Prediction default of Indian Steel Sector using MDA, Altman, Calibrated, Logit & Structural Model

Dr. Deepika Verma
Assistant Professor
Sri Aurobindo College (Eve)
Delhi University
deepika.delhiuniversity@gmail.com
Contact No – 8470977770

Abstract: *Credit risk modeling is imperative for keeping firms safe from debt trap and bankruptcy. The present study attempted to predict the default occurrence of steel sector firms using MDA, Logit function and structural model. Study developed 2 models using MDA and Logit model, further study also evaluated the Altman original model and calibrated model by applying it on sample data of selected steel sector. The developed models have also been validated on the out-of-sample data. The study obtained satisfactory statistical results pertaining to the developed models. The classification results witnessed the following accuracies for MDA, Calibrated, Altman, Logit and Structural model such as 89%, 74%, 11%, 91 and 18%. The validation accuracies obtained by mda, calibrated and logit models are 27%, 59% and 91%.*

Keywords: *Steel sector bankruptcy, Credit risk, Default prediction models, Altman, MDA, BSM model, Logistic Regression.*

Introduction

The current debt on Indian Steel industry is amounting to Rs 3 lakh crore that is a mammoth liability for any large sector industry in India. Moreover, it has been observed that lately the steel sector is the major contributor of NPA in India with majority of defaults. Since, the Steel sector is the largest sector in India and such defaults, bankruptcy can seriously hurt the economy of the nation as the employment of lakhs of people, GDP, export rely on this sector. Moreover, the maximum steel sector firms are listed in Indian stock exchange and they are major player for sensex and nifty. During the last few year the consumption of steel has grew to 39% whereas it production had reduced by 0.7% due to lack of funds. Similarly the decrease in export is witnessed. The major causes of such situation are indebtedness of the firms, higher cost of capital, bankruptcies, and default of Indian steel sector firms. According to Crisil the latest example of bankrupted steel sector firms which have been acquired under IBC 2016 are Bhushan Steel, Essar Steel, Electrosteel Steels, Monnet Ispat & Energy Ltd and Bhushan Power & Steel till March 2021. The said firms defaulted 1.7 trillion of debts and became the major contributor of NPA in Indian commercial banks. The task of identifying probability and time of default includes lengthy calculations, development of default prediction model which if not performed carefully can cause bankruptcy situation in that very firm. It is imperative for the stakeholders of the firm to have correct knowledge about the financial soundness and the credit default risk of the firms. Lately firms are engaging both short term and long term credit management regimes internally to ensure timely service of their debts. For internal credit risk management firms are using various sorts of credit risk analysis tools for predicting the potential default events in advance. The obtained empirical result of the applied default prediction models are being used to take early measures for maintaining financial liquidity and obliging the debt holder on time. The practice of internal credit risk management is not common for the developing economy's firms. However, after the implementations of Basel II the commercial banks of such economies are using statistical, financial and reduced form model for predicting the default probability of their corporate debtors. To mitigate the bankruptcy risks of any firm it is necessary to keep a hold on the causes of defaults. Such causes can be recognised using both financial and non-financial attributes of the firm. The credit risk models used this information for developing models and classifying firms as defaulting and non-defaulting. The financial statement and annual reports of the firms play a vital role to collect required data and analyse the financial position of firm. The default estimation for any public listed company can predict by tracking the market price of their share, price earnings ratios etc. The present study is drafted to predict the default events of Indian steel sector firms. The study is divided into 4 sections. First section discussed the introduction and literature review of the study. Second section explained the research methodology of the study. The research methodology enumerated sample data, source of data and defaulted prediction methods for developing default prediction models. Third section depicted the empirical results wherein the scholar discussed in detail the analysis of each empirical findings of every developed model. Fourth section of the model included the conclusion, limitations, and recommendations of the study.

Literature Review

Study aimed to develop an early warning signal model for predicting corporate default in India using MDA. Study also employed logistic regression to directly estimate the probability of default. The sample data of 542 firms for 1998 to 2004 was accessed form Crisil which later were reduced to 52 defaulted and 52 non-defaulted matched firm sample w.r.t their size of asset, year, and industry affiliation. The sample firms data is collected from Food product, paper, textile, chemical, machinery/electrical, metal, autoparts, power, and service industry. The financial ratios considered in the study are wc/ta, cash profit /ta, solvency ratios (ta/tl), operating profit/ta, total sales/ta. The Z score successfully predicted the default with high efficiency (91%) for the hold out samples by outperforming Altman 1968 and emerging financial ratio based model.

The study contributed towards the literature of the default predicting that would facilitate the banks to price risky bonds, to fix the premium for the securities and to estimate the credit risk capital (Bandyopadhyay, 2006). Study examined the prediction capability of MDA when it is incorporated along with risk assessment variable in the prediction model. The study used the sample of 56 PA and NPA of Tanzanian Bank for the year 1985 to 1994. The study developed 5 MDA models using Financial Ratios. The financial ratios used in the study are categorised into liquidity, working capital, leverage, performance and profitability and firm size. The models achieved the accuracy of 92.9% and 96.4% for categorising the PA and NPA. The selected financial are found significant predictors that could correctly predict the NPA 2 years prior. Study concluded that using MDA can correctly classify the performing and non-performing assets (Mvula Chijoriga, 2011).

The used the sample of 66 listed manufacturing companies to predict the default using financial ratios. The author selected the sample in 2 groups, group I consists of 25 paired sample firms and group II included 14 firms from diversified asset size. The financial ratios incorporated into the study are WC/TA, RE/TA, EBIT/TA, MVE/ BOOK VALUE OF TOTAL DEBT, SALES/TA. The profitability ratio contributed the most in the default prediction of the study followed by SALES/TA ratio. Study obtained accuracy of 95% with only 6% and 3% Type I and Type II errors respectively. The study also the bankruptcy prediction 2 years prior in which study obtained 72% accuracy (Altman, 1968).

The used MDA and Logistic regression to predict the default of 139 Chinese firms belong to Metal, machinery, medicines, manufacturing, real estate, electrical equipment, food beverages, paper and printing, petroleum and other chemicals etc. Study used 25 financial ratios categorised into profitability, liquidity, solvency, potentiality, activity, capital market, capital structure. Significant ratio for the logit model is net income to ta and the ratios found significant for MDA model are net sales to ta, ebit to ta, and growth rate to ta. The classification accuracies of both developed models for 1 year prior is 98% and 80% for 2 to 3 years prior. However, the logistic model outperformed the mda by providing least % of Type I and Type II errors (Liang, 2003).

This study used Altman, Taffler and logistic regression credit score model to examine the default prediction of 105 SMEs (75 non-defaulted and 30 defaulted) of European emerging economy during the period of 2009. The results obtained were in line with the existing literature where the logistic regression outperformed mda. The study used 14 financial ratios.

(Smaranda, 2014). The study conducted a comparative analysis of MDA and Logit model with respects to their advantages and disadvantages. The study summarized the contribution of past studies about the default prediction achievements of MDA and Logistic regression. Study concluded that these models are beneficial for the financial regulators, bankers and policy makers to regulate the firms and take prior corrective actions and making policies (Hassan et al., 2017).

The study tried to measure the financial health of 4 listed Indian Steel firms by taking into account their financial information from 2008 to 2015. The study applied z score model that categorised the firms into safe zone according the z score obtained by each firm (C. R. Marvadi, 2016).

The study conducted the credit risk assessment of 8 Indian Steel sector firms using Altman Z score model. study concluded that using Altman Z score firms can decide to take the decision of bankruptcy by estimating the financial risk (Ramaratnam & Jayaraman, 2011).

The study used Altman Z score model to predict the default probability and to examine the role of size of firm for predicting the default risk. The study collected the financial information for 5 years from 2009 to 2014 of 10 selected Steel firms. Study found that the size of firm (Ta/Total sales) has inverse relation with default probability of firm.

(Singla & Singh, 2017). Original Altman Z score is applied to inspect the financial health of the Indian Steel firms such as JSW, SAIL, and Steel exchange of India, Tata Steel and Visa Steel. The sample data was obtained from the CMIE Prowess database.

Findings confirmed that all the selected firms are in a safe zone (Ramaratnam & Jayaraman, 2011). The study conducted a default prediction process of selected Indian Steel firms to classify them into defaulted and non-defaulted groups for the span of time 10 years from 1991 to 2011 using 23 financial ratios. The analysis of the results illustrates that the ROI, Debtor turnover, FA turnover ratios are significant default predictors.

Additionally, the study recommended to increase ROI, improve debtor management system and to strengthen the fixed assets management for preventing the firms from failure (Pal, 2013).

Marvadi (2016) attempts to gauge the financial health of the Steel companies of India namely Tata Steel, JSW, Jindal Steel, and SAIL for the period 2008-09 to 2014-15 using the Z score model. The findings confirm that Z score worked pretty well for assessing the financial position of all the firms and indicated that the SAIL is highly profitable, Tata Steel growing upward, but JSW Steel dwindles and lies in the grey zone, and Jindal Steel grouped into the highly risky zone than can go bankrupt (C. R. Marvadi, 2016).

Ravi Singla (2017) attempts to inspect the impact of firms' size on the probability of default while estimating the probability of distress of 10 Steel companies for the period from 2009 to 2014. The study calculated the Altman Z score using sales, net profit, share capital, and total assets. The findings disclosed that the size has an inverse relation to the failure occurrence and sales are the core coefficient of default prediction (Singla & Singh, 2017). The study aspires to examine the financial health of the Indian Steel sector for 10 years from 2006 – 2016 by integrating the Altman Z score. The findings revealed that the Ferroalloys Corporation is the most solvent firm followed by Bajaj Steel, SAIL, SEIL and Uttam Galva. The firms namely Factor Steel, Jindal Steel, JSW, Tata Steel and Visa Steel are drifting towards failure (Marvadi, 2018). This study evaluates the financial ability of the Steel industry of Chhattisgarh using the Altman Z score from 2006 to 2015. The computed result exhibits that the majority of the firms are not found financially sound. This study urges to encourage the technical and managerial elevation for improving the financial health of the firms (Vadyak, 2017).

Zeitun (2007) attempts to explore the role of cash flow on the financial distress of 167 listed Jordan companies for the period 1989-2003 in an emerging market using panel data of the paired sample by employing the Logit function. The findings of the study were: the capital structure determines the probability of default, cash flow is a significant indicator of default & the financial position of the firm directly impacts the management practice (Zeitun et al., 2007).

Lieu (2008) proposed an early warning model using Logit regression for 116 (58 distressed and 58 non-distressed) listed Taiwanian firms for the horizon of 5 years from 2002 to 2007. The model provided the risk probability for 1-3 years before the event using financial ratios. The financial ratios are found to be key indicators of credit risk modeling. The result of the study is consistent to Holian & Joffe (2013) (Lieu et al., 2008).

Frade (2008) aims to create a model which can predict that 186 US issuers shall default within a year. The study used financial ratios and value of equity as the independent variables that incorporated Logistic, Altman Z score, Barclay's & bond score CRE default model. The data related to financial and market information was collected for the period 1996-2008. It is evident from the findings of the model that all the market variables are not significant predictors in a logistic regression model (Frade, 2008). A binary logistic model was developed for examining the Chinese SMEs from the year 2004 to 2007. The study concluded that only financial indicators are not enough to predict insolvency therefore, it's imperative to include qualitative indicators for eg the Type of ownership that would accelerate the model's accuracy (Wang & Zhou, 2011).

The study applied the regularization approach along with Logit to develop a default predicting model for South-Asian companies and to identify the significant predictors of default. The outcome of the study does highlight that the higher accuracy, depicts that the regularization approach is well capable to forecast and to select the default predictors for Indonesia, Singapore, and Thailand countries (Härdle & Prastyo, 2013).

Falkenstein et. Al (2000) applied Moody's KMV model for computing the distressed risk of private firms data consist of a sample data of 28104 non-distressed and 1604 distressed firms. The empirical findings unveiled that by analysing balance sheet data accompanied by market variables escalates the prediction power of public listed firms for a time horizon of 1 to 5 years (Falkenstein et al., 2000).

Moody's KMV 2003 computed the expected default frequency of 250000 company years data including default firms data of 4700-year firms by including variables such as firms asset value, default point, the book value of debts, & volatility of assets at time T. The study adhered to the assumption of BSM that the firm triggers to default when its net worth becomes zero (Kealhofer, 2003).

Hayne E Leland (2004) predicted the default probability of corporate bonds rated by Moody's for the period commences from 1970 to 2000. The model did perform the prediction quite accurately for the long term bonds than the short term bonds. The study also examined the exogenous and endogenous default boundary and its implication on the capital structure and rating methods (Leland et al., 2004).

The Merton-model approach was applied to predict the bankruptcy of individual UK companies and a group of bankrupt companies during 1990-2001. The study stated the advantages of the model for indicating failure one year prior. The study compared the model to Reduced Form model and proclaims that the Structural Model outshines the Reduced Form model for a horizon of 1 year. On the contrary the Reduced Forms model outperforms when prediction is conducted marginally (Tudela & Young, 2005).

The study employed a Structural Model for describing financial distress. The sample data was collected from 420 failed US firms from 1986 to 2001. The result signifies that a firm's volatility is the best determinant of bankruptcy for 5 years prior. Besides this, D2D is also a significant indicator of bankruptcy. The distances to default (d2d) and the probability of default at maturity (-d2) were found as the significant predictors of default (Charitou & Trigeorgis, 2005).

This study proposes an econometric method for forecasting the term structure of default probabilities for multiple future periods. The sample data comprised of 2700 US-listed companies for 1980-2004. The sample data of the bankruptcy firms was collected from Moody's default risk service and CRSP. The empirical result unveiled that the Structural Models along with macroeconomic variables can provide better estimation.

Objectives

- To develop models using MDA and Logit function for selected units in steel sector
- To predict default of Indian steel sector firms using Altman, Calibrated and Structural model.
- To validate developed MDA and developed Logit model on out-of-sample data of selected Indian steel sector.
- To compare the statistical and default prediction significance of developed and existing model.

Research Method

Sample Selection & Period of Study

The study incorporated the sample data for 15 years' time horizon from 1st April 2004 to 31st March 2019 to develop the credit risk models and to predict the default probability. The sample contains data of Indian BSE listed firms collected from Indian steel sector.

Table 1: Detailed Description of Steel Sector

Sector	Defaulted firms	Non-Defaulted Firms
Steel	33	48

Sources of Data

The present study is empirical in nature; the sample data used in the study are secondary data. The study sourced the data from a variety of sources that include company-specific data, RBI interest rate, market proxy, accounting information and market price of stock of the firms.

Company-Specific Data

The study collected the company specific information such as the accounting, market and macroeconomic data of the selected Indian firms from various sources. The accounting data was fetched from the individual financial statements of each selected firm and share price information was retrieved from the BSE website. The macroeconomic data such as interest rate and GNP index were collected from the database maintained and uploaded on the websites of RBI and World Bank. The information about the default status of selected firms is sourced from the audited annual reports of all selected firms for 15 years from 1 April 2004 to 31st March 2019.

Risk-Free Interest Rate Proxy

The sample data of the proxy interest rate of 91 days Treasury bill is collected from the database maintained by Reserve Bank of India on its website.

Market-Proxy

The information about the daily average price of shares, return on the shares, BSE index and return on BSE index of the selected firms is collected from the BSE website.

Default Prediction Methods used in the study

In light of the previous literature review, the study selected 5 default prediction methods to predict the default status of the selected firms namely MDA (Multiple Discriminant Analysis), Calibrated, Altman Original model, Logistic Regression, and Structural Model to provide the comparative analysis of the Classification results of these function. The conceptual frameworks, mathematical processes of each applied method have been discussed in detail below.

Analysis & Findings

Variables

Variables incorporated in the study to develop and validate the Default Prediction Models are of two kinds such as dependent & independent variables to predict the default probability of each selected sector.

Dependent Variable of MDA Model

Z score: It is a credit rating score that is calculated using the independent variables. The Z score categorises the sample cases into defaulted and non-defaulted groups. For categorising purpose the study shall use the centroid value of each group namely defaulted and non-defaulted. The centroid values of each group of each selected sector have been calculated after processing the sample cases on IBM SPSS Software version 22.

Independent Variables used in MDA Model

The present study has used 21 independent variables for predicting the default probabilities that belong to accounting, market and economic variables.

Table 2: Description of 21 Independent V ariables of MDA

Independent Variables		
Accounting Variables	**Market Variables**	**Economic Variables**
WC/TA	MP/EPS	LOG(TA/GNP)
RE/TA	MP/BV	SALES GROWTH/GNP GROWTH
EBIT/TA	MVE/TBD	
SALES/TA		
CA/CL		
NI/TA		
NP/TE		
TBD/TA		
EBIT/INT		
OCFR		
GRTA		
INVENTORY TURN		
FAT		
D/E		
TL/TA		
SALES GROWTH		

Dependent Variable

L Score: The L score is also a credit rating score but unlike MDA the determination of the L score is based upon simple criteria i.e. if the inverse of exponent of L score is <.5 then the firm is non-defaulted & vice versa. That's why the logistic model is called as binary Logit model because the dependent variable of the provide dichotomous result i.e. 0 and 1.

Independent Variables Used in the Logit Model

This model has incorporated 23 independent variables to predict the default probability. The Independent variables are comprised of accounting variables, market variables, economic and categorical variables. Logit model incorporated 2 qualitative variables namely X and Y along with 21 accounting, market and economic variables that are integrated into the MDA model.

Table 3: Description of 23 Independent Variables

Independent Variables			
Accounting Variables	**Market Variables**	**Economic Variables**	**Categorical Variables**
WC/TA	MP/EPS	LOG(TA/GNP)	X= 1, TL >TA and X= 0, TA>TL
RE/TA	MP/BV	SALES GROWTH/GNP GROWTH	Y=1, Avg NP for 2 years < 0 and Y=0, Avg NP for 2 years >0
EBIT/TA	MVE/TBD		
SALES/TA			
CA/CL			
NI/TA			
NP/TE			
TBD/TA			
EBIT/INT			
OCFR			
GRTA			
INVENTORY TURN			
FAT			
D/E			
TL/TA			
SALES GROWTH			

Dependent Variables of Structural Model

EDF: Expected Default Frequency is a dependent variable of Structural Model. It is a probability that a firm will default over a period of time when the market value of firm's assets falls below the book value of its Debts.

Independent Variables of Structural Model

The variables employed in the Structural Model are the Market Value of the firm's Assets, book value of the outside liability and drift rate that is used to calculate the probability of default which has been accessed from the financial statement and market-driven information.

Empirical Results

Model developed using MDA

Table 4: Developed MDA Model

Steel	Z = 0.392+11.68*EBIT/TA-24.905*NI/TA+0.044*FAT-0.159*LOG(TA/GNP)

Model developed using Calibrated Model

Table 5: Models Developed using Calibrated Model

Steel	-0.494+0.892* WC/TA +12.245* RE/TA -4.942* EBIT/TA +0.009* MVE/TBD +12.105* SALES/TA

Model developed using Altman Original

Table 6: Models Developed Using Altman

0.012*WC/TA+0.014*RE/TA+ 0.033*EBIT/TA+ 0.006*MVE/TBD+0.999*SALES/TA

Case Summary

Table 7: Summary of Cases processed from Steel Sector

Steel	In-sample	Out-of-sample
Total cases	800	269
Cases considered	746	220
Cases removed	54	49

Log Determinant

Table 8: Log Determinant

Sectors	Non-Defaulted	Defaulted	Pooled within-groups
Steel	**20.994**	**1.232**	**21.583**

One of the assumptions of the discriminant function is to have homogeneity of covariance matrices between the groups. The relatively equivalent log determinant values of the groups recommend that the covariance matrices of these groups are homogenous. Besides, homogeneity, proximity in the log determinants values of non-defaulted, defaulted and pooled with-in group indicates the robustness of the developed prediction model. The log determinant values as depicted in Table No 8 Log determinant of selected steel sector is neither equivalent nor close to defaulted, non-defaulted and pooled within group further, Table enumerates that for the log determinant values for the non-defaulted groups and pooled within-groups are closer to each other yet, these are quite distant from the defaulted groups due to the existence of higher Type II Error in the prediction results of steel sector.

Coefficients of MDA Model

Table 9: Coefficient of MDA Model

Particular	Box's M	Sig. Value of Box M	Eigen value	Canonical Correlation	Wilks' Lambda	Sig value of Wilk's lambda
Steel	2039.15	0	0.173	0.384	0.853	0

Box's M Test

To evaluate the Multiple Discriminant Analysis function's assumptions about the equality of variance-covariance matrices in dependent variable's groups (defaulted and non-defaulted) the study used Box's M Test.

Hypothesis 1

H_0: The covariance matrices are equal in both the groups namely defaulted and non-defaulted made by dependent variables of the developed models.

H_1: The covariance matrices are different in both the groups namely defaulted and non-defaulted made by dependent variables of the developed models.

The significant P-value of the box's M test of all selected sectors as depicted in Table No Coefficients contravenes the basic assumptions of the MDA function. The large sample data produces a higher value of the Box's M which generally results in a significant value of the box's M test in such instances the assumption is tested using the Log Determinants test. The Box' M value of steel sector is higher in conjunction with significant sig-value of Box's M test i.e. <.05. This is an unpleasant result that conveys the violation of the assumption of MDA. Hence, the H_0 will be rejected; this finding of the study about the Box's M test is consistent with the findings of Bandyopadhyay (2006), Altman (2000) and contrary to Suleiman (2014) and Memic (2015).

Eigen Value

The eigen value denotes the variation in the dependent variable that can be explained by the MDA model. Primarily the Eigen value is a ratio between explained and unexplained variance. The higher eigen value recommends the greater discriminatory power of MDA function that explains the variation in the dependent variable. The strong discriminant function has a higher eigen value i.e. close to 1. The present study found only .173 that conveys less prediction power of the developed models. This signifies that the variation in the dependent variable of the developed model of steel sector is explained by the developed model only with 17% accuracy.

Canonical Correlation

The Canonical Correlation gauges the association between the groups of dependent variable and discriminant function, the value of canonical correlation lies between 0 to 1. The large value of canonical correlation implies a strong association between the groups of dependent variable and developed models. Further, it signifies the high classification accuracy of the developed model. The discriminant function with a high value of canonical correlation value i.e. close to 1 is an acceptable discriminant function model. The square of Canonical Correlation is similar to R square which explains the variation in the dependent variable. When the squared value of the Canonical Correlation is more than 50% it conveys the high competence of the discriminant function. The canonical correlation values as exhibited in Table for steel sector is only 0.38. Hence, the result substantiates the less satisfactory classification ability of the developed models.

Wilk's Lambda

Wilk's lambda describes the discriminatory power of the discrimination function together with independent variables incorporated in the developed model. The Wilk's lambda ranges from 0 to 1, the smaller value signifies the higher classification accuracy of the model coupled with the significant contribution of each independent variable. Wilk's lambda always works in contrast to the canonical correlation, the higher value of the canonical correlation will lead to a lower value of Wilk's lambda which is a desirable situation for any robust model. Table No 9 Coefficients MDA model exhibits .85 along with .38 canonical correlation. This indicates lower competency of the developed model.

Hypothesis 2

H_0: There is no discriminating power in the independent variables of the developed models.

H_1: There is discriminating power in the independent variables of the developed models.

The sig value of Wilk's lambda for steel sector is <.05 that substantiates that there is a significant difference between defaulted and non-defaulted group of the dependent variable, also that the independent variables are contributing significantly well for discriminating the defaulted and non-defaulted group of dependent variable. Hence the H_0 hypothesis will be rejected, these findings of the present study concerning the hypothesis test result of each developed model and wilk's lamda value obtained for each selected sector and Complete Sample are consistent with

Altman (2000), Altman (1968) and Memic (2015).

Test of Equality of Group Means

This test measures the difference between the groups made by each independent variable.

Table: 10 Test of Equality of Group Means

Independent Variables	Steel		
	Wilks' Lambda	F	Sig.
WC/TA	0.989	8.226	0.004
RE/TA	0.991	6.508	0.011
EBIT/TA	0.96	31.081	0
MVE/TBD	0.996	2.886	0.09
Sales/TA	0.993	5.353	0.021
CA/CL	0.998	1.304	0.254
NI/TA	0.971	22.295	0
NP/TE	0.999	0.525	0.469
TBD/TA	0.988	9.409	0.002
EBIT/INT	0.999	0.621	0.431
OCFR	0.999	0.862	0.353
GRTA	0.988	8.841	0.003
INVENT. TURN	0.999	0.749	0.387
FAT	0.984	12.329	0
MP/EPS	0.999	1.087	0.298
MP/BV	0.998	1.337	0.248
D/E	0.999	0.909	0.341
TL/TA	0.996	2.643	0.104
Log (TA/GNP)	0.974	19.793	0
SG	0.993	5.508	0.019
SG/GNP Growth	1	0.218	0.641

Hypothesis 3

H_0: The mean of each independent variable between the defaulted and non-defaulted groups of developed models are equal.

H_1: The mean of each independent variable between the defaulted and non-defaulted groups of developed models are not equal.

Table no 10 Test of Equality of Group Means demonstrated the Wilk's lambda, Sig and F-value for each independent variable of the models developed for steel sector. The study has prepared the summary of the total significant independent variables of each selected sector after considering the values obtained in the test of equality of group means. Table No significant factors described the number of significant factors/independent variables of each developed model of each selected sector. By taking into account the significant variable with sig value <.05 it is inferred that the group means of each independent variable between the defaulted and non-defaulted groups are not equal. Therefore, the H_0 hypotheses will be rejected. The finding of this hypothesis test resonates with the results obtained by Sirirattanaphonkun (2012)

Significant Factors

Table No 11 Significant factors prepared by analysing the independent variables having sig value <.05. The sig value of each independent variable is accessed from Table no Test of Equality of Group Means.

Table : 11 Significant Factors

Sector	Number of significant factors
Steel	3

Source: Created by Scholar using the results of Test of equality of group means

Structure Matrix

Table No 12 Structure Matrix demonstrated the structure matrix values of each independent variable processed on IBM SPSS version 22 for the development of MDA models for each selected sector. The structure matrix values given corresponding to each independent variable indicates the contribution of these variables in the default prediction models. The independent variables having structure matrix value >.3 is considered for the model development and remaining independent variables have been dropped from the model.

Table 12 : Structure Matrix

Independent variables	Steel
CA/CL	0.101
D/E	-0.084
EBIT/Int	0.07
EBIT/TA	0.492
FAT	0.31
GRTA	0.262
Inventory Turnover	0.076
Log(TA/GNP)	-0.393
MP/BV	0.102
MP/EPS	-0.092
MVE/TBD	0.15
NI/TA	0.417
NP/TE	0.064
OCFR	0.082
RE/TA	0.225
Sales Growth	0.207
Sales Growth/ GNP Growth	-0.041
Sales/TA	0.204
TBD/TA	-0.271
TL/TA	0.143
WC/TA	0.253

Sources: Values obtained by conducting structure matrix test on SPSS

The structure matrix performs function similar to factor analysis; it displays the correlation of each independent variable to discriminant function. The predictors with a high structure matrix value contribute the most in the default prediction model to discriminate the groups of the dependent variables and to predict default. The structure matrix is more significant than the standardized canonical discriminant function coefficient to evaluate the discriminatory power of the discriminant model. The predictors/ independent variables having large structure matrix value are required to be incorporate in the model for predicting the default probability. Table No Structure Matrix shows that the steel sector has only 4independent variables with structure matrix value >.3 which shall become the significant contributors/ predictors of the default prediction model.

In-Sample Classification Result of Developed MDA, Calibrated model and Altman's model

Table 13: In-Sample Classification Result

Sectors	Models	Accuracy Rate	Type I Error	Type II Error
Steel	Developed model	89%	2%	84%
	Calibrated Model	74%	24%	43%
	Altman's Original	11%	100%	0%

Source: Prepared by Scholar using the default prediction results of MDA models obtained from SPSS and MS-Excel

Validation of the Developed Model on out-of-sample data

Table 14 : Validation Results

Sectors	Models	Accuracy Rate	Type I Error	Type II Error
Steel	Developed model	27%	74%	20%
	Calibrated Model	59%	45%	5%

Models Developed using Logit Function

Table 15: Developed Logit Model

Sectors	Logit Models
Steel	L = -3.683+1.711*WC/TA-11.539*RE/TA-0.737*MVE/TBD-4.771*SALES/TA+20.906*NI/TA+0.044*TBD/TA-1.436*GRTA+0.047*INVEN. TURN-0.303*FAT-1.854*TL/TA+0.375*LOG(TA/GNP)+1.91*Y

Cases Processed

Table 16: Summary of Cases Processed

Steel Sectors	In-sample	Out-of–sample
Total cases	770	153
Cases considered	730	123
Cases removed	40	30

Coefficients of Logit Model

Table 17: Coefficient of Logit Model

Sectors	Omnibus tests of the model coefficient (Chi-Square)	Sig Value of Omnibus tests	-2 Log likelihood	Cox & Snell R Square	Nagelker R Square	Hosmer and Lemeshow Test	Sig. value of Hosmer and Lemeshow test
Steel	183.776	0	329.205	0.223	0.441	5.049	0.752

Source: These are the results of various tests conducted on Logit models using

Omnibus Test

The Omnibus Test evaluates the significance of each independent variable of the model for predicting the default risk of the firm, for recognising the best fitting independent variables of the model, and for assessing the overall robustness of each developed model. The small value of chi-square with sig value <.05 specify the higher predictive accuracy of the developed model. In Table No Coefficients the chi-square value of the steel sector is 183.776 with 0 sig-value that suggests the robustness of the credit risk model.

Hypothesis 1

H_0: The independent variables of the developed models have no significant impact on thedependent variables.

H_1: The independent variables of the developed models have significant impact on thedependent variables.

Since the sig- values for all selected sectors in the Omnibus test given in the table are less than .05, hence it suggests the rejection of the null hypothesis. The findings of this hypothesis test are consistent with Suleiman, Suleman, Usman and Salami (2014) and Kwofiew (2015).

-2 Log likelihood

-2 Log Likelihood test examines the robustness of the model. The large values of the -2 log-likelihood depict the high robustness of the developed model. The study obtained sufficiently large value of -2 log-likelihood for steel sector i.e. 329.205. This indicated the greater classification ability of the developed model for these sectors.

Cox & Snell R Square Test

The Cox & Snell R square test provides the measure to examine the variation in the dependent variable that can be explained by the developed model. The study attained only .22 Cox & Snell R value for steel sector this refelects that the developed model explained the variations in the dependent by only 22%.

Negelker R square

Negelker R square is a Pseudo R square of the Logit model which assesses the variation in the dependent variable of the model that can be explained by the independent variables included in the logistic regression model. The study found 0.441 Negelker R square value for steel sector that is quite low. These results demonstrated that the variation in the dependent variable steel sector is explained by the independent variables by only 44% accuracy.

Hosmer and Lemeshow Test (Goodness-of-Fit Test)

The Hosmer and Lemeshow test evaluates the goodness of fit of the sample data for predicting the default probabilities. This test also indicates whether the model is specified or not which implies that how perfectly the groups of dependent variables can be classified according to the predicted probabilities. The hosmer lemeshow test is similar to the chi-square goodness of fit test of the regression. The small value of the Hosmer and Lemeshow test suggests the good fit of the sample data into the model. The insignificant sig-value value i.e. P value >.05 recommends that the data is best fitted into the specified model.

Hypothesis 2

H_0: The developed models are correctly specified and best fitting.

H_1: The developed models are not correctly specified and best fitting.

Table No 9 coefficients present that the sig-value of hosmer lemeshow test of each model developed for selected steel sector is non-significant i.e. P-value is > .05 hence; the developed model is specified and best fitting into the sample data to predict the default probability. Therefore, the study fails to reject the null hypothesis. This finding about the hypothesis test is consistent with Kwofie (2015).

Variables in the Equation Test

This test outlines the significant factors which are required to be integrated while developing the model and the factors that are to be dropped.

Table 18: Variables in Equation

Particulars	Steel		
Coefficients	**B**	**Wald**	**Sig.**
WC/TA	1.711	4.582	0.032
RE/TA	-11.539	9.785	0.002
EBIT/TA	-5.626	0.923	0.337
MVE/TBD	-0.737	5.677	0.017
SALES/TA	-14.771	13.796	0
CA/CL	0.016	0.482	0.487
NI/TA	20.906	8.604	0.003
NP/TE	-0.145	1.431	0.232
TBD/TA	0.044	4.227	0.04
EBIT/INT	0.001	0.063	0.802
OCFR	0.273	1.606	0.205
GRTA	-1.436	4.323	0.038
INVEN. TURN	0.047	7.405	0.007
FAT	-0.303	9.306	0.002
MP/EPS	0.004	2.736	0.098
MP/BV	0.058	1.042	0.307
D/E	-0.023	0.928	0.335
TL/TA	-1.854	10.131	0.001
Log (TA/GNP)	0.375	16.563	0
SG	0.013	0.002	0.966
SG/GNP Growth	0.003	1.044	0.307
X	-0.171	0.857	0.355
Y	1.91	20.701	0
Constant	-3.683	20.233	0

Source: Values obtained by author from the output of SPSS version 22

Hypothesis 3

H_0: The corresponding coefficient to each independent variable of the developed models is zero.

H_1: The corresponding coefficient to each independent variable of the developed models is not zero.

Table No 18 Variables in Equation witnessed that in the steel sector the significant variables count is quite good that is 13 with sig-value less than .05. The detail of the same is illustrated in Table No 19 Significant Factors. Hence, for the select sector the null hypotheses will be rejected. The findings of the hypothesis test are consistent with Soureshnami & Kimiagri (2013), Sirirattanaphonkun, Suluck (2012).

Significant Factors

Table No 19 Significant Factors highlights the number of significant variables included in the developed model of each sector. The significant variables have been selected from Table No Variables in Equation. The significant variables are the variables having <.05 Sig value as displayed in the table.

Table 19: Significant Factors

Sectors	Significant variables
Steel	13

Source: Prepared by Scholar using the results of Variables in the equation Table

In-sample classification result of the Logit model

Table 20: In-Sample Classification Result

Sectors	Accuracy Rate	Type I Error	Type II Error
Steel	91%	2%	62%

Validation of Model (out-of-sample classification result) of the Logit Model

Table 21: Validation Result

Sectors	Accuracy Rate	Type I Error	Type II Error
Steel	**91%**	**6%**	**100%**

Multiple Discriminant Analysis

Empirical results of the steel sector depicted that the MDA model developed for the sector included economic variables as well coupled with financial variables. The independent variables integrated into the prediction model are EBIT/TA, NI/TA, FAT and Log (TA/GNP). The trained and tested sample observations of this sector are 746 and 220 respectively. Table No 8 Log Determinant illustrated that the values of the non-defaulted and pooled within-group are quite similar yet far from the defaulted group that might be due to the high level of Type II Error in the model and it also specifies that the covariance matrices of the groups are similar. The coefficients namely Eigenvalue, Canonical Correlation and Wilk's Lambda exhibited in Table no coefficients are 0.173, .384 and .853 respectively. The coefficient values obtained in the model are inadmissible for the robustness of the model. Table No significant factors highlighted the significant factors of the model which were consistent with the variables selected in Table No 12 Structure Matrix.

Further, finding demonstrated that the developed model performed quite well followed by calibrated model for In-sample classification presented in Table No In-Sample Classification Results. Table No validation results presented a low accuracy level of the developed model along with high Type I Error this is not considered prediction worthy.

Logit Model

The default prediction model of the Steel sector is based upon the financial, market, economic and qualitative variables namely WC/TA, RE/TA, MVE/TBD, SALES/TA, NI/TA, TBD/TA, GRTA, INVENT TURN, FAT, TL/TA, LOG (TA/GNP) and Y. The observations included in the model are 730 in-sample data and 123 out-of-sample data for developing and validating the model. Table No Coefficients displayed the results of the various statistical tests conducted for evaluating the robustness and significance of the model that interpreted that the independent variables of the model had a higher contribution for predicting the credit risk. Further, the model was highly robust as the -2 Log-likelihood displayed value that at the higher side. This model failed to explain the variations in the dependent variable properly. The Hosmer and Lemeshow test substantiated that the model was specified and the sample data used for the study has a greater level of goodness of fit. This model predicted the defaulted and non-defaulted cases for both In-sample and out-of-sample data with 91% accuracy in conjunction with lesser value of Type I and Type II Errors.

Structural Model

Case Summaries

Table 22: Cases Processed

Sectors	Cases Processed
Steel	792

Prior Probabilities

Table No 23 Prior Probabilities summarises the prior probabilities of each group namely non-defaulted and defaulted. The probabilities are calculated on MS-Excel that provided information about the proportion of each group (defaulted and non-defaulted) included in each selected sector's sample data.

Table 23: Prior Probabilities

Particular	STEEL	
	CASES	PROB
NON-DEFAULTED	694	88%
DEFAULTED	98	12%
TOTAL	792	100%

Source: Calculated on MS-Excel considering the sample data

D_2 and Default Probabilities

Table 24: D_2 and Default Probabilities

Particulars	Steel	
	AVG D_2	AVG PROB
NON-DEFAULTED	18.3323	93%
DEFAULTED	11.7888	94%
TOTAL	18.2002	93%

Source: Authors calculated the values using MS-Excel

Classification Results of Structural Model

Table 25: Classification Result of Structural Model for Steel Sector

	NON-DEFAULTED	DEFAULTED	Total
NON-DEFAULTED	54	640	694
DEFAULTED	7	91	98
Accuracy Rate		18%	792
Type I Error	92%		
Type II Error	7%		
Sectors	Overall Accuracy Rate	Type I Error	Type II Error
Steel	18%	92%	7%

Analysis of the findings of Structural model

The found results depicted in Table No Classification Result of structural model for steel sector depicted that the structural model performed quite well for classifying defaulted cases than non-defaulted cases. As it's reflected in the table that out of total 98 defaulted cases structural model correctly classified 91 cases that amounts to 93% accuracy. Nonetheless, for overall accuracy the structural model provided only 18% accuracy due to higher level of Type I error.

Conclusion

The developed MDA and developed Logit model predicted the default event of the firms using the accounting variables such as NI/TA, WC/TA, EBIT/TA, RE/TA, TBD/TA, FAT which are consistent with Aguado & Benito (2013), Zmijewski (1984), Ohlson (1980), Altman (1968, 1993), While Chen and Shimerda (1981), (Casey & Bartczak, 1985), Shimerda (1981), Arlov, Rankov & Kotlica (2013), Jaffari & Ghafoor (2017). The Logit model also used one economic and one qualitative variable i.e. Log (TA/GNP) and Y for credit risk modeling that was supported by Hu & Sathye (2015).

The in-sample accuracy rate of the MDA model is 89% and calibrated model is 74% which is commensurate with the accuracy levels achieved by Karthik, Subramanyam, Srivastava, & Joshi (2018), Rashid (2014), Purohit, Mahadevan, & Kulkarni (2012), Zvaríková & Majerová (2014), Agrawal & Maheshwari (2019), Verma (2019), Verma& Raju (2019), Upadhyay (2019). However, the obtained accuracy rates of the developed MDA model are less than the level of accuracy acquired by Pongsatat et al. (2004), Bandyopadhyay A (2006), Sheikhi, Shams, & Sheikhi (2012), Sharma, Singh, &Upadhyay (2014).

The accuracy rates of out-of-Sample data for the developed MDA model is 27% whereas calibrated model obtained 59% accuracy for validation results. The acquired accuracy rates of calibrated models are similar to Agrawal & Maheshwari (2019), Sarlija & Jeger (2011), Altman & Sabato (2005), Agrawal K (2015) but less than Bandyopadhyay A (2006), Ong, Yap, & Khong (2011), Low, Nor, & Yatim (2011) and Hassan, Zainuddin, & Nordin (2018).

Altman's original model displayed poor accuracy rates i.e. 11% which is contrary to the predictive accuracy rates obtained in the studies namely Anjum (2012), Kumar & Rao (2014), Agarwal and Taffler (2005), Sarbapriya (2011), Rayalaseema and Muhammad (2012), Altman et al. (2014), Celli (2015), Karamzadeh (2013), Verma & Raju (2019), Hayes, Hodge, & Hughes (2010).

The developed logistic model outperformed the developed MDA model that attained higher predictive accuracies i.e. 91% for both in-sample and out-of-sample data. These accuracy rates are reasonably good for any robust model. The achieved accuracy rates of the developed Logit model are close to Ohlson (1980), Bandyopadhyay (2006), Agrawal & Maheshwari (2019), Sheikhi, Shams, & Sheikhi (2012), Upadhyay (2019), Ong, Yap, & Khong (2011), Moghadas & Salami (2014), Gurny & Gurny (2013), Ansari & Benabdellah (2017), Altman & Sabato (2005), Bandyopadhyay (2007), Kim & Gu (2010). Nonetheless, higher than the accuracy obtianed by Khemais, Nesrine, & Mohamed (2016), Mihalovic (2016), Memic (2015), Kwofie, Ansah, & Boadi (2015).

The overall predictive accuracies of the Structural Model attained 18% for selected steel sector which is not satisfactory due to the high level of Type I Error. However, the Type I Error was not as costly as the Type II Error according to the previous studies, yet it drops the overall classification accuracy of the model. The higher rate of Type I Error is also observed by Rao Atmanathan, Shankar, & Ramesh (2013) in the Structural Model. The results signify that the Structural Model did classify the defaulted cases with elevated accuracy but failed to recognise the non-defaulted cases in all selected sectors. The prediction accuracies acquired by the Structural Model in the present study are less than the predictive accuracies obtained by previous studies such as Karthik, Subramanyam, Srivastava, & Joshi (2018), Duan, Miao, & Wang (2014), Sharma, Singh, & Upadhyay (2014), Ko, Blocher, & Lin (1986), Mileris (2010), Hasanzedeh & Yazdanian (2017) and Bandyopadhyay (2007).

References

Altman, E. I. (1968). *Financial Ratios, Discriminant Analysis and the Prediction of Corporate Bankruptcy* (pp. 589–609).

Bandyopadhyay, A. (2006). Predicting probability of default of Indian corporate bonds: logistic and Z-score model approaches. *Journal of Risk Finance*, *7*(3), 255–272. https://doi.org/10.1108/15265940610664942.

Charitou, A., & Trigeorgis, L. (2005). Explaining Bankruptcy Using Option Theory. In *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.675704.

Duffie, D., Saita, L., & Wang, K. (2007). Multi-period corporate default prediction with stochastic covariates. *Journal of Financial Economics*, *83*(3), 635–665. https://doi.org/10.1016/j.jfineco.2005.10.011.

Falkenstein, E., Boral, A., & Carty, L. V. (2000). RiskCalc TM For Private Companies: Moody's Default Model Contact Phone RISKCALC TM FOR PRIVATE COMPANIES: MOODY'S DEFAULT MODEL Rating Methodology. *Moody's Investors Service*, *May*, 1–88. https://riskcalc.moodysrms.com/us/research/crm/56402.pdf.

Härdle, W. K., & Prastyo, D. D. (2013). *www.econstor.eu*.

Hassan, E. ul, Zainuddin, Z., & Nordin, S. (2017). A Review of Financial Distress Prediction Models: Logistic Regression and Multivariate Discriminant Analysis. *Indian-Pacific Journal of Accounting and Finance*, *1*(3), 13–23. https://doi.org/10.52962/ipjaf.2017.1.3.15.

Kealhofer, S. (2003). Quantifying credit risk I: Default prediction. *Financial Analysts Journal*, *59*(1), 30–44. https://doi.org/10.2469/faj.v59.n1.2501.

Leland, H. E., Anderson, F. R., & Hennessy, C. (2004). *Predictions of Default Probabilities in Structural Models of Debt The author thanks Dirk Hackbarth for his comments and invaluable assistance with the*. 1–31.

Liang, Q. (2003). Corporate Financial Distress Diagnosis in China/ : Empirical Analysis using Credit Scoring Models Author (s): Qi Liang Source/ : Hitotsubashi Journal of Commerce and Management , Vol . 38 , No . 1 (38) (October Stable URL/ : http://www.jstor.org/stable/. *Hitotsubashi Journal of Commerce and Management*, *38*(1), 13–28.

Marvadi, D. C. R. (2018). Assessing Corporate Financial Distress in Automobile Industry of India: An Application of Altman's Model. *Research Journal of Finance and Accounting*, *2*(3), 155–168. http://iiste.org/Journals/index.php/RJFA/article/view/335.

Mvula Chijoriga, M. (2011). Application of multiple discriminant analysis (MDA) as a credit scoring and risk assessment model. *International Journal of Emerging Markets*, *6*(2), 132–147. https://doi.org/10.1108/17468801111119498.

Pal, S. (2013). A Study on Financial Distress in Indian Steel Industry under Globalization. *IOSR Journal of Business and Management*, *14*(2), 49–53. https://doi.org/10.9790/487x-1424953.

Singla, R., & Singh, G. (2017). Assessing the Probability of Failure by Using Altman's Model and Exploring its Relationship with Company Size: An Evidence from Indian. *Journal of Technology Management for Growing Economies*, *8*(2), 167–180. https://doi.org/10.15415/jtmge.2017.82003.

Smaranda, C. (2014). Scoring Functions and Bankruptcy Prediction Models – Case Study for Romanian Companies. *Procedia Economics and Finance*, *10*(14), 217–226. https://doi.org/10.1016/s2212-5671(14)00296-2.

Tanthanongsakkun, S., Pitt, D., & Treepongkaruna, S. (2010). A Comparison of Corporate Bankruptcy Models in Australia: The Merton vs. Accounting-based Models. *Asia-Pacific Journal of Risk and Insurance*, *3*(2), 1–33. https://doi.org/10.2202/2153-3792.1042.

Tudela, M., & Young, G. (2005). *A MERTON-MODEL APPROACH TO ASSESSING*. *8*(6), 737–761.

Vadyak, A. (2017). " Altman ' s Z – Score Analysis of Chhattisgarh ' s Steel Industry ". *International Journal of Research in Management, Economics and Commerce*, *07*(2), 10–18.

Wang, W. T., & Zhou, X. (2011). Could traditional financial indicators predict the default of small and medium-sized enterprises? -Evidence from Chinese Small and Medium-sized enterprises. *Economics and Finance Research*, *4*, 72–76.

Zeitun, R., Tian, G., & Keen, S. (2007). Default Probability for the Jordanian Companies/ : A Test of Cash Flow Theory. *International Research Journal of Finance and Economics*, *8*(8), 147–161. http://ro.uow.edu.au/cgi/viewcontent. cgi? article = 1365 & context = commpapers.