

Business Bankruptcy Prediction Based on Survival Analysis Approach

Ming-Chang Lee

National Kaohsiung University of Applied Science, Taiwan

ABSTRACT

This study sampled companies listed on Taiwan Stock Exchange that examined financial distress between 2003 and 2009. It uses the survival analysis to find the main indicators which can explain the business bankruptcy in Taiwan. This paper uses the Cox Proportional Hazard Model to assess the usefulness of traditional financial ratios and market variables as predictors of the probability of business failure to a given time. This paper presents empirical results of a study regarding 12 financial ratios as predictors of business failure in Taiwan. It showed that it does not need many ratios to be able to anticipate potential business bankruptcy. The financial distress probability model is constructed using Profitability, Leverage, Efficiency and Valuation ratio variables. In the proposed steps of business failure prediction model, it used detail SAS procedure. The study proves that the accuracies of classification of the mode in overall accuracy of classification are 87.93%.

KEYWORDS

Business Failure prediction; Survival Analysis; Cox Proportional Hazard model; Logistic model

1. INTRODUCTION

Business Failure Prediction (BFP) models are estimation of the bankruptcy probability of a firm using a set of covariates, such as financial ratios, Capital turnover, Capital turnover, etc [77]. In past decades, BFP has been a topic research for business and corporate organizations. Investors or creditors, borrowing organizations and governments are increasing interest to predict of corporate bankruptcy [26]. BFP help to avoid lending to (or investing in) business likely to fail, early identification of failing business by regulatory bodies, and more accurate scoring models for rating agencies.

Bankruptcy prediction models use statistical analysis and data mining technique to enhance the decision support tool and improve decision making [68]. Statistical business failure prediction models attempt to predict the business' failure or success. The Multiple discriminant analysis (MDA) has been the most popular approaches, but there need a large number of alternative techniques available ([18], [37], [42]). Such as the data mining techniques include decision tree, neural networks (NNs), support vector machine (SVM), fuzzy system, rough set theory, genetic algorithm (GA) [68]. Various researches have demonstrated the artificial intelligence (AI) techniques such as artificial neural networks (ANNs) can serve as a useful tool bankruptcy prediction [61]. Back propagation neural network (BPNN) was used bankruptcy prediction. Before that BPNN some of the techniques followed such as k-nearest neighbor and the tree

algorithm (ID3) but offered better predictive compare than compared models. Multivariate cumulative sum (CUSUM) is a sequential procedure to predict a business' tendency towards failure.

A survival analysis (SA) technique is the term applied to a dynamic statistical tool used to analysis the time till a certain event [18]. SA uses the Cox proportional hazard model to analysis survival probability and failure times; it is one dynamic model approach [53]. SA techniques have used to examine the drivers behind the survival of Internet business ([29], [30]). Discriminant analysis (DA) and Logit analysis (LA) were found to be slightly superior predictors to the Cox proportional hazard model [27]. Nevertheless, Laitinen and Luoma [33] argued that the SA approach was more natural, flexible, and appropriate and used more information in Business Failure prediction. Keasey et al. [31] also recommended that SA techniques be used in BFP. Yap et al. [69] use financial ratio and logistic regression for evaluating company failure in Malaysia.

The models of enterprise credit risk modes include statistical model, neural network, learning vector, soft-computing, and hybrid models. Table 1 denoted as enterprise credit risk model. Enterprise credit risk evaluation models in this study are Neural networks, Bayesian classifier, Discriminant analysis, Logistic regression, K-nearest neighbor, Decision tree, Case base reasoning, Support vector machine, Software computing, Fuzzy rule-based system, Generic algorithms, Grey relation, and Hybrid models.

Table 1: Enterprise credit risk evaluation models

Category	Area	Some Approach	
Statistical model	Parametric Statistical Method	1. Discriminant analysis 2. Linear multi discriminant analysis 3. Logistic regression 4. Bayesian risk Discriminant analysis	Altman [3]; Ohlson [47]; Yap et al. [69]; Stefanescu et al. [56]; Tabachnick and Fidell [59]; Liang and Xin [36]
	Non-Parametric Statistical Method	1. K- nearest neighbor 2. Cluster analysis	Ince, and Aktan [24]; Islam et al. [25]; Lau [34]; Sun and LI [58]
Neural Networks	Machine learning	1. Multilayer perception 2. Back propagation 3. Radial function neural network 3. Probabilistic neural network 4. Self-organized competition	Islam et al. [25; Chen [13]; Lopez [39]; Mues et al. [44]; Sarkar and Sriram [50]; Stefanescu et al. [56]; Tam and Kiang [60] ; Chen [12]; Odom and Sharda [46]
Learning vector	Machine learning	Support Vector Machine	Zhou et al. [72]; Chen et al. [11] ; Kim and Sohin [32]; Shin et al. [52]
Soft-computing	Reduction attributes	1. Rough sets of reduction knowledge 2. Grey relational of reduction knowledge	Dimitras et al. [17]; Cheng et al. [14]; Bian and Mazlack [6]; Hu [23]; Liu et al.[38] ; Tung et al. [62];

		<ol style="list-style-type: none"> 3. Genetic algorithm of reduction knowledge 4. Fuzzy-Rough Sets Approach 	Wei et al. [64]; Wong et al. [65]; Zhao [72]; Xinhui and Zhong [67]
Survival analysis (SA)	Time to event data analysis	<ol style="list-style-type: none"> 1. Credit risk modeling based on SA 2. Corporate credit risk and the macro economy 	Stepanova, and Thomas [57]; Antonakis and Sfakianakis [4]; Cao et al. [8]; Sohn et al. [55]
Hybrid models	Combination of two or more methods	<ol style="list-style-type: none"> 1. Rough - K Nearest Neighbor 2. Rough Sets – Neural Network 3. Fuzzy-Rough Sets - Nearest Neighbor 4. Fuzzy- Nearest Neighbor 5. Support Vector Machine with Nearest Neighbor 6. GA-based neural network approach 7. Ant Colony Algorithm based on quick-reduct algorithm 	Tung et al. [62]; Wang et al. [63]; Wei and Zhang[64]; Wong et al. [65]; Xiao et al. [66]; Chaduhuri and De [10]; Tam and Kiang [60]; Yu et al. [70]; Zhang et al. [71]; Zhou and Bai [73]; Zhou et al. [75]

The most useful benefits to SA are: (1) In the modeling process, SA is able to take time-varying variables into account [22]. This is done through proportional hazard models [5]. (2) SA is not restricted by the assumption that the distributions of the variables in the data need to be normal [54]. (3) SA only produces positive predictions of time [21]. The time-varying has the potential to not follow a normal distribution. It needs to be positive in predictions and is influenced by time-varying variables.

The major contribution of SA methods is estimation procedures that consider changes in the value of covariates over time [35]. Thus, SA approaches to BFP different from the other approaches mentioned above [18].

3. MERHODOLOGY

3.1 Logit model

In setting up the logistic regression model, first establish the fundamental model for any multiple regression analysis. The outcome variable is assumed as a linear combination of a set of predictors. If outcome variable is Y, and a set of n predictor variables are X_1, X_2, \dots, X_n , the Logit model is: [1]

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_n X_n + \varepsilon = \beta_0 + \sum_{j=1}^n \beta_j X_j + \varepsilon \tag{1}$$

Where β_0 is the expected value of Y when X 's set 0.

β_j is the regression coefficient for each corresponding predictor variable X_j .

ε is the error of the prediction.

Defines $\pi(x)$ as the probability that Y = 1. Similarly, $1 - \pi(x)$ is the probability that Y = 0. These probabilities are written in the following form:

$$\begin{aligned} \pi(x) &= P(Y = 1 | X_1, X_2, \dots, X_n) \\ 1 - \pi(x) &= P(Y = 0 | X_1, X_2, \dots, X_n) \end{aligned} \quad (2)$$

This model for the natural logarithm of the $\frac{\pi(x)}{1 - \pi(x)}$ is:

$$\ln \frac{P(Y = 1 | X_1, X_2, \dots, X_n)}{1 - P(Y = 1 | X_1, X_2, \dots, X_n)} = \ln \frac{\pi(x)}{1 - \pi(x)} = \beta_0 + \sum_{j=1}^n \beta_j X_j \quad (3)$$

Using the inverse of the Logit transformation of (3), it obtains at the following:

$$P(Y = 1 | X_1, X_2, \dots, X_n) = \frac{e^{\beta_0 + \sum_{j=1}^n \beta_j X_j}}{1 + e^{\beta_0 + \sum_{j=1}^n \beta_j X_j}} = \frac{1}{1 + e^{-\left(\beta_0 + \sum_{j=1}^n \beta_j X_j\right)}} \quad (4)$$

Thus, (4) is a logistic regression model, the conditional mean is between 0 and 1.

Now, it will fit the logistic regression model to the data.

Firstly, it must establish a technique for estimating the parameters. The maximum likelihood is the method of parameter estimation in logistic regression model. This method constructs the likelihood function, which expresses the probability of the observed data as a function of the unknown parameters. This process will have selected the estimators in (4). For a set of observations in the data (x_i, y_i) , the contribution to the likelihood function is $\pi(x_i)$, where $y_i = 1$, and $1 - \pi(x_i)$, where $y_i = 0$. The following equation results for the contribution to the likelihood function for the observation (x_i, y_i) is $\zeta(x_i)$:

$$\zeta(x_i) = \pi(x_i)^{y_i} [1 - \pi(x_i)]^{1 - y_i} \quad (5)$$

The observations are assumed to be independent of each other so it can multiply their likelihood contributions to obtain the complete likelihood function $l(B)$. The result is given in (6).

$$l(B) = \prod_{i=1}^n \zeta(x_i) \quad (6)$$

Where B is the collection of parameters $(\beta_0, \beta_1, \dots, \beta_n)$ and $l(B)$ is the likelihood function of B .

Maximum likelihood estimates (MLE's) can be obtained by calculating the B which maximizes $l(B)$. However, to simplify the mathematics, from the logarithm of (6) before finding the value which maximizes the likelihood function. As shown in (7). $L(B)$ is denoted the log likelihood expression.

$$L(B) = \ln[l(B)] = \sum_{i=1}^n (y_i \ln[\pi(x_i)] + (1-y_i) \ln[1 - \pi(x_i)]) \quad (7)$$

It employs the techniques of calculus to determine the value of B based on maximum of $L(B)$. This is done by differentiating (3) with respect to $\beta_0, \beta_1, \dots, \beta_n$ and setting the resulting derivatives equal to zero. These equations are called likelihood estimations, and there is $n + 1$ equation. They are of the following form:

$$\sum_{i=1}^n |y_i - \pi(x_i)| = 0, \text{ for the intercept } \beta_0, \text{ and } \sum_{i=1}^n x_{ik} [y_i - \pi(x_i)] = 0, \text{ for the predictor variables, } \beta_1, \dots, \beta_n.$$

The solution can be solving by using computer programs such as SAS or SPSS. It performs the logistic regression analysis of the data for this study and will calculate the maximum likelihood estimates.

3.2 Cox's PH model [15]

According the assumption about the relationship between the hazard (or survival) function and the set (vector) of explanatory variables (X), there have various models. Thus, the general regression function can be written as $h(t) = g(t, X^T \beta)$, where X^T is the transpose of X . β is the vector of explanatory variable coefficients. In SA models, it is customary to estimate the hazard rate, and then derive the survival rate are required by using regression model. Two main types of regression models are in SA model. These types are the proportional hazards (PH) and accelerated failure time (AFT) models, both of which have fully parametric and semi-parametric version. A parametric regression model based on the exponential distribution:

$$\log_e h_i(t) = \alpha + \beta_1 x_{i1} + \beta_2 x_{i2} + \dots \beta_k x_{ik} \quad (8)$$

Or equivalently,

$$h_i(t) = \exp(\alpha + \beta_1 x_{i1} + \beta_2 x_{i2} + \dots \beta_k x_{ik}) \\ = e^\alpha \times e^{\beta_1 x_{i1}} \times e^{\beta_2 x_{i2}} \times \dots \times e^{\beta_k x_{ik}} \quad (9)$$

Where

i indexes subjects;

$x_{i1}, x_{i2}, \dots, x_{ik}$ is the values of covariates for the i^{th} subject

This model is parametric because, once the regression parameters $\alpha, \beta_1, \beta_2, \dots, \beta_k$ are specified, the hazard function $h_i(t)$ is fully characterized by the model. The constant α represents a kind of baseline hazard, in (9), since $\log_e h_i(t) = \alpha$, or equivalently, $h_i(t) = e^\alpha$ when all of the x 's are 0. Other parametric hazard regression models are based on other distributions (Gompertz and

Weibull distribution) commonly used in modeling survival data. The Cox model superseded full parametric hazard regression models, which leaves the baseline hazard function unspecified:

$$\log_e h_i(t) = \alpha(t) + \beta_1 x_{i1} + \beta_2 x_{i2} + \dots \beta_k x_{ik} \quad (10)$$

Or equivalently,

$$h_i(t) = h_0(t) \exp(\beta_1 x_{i1} + \beta_2 x_{i2} + \dots \beta_k x_{ik}) \quad (11)$$

To estimate the model parameters, the maximum likelihood estimates are derived by maximizing a likelihood function. This Cox model [17] is termed semi-parametric because while the baseline hazard can take any form, the covariates enter the model through the linear predictor

$$\eta_i = \beta_1 x_{i1} + \beta_2 x_{i2} + \dots \beta_k x_{ik} \quad (12)$$

In (12), there is no constant term (intercept) in the linear predictor; the constant is absorbed in the baseline hazard. The Cox regression model is also a proportional hazard model. Consider two observations, i and i' , that differ in their x-values, with respective linear predictors

$$\eta_i = \beta_1 x_{i1} + \beta_2 x_{i2} + \dots \beta_k x_{ik} \quad \text{and} \quad \eta_{i'} = \beta_1 x_{i'1} + \beta_2 x_{i'2} + \dots \beta_k x_{i'k} \quad (13)$$

The hazard ratios for these two observations are:

$$\frac{h_i(t)}{h_{i'}(t)} = \frac{h_0(t)e^{\eta_i}}{h_0(t)e^{\eta_{i'}}} = e^{\eta_i - \eta_{i'}} \quad (14)$$

In (14), the ratio is constant over time. Therefore, the Cox model can easily accommodate time-dependent covariates. The Cox model accounts for survival times, thus, it uses more information than the logistic model. The Cox PH model allows censored observations and incorporates survival times. A Cox PH model therefore uses more information than a logistic regression model.

3.3 Goodness-of-fit test

A set of covariates in the Cox PH model can be time-dependent (or time varying) covariates. Used SAS (or SPSS) to perform the Cox PH model analysis of the data for this paper and will calculate the maximum likelihood estimates.

Used likelihood ratio test to see the variables included in the final model are significant in explaining some of the variability in data. The Chi-Square statistic is the difference in -2 Log Likelihood (-2LL) between the final model and a reduced model. The null hypothesis is that all parameters ($\beta_1, \beta_2 - \dots, \beta_k$) of the effect are 0. This test is comparable to the usual F test for regression analysis.

The hypothesis testing is as follows:

$$H_0 : \beta_1 = \beta_2 - \dots = \beta_k = 0;$$

$$H_1 : \sim H_0$$

Where β_i is the parametric estimation of explanatory variable i

The statistic quantity of the aforesaid hypothesis testing is -2 Log Likelihood (= -2Log (L(0)-

$L(\beta)$ which observes $\chi^2(k)$, where $L(0)$ is the likelihood function value under the null hypothesis, while $L(\beta)$ is the likelihood function value in consideration of the whole model.

R^2 is an intuitive measure of how well model predicts the values of the dependent variables [69]. R^2 in the Cox regression is a pseudo measure of association between the response variable and covariates. In general, higher R^2 value means the model is fit for analysis of sampling. Since in light that maximum of 1 cannot be obtained using Cox & Snell R^2 for measurement; Nagelkerke [45] proposed a modification of Cox & Snell R^2 .

$$\text{Cox \& Snell } R^2: R_{CS}^2 = 1 - \left[\frac{L(0)}{L(\beta)} \right]^{2/N} = 1 - \exp[(L(0) - L(\beta)) * 2 / N] \quad (15)$$

$$\text{Nagelkerke } R^2: R_N^2 = R_{CS}^2 / \max R_{CS}^2 \quad (16)$$

Where $L(0)$ = the likelihood function value containing only intercept; $L(\beta)$ = the likelihood function value in consideration of the whole model; N = sample size, $\max R_{CS}^2 = 1 - [L(0)]^2$

4. COX MODEL PREDICTIVE ABILITY

The most important characteristics of a BFP model are its production of accuracy.

Type I error refer to the situation when actual failure company is classified as non failures company, and Type II error refer to non failure company is classified as non failures company. Type I error is more important than Type II error. The objectives of predictive of accuracy should be to reduced Type I error while keep Type II error. The reason for this is that Type II error only creates a lost opportunity cost from not dealing with a successful business, for example, missed potential investment gains.

In contrast, due to involvement with a business that will fail, Type II error results in a realized financial loss, for example, losing all money invested in an impending bankrupt business [18]. The method used for calculating the accuracy of classifying distressed companies and non-distressed companies is illustrated in Table 2, in which C denotes the number of Type I error, that is the number of distressed companies in the sample based on actual observation that were misclassified as a non-distressed company. B denotes the number of Type II error that is the number of non-distressed companies in the sample based on actual observation that were misclassified as a distressed company. A and D represent respectively the number of non-distressed and distressed accurately classified by the models [34]. By determining accuracy of classification, we can learn about whether the constructed model is the optimal prediction model.

Table 2: Robustness of model

No-distressed Observed value	Non-distressed company	Distressed company	Accuracy of Classification
Non-distressed company	A	B	E
Distressed company	C	D	F
Overall accuracy of classification			G

Note: 1. The accuracy of classification of non-distressed company is expressed by $E = A / (A + B)$

2. The accuracy of classification of distressed company is expressed by $F = D / (C + D)$

3. The overall accuracy of classification is $G = (A + D) / (A + B + C + D)$

5. EMPIRICAL RESEARCH

In this section, the study first performs descriptive statistic of the sampling and Covariates, and follows by the construction of business failure prediction model based on Cox model and analysis of empirical results. In order to better analyze the effect of the Cox model predicted, we random select in the stock market listed company's traditional manufacturing in Taiwan. In sub-section, goodness-of-fit test is carried out and robustness of the model is examined using accuracy of classification.

The proposed steps of business failure prediction model are:

- Step 1: Definition of variables
- Step 2: Sampling and data
- Step 3: Reduced the number of financial ratio
- Step 4: Goodness-of-fit test
- Step 5: Robustness of model in prediction accuracy

5.1 Selection of Variables

The main goal of this research is to assess the empirical classification and prediction accuracy of the COX SA model when applied to BFP. Karels and Prakash [28] suggested a careful selection of ratios to be used in the development of bankruptcy prediction model. A set of covariates used in this study includes a combination of financial ratios and market variables [20]. In financial reporting analysis, [19] suggest five factors for evaluation enterprise financial failure. Financial ratios have been widely used in explaining the possibility of business financial distress ([3], [7], [9], [43], [47], [48], [49], and [76]). Table 3 is The 12 ratios selected in this study.

Table 3: The 12 ratios selected in this study

	Category	Covariate	Code	Definition
1	Profitability	EBIT margin	EBT	EBIT/operating revenue
2		Return to Equity	ROE	Net income/ Total equity
3		Return on Assets	ROA	Net income/ Total assets
4	Liquidity	Current ratio	CUR	Current assets/ current liabilities
5		Quick ration	QUK	Quick assets/ current liabilities
6	Leverage	Debt ratio	DET	Total liabilities/ Total assets
7		Debt to Equity ratio	DER	Total liabilities/ Total equity
8	Efficiency	Fixed Asset turnover	FAT	Revenue/Asset
9		Capital turnover	CAT	Operating revenue/ operating invest capital
10	Valuation ratio	Price to Sales ratio	PSR	Stock price per share/ Sales per share
11		Price earnings ratio	PER	Stock price per share/ Earnings per share
12		Price to book value	PBV	Stock price per share/ Equity per share

Table 3 shows the details and definition of covariates used in this study. 12 financial ratios are used in this study. The Profitability ratios include EBIT margin (EBT), Return to equity (ROE), and Return on assets (ROA). Current ratio (CUR) and Quick ratio (QUK) will be used in this study in order to measure the liquidity of the firms. Two types of Leverage ratios are Debit ratio (DET) and Debit to equity ratio (DER), two types of Efficiency ratios are Fixed asset turnover (FAT) and Capital turnover (CAT) and three types of Valuation ratios are Price to sales ratio (PSR), price earnings ratio (PER), and price to book value (PBV).

5.2 Data collection and Sample

The sample in this research is random selection in the stock market listed company's traditional manufacturing in Taiwan. In order to consider the survival problem, the choice of listed companies listed on the Taiwan Stock Exchange using annual data on financial ratios for the period 2003-2009. In order to better analyze the effect of the Cox model predicted, this study estimated that from 2003 to 2009 sample was divided into estimating samples and forecasting samples. This paper select sample listed companies from 2003 to 2006 for estimating sample. There are 56 financially distressed company and 154 activity listed companies in the analysis. This paper select sample listed companies from 2007 to 2009 for forecasting sample. It randomly selected 46 financially distressed company and 128 activity listed companies in the analysis as forecasting samples.

5.3 Reduced the number of financial ratio

There are two ways to reduce the large number of financial ratio (1) Pearson correlation (2) The model accepted has a good fit and that the multi-linearity level is acceptable.

According to Pearson correlation, the correlation between CUR and QUK is 0.9876, which are statistically significant with p-value less than 0.0001. This means the positive relationship

between these pair of variable. Based on the likelihood ratio resulted from Pearson correlation, the covariates QUK are selected into Cox proportional hazards model. Table 4 is the Cox proportional hazards model.

Table 4: The Cox proportional hazards model

Variable	D F	Parameter Coefficient	Standard error	Chi- square	pr> chisq	Hazard Ratio
EBT	1	0.3272	0.0125	0.0785	0.6542	1.387
ROE	1	-0.2631	0.1345	5.9682	0.0164**	0.769
ROA	1	-0.5012	0.1877	6.5896	0.0068**	0.606
QUK	1	-0.0165	0.0112	2.0125	0.0781	0.984
DET	1	0.2958	0.1432	5.2146	0.5606	1.344
DER	1	-0.1285	0.1245	0.3352	0.0085**	0.879
FAT	1	0.1421	0.2198	1.5428	0.0109**	1.153
CAT	1	0.0026	0.0156	0.5976	0.1875	1.003
PSR	1	0.2415	0.0968	0.3524	0.2432	1.273
PER	1	0.1243	0.1265	0.1548	0.1861	1.132
PBV	1	0.2341	0.0065	0.2382	0.0235**	1.264

** Significant at 5 percent

Using Cox proportional hazards model with financial ratios, the proportional hazards model are represented in Table 4. In SAS software, PROC RHREG is used to fit the Cox proportion hazards model and to analyze the effects of the financial on the survival of the company. Table 4 is denoted as the coefficient estimation, the Standard error, Chi-square tests with the relative p-value for testing the null hypothesis that the coefficient of each covariate is equal to zero. Hazard ratio is obtained by computing e^{β} , where β is the coefficient in a proportional hazard model. By considering the p-value, six covariates are highly significant at 5 percent. These ratios are EBT, QUK, DET, CAT, PSR and PER with the coefficient 0.3272, -0.0165, 0.2958, 0.0026, 0.2415 and 0.1243 respectively. Therefore, the early warning indicators are ROE, ROA, DER, FAT and PSV. ROE and ROA are negative value indicating that an increase in either covariate decreases the hazard of entering into financially distressed. Hazard ratio of ROE is 0.769 ($e^{-0.2631} = 0.769$). It means that an increase of one unit in ROE implies 0.769 decreases in risk financial distress. For the sample in this study, profitability (EBT), liquidity (QUK), leverage (DET), Efficiency (CAT) and Valuation (PSR, PER) have never found statistically significant in the model. The model is shown as follow:

$$\text{Log } h(t) = \beta_1 \text{ROE}(t) + \beta_2 \text{ROA}(t) + \beta_3 \text{DER}(t) + \beta_4 \text{FAT}(t) + \beta_5 \text{PSV}(t)$$

5.4 Goodness-of-fit test

One measure of overall goodness-of-fit test is partial likelihood-ratio test. In SAS software, PROC RHREG is used to obtain the likelihood ratio chi-square statistic from the model fit statistics Table 5. In Table 5, the output produces includes the value of -2log likelihood for fitting, AIC (Akaike Information Criterion) and SBC (Schwartz Bayesian Criterion) for fitting a model without covariance and fitting a model with covariates.

Akaike [2] introduced the concept of the information criteria as a tool optimal model selection. AIC is a function of the number of observations n , the sum of square errors (SSE), and the number of independent variables $k \leq p + 1$ where k includes the intercept, as shown in (17).

$$AIC = n \cdot \ln\left[\frac{SSE}{n}\right] + 2k \tag{17}$$

The first term in (17) is a measure of the model lack of the fit while the second term ($2k$) is a penalty term for additional parameters in model.

Schwartz [51] derived from a Bayesian modification of the AIC criterion to develop a SBC model. SBC is a function of the number of observations n , the SSE, and the number of independent variables $k \leq p + 1$ where k includes the intercept, as shown in (18).

$$SBC = n \cdot \ln\left[\frac{SSE}{n}\right] + k \ln n \tag{18}$$

Table 5: Goodness-of-fit test

The PHREG Procedure		
Model Fit Statistics		
Criterion	Without Covariates	With Covariates
-2 LOG L	270.544	253.390
AIC	270.544	258.428
SBC	270.544	266.512

Testing Global Null Hypothesis: Beta = 0

Test	Chi-square	DF	Pr > chi-square
Likelihood Ratio	17.154	4	<0.001
Score	19.218	4	<0.001
Wald	18.356	4	<0.001

The chi-square of likelihood ratio is 17.145 (270.544-253.390). This statistic is also shown in the Table 5 Testing Global Null Hypothesis: Beta = 0. The likelihood-ratio test, Score test and Wald test equals 17.154, 19.218, 18.356 respectively, with 4 degree of freedom. Thus, the null hypothesis is rejected ($p < 0.001$).

Another measure of model performance may be some measure analogous to R^2 , as shown in the formula below. Keep in the mind that this measure does not explain the proportion of variability of the response variable by the explanatory variables as in the linear regression.

$$R_{CS}^2 = 1 - \left[\frac{L(0)}{L(\beta)}\right]^{2/N} = 1 - \exp[(L(0) - L(\beta)) * 2 / N] = 0.406$$

$$R_N^2 = R_{CS}^2 / \max R_{CS}^2 = R_{CS}^2 / (1 - [L(0)]^2) = 0.63$$

The validation by Cox & Shnell R^2 and Nagelkerke R^2 shows that the explanatory variables of the prediction model process explanatory power for the incidence of financial distress.

After we have settled on assessing the adequacy of the model that seems a good-fit, we can carry out statistical inference of a fitted model. The output below is produced by running PROC PHREG with 5 covariates, ROE, ROA, DER, FAT and PBV. The RL (RISKLIMITS) option in the Model statement provides 95% confidence interval for the hazard ratio estimates. Table 6 is the PHREG procedure.

Table 6
The PHREG Procedure

Analysis of Maximum Likelihood Estimates								
Variable	D F	Parameter Coefficient	Standard error	Chi- square	pr> chisq	Hazard Ratio	95% Ratio Confidence limits	Hazard
ROE	1	-0.2584	0.1425	5.869	0.0158**	0.769	0.479	1.049.
ROA	1	-0.4952	0.1860	6.5765	0.0124**	0.606	0.324	0.888
DER	1	-0.1308	0.1230	0.3254	0.0105**	0.879	0.456	1.302
FAT	1	0.1546	0.2065	1.6488	0.0209**	1.153	0.568	2.738
PBV	1	0.2438	0.0078	0.2412	0.0242**	1.264	0.642	2.886

** Significant at 5 percent

Results of the analysis indicate that five covariates appear to add significantly to the model. The p-value of the parameter estimates for the regression coefficients are highly significant for ROE, ROA, DER, FAT and PBV. The coefficient signs of ROE, ROA DER covariates are negative indicating that an increase in either covariate decreases the hazard of entering into financially distressed. For example, Hazard ratio of ROE covariate is 0.769 ($e^{-0.2584} = 0.769$). It means that an increase of one unit in ROE covariate implies 0.769 decreases in risk financial distress. The interpretation of the estimated hazard ratio of ROE is 0.769. It means that an increase of one unit in the ratio of Net income to Total equity will shrink the hazard rate by 23.1% (1- 0.769). The interpretation of the estimated hazard ratio of (FAT) is that financial companies in this study fail at about 1.153 times the rate of those in non-financial sector. The 95% confidence interval for hazard ratio suggests a sector as low as 0.568 or as high as 2.738. The interval width equals 2.170 (2.738-0.568). This interval also includes the point estimate of 1.153 and does not contain the null value of 1.

5.5 Accuracies of classification on the model

The fit of PH model used in this study is validated by comparing the predicted value of each sample with the cutoff value [40]. If the predicted value is below this cut value, the sample is classified as a distressed company; otherwise, the company is classified as non-distressed company. According to the suggestion of Martin [41], this study uses empirical cutoff value which is the percentage of financial distressed samples in total sample at 0.264 (46/174). The accuracies of classification of the model are compiled in Table 7.

Table 7: The accuracies of classification of the mode

No-distressed Observed value	Non-distressed company	Distressed company	Accuracy of Classification
Non-distressed company	115	13	115/128 (89.84%)
Distressed company	8	38	38/ 46 (82.60%)
Overall accuracy of classification			153/174 (87.93%)

Therefore, Type I error is $13 / 128 = 10.64\%$ and Type II error is $8 / 46 = 17.50\%$ and the overall accuracy of classification is 87.93%

6. CONCLUSION

In this paper, the listed companies on the Taiwan Stock Exchange that experienced data between 2003 and 2009 are employed as distressed data set. In order to better analyze the effect of the Cox model predicted, this study estimated that from 2003 to 2009 sample was divided into estimating samples and forecasting samples. This paper selected 56 distressed companies and 154 non-distressed companies for estimating samples data between 2003 and 2006; 46 distressed companies and 128 non-distressed companies for forecasting samples data between 2007 and 2009.

Form the proposed steps of business failure prediction model; the financial distress probability model is constructed using Profitability, Leverage, Efficiency and Valuation ratio variables. Step 1 select the financial ratios for using in the development of bankruptcy prediction model. Step 2 consider the survival problem, the choice of listed companies listed on the Taiwan Stock Exchange using annual data on financial ratios for the period 2003-2009. In Step 3, there are two ways to reduce the large number of financial ratio (1) Pearson correlation (2) The model accepted has a good fit and that the multi-linearity level is acceptable. In SAS software, PROC RHREG is used to fit the Cox proportion hazards model and to analyze the effects of the financial on the survival of the company. Step 4, one measure of overall goodness-of-fit test is partial likelihood-ratio test. In SAS software, PROC RHREG is used to obtain the likelihood ratio chi-square statistic from the model fit statistics. The validation by Cox & Shnell R^2 and Nagelkerke R^2 shows that the explanatory variables of the prediction model process explanatory power for the incidence of financial distress. In Step 5, consider the robustness of model in prediction accuracy, in this study the accuracies of classification of the mode in overall accuracy of classification is 87.93%.

ACKNOWLEDGEMENTS

I would like to thank the anonymous reviewers for their constructive comments on this paper.

REFERENCES

- [1] Adiana, NHA., Halim, A., Ahmad, H., and Rohani, M. R. (2008), "Predicting corporate failure of Malaysia's listed companies: Comparing multiple discriminant analysis, logistic regression and the Hazard model". *International Research Journal of Finance and Economic*, Vol. 5, pp. 202-217.
- [2] Akaike, H. (1973), "Information theory and an extension of the maximum principle". *Second International symposium on Information theory*, pp. 267-281.
- [3] Altman, E. I., (1968), "Financial Ratios, Discriminant analysis and the prediction of corporate bankruptcy". *Journal of Finance*, Vol. 23, No. 4, pp.589-609.
- [4] Antonakis, A. C. and Sfakianakis, M. E.,(2009), "Assessing naïve Bayesian as a method of screening credit applications". *Journal of applied Statistics*, Vol. 36, No. 5, pp. 537-545.
- [5] Barros, C. P., Butler, R. and Correia, A. (2010), "The length of stay of golf tourism: A survival analysis". *Tourism Management*, Vol. 31, No.1, pp. 13-21.
- [6] Bian, H. and Mazlack, L.,(2003), "Fuzzy-rough nearest-neighbor classification approach", 22nd International Conference of the North American, 24-26 July, 2003.
- [7] Bongini , P., Ferri, G. and Hahm, H.(2000), "Corporate bankruptcy in Korea: Only the strong survival?". *Financial Review*, Vol. 35, pp. 71-111.
- [8] Cao, R. Vilar, J. M. and Devia, A., (2009), "Modeling consumer credit risk via survival analysis". *Statistics & Operational Research Transactions*, Vol. 33, No. 1, pp. 3-30.
- [9] Catanach, A. H. and Perry, S. H. (2001), "An evaluation of survival model's contribution to thrift institution distress prediction". *Journal of Managerial Issue*, Vol.13,No.4, pp. 401-417.
- [10] Chaduhuri, A, and De, K., (2011), "Fuzzy support vector machine for bankruptcy prediction", *Applied Soft Computing*, Vol. 11, No. 2, pp.2472-2486.
- [11] Chen W. Ma, C. and Ma, L.,(2009), "Mining the customer credit using hybrid support vector machine techniques". *Expert Systems with Application*, Vol. 36, No. 4, pp. 7611-7616.
- [12] Chen, D., (2009), "Using neural networks and data mining techniques for financial distress prediction model". *Expert Systems with Application*, Vol. 36, No. 2, pp.4075-4086
- [13] Chen, M. Y., (2011), "Prediction corporate financial distress based on integration of decision tree classification and logistic regression". *Expert Systems with Application*, Vol. 38, No. 9, pp.11261-11272.
- [14] Cheng, J. U., Yeh, C. H. and Chiu, Y. W.,(2007), "Improving business failure prediction using rough sets with Non-financial variables". *LNCS*, Vol. 4431/2007, 2007, pp. 614-621.
- [15] Cox, D. R. & D. Oakes, (1984), "Analysis of Survival Data". London: Chapman and Hall.
- [16] Cox, D. R. (1972), "Regression models and life-tables", *Journal of Royal Statistical Society B*, Vol. 34, pp. 187-220.
- [17] Dimitras, A. L., Slowinski, R., Susmaga, R. and Zopounidis, C., (1999), "Business failure Prediction using rough sets". *European Journal of Operational Research*, Vol. 114, pp. 263-280
- [18] Gepp, A. and Kumar, K. (2008), "The role of survival analysis in financial distress prediction". *International Research Journal of Finance and Economic*, Vol. 16, pp.13-34.
- [19] Gibson, C. H. (2006), "Financial Reporting and analysis: using financial associating information". South-Western College Pub, 10 editions, April, 27.
- [20] GIGS (2006): Global Industry Classification Standard (GICS@), Standard & Poor's, August, 2006
- [21] Gokovali, U., Bahar, Q. and Kozak, M. (2007), "Determinants of length of study: A practical use of survival analysis". *Tourism Management*, Vol. 28, No.3, pp.736-746.
- [22] Golub, J. (2007), "Survival analysis and the European Union decision-making". *European Union Politics*, Vol. 8, No. 2, pp. 155-179.
- [23] Hu, X. T., Lin, T. Y. and Han, J., (2003), "A new rough sets model based on database systems", *Lecture Notes in Artificial intelligence*, Vol. 2639, pp. 114-121
- [24] Ince, H. and Aktan, B., (2009), "A comparison of data mining techniques for credit scoring in banking: a managerial perspective". *Journal of Business Economics and Management*, Vol. 3, No. 2, pp. 233-240
- [25] Islam, M. J., Wu, Q. M. J., Ahmadi, M. and Sid-Ahmed, M. A.,(2007), "Investigating the performance of Naïve- Bayes Classifier and K-NN classifiers". *International Conference on Convergence*

- Information Technology, IEEE Computer Society. 2007.
- [26] Javanthi, J., Suresh Joseph, K. and Vaishnavi, J. (2011), "Bankruptcy prediction using SVM and Hybrid SVM survey". *International Journal of Computer Applications*, Vol. 34, No.7, pp. 39-45.
- [27] Jones, S. and Hensher, D. A. (2004), "Predicting firm financial distress: A mixed logit Model". *Accounting Review*, Vol.79, No. 4, pp. 1011-1038.
- [28] Karels, G. V. and Prakash, A. J. (1987), "Multivariate normality and forecasting of corporate bankruptcy". *Journal of Business Finance and account*, Vol. 14, No. 4, pp. 573-592.
- [29] Kauffman, R. and Wang, B. (2001), "The success and failure of dotcoms: A multi-method survival analysis". In proceeding of the 6th INFORMS Conference on Information System and Technology (CIST), Miami, FL, USA
- [30] Kauffman, R. and Wang, B. (2003), "Duration in the digital economy: Empirical bases for the survival of internet firms". In 36th Hawaii International Conference on System Sciences (HICSS), Hawaii.
- [31] Keasey, K., McGuinness, P. and Shot, H. (1990), "Multilogit approach to predicting corporate failure: Further analysis and the issue of signal consistency". *Omega*, Vol. 18, No. 1, pp. 85-94.
- [32] Kim, H. and Sohin, S. Y., (2010), "Support vector machines for default prediction of SMEs based on technology credit". *European Journal of Operational Research*, Vol. 201, No. 3, pp. 838-846.
- [33] Laitinen, E. K. and Luoma, M. (1991), "Survival analysis as a tool for company failure prediction", *Omega*, Vol. 19, No. 6, pp. 673-678.
- [34] Lau, S. T., Cheng, B. W., and Hsieh, C. H., (2009), "Prediction model building with clustering-launched classification and support vector machines in credit scoring". *Expert Systems with Applications*, Vol. 36, No. 4, pp. 7526-7556.
- [35] LeClere, M. J. (2000), "The occurrence and timing of events: Survival analysis applied to the study of financial distress". *Journal of Accounting Literature*, Vol. 19, pp. 158-189.
- [36] Liang, Y. and Xin, H., (2009), "Application of Discretization in the use of Logistic Financial Rating". *International Conference on Business Intelligence and Financial Engineering*, 24-26 July 2009, pp. 364-368
- [37] Lin, L. and Piesse, J. (2004), "The identification of corporate distress in UK industrials: A conditional probability analysis approach". *Applied Financial Economics*, Vol. 14, pp. 73-82.
- [38] Liu, K., Lai, K. K., and Guu, s. M., (2009), "Dynamic credit scoring on consumer behavior using Fuzzy Markov model". *Fourth International Multi-Conference on Computing in the Global Information Technology*. IEEE Computer Society, 2009.
- [39] Lopez, R. F., (2007), "Modeling of insurers' rating determinants. An application of machine learning techniques and statistical models". *European Journal of Operational Research*, Vol. 183, No. 2. PP.1488-1512.
- [40] Lu, Y. C. and Chang, S. L., (2009), "Corporate governance and quality of financial information on prediction power of financial distress of listed companies in Taiwan". *International Research Journal of Finance and economics*, Vol. 32, pp. 114-138.
- [41] Martin, D. (1977), "Early warning of bank failure: a Logit regression approach". *Journal of Banking and Finance*, Vol. 1, pp. 249-276.
- [42] Mcleay, S. and Omar, A. (2000), "The sensitivity of prediction models to the non-normality of bound and unbounded financial ratios", *The British Accounting Review*, Vol. 32, No. 2, pp. 213-230.
- [43] Mossman, C. E., Bell, G. G. Swartz, L. M. and Turtle, H. (1998), "An empirical comparison of bankruptcy models". *Financial Review*, Vol. 33, No. 2, pp. 35-53.
- [44] Mues, C. Baesens, B., Files, C. M., and Vanthienen, J., (2004), "Decision diagrams in machine learning: an empirical study on real-life credit-risk data". *Expert Systems with Applications*, Vol. 27, No. 2, pp. 257-264
- [45] Nagelkerke, N. J.D. (1992) "aximum Likelihood Estimation of Functional Relationships, Pays-Bas". *Lecture Notes in Statistics*, Vol. 69, pp. 110
- [46] Odom, M., Sharda, R., (1990), "A neural network model for bankruptcy prediction". *IEEE INNS International Joint Conference on Neural Networks*, Vol. 12, 1990, pp.163-168.
- [47] Ohlson, J. A. (1980), "Financial Ratios and the Probabilistic Prediction of Bankruptcy". *Journal of Accounting Research*, (Spring,1980), Vol. 14, No. 1, pp. 109-131.

- [48] Rommer, A. D. (2005), "A comparative analysis of the determinants of financial distress in French, Italian and Spanish firms". Working paper, Danmarks National Bank, Copenhagen. 18, May, 2005.
- [49] Routledge, J. and Gadenne, D. (2000), "Financial distress, reorganization and corporate performance". *Accounting and Financial*, Vol. 40, No. 3, pp. 233-259.
- [50] Sarkar, S., and R. S. Sriram., (2001), "Bayesian Models for Early Warning of Bank Failures". *Management Science*, Vol. 47, No. 11, pp.1457-1475.
- [51] Schwartz , G. (1978), "Estimating the dimension of a model". *Annals of Statistics*, Vol. 6, pp. 461-464.
- [52] Shin, K. S., Lee, T. S. and Kim, H. J.,(2005) , "An application of support vector machines in bankruptcy prediction model". *Expert system Application*, Vol. 28, No. 1, pp. 127-135
- [53] Shumway, T. (2001), "Forecasting bankruptcy more accurately: A simple hazard model". *Journal of Business*, Vol. 74, pp. 101-124.
- [54] Sloot, T. and Verschuren, P. (1990), "Decision-making sped in the European community". *Journal of Common Market Studies*, Vol. 29, No.1, pp. 75-85.
- [55] Sohn , S. Y.and Shim, H. W.,(2006), "Reject inference in credit operations based on survival analysis". *Expert Systems with Applications*, Vol. 31, No. 1, pp. 26-29.
- [56] Stefanescu , C. Tunaru, R., and Tunbull, S., (2009), "The credit rating process and estimation of transition probabilities: a Bayesian approach", *Journal of Empirical Finance*, Vol. 16, No. 2, pp. 216-234
- [57] Stepanova, M. and Thomas, L. C.,(2002), "Survival analysis methods for personal loan data". *Operations Research*, Vol. 50, No. 2, pp. 277-289.
- [58] Sun, J. and LI, H. L., (2009), "Financial distress prediction based on serial combination of multiple classification". *Expert Systems with Application*, Vol. 36, No. 4, pp.8659-8666
- [59] Tabachnick, B. G. and Fidell, L. S., (2000), "Using Multivariate Statistics". Allyn and Bacon Press, UK. 2000
- [60] Tam , K. Y. and Kiang, M., (1992), "Managerial applications of neural network: the case of bank failure predictions". *Management Sciences*, Vol. 38, pp.927-947
- [61] Tan, C. N. W. and Dihardjo, H. (2001), "A study on using artificial neural networks to develop an early warning predictor for credit union financial distress with comparison to the probit model". *Managerial Financial*, Vol. 27, No. 4, pp. 56-77.
- [62] Tung , W. L., Quek, C., Cheng, P. and EWS, G.,(2004), "a novel neural-fuzzy based early warning system for predicting bank failures". *Neural Networks*, Vol. 17, No.4, PP. 567-587.
- [63] Wang, Y., Wang, S. and Lai, K. K., (2005), "A new fuzzy support vector Machine to evaluate credit risk". *IEEE Transactions Fuzzy Systems*, Vol. 13, No. 6, pp. 820-831.
- [64] Wei , L. L. and Zhang, W. X., (2003), "Probabilistic rough characterized by Fuzzy Sets". *Lecture Notes in Artificial Intelligence*, 2369, pp. 173-180
- [65] Wong, S. K.M., Zialo, W. and Li, Y. R., (1986), "Comparison of rough set and statistical methods in inductive learning". *International Journal of Man-Machine Studied*, Vol. 24, pp. 53-72.
- [66] Xiao, Z., Yang, X., Pang, Y. and Dang, X.,(2011), "The prediction for listed companies' financial distress by using multiple prediction network with rough set and Dempster -Shafer evidence theory". *Knowledge – Based Systems*, Vol. 26, PP.196-206.
- [67] Xinhui, C. and Zhong, Q., (2009), "Consumer credit scoring based on multi-criteria fuzzy logic, International Conference on Business Intelligence and Financial Engineering". IEEE Computer Society. 2009.
- [68] Yang, Z., You, W. and Ji, G. (2011), "Using partial least squares and support vector machines for bankruptcy prediction, *Expert Systems with Applications*". Vol.38, No.7, pp. 8336-8342.
- [69] Yap, BCF., Munaswamy, S., and Mohamad, Z.,(2012), "Evaluating Company Failure in Malaysia Using Financial Ratios and Logistic Regression". *Asian Journal of Finance & Accounting*, Vol. 4, No. 1, pp. 330-344.
- [70] Yu, L., Wang, S., Wen, F., Lai, K. K., and He, S.,(2008), "Designing a hybrid intelligent mining system for credit risk evaluation". *Journal of Systems Science and Complexity*, Vol. 21, No. 4, pp. 527-539.
- [71] Zhang, D., Hifi, M. Chen, Q. and Ye, W.,(2008), "A hybrid credit scoring based on genetic

- programming and support vector machine”. 4th International Conference on Natural Computation, IEEE Computer Society, 2008, pp. 8-12
- [72] Zhao, H. (2007), “A multi-objective genetic programming approach to developing Pareto optimal decision tree”. *Decision Support Systems*, Vol. 43, No. 3, pp. 809-826.
- [73] Zhou, J. and Bai, T., (2008), “Credit risk assessment using rough set theory and GA-based SVM”. *The 3rd International Conference on Grid and Pervasive Computing- Workshops*, IEEE Computer Society Kunming, 2008, pp. 320-325.
- [74] Zhou, L. and Lai, K. K. and Yu, L.,(2009), “Credit scoring using support vector machine with direct search for parameters selection”. *Soft Computing*, Vol. 13, No. 2, pp. 149-155.
- [75] Zhou, L., Lai, K. K., and Yu, L.,(2010), “Least squares support vector machines ensemble models for credit scoring”. *Expert Systems with Applications*, Vol. 37, No. 1, pp. 127-133.
- [76] Zulkarmain, M. S., Mohamad, A. A. H., Annuar, M. N. and Zainal, A. M. (2001), “Forecasting corporate failure in Malaysian industrial sector firms”. *Asia Academy of Management Journal*, Vol. 6, No.1, pp. 15-30.
- [77] Bunyaminu, A. and Bashiru, S. (2014), Corporate failure prediction: A fresh technique for dealing effectively with normality based on quantitative and qualitative approach, *International Journal of Financial Economic*, Vol. 3, No. 1, pp. 1-12.

Author

Ming-Chang Lee is Assistant Professor at National Kaohsiung University of Applied Sciences. His qualifications include a Master degree in applied Mathematics from National Tsing Hua University and a PhD degree in Industrial Management from National Cheng Kung University. His research interests include knowledge management, parallel computing, and data analysis. His publications include articles in the journal of *Computer & Mathematics with Applications*, *International Journal of Operation Research*, *Computers & Engineering*, *American Journal of Applied Science and Computers*, *Industrial Engineering*, *International Journal innovation and Learning*, *Int. J. Services and Standards*, *Lecture Notes in computer Science (LNCS)*, *International Journal of Computer Science and Network Security*, *Journal of Convergence Information Technology* and *International Journal of Advancements in computing Technology*.

