ISSN:2229- 6247
Deepika Verma | International Journal of Business Management and Economic Research(IJBMER), Vol 15(1),2024, 2327- 2338

Predicting default of chemical sector using MDA, Altman, Calibrated, Logit, and Structural Model

Dr. Deepika Verma
Assistant Professor
Sri Aurobindo College Eve
Delhi University
deepika.delhiuniversity@gmail.com

Abstract
Big listed companies like Andhra Cement, Gayatari bioorganic and madras fertilizers are not escapes from debt trap. Therefore, an attempt has been made in the present study to predict the default occurrence of selected chemical 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 textile sector. The developed models have also been validated on the out-of-sample data. The study obtained satisfactory statistical results pertaining to the MDA and Logit developed models but not from Structural Model and Altman Z score model. Additionally, the classification results witnessed the following accuracies for MDA, Calibrated, Altman, Logit and Structural model such as 90%, 87%, 16%, 93% and 25%. The validation accuracies obtained by mda, calibrated and logit models are 16%, 78% and 82%.
Keywords: Financial distress, default prediction, credit risk modeling, BSM model, logit model, mda model, structural model

INTRODUCTION

Credit risk modelling is crucial among financial institutions, banks and regulators. Because mounting incidences of defaults and growing NPA have severely impacted on the banking sector of developed and developing economies like India (World bank, 2010). Many banks either have become bankrupted or in chronic financial distress due to the expanding credit risk (Eken, et al., 2012).

Study of credit risk management has received serious attention during last decade across the world due to the bankruptcy new of big USA firms namely Enron, Worldcome and the collapse of lehman brother in 2008 that erupted global crises. S&P reported 496 USA listed firms that defaulted in the on the outstanding debt of 1 trillion dollar which has dwarfed the earlier reported defaults in USA.

Firms fund their capital intensive project by raising external finance from financial institutions, banks and by issuing bonds. As in India the bonds are less popular method of financing, the firms target public sector banks for seeking hefty amount of loan. Hence, banks have to predict the financial position of the firms before sanctioning loan. For the same the banks evaluate the financial statement of the firms and also assess the qualitative measures that impact the profitability and liquidity position of the firms. Predicting the loss given default of firms are also required for bank for making provisioning in their accounts by keeping the minimum required safety capital.

Banks use credit risk model for developing the default prediction to forewarn themselves and the other stakeholders of the firms. It is a task of the internal credit risk manager of the banks to make such an accurate default prediction model that can predict the time, frequency and likelihood of default accurately in advance. There are various risks involved in the business operations like business risk, market risk, foreign currency risk, and credit risk. Amongst all risks the most crucial risk is credit risk that can take the firm towards insolvency. Credit default arises when firm fails to oblige its debts on due date.

Fitzpatrick (1932), initiated the process of default prediction by evaluating firms specific financial ratios for determining the firm's financial status which is latter followed by Smith and Winakor (1935), Merwin (1942); Chudson (1945). Smith and Winakor (1935) attempted to classify 183 distressed firms across the various industries and found that wc/ta is a better predictor than cash/ta and ca/cl. Whereas, Jackendoff (1962) supported ca/cl and wc/ta in the classification of profitable and unprofitable USA firms. Lately few scholars namely Merton (1974), Black and Scholes (1973) included macroeconomic variable in the default prediction study such as GDP, interest rate and stock index return. later some attempts have been made to predict default based on macroeconomic variables. In this regard, some macro-economic indicators have been used in the past literature, including: interest rate, stock index return and GDP.

Altman (1968) applied MDA first time to classify the defaulted and non-defualted U.S firms which has become a benchmark that is still relevant and being used by scholars, bankers and to predict firm default. Subsequent in 80's Ohlson (1980) introduced Logit function in study of default prediction using firm-specific independent variables. Later on Zmijewski (1984) applied probit model for determining the determinant of defaults in U.S firms. Shumway (2001) introduced multivariate statistical function for estimating corporate default. Few studies conducted by applying support vector machines (SVM) for predicting the default of US firms.

As it is witnessed from the above paragraph that majority of the studies have been conducted so far in the developed countries like USA. Hence, the present study attempted to classify the defaulted and non-defaulted sample of selected chemical sector. The current study is divided into four sections. First section deals with the introduction, literature review, objectives, and hypothesis of the study. Second section covers the research methodology which constitutes the sample data, source of data, variables and default prediction methods. The third section elaborates the empirical results that comprises of developed models, statistical finding, in-sample and out-of-sample classification results and the detailed analysis of the findings of the developed models. The fourth section describes the conclusion of the study in which the obtained findings have been compared to the early default prediction studies.

LITERATURE REVIEW

MDA Model

Beaver, W. (1966) predicted the distress using a Univariate model based on 30 financial ratios for the 79 pair of distressed and non-distressed firms. The study found that WC/TA ratio and NI/TA ratio are the best discriminators for the distress prediction (Beaver, 1966).

The study investigates the use of discriminant analysis for multi-level classification on large datasets. This study unveils that the discriminant analysis gives a fast, effective and accurate alternative for multi-level classification. The result achieved using LDA is comparable to SVM and less time consuming than the other approaches (Li et al., 2006).

An internal credit scoring system was developed to rate the external bonds and to assess the probabilities of default. Z-score model was applied along with financial ratios, which was later validated on Steel companies that provided 85-90 % accuracy. The study concluded that the model is accurate, simple, accessible, however not perfect since it has Type II Errors (Altman, 2006).

Suzzane Hayes (2010) aimed to develop Z score for public sector retail bankrupt companies for up to 2 succeeding years. Altman's Z score successfully predicted all companies' financial health except 2 but its accuracy is lower than Z". Study states that Z" score is effective for public non-manufacturing firms unlike Altman's Z (Hayes et al., 2010).

Abdul Rashid (2011) attempts to identify which financial ratios are important predictors of bankruptcy in the non-financial sector of Pakistan. The sample selected from the companies that went insolvent and delisted from the Karachi stock exchange during 1996-2006. The study used 24 financial ratios which reflect the following features of the firm such as profitability, liquidity, leverage, and turnover. The ratios were assessed for 5 years before insolvency. The result draws the inference that the ratios namely Sales/TA, EBIT/CL, and cash flow ratio were identified as better predictors of bankruptcy (Rashid, 2011).

An investigation was conducted to check whether the inclusion of risk assessment variables in the MDA model improved the bank's ability to classify customers and predict the firm's financial performance. The study was based on the recent financial calamity of 2009. The financial information was gathered for the period 1985-1994 of 100 customers from the National Bank of Commerce. The outcome of the model signifies that the MDA model has higher predictive and classification accuracy when the model integrates both qualitative and quantitative variables (Mvula Chijoriga, 2011).

Ohlson (1980) developed a credit risk model using statistical method called the Conditional Logit model that does not need to meet the assumptions required for MDA. The model attained 88% accuracy on the sample data of 105 listed firms (Zvaríková & Majerová, 2014). The Logit model was introduced by Martin (1977) that classified the distressed and non-distressed banks. Later on Andersen (2008) applied the Logit approach to determine the most appropriate predictors of Norwegian bank failure. The study incorporated 23 financial and non-financial variables out of which 6 variables are found to be the best fitting.

The assumptions of MDA such as the normal distribution of variables and equal variance and covariance matrices of defaulted and non-defaulted firms have been violated in many studies which paved a way for the Logit or O-score model (Ohlson, 1980). According to Thomas, Edelman, & Cook (2002) Logit is the most used statistical method in the field of prediction of default where the dependent variable is binary. The binary result of the conditional Logit model describes the default probability and provides a list of significant variables (Balcaen & Ooghe, 2006). Studies that used the Logit model are Kwofie, Ansah, & Boadi (2015), Bartual, Garcia, Guijarro, & Romero-Civera (2012), Bewick, Cheek, & Ball (2005), Bandyopadhyay (2007). Thereafter, the Multi-period Logit framework was brought up by Shumway (2001) which included time-varying variables for predicting failure. This model stood out against the single period Logit model.

Jones & Hensher (2004) & Train (2002) advocated the use of the Mixed Logit model to label the firms into non-failed, insolvent firms, and the firms filed for bankruptcy. This Model grouped the firms with high accuracy, result of the model shows that the mixed Logit model stood out in the prediction and classification of firms into appropriate categories. The Mixed Logit is the most recently developed technique. Wooldridge (2009) criticized the Logit model for over prediction of the bankruptcy risk.

Logit model

A corporate failure prediction study of 105 bankrupted and 2058 non-bankrupted firms was conducted using a conditional Logit model. The study developed 3 models where, the first model predict bankruptcy within one year, the second within 2 years, the third model predicted the bankruptcy within one or two years for the period 1970-76. The study denoted the size of the firm as the key predictor of financial distress. The findings of the study unearthed that financial factors surge the predictive power of the model. Further, the results of the study was validated by Memic & Rovcanin (2012) (Ohlson, 1980).

Lau (1987) develops a model which can predict the probability that a firm shall enter into every five financial states such as 0: financial stability; 1: Omitting and reducing dividend payments; 2: technical default and default on loan payment; 3: protection under chapter X or XI of the bankruptcy act; and 4: bankruptcy and liquidation. The result of the study exhibits that the Multinomial Logit model is robust to perform the estimation (Lau, 1987)

The default risk of Norwegian limited companies that belongs to the Agriculture, Construction, Industry and Service sector for the period 1995-1999 was estimated using Logistic regression by integrating financial ratios into the model. The findings inferred that model is static, helpful for a short time horizon only (Westgaard & Van der Wijst, 2001).

This study has reviewed various tests of logistic regression namely the hosmer lemeshow test, R square test, wald test which examines the goodness of fit, the utility of the model and measures the importance of individual coefficients. The model was applied to medical research to investigate that how death and survival of patients can be predicted by logistic regression which provides binary outcomes i.e 0 and 1 (Bewick et al., 2005).

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).

Structural Model

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 (Duffie et al., 2007).

This study investigates the performance of the indicators generated using the Merton Model to predict the bankruptcy of corporate in Australia for the period 1990-2003 by applying a multiperiod Logit model. The sample data of the failed companies was obtained from www. Delisted.com. The study exhibits that the Merton model significantly predicts bankruptcy. The study revealed that the TL/TA ratio and idiosyncratic standard deviation of stock returns are remarkable indicators of the failure (Tanthanongsakkun et al., 2010).

Tarashev (2011) attempts to examine the performance of various Structuralcredit risk models. The study recommends that leverage ratio, default recovery rate and risk-free ROR impacts the prediction power of

model. The findings suggest that the appropriate model to predict the default is an endogenous model group that provides impartial prediction. This study also substantiate that the Structural Model unveils material information about the time pattern of default rates (Tarashev, 2011).
Ahmad & Wahab (2012) documented some of the distinguished attributes and assumptions of Moody's Structural Model such as the default triggers when a firm's asset drops below the threshold limit or when the firm's net worth reached zero even before the maturity of the debt. The default position of the firm also get impacted by the market variables like market value asset of the firm; its equity, its market volatility etc. The study further exemplified the role of macroeconomic variables and their interdependence for instance recession brings more default occurrence than the boom. The 2008-09 financial crises is the classic example that ends up spreading the epidemic of bankruptcies which culminate into the growth of NPAs (Ahmad & Wahab, 2012).
The study aims to assess the performance of credit scoring and Merton based model for predicting insolvency of 246 UK SMEs from 2001 to 2004. The performance of the models was tested for 4 years using AUROC. The Merton model is used to calculate DD and EDF in the study. The credit scoring model performed better with the sample group by incorporating a sufficient number of bankrupt firms consequently, the Merton performed quite well with higher acceptance rates (Lin et al., 2012).
The study employed a hybrid model which is an amalgamation of option and accounting-based models. The sample data consists of financial information collected from Compustat annual file for the span from 1970 to 2006. This study witnessed that the option-based model performed better than the accounting-based model for discriminating companies. The hybrid model defeated both option based and the accounting-based model (Tsai et al., 2012).
The study inscribed that according to the theoretical framework of Structural Models default occurs when the market value of assets of the firm drops down a certain solvency boundary. Nevertheless, they would be wrong in prediction and classification. Hence the study evidence that the application of the empirical parameters is advantageous to boost the model's predictive competency remarkably. The sample comprises of bond issuers who defaulted from 1997 to 2005 (Davydenko, 2013).
The study used financial ratios with the Altman Z score model to predict the solvency of Indonesian chemical sector and to establish a causal relationship between ratios, financial health and the price of their shares. The conclusion illustrated that the stock prices are being significantly influenced by the EBIT to TA ratio (Lestari et al., 2016).
This study investigates the effects of the sensitivity variable, industry beta on the probability of default of a firm using Logit and MDA function to classify the firms. The results of the study unveiled that both the variables performed outstandingly for predicting the default. Additionally, findings addressed that the increasing sensitivity to industry factors triggers default. The study advises lenders and investors to keep a check on the sensitivity of a firm to such changes (Agrawal & Maheshwari, 2019).

OBJECTIVES

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

RESEARCH METHODOLOGY

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 Chemical firms collected. **In-Sample data** is used to develop the model and second part of the sample data called **out-of-sample** is used to validate the developed models.

Table No 1 Description of selected Chemical Sector

Sectors	Defaulted Firms	Non-Defaulted Firms
Chemicals	18	29

Data Sources

The study collected the company specific information such as the accounting, market and macroeconomic data of the selected Indian chemical firms from various sources. The accounting data was fetched from the individual financial statements of each selected chemical sector 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 chemical sector firms is sourced from the audited annual reports of all selected chemical sector firms for 15 years from 1 April 2004 to 31st March 2019.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. The information about the daily average price of shares, return on the shares, BSE index and return on BSE index of the selected chemical sector 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 chemicals sector 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.

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 selected chemicals sector firm 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 No 2 Description of Independent Variables 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 of Logit Model

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 No 3 Description of Independent Variables for Logit Model

Categorical Variables
X= 1, TL >TA and X= 0, TA>TL
Y=1, Avg NP for 2 years < 0 and Y=0, Avg NP for 2 years >0

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

Models developed using MDA

Developed MDA Model

Z = -1.515+0.648*WC/TA+0.115*CA/CL+0.555* NI/TA+1.199*NP/TE-1.647*TBD/TA+0.047*FAT

Source: developed by author using SPSS version 22

Models developed using Calibrated Model

Model developed using Calibrated Model

-0.943+2.717* WC/TA +3.022* RE/TA +0.043* EBIT/TA + 0* MVE/TBD +4.81*SALES/TA

Source: developed by author using SPSS version 22

Models Developed Using Altman

Model Developed Using Altman

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

Source: developed by author using Altman (1968)

Description of Sample Data

Table No 4 Summary of Cases processed from Chemical Sector

Sector	In-sample	Out-of–sample
Chemicals		
• Total cases	495	223
• Cases considered	449	212
• Cases removed	46	11

Log Determinant

Table No 5 Log Determinant

Sector	Non-Defaulted	Defaulted	Pooled Within-groups
Chemicals	23.36	-15.485	23.89

Source: developed by author using SPSS version 22

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 5 Log determinant of selected chemical sector firm are neither equivalent nor close to defaulted, non-defaulted and pooled within group. Table No 5 Log determinant also enumerates that for selected chemical sector the log determinant values for the non-defaulted groups and pooled within-groups are closer to each other yet, this is quite distant from the defaulted groups due to the existence of higher Type II Error in the prediction results of selected chemical sector firm.

Coefficients of MDA Model

Table No 6 Coefficient of MDA Model

Particular	Box's M	Sig. Value of Box M	Eigenvalue	Canonical Correlation	Wilks' Lambda	Sig value of Wilk's lambda
Chemicals	1945.98	0	0.31	0.486	0.763	0

Source: developed by author using SPSS version 22

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.
The significant P-value of the box's M test of selected chemical sector as depicted in Table No 6 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 selected chemical sector is higher as displayed in Table No 6 Coefficients 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) however, it is contrary to Suleiman (2014) and Memic (2015). Nonetheless, the developed model was found robust even if it violates the box' M test condition because of the large number of sample cases considered in the study that makes the Box's M test less relevant for the default prediction.

Eigen Value
The eigenvalue denotes the variation in the dependent variable that can be explained by the MDA model. Primarily the Eigenvalue is a ratio between explained and unexplained variance. The higher eigenvalue recommends the greater discriminatory power of MDA function that explains the variation in the dependent variable. The strong discriminant function has a higher eigenvalue i.e. close to 1. The present study found lower eigenvalue for the selected chemical sector i.e. .31 as depicted in Table No 6 Coefficients that conveys the lower prediction power of the developed models. This indicates that the variation in the dependent variable is explained by developed model by 31% 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 selected chemical sector is .48 which is less than .50. This lower canonical correlation values substantiate the average 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 Coefficients exhibits the value of the wilk's lambda for selected chemicals sector firm. Table demonstrated that the wilk's lambda value i.e. .76 this is not appreciable finding for discriminating the defaulted and non-defaulted cases of the selected chemical sector firm.

Hypothesis 2
H_0: There is no discriminating power in the independent variables of the developed models.
Since the sig value of Wilk's lambda for selected chemical sector is <.05, this 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 selected chemical sector are consistent with Altman (2000), Altman (1968) and Memic (2015).

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

Table No 7 In-Sample Classification Result of Chemical Sector

Sectors	Models	Accuracy Rate	Type I Error	Type II Error
Chemicals	Developed model	90%	3%	69%
	Calibrated Model	87%	2%	88%
	Altman's Original	16%	94%	0%

Source: developed by author using SPSS version 22 and Altman (1968)

Findings and Discussion

The developed model outperformed with 90% accuracy rate in Chemical sector followed by calibrated that achieved 87% predictive accuracy. The Altman model classification results did not show impressive classification accuracy for selected chemical sector. The sector witnessed 94% Type I Error in the Altman model whereas, the calibrated and developed model had displayed only 2% and 3% Type I Error. The highest Type II Error of 88% is found in the calibrated model and 0% Type II Error is obtained by Altman model for selected Chemicals sector. The developed model performed well concerning the higher prediction accuracy and least type ii error for the selected chemical sectors.

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

Table No 8 Validation Results

Sectors	Models	Accuracy Rate	Type I Error	Type II Error
Chemicals	Developed model	16%	91%	11%
	Calibrated Model	78%	22%	0%

Source: developed by author using MS-Excel

Findings & Discussion

Table No 8 Validation Results summarises the results achieved by the study for validating the developed and calibrated model on the out-of-sample data of the selected chemicals sector firms. The study attained higher accuracies of 78% in calibrated mode whereas the developed model could only achieve the accuracy of 16% for the selected sector. Calibrated model performed surprisingly well w.r.t predictive accuracies and errors. The calibrated model outperformed the developed model by obtaining only 0% type ii error and only 22% type i error whereas, the developed model acquired 91% and 11% type I and type ii errors respectively.

Models Developed using Logit Function of Chemical Sector

$$L = -0.77 - 0.214 * FAT - 0.282 * LOG(TA/GNP) + 2.582 * Y$$

Source: developed by author using SPSS version 22

Description of Sample Data

Table No 9 Summary of Cases Processed

Sectors	In-sample	Out-of–sample
Chemicals		
• Total cases	493	223
• Cases considered	447	177
• Cases removed	46	46

Coefficients of Logit Model

Table No 10 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
Chemicals	124.874	0	167.278	0.243	0.508	9.04	0.339

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 10 Coefficients of Logit Model the chi-square value of the selected chemical sector is 124.874 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 the dependent variables.

Since the sig- values for selected chemical sector in the Omnibus test given in the table is 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 value of -2 log-likelihood is attained for selected chemical sector is 167.278 which is sufficiently high for this sample size. This indicated the greater classification ability of the developed model for the chemical 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. Study obtained only 0.243 cox and snell r square value, this signifies that the developed model explained the variations in the dependent by only 24%.

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 0.508 Negelker R square value for selected chemical sector which below average. This result demonstrated that the variation in the dependent variable of the model developed for selected chemical sector is explained by the independent variables of the developed model by only 51%.

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.

Table No 10 coefficients of logit model presents that the sig-value of hosmer lemeshow test of model developed for selected chemical 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).

In-sample classification result of the Logit model

Table No 11 In-Sample Classification Result

Sectors	Accuracy Rate	Type I Error	Type II Error
Chemicals	93%	1%	60%

Source: developed by author using SPSS version 22

Findings and Discussion

Table No 11 In-sample Classification Results presents the empirical results of Logit Models developed for selected Indian chemical sector. The result consists of Accuracy rate, Type i and Type ii error. The large value of the accuracy rate with minimum Type i and Type ii error suggests the higher predictive competency of the developed model. The accuracy rate of the Logit model is at the higher side for Chemical sector i.e. 93% in conjunction with 1% Type i Error. However, the developed logit model is plagued with higher type ii error i.e. 60%

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

Table No 11 Validation Results

Sectors	Accuracy Rate	Type I Error	Type II Error
Chemicals	82%	16%	38%

Table No 11 validation results displays the validation result of the out-of-sample data that is employed in the study to check the validity of the developed Logit model. The accuracy rate and errors explain the robustness of the model and its validity to apply the model to the varied generic sample data. The validation result shows accuracy rate i.e. 82% for the selected chemical sector which is satisfactorily high. Nonetheless, the model depicted type ii error and type I error.

Analysis of Empirical Results
Multiple Discriminant Analysis
The model was developed using 6 independent variables comprised of financial ratios only such as WC / TA, CA / CL, NI / TA, NP / TE, TBD / TA, Fixed Asset Turnover. The study used 449 and 212 cases for the in-sample and out-of-sample classification results. The log determinant of the non-defaulted and pooled within-groups are relatively closer to each other which specify that the model contains more non-defaulted cases than the defaulted cases; this culminates in the higher level of Type II Error. Table No 6 Coefficients depict the eigenvalue and the canonical correlation value at the lower side whereas, the Wilk's Lambda value at the higher side. This suggests that the model is not robust enough to classify the groups correctly. The empirical results demonstrated that the calibrated model classified the groups with the highest accuracy level in conjunction with less Type i error and exorbitant Type ii error. The performance of the developed model outpaced the Altman's original model which exhibited substandard results. The validation results included the comparison of developed and calibrated model only. Table no validation results presented that the developed model failed to provide satisfactory prediction results. Nonetheless, the calibrated model can be considered for the default prediction.

Logit Model
This model encapsulates the accounting, economic and qualitative variables namely FAT, LOG (TA/GNP) and Y. The model was built and tested upon 447 and 177 observations respectively. The values of the coefficients namely Omnibus tests with sig value, -2 Log-likelihood, Cox & Snell R square, Nagelkar R square and Hosmer and Lemeshow test are 124.874; P-value 0, 167.278, 0.243, 0.508 & 9.04; P-value .339 respectively. These findings convey that the variables incorporated into the model are quite significant to predict the credit risk, and model is robust. Nonetheless, the model cannot appropriately explain the variations in the dependent variable. Study found that the independent variables of the model can explain the variations in the dependent variable with 50% accuracy. The small value of the Hosmer Lemeshow test explains the high level of goodness of fit of the include data and indicates that the model is specified to predict the default risk. The model classified the groups correctly with 93% accuracy in conjunction with value of fewer Type I and Type II Errors as highlighted in Table no In-sample classification results. Table No 11 Validation Results outlines the 82% accuracy of the developed model coupled with lower Type I and Type II Errors.

Structural Model

Table No 12 Cases Processed

Sectors	Cases Processed
Chemicals	563

Classification Result of Chemicals Sector by Structural Model

Table No 13 Classification Results of Structural Model

	NON-DEFAULTED	DEFAULTED	Total
NON-DEFAULTED	92	418	510
DEFAULTED	4	49	53
Accuracy Rate		25%	
Type I Error	82%		
Type II Error	8%		

Source: Classification Results derived using MS-Excel

Analysis of the structural model results
The found results depicted in Table No 13 Classification Result of structural model for chemical 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 53 defaulted cases structural model correctly classified 49 cases that amounts to 94% accuracy. Nonetheless, for overall accuracy the structural model provided only 25% accuracy due to higher level of Type I error. The Type I Error is the most troublesome error found in the Structural Model; due to the high percentage of Type I Error the classification accuracy of the model becomes smaller. The higher value of Type I Error also signifies that the structural model is most compatible to classify the defaulted cases than non-defaulted cases.

CONCLUSION AND DISCUSSION

The classification results witnessed the following accuracies for MDA, Calibrated, Altman, Logit and Structural model such as 90%, 87%, 16%, 93% and 25%. The obtained accuracy rate of MDA is which is commensurate with the accuracy levels achieved by Jayadev (2006), Slefendorfas (2016), Jaffari & Ghafoor (2017), Abid, Masmoudi, & Ghorbel (2016), Altman E. I. (2006), Memic (2015), Liang Q. (2003), Hassan, Zainuddin, & Nordin (2018), Bartual, Garcia, Guijarro, & Romero-Civera (2012), Thai, Goh, HengTeh, Wong, & ong (2014). However, the obtained accuracy rates of the developed MDA model are less than the level of accuracy acquired by Pongsatat et al. (2004), Pang & Kogel (2013), Salehi & Abedini (2009), Desai & Joshi (2015), Chijoriga (2011), Kumar &Rao (2014).

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).
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).

The validation accuracies obtained by mda, calibrated and logit models are 16%, 78% and 82%. Altman's original model depicted a lower accuracy level for selected chemicals sector this substantiates the irrelevance of the Altman (1968) Original model. The calibrated model performed considerably well for both in-sample and out-of-sample data. The calibrated model outperformed the developed MDA model concerning the accuracy rate for selected sectors in out-of-sample data this indicates that the independent variables used by Altman (1968) are still relevant.

The developed, calibrated and Altman's original models have experienced a considerable amount of misclassifications that are quantified as Type I and Type II Errors. Altman's original model has encountered maximum Type I Error for in-sample data. However, Altman original model experienced minimum Type II Error; this suggests that Altman's original model misclassifies the non-defaulted cases as defaulted. The Type I Error values are negligible in the developed model for the In-sample classification results. However, there are considerable rate of Type II Error found in selected Chemical sector that witnessed 69% Type II Error. The values of Type I Error for the out-of-sample data cases for sectors namely Chemicals is 71%. The Type II Errors values for the validation cases are quite low for all the selected sectors except Complete Sample.

The classification accuracy of the developed Logit model for in-sample data for the selected chemical sector is 93%; this is significantly high in comparison to the developed MDA, calibrated and Altman's original model. There is no acute misclassification problem with the Logit model specifically rate of Type I Error attained by developed models for selected chemical sector is at minimum level. However, the study witnessed substantially high rate of Type II Error i.e. 60%. The validation results of the Logit model are also remarkable i.e. 82%. The values of Type I Error are not troublesome for all selected sectors i.e. 16%. Nonetheless, the values depicted for the Type II Errors are quite high i.e. 38%.

The classification results of the Structural Model witnessed undesirable results since the overall accuracy attained by the Structural Model is at the lower side in contrast to the default events probabilities discussed in the immediately above point. The overall accuracy of all selected chemical sector is 25%, it signifies that the Structural Model is competent to predict defaulted cases accurately. However, it misclassifies the non-defaulted cases as defaulted due to which its overall accuracy drops. The overall predictive accuracies of the Structural Model attained for selected sectors and Complete Sample were 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

Agrawal, K., & Maheshwari, Y. (2019). Efficacy of industry factors for corporate default prediction. *IIMB Management Review*, *31*(1), 71–77. https://doi.org/10.1016/j.iimb.2018.08.007

Ahmad, N., & Wahab, N. A. (2012). Estimates of Distance-To-Default (Dtd) Risk Indicator for Selected Sukuk in Malaysia. *3rd INTERNATIONAL CONFERENCE ON BUSINESS AND ECONOMIC RESEARCH (3rd ICBER 2012) PROCEEDING*, *March*, 2634–2649.

Altman, E. I. (2006). Estimating Default Probabilities of Corporate Bonds over Various Investment Horizons. *CFA Institute Conference Proceedings Quarterly*, *23*(1), 65–71. https://doi.org/10.2469/cp.v23.n1.3549

Balcaen, S., & Ooghe, H. (2006). 35 years of studies on business failure: An overview of the classic statistical methodologies and their related problems. *British Accounting Review*, *38*(1), 63–93. https://doi.org/10.1016/j.bar.2005.09.001

Bandyopadhyay, A. (2005). *Mapping Corporate Drift towards Default : A Study of Distance to Default of Indian Corporates Mapping Corporate Drift towards Default : A Study of Distance to Default of Indian Corporates*. *87*(3203), 0–24.

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

Beaver, W. H. (1966). of Failure Financial Ratios as Predictors. *Journal of Accounting Research*, *4*(1966), 71–111.

Bewick, V., Cheek, L., & Ball, J. (2005). Statistics review 14: Logistic regression. *Critical Care*, *9*(1), 112–118. https://doi.org/10.1186/cc3045

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

Davydenko, S. A. (2013). Insolvency , Illiquidity , and the Risk of Default. *Working Paper*, *February*, 1–44.

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

Frade, J. (2008). Credit Risk Modeling: Default Probabilities. *Florida State University*. https://stat.fsu.edu/~jfrade/HOMEWORKS/STA5168/FRADE_STA5168_paper.pdf

Hayes, S. K., Hodge, K. A., & Hughes, L. W. (2010). A Study of the Efficacy of Altman's Z To Predict Bankruptcy of Specialty Retail Firms Doing Business in Contemporary Times. *Economics & Business Journal*, *3*(1), 120–134.

Lau, A. H.-L. (1987). A Five-State Financial Distress Prediction Model. *Journal of Accounting Research*, *25*(1), 127. https://doi.org/10.2307/2491262

LELAND, H. E. (1994). Corporate Debt Value, Bond Covenants, and Optimal Capital Structure. *The Journal of Finance*, *49*(4), 1213–1252. https://doi.org/10.1111/j.1540-6261.1994.tb02452.x

Lestari, S. D., Oktaviani, R. F., & Arafah, W. (2016). Financial Distress Prediction With Altman Z-Score And Effect On Stock Price: Empirical Study On Companies Subsectors Chemical Listed In Indonesia Stock Exchange Period. *International Journal of Business and Management Invention ISSN*, *5*(8), 30–39.

Li, T., Zhu, S., & Ogihara, M. (2006). Using discriminant analysis for multi-class classification: An experimental investigation. *Knowledge and Information Systems*, *10*(4), 453–472. https://doi.org/10.1007/s10115-006-0013-y

Lieu, P. T., Lin, C. W., & Yu, H. F. (2008). Financial early-warning models on cross-holding groups. *Industrial Management and Data Systems*, *108*(8), 1060–1080. https://doi.org/10.1108/02635570810904613

Lin, S. M., Ansell, J., & Andreeva, G. (2012). Predicting default of a small business using different definitions of financial distress. *Journal of the Operational Research Society*, *63*(4), 539–548. https://doi.org/10.1057/jors.2011.65

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

Ohlson, J. A. (1980). Ratios financieros y la predicción probabilística de la bancarrota. *Journal of Accounting Research*, *18*(1), 109. https://www.jstor.org/stable/10.2307/2490395?origin=crossref

Rashid, A. (2011). Predicting Bankruptcy in Pakistan. *Theoretical and Applied Economics*, *18*(9), 103–128.

Sharma, D. C. ., Singh, D. R. K., & Upadhyay, R. K. (2014). Predicting Probability of Loan Default. *Primax International Journal of Commerce and Management Reserach*, *II*(3), 5–13.

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

Tarashev, N. A. (2011). An Empirical Evaluation of Structural Credit Risk Models. *SSRN Electronic Journal*, *179*. https://doi.org/10.2139/ssrn.846304

Tsai, L. K., Tserng, H. P., Liao, H. H., Chen, P. C., & Wang, W. P. (2012). Integration of accounting-based and option-based models to predict construction contractor default. *Journal of Marine Science and Technology (Taiwan)*, *20*(5), 479–484. https://doi.org/10.6119/JMST-010-1112-1

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

Westgaard, S., & Van der Wijst, N. (2001). Default probabilities in a corporate bank portfolio: A logistic model approach. *European Journal of Operational Research*, *135*(2), 338–349. https://doi.org/10.1016/S0377-2217(01)00045-5

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

Zvaríková, K., & Majerová, J. (2014). Financial Literacy in the Slovak Republic. *Procedia - Social and Behavioral Sciences*, *110*, 1106–1115. https://doi.org/10.1016/j.sbspro.2013.12.957