Journal of Finance and Economics
Volume 4, No. 2 (2016), 46-57
ISSN 2291-4951 E-ISSN 2291-496X
Published by Science and Education Centre of North America

A Comparative Study with Quantile Regression and Back Propagation Neural Network for Credit Rating

Shin-Yun Wang[1*], He-Shun Syu[2]

[1] Department of Finance, National Dong Hwa University, Hualien, Taiwan

[2] Chang Hwa Commercial Bank, Taipei, Taiwan

*Correspondence: Shin-Yun Wang, Department of Finance, National Dong Hwa University. 1, Section 2, University Rd. Shou-Feng, Hualien 974, Taiwan. Tel: 886-3-8633136; E-mail: gracew@mail.ndhu.edu.tw

Received: March 21, 2016 Accepted: April 28, 2016 Online Published: July 28, 2016

DOI: 10.12735/jfe.v4n2p46 URL: http://dx.doi.org/10.12735/jfe.v4n2p46

Abstract

In this study, we use the quantile regression and the back propagation neural network to construct a credit rating model for companies listed in Taiwan Stock Exchange and Over-The-Counter. The data we use is from 1997 to 2013 in Taiwan. The data in the period from 1997 to 2005 is in sample and the data in the period from 2006 to 2013 is out of sample. TCRI established by TEJ is used as a dependent variable to analyze the relationship between 12 financial ratios and credit rating. Our results show that the average forecasting correction rate based on the propagation neural network, which is about 70%, is higher than that based on the quantile regression, which is about 60%. However, investors and financial institution are mainly concerned about the companies facing bankruptcy so they are more interested in which companies bear higher risk. In this case, the quantile regression can provide higher forecasting correction rate for low-credit-ranking companies, which is about 80%, than that provided by the back propagation neural network, which is about 55%.

JEL Classifications: C52, C21

Keywords: credit rating, quantile regression, back propagation neural network

1. Introduction

It is important to forecast the bankruptcy among enterprises due to a lot of crisis punching the market. In 1997, many countries like Indonesia, Thailand, and Korea etc. were facing serious

How to cite this paper: Wang, S. Y., & Syu, H. S. (2016). A comparative study with quantile regression and back propagation neural network for credit rating. *Journal of Finance and Economics*, *4*(2), 46-57. http://dx.doi.org/10.12735/jfe.v4n2p46

bankruptcy and suffering from great loss in the Asian economic crisis. Before the dot-com bubble in 2000, a lot of internet companies had been established even without clear business model for revenue. Until many companies could not pay the loan, the collapse of the bubbles happened. In 2008 financial crisis, global market faced the extreme challenge because the poor quality of the subprime mortgage which obtains investment grade was sold to investors. One year later, the governments of Iceland and Greece had the overload debt and they could not repay for creditors. Above all events, many companies were default or bankruptcy so the credit rating is important to measure whether the company is good or bad.

A proper credit rating model can reflect the default risk of companies accurately. High credit rating score stands for low default risk so the companies with high credit rating scores can hold better reputation and are easier to get financed from the capital market. Due to information asymmetry, it's more difficult for a company without any credit rating to raise fund in the market because investors can't get any information about the company for investment decision making. Therefore, credit rating can not only decrease the information asymmetry to investors but also lower the cost of financing for companies. Furthermore, credit rating can also provide a good reference of the company financial status to strengthen government's supervision and management. This study compares the quantile regression with the back propagation networks (BPN) and tries to conclude the best method for credit rating. With more accurate credit rating, enterprises can adjust their strategy accordingly and investors can also have more sufficient information for their investment policy.

Credit rating receives affirmation in the international community. Moody and S&P (Standard & Poor) are famous in the global and their rating results are used extensively. However, Bear Stearns rated as A by Moody and S&P went bankrupt during 2008 financial crisis so there is still room to improve exiting rating systems. If there is a more effective rating model to reflect the company financial status timely, it could improve market efficiency as well as help the government to strengthen its supervision and management. In these years, there are many methods established for credit rating model, such as linear probability model, Discriminant analysis, Probit model, and Logit model. Recently, neural network which can simulate artificial intelligence was also proposed. However, the proposed approaches usually provide averaging results only and can't provide further information on extreme cases. Therefore, the credit rating model based on these approaches can't provide sufficient information to investors and may mislead them for risk control. This study attempts to use quantile regression approach which was introduced by Koenker and Bassett (1978) to establish credit rating model because quantile regression approach which could provide observations on all cases including extreme ones in different quantiles is more robust to outliers and provides another option to measure credit risk.

In the past, most of studies conclude their observations based on the analysis on single major approach. However, even though there are already a lot of studies on credit rating, it's still difficult for people to judge which approach can provide more authentic results because there is no common sample pool and assumptions for the comparison across different studies. This study compares two approaches, quantile regression and BPN, for credit rating and analyzes the potential impacts when they are used for credit rating. Hopefully, it could shed light on which approach is more reliable for credit rating.

2. Literature Review

The earliest application of ordinary least squares (OLS) to bond rating model can be traced back to the study of Fisher (1959). Altman (1968) utilized multiple discriminant analysis (MDA) in Z-score. Altman selected, from 1946 to 1965, 33 bankruptcy firms as samples and another 33 financially healthy firms as matched samples. 22 accounting ratios were selected to extract liquidity, profitability, leverage, solvency, and activity by MDA. The results showed that $Z = 2.675$ is the critical point. Firms are in the safe zones when $Z > 2.675$ and contrarily firms are in the distress zones when Z < 2.675. Besides, the bankruptcy probability of a company in the distress zones is 95% in one year and 70% in two years. Therefore, the bankruptcy rating accuracy descends with time. Meanwhile, Deakin (1972), Pinches and Mingo (1973), and Blum (1974) also utilized MDA in the models but they chose different accounting ratios in their models. The common result was that the bankruptcy rating for a company based on MDA has high accuracy in one year.

Ohlson (1980) used 105 manufacturing firms which confronted default from 1970 to 1976 and 2058 firms with healthy finance as samples. These studies utilized logistic regression analysis to establish credit rating model for bankruptcy prediction in recent three years. The results showed that the accuracy is 92.84% for three years predicting. Similarly, Dambolena and Khoury (1980) used logit model to analyze the probability of the failure. The results presented that the accuracy is 82.6% for predicting bankruptcy in five years. Zmijewski (1984) used probit model to construct financial distress prediction model. Dutta and Sheckhar (1988) applied neural network for bond rating and the accuracy is 83.3%. Surkan and Singleton (1990) used BPN in credit rating and the accuracy is 88%. Additionally, these studies also showed BPN has better prediction accuracy than MDA. Kim, Weistroffer and Redmond (1993) used linear model, MDA, logit model, and neural network to develop rating model and found that neural network is the most accuracy among all models. Chaveesuk, Srivaree-Ratana and Smith (1999) used logit model, BPN, and support vector machines to establish rating model and found that these models have great accuracy. Huang, Chen, Hsu, Chen and Wu (2003) used Support vector machines and BPN to research the rating of US and Taiwan firms and results showed that the accuracy is approximately 80%. Niemann, Schmidt and Neukirchen (2008) indicated that financial data is not normal distribution. Unless the data is processed by Box and Cox, it can’t be used in statistical models appropriately.

Quantile regression was introduced by Koenker and Bassett (1978). The advantage is that it does not define samples as normal distribution. When the distribution of samples is not a normal distribution, the results of quantile regression could reasonably explain the margin effects in the different quantiles. Kordas (2002) used Binary quantile regression to model firm’s credit. The sample of this study included 300 defaults and 800 non-defaults. The results showed that binary quantile regression is better than probit model. Whittaker, Whitehead and Somers (2005) also used quantile regression to model consumer credit scoring. In recent years, quantile regression is used extensively in the research included stock market, labor market, and medical field, but few credit rating researches used quantile regression. This study attempts to use quantile regression to establish credit rating model and choose appropriate financial variables according to related researches to evaluate company’s credit. The purpose is to use quantile regression to analyze the tails effectively and accurately. Based on the works in above literatures, we choose quantile regression and BPN for comparison to distinguish which model is better and can provide more information to the market.

3. Methodology

The data includes listed and OTC companies issued by Taiwan Economic Journal (TEJ). Because of the specialty, financial industry companies are not included. The research period is from 1997 to 2013. Taiwan Corporate Credit Risk Index (TCRI) issued by TEJ is from 1 to 9. The dependent variable, TCRI, is assigned as 3, 2 and 1, respectively if Taiwan TCRI score is from 1 to 3, from 4 to 5 and from 6 to 9.

3.1. Independent Variables

According to Altman (1968), Ohlson (1980), Huang *et al*. (2003), and Niemann *et al.* (2008), this study uses some financial ratios to estimate credit rating. There are nine variables in the model such as sales over total asset (STA), days-account receivable turnover (TDCP), account receivable turnover (ART), working capital over total asset (WCTA), quick ratio (QR), earnings before interest and tax over total asset (EBITTA), inventory turnover (IT), equity over total liability (ETL), and natural logarithm of total assets (LNTA). They could estimate profitability, return on equity, activity, liquidity, solvency, and leverage.

3.2. Quantile Regression

Quantile regression, which is superior to ordinary least squares (OLS), was proposed by Koenker and Bassett in 1978. Especially for analysis of financial data, which has fat-tail distribution, quantile regression could be used to observe the margin effect of independent variables affecting the dependent variables in each quantile. Since OLS uses normal distribution, which may not be true in every case, to describe the average margin effect, there is limitation to use it for credit rating, especially for extreme value measuring. Quantile regression is based on conditional quantile function; it could obtain the slope of the endogenous variables in different quantiles when given the exogenous variables. Due to distribution-agnostic assumption in residual term, quantile regression is more robust. The method of quantile regression is introduced as follows.

Given ϕth quantile ($\varphi \in (0,1)$), there is a linear relationship between conditional quantile of y_i and x_i, where i=1,…,n, y_i is the credit rating and x_i represents exogenous variable including company's financial variable. Quantile regression could be written as:

$$y_i = x_i'\beta_\phi + u_{\phi i} \tag{1}$$

Under the assumption of equation (1),

$$Quant_\varphi(y_i|x_i) = \inf\{y : F_i(y|x)\varphi\} = x_i'\beta_\varphi \tag{2}$$

$$Quant_\varphi(u_{\varphi i}|x_i) = 0 \, ,$$

where $Quant_\phi(y_i|x_i)$ means the conditional quantile of y_i in ϕth quantile, given x_i. Furthermore, we could estimate $\hat{\beta}_\varphi$ while φ varies from 0 to 1. Besides, there is a special case, median regression when ϕ is 0.5.

$$\mathrm{Min}\sum_{i}^{n} \rho_\varphi(y_i - x_i'\beta_\varphi) \, , \tag{3}$$

where $\rho_\varphi(y_i - x_i'\beta_\varphi)$ is an ABS (absolute value) function and could be defined as

$$\rho_\varphi(u) = \varphi u \ \text{ if } \ u \ge 0 \ \text{ or } \ \rho_\varphi(u) = (\varphi - 1)u \ \text{ if } \ u < 0 \tag{4}$$

$$\hat{\beta}_\varphi = \min\left[\sum_{u \ge 0} \varphi \left|y_i - x_i'\hat{\beta}\right| + \sum_{u<0} (1-\varphi)\left|y_i - x_i'\hat{\beta}\right|\right] \tag{5}$$

If $\varphi = 0.5$, equation (5) could be rewritten as $\hat{\beta}_\varphi = \min\left[0.5 \cdot \sum_u \left|y_i - x_i'\hat{\beta}\right|\right]$, which is LAD (Least absolute deviation) estimator. This regression is called median regression which is a special case in 0.5 quantile. The basic concept of quantile regression gives estimators different weight in different quantile. Therefore, when there are extreme values which exist in the tail, quantile regression model is more robust than OLS. This study will use quantile regression to examine the relationship between credit rating and financial variables. From the above description, it is obvious that the quantile regression considers the conditional probability distribution of explained variables. In order to use quantile regression to analyze data, the dependent variables must be ordered. When the variable of credit rating is ordered from low to high, low quantile stands for poor credit and high quantile stands for preferred credit.

3.3. Neural Network

Neural network is a kind of model which equips the structure of the biological neural network. NN uses a bulk of artificial neuron to calculate. In most of cases, NN could change internal structure based on the external information. NN is not a linear statistic model and usually used to model the complex relationship between input and output. Due to the weakness of depending historical financial data for credit rating, the outcomes must be modified by artificial concept. Hence, NN model could imitate people's consideration for credit rating model development.

3.3.1. Base Model

In the sample NN model, biological neuron obtains the information from outside and the information is handled by neural nucleus. The processing procedure gives the messages different weights based on their importance. When transferring through the function of artificial neuron, we can get the output in equation (6).

$$y = f(\sum_{i=1}^{n} w_i \cdot x_i - u) \tag{6}$$

w : weight; u : bias; f : transfer function.

3.3.2. Basic Structure

The basic structure of NN equips multi-layer structure and multilayer feedforward network, which is composed by input layer, hidden layer, and output layer, is the most popular one. Different layers are connected with each other. Input layer is responsible for gaining a bulk of non-linear messages. Output layer produces outcomes by transferring, analyzing, and weighting data. Hidden layer which connects many neurons in each layer lies between input layer and output layer. There could be more than one hidden layers but according to Zhang, Patuwo and Hu (1998), single hidden layer is already sufficient to describe complex linear relationship of the data in the model. In this study, multilayer feedforward network is applied with single hidden layer and there are research variables and financial variables in the input layer. In the hidden layer, different weights are applied to the variables which have no relationship with each other. Finally, the output layer presents the outcome.

3.3.3. Back Propagation Neural Network (BPN)

BPN, a kind of multilayer feedforward network, can enhance the prediction accuracy through the process of supervised learning. Supervised learning revises the target value by adjusting the weights repeatedly until the bias is close to zero. Although the prediction accuracy of BPN is high, there are some shortcomings. Its learning process could take a long time to converge, e.g. hundreds or thousands cycles. In addition, there are no clear rules to determine the number of applied neurons in hidden layer and how to set the learning speed. So far the best way is try-and-error to find the best setting for the outcome.

3.3.4. Predication Accuracy Calculation

In order to calculate the prediction accuracy, classification of the prediction values, t, is needed. A cut-off score, t_c, is used for the classification in this study. When $t > t_c$, the prediction value is assigned to a group. Contrarily, the prediction value is assigned to another group. The determination of t_c is shown in equation (7).

$$t_c = \frac{(n_2\overline{t_1} + n_1\overline{t_2})}{(n_1 + n_2)} \tag{7}$$

The cut-off score considers the different sample number in each group and assign each group different weights to reduce the error of category. Finally, this study uses the confusion table to present the outcomes. Confusion table is mainly used for the prediction accurate rate calculation.

4. Empirical Results

This part includes descriptive statistics, analytical results of quantile regression, and analytical results of BPN. First, the descriptive statistics are used to demonstrate the tendency difference between in-sample period (1997~2005) and out-of-sample period (2006~2013). Second, the results of quantile regression are analyzed to verify whether the tendency of the dependent and independent variables matches the intuition. Then how we establish BPN and its basic structure is described. Finally, a confusion table is used for the prediction accuracy comparison between quantile regression and BPN.

4.1. Descriptive Statistics

In-sample includes listed and OTC companies in Taiwan from 1997 to 2005. Out-of-sample includes listed and OTC companies in Taiwan from 2006 to 2013. On the other hand, financial and insurance industries are more specific, so this study excludes them. Total sample of in-sample are 6490. On the other hand, Total sample of out-of-sample are 8938.

Table 1. The descriptive statistics of in-sample period

	Samples	Average	Median	Minimum	Maximum	STDEV.
RETA	6490	0.0545	0.0595	-1.9820	0.6560	0.1223
ROE	6490	0.0746	0.0824	-1.7854	3.1782	0.1580
EPS	6490	1.4370	1.0900	-13.0200	31.6700	2.3993
STA	6490	0.9041	0.7798	0.0116	5.3309	0.5785
TDCP	6490	74.9140	70.3800	0.3000	970.1500	43.6880
ART	6490	9.6922	5.1900	0.3800	1225.2000	34.2810
WCTA	6490	0.2093	0.1992	-0.4044	0.8417	0.1887
QR	6490	135.0700	100.4600	0.7500	3094.0000	135.7100
EBITTA	6490	0.0670	0.0637	-0.7023	0.5830	0.0874
IT	6490	9.4390	5.0800	0.0400	976.5000	31.9850
ETL	6490	1.8260	1.2894	0.0185	47.0530	1.9374
LNTA	6490	14.9510	14.7690	11.7530	20.0290	1.3119

Table 2. The descriptive statistics of out-of-sample period

	Samples	Average	Median	Minimum	Maximum	STDEV.
RETA	8938	0.0310	0.0631	-13.2930	0.6369	0.3553
ROE	8938	0.0539	0.0779	-7.9104	1.1851	0.2540
EPS	8938	1.7989	1.2600	-18.0700	83.0900	3.4929
STA	8938	0.9993	0.8626	0.0038	8.9789	0.6695
TDCP	8938	72.4360	67.5400	0.0300	1046.9000	43.7340
ART	8938	16.2160	5.4000	0.3500	11908.0000	199.5700
WCTA	8938	0.2812	0.2760	-0.5260	0.9922	0.2060
QR	8938	185.9300	120.5800	0.5000	15741.0000	342.4800
EBITTA	8938	0.0568	0.0598	-2.0996	1.0065	0.1084
IT	8938	47.8030	5.3700	0.0100	82264.0000	1098.6000
ETL	8938	2.2730	1.4143	0.0177	65.6130	3.1177
LNTA	8938	15.1600	14.9710	10.7490	21.2720	1.4149

From table 1 and table 2, we can observe that TDCP (days-account receivable turnover), ART (account receivable turnover), QR (quick ratio) and IT (inventory turnover) have large variation between in-sample period and out-of-sample period. ART and IT have a right-skewed tendency while the retained earnings of companies and ROE have a decreasing tendency. When companies have higher WCTA (working capital over total asset) and QR (quick ratio), companies are more conservative. In addition, the leverage and ETL (equity over total liability) have a rising tendency. This means that companies has lower debt and raises equity in finance.

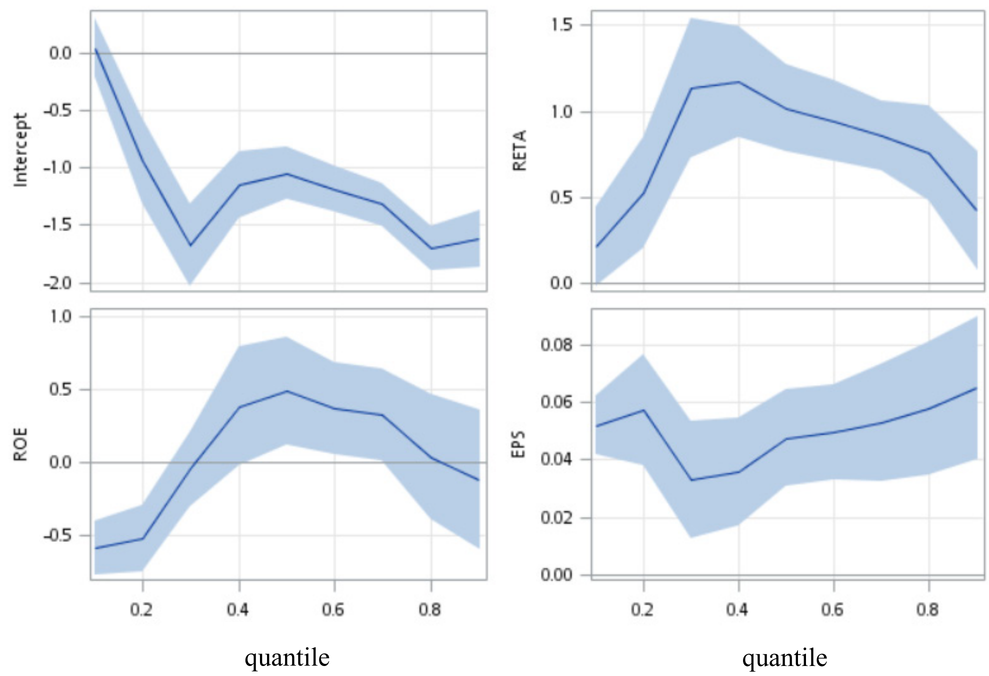
4.2. Analytical Results of Quantile Regression

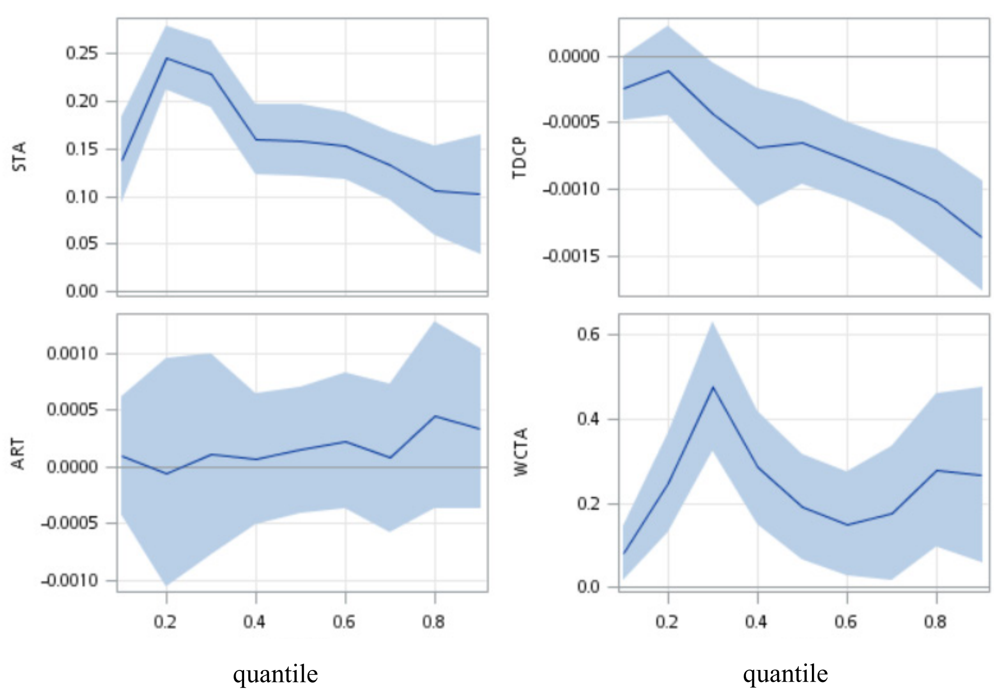
Table 3. The result of quantile regression

φ	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
const	0.040	-0.945**	-1.678**	-1.147**	-1.05**	-1.188**	-1.321**	-1.698**	-1.620**
	0.30	-5.01	-9.22	-7.77	-9.02	-11.77	-13.68	-17.14	-12.54
RETA	0.212	0.528**	1.135**	1.170**	1.018**	0.943**	0.860**	0.753**	0.422*
	1.79	3.19	5.49	7.11	7.75	7.78	8.32	5.31	2.39
ROE	-0.588**	-0.524**	-0.049	0.380**	0.486**	0.368*	0.327*	0.036	-0.122
	-6.15	-4.46	-0.37	1.82	2.59	2.31	2.02	0.16	-0.49
EPS	0.052**	0.057**	0.033**	0.036**	0.048**	0.050**	0.053**	0.058**	0.065**
	9.80	5.82	3.19	3.73	5.55	5.89	5.13	4.92	5.14
STA	0.137**	0.2454**	0.228**	0.160**	0.159**	0.152**	0.133**	0.106**	0.102**
	5.97	14.36	12.68	8.33	8.35	8.45	7.17	4.44	3.15
TDCP	-0.000*	-0.000	-0.000*	-0.001**	-0.001**	-0.001**	-0.001**	-0.001**	-0.001**
	-2.01	-0.63	-2.24	-3.01	-4.08	-5.27	-5.79	-5.40	-6.44
ART	0.000	-5.305	0.000	0.000	0.000	0.000	0.000	0.001	0.000
	0.38	-0.1	0.24	0.24	0.52	0.74	0.23	1.08	0.93
WCTA	0.080*	0.247**	0.477**	0.285**	0.191**	0.150*	0.176*	0.278**	0.266*
	2.4	4.16	6.03	4.12	2.96	2.37	2.16	2.97	2.48
QR	0.001**	0.001**	0.001**	0.000**	0.000*	0.001**	0.001**	0.001**	0.000*
	9.79	4.69	3.30	2.46	2.39	3.10	2.94	2.90	1.99
EBITTA	1.628**	1.784**	1.310**	0.590*	0.171	0.365	0.439	0.892**	1.3102**
	6.32	6.10	4.81	1.97	0.63	1.53	1.78	2.77	3.70
IT	0.000	0.001	0.000	0.000	0.000	0.000	0.001	0.000	0.0005
	1.24	2.02	1.64	1.15	1.66	1.07	0.73	0.32	0.36
ETL	0.019**	0.032**	0.029**	0.035**	0.035**	0.035**	0.036**	0.027**	0.039**
	3.07	3.32	2.65	4.37	4.84	4.24	3.2	2.63	3.01
LNTA	0.050**	0.116**	0.183**	0.171**	0.174**	0.191**	0.208**	0.246**	0.258**
	6.8	9.54	16.00	20.47	25.81	30.79	34.97	41.54	34.67

This table describes the result of quantile regression including financial variables against TCRI. This study divides 9 quantiles, ϕ is from 0.1 to 0.9.**5% significance level *1% significance level.

Table 3 shows the analytical results of quantile regression using 6490 samples from 1997 to 2005 for nine quantile values with φ = 0.1, 0.2,…, 0.9. The results show that most of the variables have a positive effect on the credit rating at 5% significance level and it's consistent with our expectation. However, the ROE have a negative effect on credit rating for 0.1 and 0.2 quantile at 1% significance level. This means that the market would request higher return when companies bear higher credit risk. In addition, although ART and IT are not significant for each quantile at 5% significance level, the regression results show that they have a positive effect on credit rating and it's also consistent with our expectation.





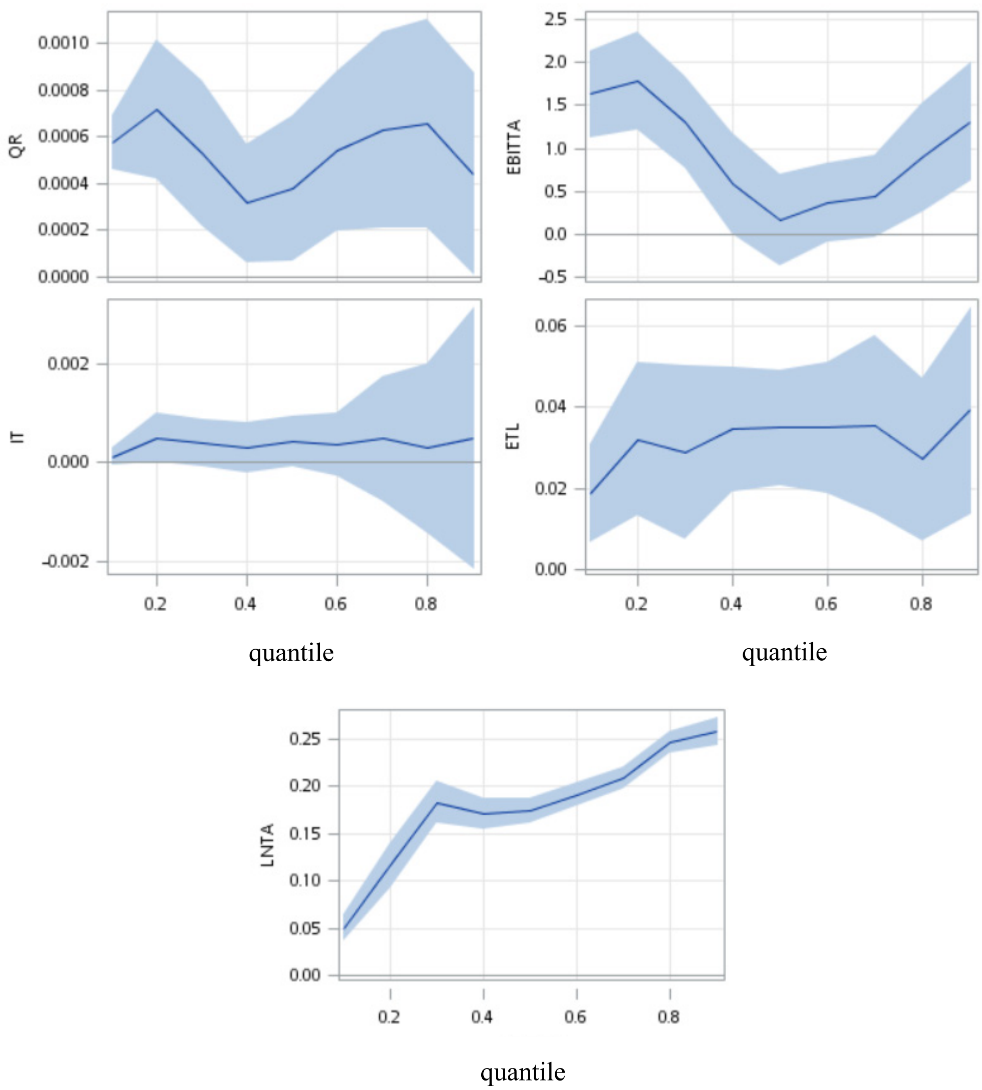


Figure 1. The graph of quantiles and the intercept of variables (contd.)

From Figure 1, we could observe that the regression coefficient of ROE increases from negative to positive with the quantile value from 0.1 to 0.9. This means that ROE has a negative effect on credit rating when the credit rating is poor and ROE has a positive effect on credit rating when the credit rating is good. Although the regression coefficient of ROE is negative in 0.3 and 0.9 quantile, it is not significant. For ART, the graph shows that the regression coefficient increases with the quantile value from 0.1 to 0.9. This means that the marginal effect of ART on credit rating grows with ART. For the scale of enterprise, the regression coefficient of LNTA grows larger and larger with the quantile value from 0.1 to 0.9. This means that the marginal effect of the enterprise scale on credit rating grows with the enterprise scale.

4.3. Analytical Results of Back Propagation Neural Network

Due to the learning process, BPN would modify the parameters of the model based on new input data continuously to achieve the best modeling accuracy. Hence, when establishing BPN model, we need to set initial values for the parameters and then try-and-error for the best condition. In this study, there are 12 variables in the input layer and there is single hidden layer with 10 neurons. In the output layer, credit rating is used. Bayes rule method is applied in the training function to enhance the generalization ability of the network as well as shorten the learning time. Batch gradient descent with momentum algorithm is applied in the learning function so that it could response the local gradient change and the latest modeling error tendency.

Table 4 compares the analytic results of quantile regression and BPN. It shows that quantile regression has better prediction accuracy in the group of $Y_i = 1$ in the quantile 0.1 and 0.2. The accuracy is 97.46% and 89.49% for the in-sample data, and 93.23% and 78.62% for the out-of-sample data. Although the total accuracy is low in the quantile 0.1 and 0.2, quantile regression has better capability to predict which company is a bad company in this range. In the quantile from 0.3 to 0.7, the accuracy is good in the group of $Y_i = 2$. The best accuracy of 97.07% and 94.90% is achieved in the quantile 0.5 and 0.4 for in-sample data and out-of-sample data, respectively. The total accuracy is from 58% to 62% for in-sample data and from 61% to 66% for out-of-sample data. In the quantile 0.8 and 0.9, quantile regression has better prediction ability in the group of $Y_i = 3$ with the accuracy of 65.85% and 88.24% for the in-sample data and 87.85% and 97.91% for the out-of-sample data. This means quantile regression has better capability to predict which company is good company. Finally, BPN has better prediction accuracy in the group of $Y_i = 2$ with the accuracy of 93.84% for in-sample data and 93.2% for out-of-sample data. Although its prediction accuracy is low in the groups of $Y_i = 1$ and $Y_i = 3$, the total accuracy is the best with 71.2% for in-sample data and 75.02% for out-of-sample data.

Table 4. The quantile regression compares with BPN

model	TCRI / Sample	$Y_i = 1$	$Y_i = 2$	$Y_i = 3$	Total
0.1	In-sample	97.46%	20.11%	0.60%	37.15%
	Out-of-sample	93.23%	32.30%	1.69%	43.53%
0.2	In-sample	89.49%	55.64%	2.23%	53.81%
	Out-of-sample	78.62%	72.57%	5.52%	61.37%
0.3	In-sample	56.76%	86.14%	6.62%	61.66%
	Out-of-sample	52.91%	90.82%	16.34%	65.83%
0.4	In-sample	26.06%	96.62%	11.01%	59.66%
	Out-of-sample	35.52%	94.90%	23.49%	63.63%
0.5	In-sample	16.51%	97.07%	17.93%	58.72%
	Out-of-sample	22.73%	94.09%	35.41%	62.74%
0.6	In-sample	12.89%	96.30%	28.13%	59.45%
	Out-of-sample	17.28%	91.70%	49.13%	62.59%
0.7	In-sample	10.45%	94.02%	40.70%	61.78%
	Out-of-sample	13.50%	85.85%	64.83%	61.47%
0.8	In-sample	4.97%	87.68%	65.85%	60.62%
	Out-of-sample	9.36%	72.21%	87.85%	57.55%
0.9	In-sample	1.64%	70.27%	88.24%	55.29%
	Out-of-sample	4.86%	48.47%	97.91%	45.73%
BPN	In-sample	58.00%	93.84%	31.80%	71.20%
	Out-of-sample	54.44%	93.20%	55.35%	75.02%

Note: This table presents the accuracy of quantile regression and BPN with three groups

5. Conclusion

This study uses financial variables to establish credit rating model and compares quantile regression with BPN. The data includes listed and OTC companies in Taiwan from 1997 to 2013 but companies in finance and insurance industry are excluded due to the specialty. TCRI, presented in TEJ, is the credit rating variable as dependent variable. According to the references, 12 financial variables are chosen as independent variables to construct credit rating model in this study. The prediction accuracy of quantile regression, 93.23% and 78.62% for out-of-sample data, is better than BPN for companies with poor credit rating, $Y_i = 1$, in the quantile 0.1 and 0.2. In addition, the accuracy of quantile regression, 87.85% and 97.91% for out-of-sample data, is better than BPN for companies with good credit rating, $Y_i = 3$, in the quantile 0.8 and 0.9. However, the total accuracy of BPN is better than quantile regression and has the best prediction accuracy for the median credit group of $Y_i = 2$ among three credit groups.

In the quantile from 0.3 to 0.7, quantile regression has similar credit prediction capability to BPN with good accuracy for the group of $Y_i = 2$ and they both have worse credit prediction capability for the group of $Y_i = 1$ and $Y_i = 3$ among three credit groups. BPN has about 70% total accuracy which is higher than the total accuracy, 60%, of quantile regression. Hence, BPN is better than quantile regression in total accuracy. However, investors sometimes care more about high-risk companies rather than low-risk companies because they may suffer from great investment loss due to high-risk companies. Therefore, the main purpose of credit rating is to identify high-risk companies to prevent investors from great loss and provide guidance to the government for supervision policy adjustment. In this case, quantile regression is a better credit rating model than BPN even though the total accuracy could be sacrificed.

References

[1] Altman, E. I. (1968). Financial ratios, discriminate analysis and the prediction of corporate bankruptcy. *Journal of Finance*, *23*(4), 589-609. doi:10.1111/j.1540-6261.1968.tb00843.x

[2] Blum, M. (1974). Failing company discriminant analysis. *Journal of Accounting Research*, *12*(1), 1-25. doi:10.2307/2490525

[3] Chaveesuk, R., Srivaree-Ratana, C., & Smith, A. E. (1999). Alternative neural network approaches to corporate bond rating. *Journal of Engineering Valuation and Cost Analysis*, *2*(2), 117-131.

[4] Dambolena, I. G., & Khoury, S. J. (1980). Ratio stability and corporate failure. *Journal of Finance*, *35*(4), 1017-1026. doi:10.1111/j.1540-6261.1980.tb03517.x

[5] Deakin, E. B. (1972). A discriminant analysis of predictors of business failure. *Journal of Accounting Research*, *10*(1), 167-179. doi:10.2307/2490225

[6] Dutta, S., & Shekhar, S. (1988). Bond rating: A non-conservative application of neural networks. *Proceedings of IEEE International Conference on Neural Networks*, *2*, 443-450. doi:10.1109/ICNN.1988.23958

[7] Fisher, L. (1959). Determinants of risk premiums on corporate bonds. *Journal of Political Economy*, *67*(3), 217-237.

[8] Huang, Z., Chen, H., Hsu, C. J., Chen, W. H., & Wu, S. (2003). Credit rating analysis with support vector machines and neural networks: A market comparative study. *Decision Support Systems*, *37*(4), 543-558. doi:10.1016/S0167-9236(03)00086-1

[9] Kim, J. W., Weistroffer, H. R., & Redmond, R. T. (1993). Expert systems for bond rating: A comparative analysis of statistical, rule-based and neural network systems. *Expert Systems*, *10*(3), 167-172. doi:10.1111/j.1468-0394.1993.tb00093.x

[10] Koenker, R. W., & Bassett, G. (1978). Regression quantiles. *Econometrica*, *46*(1), 33-50. doi:10.2307/1913643

[11] Kordas, G. (2002). Credit scoring using binary quantile regression. In Y. Dodge (Ed.), *Statistical data analysis based on the L_1-Norm and related methods* (pp.125-137). Switzerland: Springer Basel AG. doi:10.1007/978-3-0348-8201-9

[12] Niemann, M., Schmidt, J. H., & Neukirchen, M. (2008). Improving performance of corporate rating prediction models by reducing financial ratio heterogeneity. *Journal of Banking and Finance*, *32*(3), 434-446. doi:10.1016/j.jbankfin.2007.05.015

[13] Ohlson, J. A. (1980). Financial ratio and the probabilistic prediction of bankruptcy. *Journal of Accounting Research*, *18*(1), 109-131. doi:10.2307/2490395

[14] Pinches, G. E., & Mingo, K. A. (1973). A multivariate analysis of industrial bond rating. *Journal of Finance*, *28*(1), 1-18. doi:10.1111/j.1540-6261.1973.tb01341.x

[15] Surkan, A. J., & Singleton, J. C. (1990). *Neural networks for bond rating improved by multiple hidden layers*. Paper presented at the Proceedings of the IJCNN International Joint Conference on Neural Networks, 17-21 June. San Diego, CA, USA. doi:10.1109/IJCNN.1990.137709

[16] Whittaker, J., Whitehead, C., & Somers, M. (2005). The Neglog transformation and quantile regression for the analysis of a large credit scoring database. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, *54*(5), 863-878. doi:10.1111/j.1467-9876.2005.00520.x

[17] Zhang, G., Patuwo, B. E., & Hu, M. Y. (1998). Forecasting with artificial neural networks: The state of the art. *International Journal of Forecasting*, *14*(1), 35-62. doi:10.1016/S0169-2070(97)00044-7

[18] Zmijewski, M. E. (1984). Methodological issues related to the estimation of financial distress prediction models. *Journal of Accounting Research*, *22*, 59-82. doi:10.2307/2490859