Indian Journal of Finance and Economics
Vol. 2, No. 2, **2021** : pp. 149-161

URL : www.arfjournals.com

ARF INDIA
Academic Open Access Publishing
www. arfjournals. com

A COMPARISON OF ACCELERATED FAILURE TIME MODELS AND THE COX PROPORTIONAL HAZARD MODEL IN THE ANALYSIS OF COMPANIES UNDER NATIONAL STOCK EXCHANGE, INDIA

Pal Deka[1] and Manash Pratim Barman[2]

[1]*Research scholar, Dibrugarh University, Dibrugarh. E-mail: pdeka7863@gmail.com*
[2]*Associate Professor, Dibrugarh University, Dibrugarh. E-mail: manashstats@gmail.com*
**Corresponding author*

Received : 22 October 2021; Revised : 10 November 2021; Accepted : 24 November 2021; Published : 30 December 2021

Abstract: Survival time data can be modeled with the help of parametric, semi-parametric and non-parametric models. Due to incompleteness of survival data, the distribution of survival data is not known and is not symmetrically distributed. Here, in this study an attempt is made to best fit the survival time data. A sample of 305 companies listed under National Stock Exchange, India is considered out of 1328 companies comprise of various industrial sectors. The analysis of the study is carried out with parametric and semi-parametric survival models. The Akaike Information Criterion (AIC) is used as criteria to find the best fitted model and the Cox-Snell Residual plot is used to check the goodness of fit of the model. The result of the study shows that the Cox Proportional hazard model is the best fit among the survival analysis models for company's stock exchange data.

Keywords: Share price; Cox Proportional hazard model; threshold value

1. INTRODUCTION

Survival analysis refers to analysis of 'Survival data' or 'time to an event' data. In other words it is a collection of procedures/techniques for analysis of survival time data. Survival time is a random variable and it denotes time to occurrence of a particular event. In survival analysis there are various events for which survival time is calculated such as Death, occurrence of a disease, recovery, relapse after treatment etc. Survival analysis has huge application in the field medicine & health related sciences. For mechanical related component also survival analysis can be used. The survival data used for survival analysis suffers from incompleteness. The standard

To cite this article:

Pal Deka and Manash Pratim Barman (2021). A Comparison of Accelerated Failure Time Models and the Cox Proportional Hazard Model in the Analysis of Companies under National Stock Exchange, India *Indian Journal of Finance and Economics*, Vol. 2, No. 2, pp. 149-161.

statistical methods or techniques use for analysis of data is not applicable in case of survival analysis data. The distribution of survival data is not known and is not symmetrically distributed. There are various types of procedures proposed for analysis of such data in survival analysis namely, non parametric methods, parametric methods and semi parametric methods. The two important functions in survival analysis are survival function and hazard function. Survival function is defined as the probability of surviving beyond any specified time. Survival function is a monotonic decreasing function of time. It is denoted as S (t). The distribution of survival time is not known in many cases, but can be approximated to a distribution with the help of survival function. The shape of the survival function displayed through the graphs are defined by a particular probability distribution such as exponential, Weibull, log logistic etc. Survival functions that are defined by parameters are known as parametric. The non parametric method for modeling survival functions are Kaplan Meier and Cochran-Mantel-Haenzel method or Log-Cox test. The well known semi-parametric method is Cox regression method. Parametric models are considered more preferable in the situations when the correct form of the parametric model is exactly known. The main objective of this study is to find the best fitted survival model of survival time to an event data of companies under National Stock Exchange (NSE).

2. INTRODUCTION OF THE MODELS

2.1. Accelerated Failure Time Models

There are different parametric survival models in survival analysis for fitting survival time data such as Weibull, Log-logistic, Exponential, Log-normal etc. Linear regression, Logistic regression and Poisson regression models are used when dealing with normal (uncensored) data. These parametric models follow certain distributional assumptions. But while dealing with survival time data, these techniques cannot be used. In parametric survival model, survival time (variable of interest) is also random variable and follows a known distribution i.e. Weibull, Exponential, Log-logistic, Log-normal, generalized gamma. The AFT model assumption is given as follows

$$S\,(t/x) = S_0\,(\exp(\beta'\,x)t);\ t \geq 0 \tag{1}$$

Where $S(t/x)$ is survival function at time t and $S_0(\exp(\beta')t)$ is baseline survival function at time t. $\beta' = (\beta_1, \beta_2 \ldots\ .\ \beta_k)$. From equation (1) it can be state that the survival function with covariate 'x' vector at time t is same as the baseline survival function at time t. This $\exp(\beta' x)$ is termed as acceleration factor respectively. The following Accelerated Failure time models are discussed below:

Exponential model: The Exponential model is one of the simplest parametric survival models. It satisfies both Proportional hazard assumption and Accelerated failure time model. The survival time to event data following Exponential distribution has constant hazard rate. It is one of the most commonly use parametric model. The constant baseline hazard in the model is defined as 'λ'. The survival function is given as,

$$S(t) = \exp(-t/\beta);\ \beta \geq 0,\ t \geq 0 \quad \text{(i)}$$

And the hazard function is also defined as,

$$h(t) = \frac{1}{\beta};\ t \geq 0,\ \beta \geq 0 \quad \text{(ii)}$$

The mean and variance of Exponential distribution with single parameter 'λ' is defined as $\frac{1}{\lambda}$ and $\frac{1}{\lambda^2}$. If 'T' is survival time to an event following Exponential distribution is given as T~ exp (λ). The Exponential distribution has relation with extreme-value distribution.

Weibull model: The Weibull model is also one of common parametric survival models. The hazard considered in the model is not constant over time. The survival time data following weibull distribution may generate two types of curve i.e. increasing Weibull and decreasing Weibull. The distribution consists of two parameters 'α' and 'ρ', where 'α' and 'ρ' is the scale and shape parameter. Due to inclusion of shape parameter it is more flexible than the exponential model. The value of 'ρ' determines whether the hazard is increasing (ρ>1), decreasing (ρ<1) and constant (ρ=1). The hazard function of the model is given as,

$$h(t) = \lambda\rho t^{\rho-1}\ \exp^{\beta x};\ t \geq 0,\ \lambda \geq 0 \quad \text{(i)}$$

The equation (*i*) states that the hazard function of the Weibull survival model at time t with x covariate is the product of baseline hazard and exponential expression of β coefficients. To fit survival time data following Weibull distribution, log-cumulative hazard plot should be a straight line. The Weibull distribution is related with extreme-value distribution.

Log-Normal model: The Log-Normal distribution is also help for fitting survival time to an event data. If t is denoted as survival time to an event following Log-Normal distribution the Log (t) follows Normal distribution. Similar to Normal distribution, the density function of Log-Normal distribution also consists of two parameters i.e. 'μ' and 'σ'. The shape of Log-Normal distribution is same with the Log-Logistic distribution. It accommodate Accelerated failure time model but not Proportional hazard model. If t is survival time to an event, the formula of survival function of the model is given as

$$S(t) = 1 - \Phi\left(\frac{\ln(t)}{\sigma}\right); t \geq 0 \quad \text{(i)}$$

The hazard function of the lognormal distribution is defined as

$$h(t) = \frac{\left(\frac{1}{t\sigma}\right)\Phi\left(\frac{\ln t}{\sigma}\right)}{\Phi\left(\frac{-\ln T}{\sigma}\right)}; t \geq 0 \quad \text{(ii)}$$

Log-Logistic model: The Log-Logistic survival model is one of the parametric accelerated failure time models. The Log-Logistic distribution consists of two parameters i.e. scale and shape. The shape of the distribution is similar to Log-Normal distribution. The survival time to event data is said to follow Log-Logistic distribution if the hazard rate increases initially and then decreases after some period of time. The following distribution can be used in place of the Weibull distribution. If t is survival time then the survival function is given as,

$$S(t) = [1 + \left(\frac{t}{\alpha}\right)^{\beta}]^{-1}; t \geq 0 \quad \text{(i)}$$

Where 'α' and 'β' are scale and shape parameters.

Similarly, the hazard function is defined as,

$$h(t) = \frac{\left(\frac{\beta}{\alpha}\right)\left(\frac{t}{\alpha}\right)^{\beta-1}}{1 + \left(\frac{t}{\alpha}\right)^{\beta}}; t \geq 0 \quad \text{(ii)}$$

2.2. The Cox Proportional Hazard Model

The survival pattern of an event depends on several factors which influence the variable of interest i.e. time to an event. Methods like Kaplan-Meier, Mantel-Haenzel, Log-Cox test etc. doesn't consider the effect of covariates (explanatory variables) on survival time. These methods are considered as non parametric methods in survival analysis. The covariates may be qualitative/ quantitative measured in nominal, ordinal or metric (ratio & interval) scale. In general, in case of explanatory variables, linear regression or logistic regression model is appropriate to be fitted. In survival analysis, the survival data is censored and standards statistical techniques cannot be applied. Logistic regression is considered to be appropriate in case of binary dependent variables. The Cox regression model helps us to measure the effect of various covariates on survival time of an event. The Cox model is product of baseline hazard which involves t and the exponential expression which is independent of t and involves regression coefficients and covariates is given below:

$$h(t,x) = h_0(t)e^{\Sigma\beta x_i};\ t \geq 0,\ x \geq 0 \qquad \text{(i)}$$

Where $h_0(t)$ is denoted as baseline hazard and exponential part of the model consist of 'β' parameters to be estimate and x_i's explanatory variables.

In the model, the functional form of baseline hazard is not specified function, therefore it is also known as semi parametric model. In contrast to a parametric model the functional form is specified with unknown parameters. Some of the parametric models follow respective distribution such as Weibull, Exponential, log-normal etc. *The* x_i*'s here are* time independent variables. It is also possible to fit the model for time dependent variable. The model for such variables is called the extended Cox model. One of the properties of Cox model is that if all the explanatory variables are equal to zero then the formula is left with baseline hazard. Exponential part of the formula becomes unity. Although the functional form of baseline hazard is not known, the model gives appropriate estimates of regression coefficients, hazard ratio and an adjusted survival curve can be obtained. The regression coefficients estimated are non negative due to the exponential part of the formula. As hazard function should range between zero and positive infinity, the hazard are always non negative. Maximum likelihood estimation method is used to estimate the unknown parameters. In many data situations where the decision to use correct parametric model is in doubt, using Cox model will give reliable results. The results obtained through Cox model will approximate to results which will be given by any parametric model. Hazard function and survival function can be obtained from Cox model without estimating the baseline hazard. The Cox model is preferred over Logistic model when information about survival time is available and censoring is present in data. Logistic model ignores such information. One of the important assumptions of Cox hazard model is proportionality hazard assumption. The Proportionality assumption specifies that hazard ratio is constant over time. In other words hazard for two different individuals are proportional and proportionality constant is independent of time. Hazard ratio is defines as the ratio of hazard for one individual to the hazard for other individual with different set of explanatory variables. The final expression is independent of time i.e. constant. It can also be plotted graphically. If the proportionality hazard assumption is not fulfilled, Cox Stratified model is used.

1. Literature Review

The research paper by a graduate student Jiayi Ni (2008) of Ball State University studies the survival time of companies which were listed in Shanghai Security market. She tried to study the breakdown of companies stocks and factors associated with change in share prices. The companies whose share prices drop below 40 % were censored. The duration of study period was 8 months.

Andre Paul Lamberto and SubhrenduRath (2010) of Curtin University investigate the survival of Initial Public Offerings (IPO) in Australia. They had used the risk factors that are listed in prospectus issued by the respective firms. They also discussed the failure rate of IPO's in Australia from the first five years of listed date and from first seven years too. Survival of Australian firm's are different from U.S capital market is also figured in this study.

Mrunal Joshi (2013), Assistant Professor, B.R.C.M. College of Business Administration, Surat, Gujarat, studies variables affecting Indian stock market. Some factors like flow of foreign investments, deflation, liquidity, etc. are considered as factors.

Alhassan Bunyaminu (2015) examined business failure using both financial (ratios) and non-financial variables of companies listed on Ghana Stock Exchange. The study includes both the type of quantitative and qualitative variables to determine business failure. The researcher used the Cox regression model technique and Generalized Linear regression modeling to predict business failure with an appropriate degree of accuracy.

E. Geetha & Ti. M. Swaaminathan (2015) of PG and Research Department of Commerce, Pachaiyappas College, Tamil nadu studies the factors influencing market price of companies listed in stock exchanges. They undertook a sample of four automobile and IT companies each listed in both and National Stock Exchange and Bombay Stock Exchange respectively. Common factors such as Earning per share, Book value, Dividend per share and P/E ratio are considered from the annual report of the companies.

Vandana Gupta (2017), Finance, Fore School of Management, New Delhi, India identified the independent variables which explains default risk for

Indian listed companies. The Cox proportional hazard model is used to test the impact of financial ratios, capital market ratios, macro-economic variables, size & age of companies etc. For this purpose a dataset of 859 companies panning across 10 sectors are considered.

4. METHODOLOGY

As the objective of the study comprise of fitting survival time to an event data of companies listed under National Stock Exchange, India with appropriate survival analysis models and determining best fitted model with the help of goodness of fit tests. A total of 1328 companies/industries comprise of various sectors, which were listed in National Stock Exchange, India have been considered. Out of which a sample of 305 companies were drawn by using simple random sampling method in MS Excel respectively. Simple random sampling technique enables the inclusion of each and every company from different industrial sectors with equal probability without being biased. The sample size for the study is calculated by using the Taro Yamane (Yamane, 1973) formula for finite population. The formula is given as follows:

$$n = \frac{N}{1 + N(e)^2}$$

Where, n = required sample size

N = Total number of companies i.e. population size

e = margin of error (%)

Here, we had considered margin of error as 5% while determining the size of the sample. The companies were observed for financial year 2018 i.e. from 1st April, 2017 to 31st March, 2018. The share price of the shares on 1st April 2017 is considered as the base price. With the help of right censoring technique certain companies which failed to achieve a 10% increase in share price value are considered as censored. By going through stock return index (yearly) of credit rating agencies for stocks listed in National stock exchange, India, the return percentage values differ among the companies. In this study, the threshold percentage return is considered as 10%. This percentage value is displayed as annual average return by some credit rating agencies. Further, a group of 5 explanatory variables or factors is considered while gone through earlier research studies viz. Earnings per Share (EPS), Book Value, Operating Profit Margin, Net Profit Margin, Total Asset turnover ratio, Return on Capital Employed (ROCE) (Gupta *et al.* 2017) respectively. The survival models are fitted with the help of these explanatory variables. The information about the factors or explanatory variables is taken from the annual report of the sample companies. The annual report consists of financial ratios, balance sheet, profit & loss statement etc. The data for the fiscal year 2018 of the factors is obtained from an easy to access and useful website **moneycontrol.com**.

5. RESULTS

For fulfilling the objective of the study, the analysis of the data is carried out with parametric and semi-parametric models i.e. the Cox Proportional hazard model and Accelerated failure time models such as Exponential, Weibull, Log-Normal and Log-Logistic.

5.1. Result of the Exponential Model

Table 5.1
Summary of Exponential Model Analysis

Variables	*Coefficients*	*Standard Error*	*Z*	*P-Value*
Intercept	5.236	0.126	41.37	0.0002
Earnings Per Share	0.025	0.0023	1.11	0.2664
Book Value	-0.0098	0.00049	-2.00	0.0454
Operating profit Margin	-0.010	0.0030	-3.34	0.00083
Total asset turnover ratio	-0.158	0.0516	-3.06	0.0022
Return on Capital employed	0.0014	0.0064	0.22	0.8244

Goodness of fit of the model

Table 5.2
Output of Goodness of fit test

	Overall (score)		
-2 Log Likelihood	*Chi-square*	*d.f*	*Sig.*
-1516.6	23.63	5	0.0002

Interpretation of Exponential Model analysis

From table 5.2 it can be observed that the fitted regression model is statistically significant (p-value<0.05).Thus from table, it can be concluded that the regression model fits well with the data i.e., the fitted model is meaningful with the independent variables (factors).

From the results obtained in table 5.1 it shows that Book Value (p-value=0.045), Operating Profit margin (p-value=0.00083) and Total asset turnover ratio (p-value=0.0022) of a company has significant influence on achieving 10% percent increase in share price value of that company. The remaining explanatory variables i.e. Earnings per share (p-value=0.2664) and Return on capital employed (p-value=0.8244) indicates that these variables has least influence on achieving 10 % increase of share price.

5.3. Results of the Weibull Model

Table 5.3
Summary of Weibull Model Analysis

Variables	*Coefficients*	*Standard Error*	*Z*	*P-Value*
Intercept	5.124	0.2013	25.46	0.0002
Earnings Per Share	0.0039	0.0037	1.06	0.2891
Book Value	-0.0011	0.0008	-1.49	0.1359
Operating profit Margin	-0.0142	0.0048	-2.93	0.0034
Total asset turnover ratio	-0.195	0.086	-2.25	0.0243
Return on Capital employed	0.0036	0.010	0.36	0.7223

Goodness of fit of the model

Table 5.4
Output of goodness of fit test

-2 Log Likelihood	*Overall (score)*		
	Chi-square	*d.f*	*Sig.*
-1463.7	14.68	5	0.012

Interpretation of the Weibull Model analysis

From table 5.4 it can be observed that the fitted regression model is statistically significant (p-value<0.05).Thus from table, it can be concluded that the regression model fits well with the data i.e., the fitted model is meaningful with the independent variables (factors).

From the results obtained in table 5.3 it shows that Operating Profit margin (p-value=0.0034) and Total asset turnover ratio (p-value=0.0243) of a company has significant influence on achieving 10 % increase in share price value of that company. The remaining explanatory variables i.e. Earnings per share (p-value=0.2664), Book value (p-value=0.1359) and Return on capital employed (p-value=0.7223) indicates that these variables has least influence on achieving 10 % increase of share price.

5.5. Results of the Log-Normal Model

Table 5.5
Summary of Log-Normal Model Analysis

Variables	*Coefficients*	*Standard Error*	*Z*	*P-Value*
Intercept	4.1313	0.2099	19.68	0.002
Earnings Per Share	0.0068	0.0043	1.57	0.115
Book Value	-0.0009	0.0092	-1.06	0.290
Operating profit Margin	-0.0154	0.0051	-2.98	0.0029
Total asset turnover ratio	-0.1502	0.1055	-1.42	0.154
Return on Capital employed	0.0092	0.0115	0.80	0.423

Goodness of fit of the model

Table 5.6
Output of goodness of fit test

-2 Log Likelihood	*Overall (score)*		
	Chi-square	*d.f*	*Sig.*
-1445	11.12	5	0.049

Interpretation of Log- Normal Model analysis

From table 5.6 it can be observed that the fitted regression model is statistically significant (p-value<0.05).Thus from table, it can be concluded that the regression model fits well with the data i.e., the fitted model is meaningful with the independent variables (factors).

From the results obtained in table 5.5 it shows that Operating Profit margin (p-value=0.0029) of a company has significant influence on achieving 10 % increase in share price value of that company. The remaining explanatory variables i.e. Earnings per share (p-value=0.115), Book value (p-value=0.290), Total asset turnover ratio (p-value=0.154) and Return on capital employed (p-value=0.423) indicates that these variables has least influence on achieving 10 % increase of share price.

5.7. Results of the Cox Proportional Hazard Model

Table 5.7
Summary of the Cox Proportional hazard model

Variables	*Coefficients*	*Standard Error*	*Wald statistic*	*P- Value*
EPS	-.0026	.002	1.232	0.267
Book value	.0006	.001	1.640	0.200
Operating profit margin	.0085	.003	7.762	0.005
Total asset turnover ratio	.1112	.055	4.012	0.045
Return on capital employed	-.00028	.007	.192	0.661

Goodness of fit of the model

Table 5.8
Output of goodness of fit test

-2 Log Likelihood	*Overall (score)*		
	Chi-square	*d.f*	*Sig.*
2630.015	11.172	5	0.048

Interpretation of Cox regression Model analysis

From table 5.8 it can be observed that the fitted regression model is statistically significant (p-value<0.05).Thus from table, it can be concluded that the regression model fits well with the data i.e., the fitted model is meaningful with the independent variables (factors).

From the results obtained in table 5.7 it depicts that Operating Profit margin (p-value=0.005) and Total asset turnover ratio (p-value=0.045) of a company has significant influence on achieving 10% increase in share price value of that company. The remaining variables i.e. Earnings per share (p-value=0.267), Book value (p-value=0.200) and Return on capital employed (p-value=0.661) indicates that these variables has least influence on achieving 10 % increase of share price.

5.9. Goodness of Fit Tests Analysis

Table 5.9
Output of Goodness of fit on the basis of AIC

Models	*Exponential*	*Weibull*	*Log-Normal*	*Cox regression*
AIC	2913.68	2941.99	2903.99	2640.01

From table 5.9 it shows that the Akaike Information Criterion (AIC) for Accelerated Failure time models is more or less similar. The AIC (2640.01) for the Cox Proportional hazard model is lowest in comparison to other models respectively. Therefore, the Cox Proportional hazard model best fit the survival time of stock data listed under National Stock Exchange.

5.10. Cox-Snell Residual Plot for the Cox Proportional Hazard Model

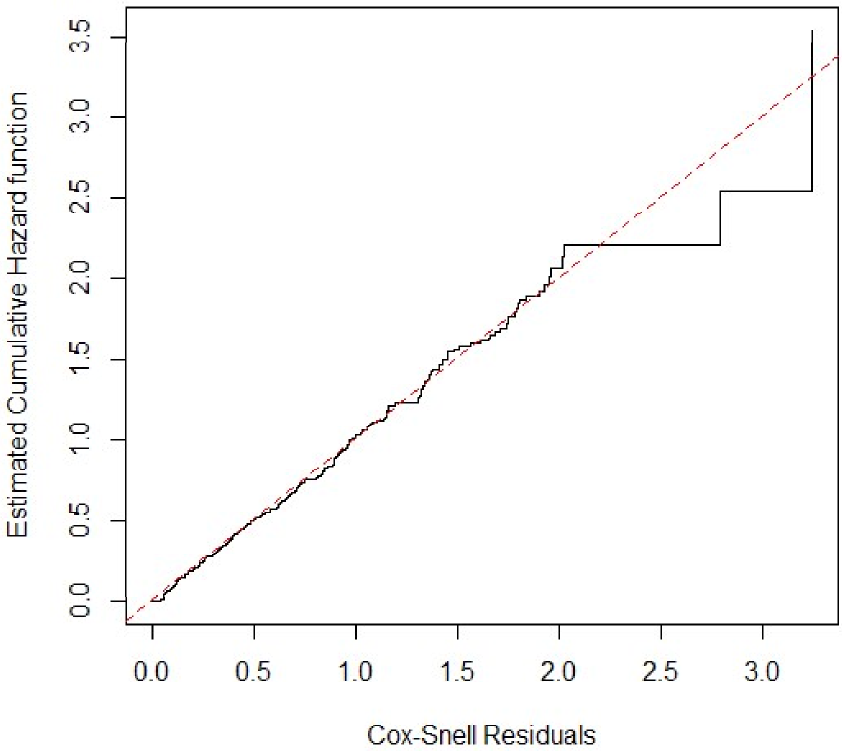

Figure 5.10: Cox-Snell Residual Plot for the Cox PH model

6. CONCLUSION

The study attempts to fit Semi-parametric and Parametric survival models of survival time to event data. From the results of the goodness of fit tests (table 5.9), the Cox Proportional hazard model has lowest AIC (2640.01). Therefore, the Cox Proportional hazard model is considered as best fitted model for studying survival time of companies under NSE. The earlier research studies may have different best fit model based on the objective of the study. From the results of table 5.7, it shows that an increase in Operating profit margin and Total Asset turnover ratio of the companies listed under National Stock Exchange(NSE) will increase the potential for achieving 10 % increase in share price value of companies under NSE. It provides information that share price movement depends on how well companies are utilizing its assets. According to the results of the study, the companies profit produced from its business operations also influence share price value.

References

Bunyaminu Alhaman (2015). "Business Failure Prediction: An Empirical study based on Survival Analysis and Generalized Linear Modeling (GLM) Techniques". *International Journal of Financial Economics*, Vol. 4, No. 3, 2015, pp. 135-149.

B. Garima, S. Balwinder (2014). "Survival Pattern of Initial Public Offerings in India: An Emprical Study". *Amity Business Review,* Vol. 15, No. 2, 2014.

Geetha.E, Swaaminathan. Ti. M. (2015). "A study on the factors influencing stock price A Comparative study of Automobile and Information Technology Industries stocks in India". *International Journal of Current Research and Academic Review,* ISSN 2347-3215, Vol. 3 , No. 3, 2015, pp. 97-109.

Gupta Vandana (2017). "A Survival Approach to Prediction of Default Drivers for Indian Listed Companies". *Theoretical Economics Letters*, ISSN 2162-2086 (Online), ISSN 2162-2078 (Paper), 7, pp. 116-138.

Hou, O.Y., Chuang, S.S. (2007). "Examining Financially Distressed Companies in Taiwan: Application of Survival Analysis". *The Business Review*, Vol. 8, pp. 231-242.

Joshi Mrunal (2013). "Factors Affecting Indian Stock Market". *International Journal of Contemporary Research in Management, Engineering and Health Science,* Vol. 1, No. 2, pp. 37-45.

Kumar, R.G. and Kumar, K.(2012). "A Comparison of Bankruptcy Models". *International Journal of Marketing, Financial Services and Management Research*, Vol. 1, pp. 76-86.

Lamberto Paul Andre, Rath Subhrendu (2010). "The Survival of Initial Public Offerings in Australia". *The International Journal of Business and Finance Research,* Vol. 4, No. 1, 2010.

Lee Chang Ming (2014). "Business Bankruptcy Prediction Based on Survival Analysis Approach". *International Journal of Computer Science & Information Technology,* Vol. 6, No. 2, 2014.

M. G. Dmitri, Golder N. Peter (2008). "Using Stock Prices to Predict Market Events: Evidence on Sales Takeoff and Long-Term Firm Survival". *Marketing Science*, ISSN-0732-2399, Vol. 27, No. 4, pp. 717-729.

Ni Jiayi (2008). "Application of Cox Proportional Hazard Model to the Stock Exchange Market". *B.S. Undergraduate Mathematics Exchange*, Vol. 6, No. 1.

Stepanova, M. and Thomas, L. C.(2002). "Survival analysis methods for personal loan data". *Operations Research*, Vol. 50, No. 2, pp. 277-289.

Yamane Taro (1973). Statistics: an introductory analysis. New York: Harper and Row. Yash.

www.moneycontrol.com