial data with respect to the four commercial banks, which are listed in Table 13. In addition, to provide a management direction with respect to internal independent variables of banks, 2020 optimal results obtained in Section 4 are attached.

Table 13. Actual and optimal values of variables for four commercial banks in 2021

Variable	Bank of Ningbo	CMB	CCB	ICBC	Optimal results
ROE	0.1663	0.1696	0.1255	0.1215	0.2052
EC	18.74%	18.05%	57.03%	34.71%	0.3000
DA	92.56%	90.64%	91.36%	90.69%	0.9200
AS	14.5164	16.0400	17.2251	17.3757	17.9984
NPL	0.77%	0.91%	1.42%	1.42%	5.1338e-15
AT	0.0262	0.0358	0.0272	0.0268	0.0400

Table 13 suggests, CCB and ICBC have higher NPL compared to Bank of Ningbo and CMB, and they are all higher than the optimal NPL. At the same time, asset turnover ratios of the four commercial banks are lower than the optimal AT. CMB has the highest AT and the largest weighted average ROE. In a word, the performance of the four banks can still be improved by adjusting their internal independent variables.

Inspired by the idea of autoregressive moving average model in time series analysis [20], we consider predicting future bank performance by using fitted performance of the last one year and mean fitted performance of the last three years, which are called one-year predicted performance (OPP) and three-year mean predicted performance (TMPP), respectively. It should be noted, they are calculated by using regression equation (6), (8), (10), in which independent variables are substituted with corresponding actual value in 2021. To compare the predicted effect, we also present Actual value of weighted average ROE (AROE) for four commercial banks. At the same time, we compare them with one-year optimized performance (OOP) for 2020, and three-year mean optimized performance (TMOP), which are calculated according to optimal results of the last three years 2018, 2019 and 2020. They are all shown in Table 14.

Table 14. Comparison of OPP, TMPP, AROE, OOP and TMOP for four commercial banks

Bank name	OPP	TMPP	AROE	OOP	TMOP
Bank of Ningbo	0.1420	0.1537	0.1663	0.2052	0.2291
CMB	0.1337	0.1528	0.1696	0.2052	0.2291
CCB	0.1166	0.1239	0.1255	0.2052	0.2291
ICBC	0.1210	0.1268	0.1215	0.2052	0.2291

From Table 14, we can observe that, in general, the TMPP of the four commercial banks is closer to corresponding actual value than OPP. For a particular bank, the OPP may be better than TMPP, such as ICBC. Therefore, in practice, when predicting bank performance, we should decide according to the specific situation. If the bank performance has an obvious upward or downward trend, we should choose the OPP, and if the bank performance shows a volatile trend, we should choose the TMPP. Obviously, actual performance of the four banks in 2021 has not reached the optimal level, which means that there is still room for improvement.

6. Concluding Remarks

This article takes China's 25 A-share listed banks as the research object, collects their annual financial report data from 2016 to 2020, then establishes nonlinear regression models between the weighted average ROE and main influencing factors. After fitting nonlinear regression equations, we take these nonlinear regression equations as objective functions and solve nonlinear programming problems, then obtain the optimized weighted average ROE for five consecutive years, which provides a reference to optimize future bank performance. To predict the future performance of a bank, we can consider using TMPP or OPP, depending on the situation. At the same time, banks can optimize their own performance by adjusting internal influencing variables.

From the data collected in the annual financial report, the debt to asset ratio of banks is generally between 90% and 95%, and most of them exceed 92%. The asset size of large state owned banks and small city banks is quite different, and the asset turnover ratio is between 1% and 4%, most of which are around 2%. This may be because banks have many long-term loans and non-performing assets, resulting in very slow capital turnover. The higher the debt to asset ratio, the greater the asset size and the asset turnover ratio, the lower the non-performing loan ratio, the better the bank's performance. Banks should appropriately increase the debt to asset ratio, reduce non-performing loans, make internal adjustments to improve their performance on the premise of not exceeding the limit.

To explain fluctuations in bank performance, we identify five influencing variables in this article. Our main purpose is to demonstrate methods of prediction and optimization using these five variables. In future research or practice, to explain bank performance in a more comprehensive or targeted way, scholars can increase or decrease the number of influencing variables.

Although this article only focuses on China's 25 A-share listed banks, the methodology provided is applicable to all industries. As far as the conclusion, it is the common goal of all companies to properly control the equity concentration, avoid excessive concentration or divergence of equity, strengthen risk supervision and management, and improve asset quality and asset operation capabilities. In addition, this article mainly studies bank performance by establishing multivariate nonlinear regression model and optimizing objective function, which provides a research idea for the improvement of bank performance. In this process, researchers can also build linear regression models based on actual data and solve linear programming problems to obtain optimized performance.

Regression analysis studies the causal relationship between economic variables on a static time section. Time series analysis considers the changing laws of variables over time by arranging data in chronological order, and present the fluctuations of the series. Panel data analysis synthesizes the above two analysis methods, since dynamic changes and the relationship between variables on the cross-section are considered at the same time, which increases the difficulty of analysis. This paper does not expand it here, leaving it as a topic for future research.

Conflicts of Interest

The authors declare that there is no conflict of interest regarding the publication of this paper.

References

[1] H. Abdullah, and T. Tursoy, "Capital Structure and Firm Performance: Evidence of Germany Under Ifrs Adoption," *Review of Managerial Science*, vol. 15, pp. 379-398, 2021.

[2] K. S. Adesina, "How Diversification Affects Bank Performance: the Role of Human Capital," *Economic Modelling*, vol. 94, pp. 303-319, 2021.

[3] R. W. Banz, "The Relationship Between Return and Market Value of Common Stocks," *Journal of Financial Economics*, vol. 9, no. 1, pp. 3-18, 1981.

[4] A. N. Berger, and E. B. D. Patti, "Capital Structure and Firm Performance: A New Approach to Testing Agency Theory and An Application to the Banking Industry," *Journal of Banking & Finance*, vol. 30, no. 4, pp. 1065-1102, 2006.

[5] V. Bouvatier, L. Lepetit, and F. Strobel, "Bank Income Smoothing, Ownership Concentration and the Regulatory Environment," *Journal of Banking & Finance*, vol. 41, no. 41, pp. 253-270, 2014.

[6] I. Busta, E. Sinani, and S. Thomsen, "Ownership Concentration and Market Value of European Banks," *Journal of Management & Governance*, vol. 18, pp. 159-183, 2014.

[7] S. Daly, and M. Frikha, "Determinants of Bank Performance: Comparative Study Between Conventional and Islamic Banking in Bahrain," *Journal of the Knowledge Economy*, vol. 8, pp. 471-488, 2017.

[8] I. Ducassy, and A. Guyot, "Complex Ownership Structures, Corporate Governance and Firm Performance: The French Context," *Research in International Business and Finance*, vol. 39, pp. 291-306, 2017.

[9] P. M. Fairfield, and T. L. Yohn, "Using Asset Turnover and Profit Margin to Forecast Changes in Profitability," *Review of Accounting Studies*, vol. 6, pp. 371-385, 2001.

[10] B. B. Francis, I. Hasan, A. M. KÜLlÜ, and M. Zhou, "Should Banks Diversify Or Focus? Know Thyself: the Role of Abilities," *Economic Systems*, vol. 42, pp. 106-118, 2018.

[11] M. M. Hanafi, F. Santi, and Muazaroh, "The Impact of Ownership Concentration, Commissioners on Bank Risk and Profitability: Evidence From Indonesia," *Eurasian Economic Review*, vol. 3, no. 2, pp. 183-202, 2013.

[12] F. S. Hillier, and G. J. Lieberman, *Introduction to Operations Research*, Mcgraw-Hill Education, 2015.

[13] J. Huang, and Z. J. Wang, "A Study on the Relationship Between the No.1 Large Shareholder's Shares Holdings and Company's Performances (in Chinese)," *East China Economic Management*, vol. 21, no. 2, pp. 26-30, 2007.

[14] S. Kendo, and J. Tchakounte, "Impact of Asset Size on Performance and Outreach Using Panel Quantile Regression with Non-Additive Fixed Effects," *Empirical Economics*, vol. 62, pp. 65-92, 2022.

[15] J. Kjosevski, and M. Petkovski, "Macroeconomic and Bank-Specific Determinants of Non-Performing Loans: the Case of Baltic States," *Empirica*, vol. 48, pp. 1009-1028, 2021.

[16] X. H. Liu, J. S. Sun, F. Yang, and J. Wu, "How Ownership Structure Affects Bank Deposits and Loan Efficiencies: An Empirical Analysis of Chinese Commercial Banks," *Annals of Operations Research*, vol. 290, pp. 983-1008, 2020.

[17] A. D. Massis, J. Kotlar, G. Campopiano, and L. Cassia, "Dispersion of Family Ownership and the Performance of Small-to-Medium Size Private Family Firms," *Journal of Family Business Strategy*, vol. 4, no. 3, pp. 166-175, 2013.

[18] J. J. Mcconnell, and H. Servaes, "Additional Evidence on Equity Ownership and Corporate Value," *Journal of Financial Economics*, vol. 27, no. 2, pp. 595-612, 1990.

[19] Q. T. Phung, H. V. Vu, and H. P. Tran, "Do Non-Performing Loans Impact Bank Efficiency?" *Finance Research Letters*, vol. 46, Article 102393, 2022.

[20] R. H. Shumway, and D. S. Stoffer, "Time Series Analysis and Its Applications*:* With R Examples," 4th Ed., *Springer Texts in Statistics*, Springer, Cham, 2017.

[21] C. Stickney, and P. Brown, "Financial Reporting and Statement Analysis," 4th Ed., New York: Harcourt Brace & Co, 1999.

[22] Y. Tan, "Investigating the Performance of Chinese Banks: Efficiency and Risk Features," *Palgrave Macmillan London*, 2016.

[23] A. Trujillo-Ponce, "What Determines the Profitability of Banks? Evidence from Spain," *Accounting and Finance*, vol. 53, pp. 561-586, 2013.

[24] D. Tsuruta, "Leverage and Firm Performance of Small Businesses: Evidence from Japan," *Small Business Economics*, vol. 44, no. 2, pp. 385-410, 2015.

[25] F. Wu, "Analysis on China's Listed Commercial Banks' Capital Structure and Corporate Value─Based on Threshold Panel Model (in Chinese)," *Exploration of Financial Theory*, no. 5, pp. 36-43, 2019.