Do borrowers' financial statement information still holds relevance to manage credit risk in banks?

Dr. Lalit Arora (Corresponding author)
Faculty - State Bank Academy
State Bank of India
Plot No. 77, Sector - 18, Gurugram - 122015 (Haryana) - INDIA
Email: lalitvijayarora@gmail.com

Manish Yadav
Chief Manager & Research Officer - State Bank Academy
State Bank of India
Plot No. 77, Sector - 18, Gurugram - 122015 (Haryana) - INDIA
Email: mailtomanish@gmail.com

ABSTRACT
Managing credit risk in banks has been challenging for banking regulators and policy makers. This study encompasses different aspects of borrowing firms - liquidity, repayment capacity, solvency and profitability in predicting default by borrowers. Using data of non-performing and the standard (good) accounts from major Indian commercial banks from April 2013 to March 2020, we apply logistic regression, decision tree and random forest algorithms to Altman's Z-Score and other financial variables chosen from banks' credit analysis structure to predict borrower default. We report substantial ability of Z-Score and cash flow from operations to predict delinquency. The results suggest that solvency and liquidity interaction, especially in manufacturing firms is significant and thereby provides an appropriate approach to develop credit risk models. We propose that inclusion of Z-Score in the credit underwriting framework of banks could possibly improve their ability to predict these defaults. In the process, we aim to strengthen the credit risk assessment framework of banks.
Key words: Altman's Z score, Credit risk, Default, NPA, Logistic regression.

INTRODUCTION
After numerous painful experiences leading to huge amounts of nonperforming assets (NPAs), banks are now convinced of the lending mistakes that they have made in the past in assessing borrowers' repayment capacity. These mistakes also includes banks' over-optimism regarding borrowers' future prospects, coupled with their strong balance sheets. Moreover, increasing competition among banks has resulted in more liberal credit policies and lower credit standards. There are explanations to rationalize such practices which include agency problems - rewarding managers more in terms of growth instead of profitability (Jiménez & Saurina, 2006) or increasing social presence of banks (Williamson, 1963). However, experiences of banking supervisors and the theoretical developments in the field have dramatically evolved the way banks assess credit risk. This includes new early warning systems, measures of credit concentration risk, more sophisticated risk management models like Risk Adjusted Return on Capital (RAROC) and accurate assessment of risk arising due to off-balance sheet items (Altman &

Saunders, 1997). Traditional methodology to minimize credit risk involved a loan-by-loan analysis (Suresh & Paul, 2014). However, the advent of credit scoring models in the early 1960s laid the foundation for using analytical frameworks in taking credit related decisions. In fact, in the last two decades, there have been important advances in both practice and policy in terms of credit risk modeling. For example, models like CreditRisk+ (sponsored by Credit Suisse) and CreditMetrics (sponsored by JP Morgan) have gained popularity, ever since their release in 1997 (Gordy, 2000). Not only have these advancements benefitted senior bank officials, amendments in the banking regulatory framework have also been made due to these models. Emergence of RAROC as a performance measure in the 1980s enabled banks to calibrate credit scores to the probability of default (PD), enabling them thereby to estimate expected losses. Seminal studies by Beaver (1966), Altman (1968), and Ohlson (1980) initiated discussions on having robust credit rating models decades ago. As a matter of fact, their models have been applied on different markets, industries and time periods [Taffler (1982, 1984); Pantalone and Platt (1987), Betts and Belhoul (1987) and Piesse and Wood (1992)]. Pai et al. (2006) while predicting industrial sickness proposed the Principle Component Analysis (PCA) - Multiple Discriminant Model (MDA) model to be a better predictive technique than neural network techniques.

Undoubtedly, financial vulnerability of borrowers and its predictability has received considerable attention in scholarly literature. Particular attention has been given to techniques, which involve examination of borrowers' financial statements. However, despite the focus, predictive models seem to have been inadequate; a robust measure of assessing borrowers' default risk still remains thereby in the black-box. Earlier approaches of credit risk analysis were based either on: expert's subjective analysis, limiting exposures in certain areas, and migration analysis (using transitions of homogeneous loans). Essentially, bankers today, rely on the 5C framework of credit, which includes borrowers' repayment capacity (earnings volatility), capital (net worth), collateral (additional security), credit (historical and proposed), and the character (track record). Even other financial institutions rely on banks' so-called expert systems of assessing credit risk in corporate loans, thus making these parameters as essential benchmarks of assessing the borrowers' credit worthiness. Though credit decisions take into account several qualitative factors, financial health of borrowers does significantly influence such decisions. This study examines the efficacy of majorly used ratios in banks in assessing the financial health of corporate borrowers. It also incorporates Altman's Z-Score (henceforth Z-Score) as an important predictor of default by borrowers. Since a precise measure of assessing default probability is central to the credit underwriting process in banks, an appropriate model to assess credit risk would be of significant value to lenders.

Risk management practices, for their effective implementation, require assessment of probability of default by borrowers. Considering a dataset of standard accounts, and those that became delinquent over a six-year horizon (i.e. 2013 to 2019), we examine the tendency of banks targeting 'bad' accounts due to their credit decisions based on misleading numbers. Specifically, we compare the efficacy of credit risk assessment of borrowers by using financial statement information in contrast to that of a simple measure of default predictability in the form of Z-Score. An indicator variable (equal to 1) that represents a default, for us serves as the dependent variable. Independent variables on the other hand, include banks' benchmarks of borrowers' financial health - profitability, measured by earnings before interest, depreciation, amortization and taxes (EBIDTA margin); solvency, measured as a ratio of total outside liabilities to adjusted tangible net worth (TOL/ Adj. TNW); debt serviceability, in the form of interest coverage ratio (ICR) measured as EBIDTA/ Outstanding interest for the year; and liquidity, assessed by the level of cash flow from operations (CFO). Using these variables, we compare three modeling techniques - logistic regression, decision tree, and random forest algorithms. By and large, the results indicate that Z-Score outperforms other financial health indicators in predicting a default by 'corporate borrowers'. The study highlights the importance of incorporating Z-Score in banks' formal credit underwriting structure, coupled with their foresight, based on the current examination of ratios.

The remainder of the study is as follows: section two presents linkage with literature on credit risk. The methodology followed for this study is discussed in section three; while section four comprise results and related discussions. The study concludes in section five, encompassing implications and limitations.

FINANCIAL DISTRESS MODELS IN CREDIT RISK MANAGEMENT

Beaver (1966) was the first to recognize the characteristics of failing firms in comparison to a paired sample of healthy ones. The study found that financial ratios were significant predictors of a firm's failure, way before it actually happens. Recent studies include Gerantonis et al. (2009) who reported Z-Score to perform well in predicting firm failures. Categorizing variables as loan characteristics, borrowers' credit history, and personal details, Liu et al. (2018) reported significant ability of these characteristics towards default prediction. Altman et al. (2020) tested the default prediction model for SMEs to predict mini-bond issuers in Italy, and reported the model to be effective for firms, which opted out from bank borrowing.

Though the utility of Z-Score in testing financial distress in businesses in unarguable, it has had its share of criticisms too. Deakin (1977) argued that financial ratios do not follow the normal distribution, even after performing data transformation. In sensitivity evaluation of financial distress prediction models, Hamer (1983) examined the models of Altman (1968); Deakin (1972); Blum (1974); and Ohlson (1980) and highlighted statistically different covariance in each model.

Generalized linear models or multiple logistic regression models have also been popular. Ohlsons O-Score (1980) for instance, has been based on generalized linear models with a logit link function, also referred to as logit analysis. Neural network models are powerful and popular alternatives, with the ability to incorporate a very large number of features in an adaptive nonlinear model (Wilson and Sharda, 1994). In India, prediction models have been developed by Kaveri (1980), Srivastava (1986), and Yadav (1986). In fact, Yadav (1986) developed discriminant model by using financial ratios, which covers the financial characteristics of a firm. Regardless of the advantages or the disadvantages of a predictive model, the very idea of developing such models to predict financial distress and failure itself is welcome all-over, for a model could help in detecting the likelihood of forthcoming sickness, thereby facilitating banks to predict borrower defaults. Thus, bankruptcy models can be used as early warning signals, such that, corrective action may be undertaken immediately by the management.

Beaver (1966) found that the cash flow to debt ratio was the best single ratio predictor of distress in his univariate discriminant analysis. Altman's Z-Score model used multivariate discriminant analysis to select the five most significant variables for measuring the financial distress of firms. Ohlson's O-Score model used a logit analysis to generate a one-year prediction model, and his academic descendants frequently referred to his discrete variables as a proxy for the probability of financial distress. Altman (1968) collected data of 33 failed firms and 33 matching firms, during the period 1946-1965, to find discriminating variables for bankruptcy prediction. Herein, he evaluated 22 potentially significant variables of these 66 firms by using multiple discriminant analysis to build a discriminant function with five variables.

Assessments of borrowers' financial health encompass multiple aspects including the Zeta model (Altman et al., 1977), industry relative accounting ratios (Platt & Platt, 1991) and first-to-default basket contracts (Nicolo & Pellizon, 2006). Notably, these models can be potential tools to examine defaults (probabilities). However, their efficacy depends upon the loan default database of significant size. Neural network analysis, essentially similar to non-linear discriminant analysis, is a newer approach to predict defaults. Expansion of off-balance sheet instruments (options, forwards, futures, swaps) has been a profound development in the area (refer Brewer & Koppenhaver, 1992; Saunders, 1994; Jagtiani et al., 1995). Kaminskaya & Ivanchenko (2011) reported banks to make limited use of statistical and scoring methods in assessing credit risk. They stressed on the use of inherent credit risks and individual credit risks in examining borrowers' credit worthiness. Missing or uncertain information about the borrowers coupled with lenders' inexperience, generally results in inaccurate credit rating scores (Moscato et al., 2021). Defining borrower's solvency ratios for maximum and minimum credit risk, Bobyl (2014) used a neuro-fuzzy scoring model. He contended the application of neural networks in scoring models of banks as effective tools - especially for statistical model creation with insufficient historical data or while using qualitative factors in the model, which could be estimated only by expertise. De Salceda Ruiz (1997) argued in favor of models which included leverage and liquidity as better predictors of financial stress. Bao et al. (2019) used the GARCH and Extreme Value Theory (EVT) models to evaluate banks' credit risk. They claimed that their model could improve the deviations caused by small samples. Crouhy et al. (2000) conducted a comparative analysis of credit risk models which included CreditMetrics (JP Morgan), KMV Model, CreditRisk+ (Credit Suisse Financial Products), and CreditPortfolioView (McKinsey). They also reported that there was no evidence of such models performing better than a simple Baysian model. In a major finding, Ericsson & Renault (2006) reported that levels of liquidity spreads are positively correlated with credit risk, especially in terms of decreasing functions of time to maturity.

Other empirical and theoretical studies have pointed towards the interactions between various firm-level variables and their ability to predict bankruptcy. Despite strong evidence of close link between financial ratios and firm financial stress predictability, a comprehensive score in the form of Z-score, vis-a-vis its interaction with other firm ratios has been less calibrated. We aim to tackle this issue by developing the borrowers' stress testing framework for banks, as discussed in the following sections.

DATA AND EMPIRICAL MODEL

Data set

The underlying data contained in the dataset is confidential with strict terms and conditions surrounding its usage to ensure the privacy of institutions involved in the study. The unit of analysis in our models comprises the borrowing firm's account. Since, some results would allow the identification of banks considered for the study, they have been prohibited to be published. Thus, our final dataset comprised 69 firms, which included 36 accounts that have defaulted, along with 33 standard (non-defaulting) accounts over a period 2013-14 to 2019-20. Notably, sample Indian firms that defaulted prior to the ongoing Covid-19 pandemic have been considered. Standard accounts have been considered as on March 31, 2020. Financial year closing values of the variables used for the purpose of computing Z-score and other ratios were considered. Data pertaining to three years prior to default year (t – 1, t – 2, t – 3) was considered for calculating Z-scores and other variables used in the model. The average value of the three year data prior to the default year was considered for running the logistic regression model. Notably, the same analogy has been considered for standard accounts. Descriptive statistics for the predicting variables (prior to standardization) are shown in table I. We standardized data by dividing the demeaned values by standard deviation. This was done keeping in view two important aspects: first, was to avoid the impact of outliers in the data set. Second, since the data included ratios as well as absolute values of various attributes, uniformity in the data set had to be ensured. Importantly, the descriptive statistics of the standardized data have not been tabulated.

Table I. Descriptive statistics for predictors

Default	Total Sample	CFO		ICR		TOL/ Adj. TNW		Z Score		EBIDTA Margin	
		M	SD	M	SD	M	SD	M	SD	M	SD
Yes	36	80.016	138.555	1.288	0.982	2.876	7.139	2.053	1.728	8.784	5.819
No	33	1099.209	1675.767	8.654	14.480	1.166	0.950	7.520	7.888	9.322	18.936
Summary	69	567.456	1262.693	4.811	10.626	2.058	5.234	4.667	6.195	9.041	13.647

Note: CFO values are in Rs. million. M and SD represent mean and standard deviation respectively.

Selection of attributes

Although studies in the past have used variables to benchmark our attributes (Hamer, 1983 and Gerantonis et al., 2009), we included variables that adequately represent industry standards. CFO, ICR, TOL/ Adj. TNW, Z-score and EBIDTA margin were the predictors. Moreover, variables were carefully chosen from banks' credit analysis structure, which examined different aspects of borrowing firms - liquidity, repayment capacity, solvency and profitability. Notably, joint modeling, using these dimensions of firm performance leads to an accurate representation of firm stress levels, vis-a-vis its outcomes (default). Equation (1) depicts the calculation of Z-score.

$$Z = 1.2X_1 + 1.4X_2 + 3.3X_3 + 0.6X_4 + 0.99X_5 \qquad (1)$$

Where, X_1 is the measure of short term health of firms, calculated as net working capital to total assets (NWC/TA); X_2 measures cumulative earnings (losses) of firms measured as retained earnings to total assets (RE/TA); X_3 depicts the firm's ability to generate profits, measured as earnings before interest and tax to total assets (EBIT/TA); X_4 is a measure of investor confidence that has been estimated as a ratio of market value of equity to book value of debt (MV_e/BV_d); and X_5 indicates the efficiency with which firms use their assets, measured as sales to total assets (Sales/TA).

Dependent variable

Our logistic regression model requires default/ no default as dependent variables. Thus, the dependent variable in our study represents the delinquency status. As per India's banking regulator's guidelines, delinquency has been defined as an account with past dues of 90 days or more. Importantly, this is a conservative definition of default, as the chances of accounts, which can recover post this period, are less.

Empirical model

We compared three basic models of classifying delinquency of borrowers - logistic regression model, decision tree, and random forest model. We used Z-score as a proxy to capture multiple aspects of firms' financial position and its ability to predict default. Thus, we closely evaluated the role of Z-score and the appropriateness of the variables by using decision tree and random forest algorithms. Considering the theoretical contemplations and empirical results of previous studies, we expected a direct effect of Z-score on the default probability of firms. Thus, we assumed that higher Z-scores would indicate lower default rates, in which case, the logistic regression model would take the following form:

$$logit\ (default) = \ln\left[\frac{\pi}{1-\pi}\right] = \alpha + \beta_1 CFO + \beta_2 ICR + \beta_3 \frac{TOL}{Adj.\ TNW} + \beta_4 Z\ Score + \beta_5 EBIDTA \qquad (2)$$

where, π denotes event (default) probability, α is the intercept and *β*s are the regression coefficients. We followed the maximum likelihood approach to estimate α and β values as against the weighted least square approach used by Haberman (1978) and Schlesselman (1982). Null hypothesis underlying the model was that all *β*s equal zero (β=0). Then, we used the open source software package Python to run the models, as it offers a wide range of algorithms. As a test of appropriateness of the selected variables, we encompassed decision tree analysis. Notably, decision trees act as powerful models, which partition the space *X*, with specific predictions of *y* (*y* being 0 or

1). The models partition the space into k mutually exclusive parts (R_1……..R_k). Thus, the output model of the decision tree takes the form:

$$f(\mathrm{x}) = \sum_{m=1}^{k} c_m I[\mathrm{x} \in R_m] \qquad (3)$$

where, $c_m \in \{0,1\}$ and I depicts an indicator function (Hastie et al., 2009). Since the typical partition happens through hierarchical tests in a series form, thus the name decision ‘tree’. Major benefit of decision tree models is their interpretability. Therefore, decision trees have been proved to be successful due to their out-of-sample classification performance (Butaru et al., 2016); however this too, has had its oppositions. Dietterich (2000) and Hastie et al. (2009) for instance, contended that decision tree models do not achieve significant results in out-of-sample classification. To further strengthen the results and determine the possible improvements, we used a state-of-art technique: random forest (refer Breiman & Cutler, 2004). Random forests have resulted in enormous success for out-of-sample learning algorithms (Criminisi et al., 2012).

MODEL RESULTS AND DISCUSSION

Our delinquency model aims to classify the firm level attributes as ‘good’ or ‘bad’ in predicting default. Thus, our performance measure must reflect the accuracy of our model to classify accounts in these categories. A common performance measure of models with binary classification is the precision and recall calculation.

		Model prediction	
		Default	Standard
Actual outcome	Default	TP (6)	FN (1)
	Standard	FP (3)	TN (8)

Note: TP - True positive; TN - True negative; FP - False positive; FN - False negative

For the purpose of understanding the model outcomes, we considered occurrence of default as a positive outcome, while no default was considered as a negative outcome. We defined precision as delinquent accounts that were correctly predicted, divided by the total number of predicted delinquent accounts (TP/TP+FP). Recall represented a measure of delinquent accounts predicted correctly to the actual number of delinquent accounts (TP/TP+FN). Total data set comprising 69 firms was segregated into train data (51 firms) and test data (18 firms). Notably, both datasets included defaulting and standard accounts. Then, the accuracy score of the model was calculated using test data as: [(6+8)/ (6+3+1+8)] = 77.78%. The threshold probability value (PV) of 0.60 was considered. Thus, all PV ≥ 0.60 were considered as default. However, precision or recall alone is insufficient to explain model accuracy. Thus, we considered F-measure, a statistic that combines precision and recall, the value of which was noted to be 0.752. Like precision and recall, values closer to 1.0 indicate perfect F-measure. We calculated F-measure as: F = [(2 x Recall x Precision)/ (Recall + Precision)]. As a performance measure relative to random classification, we used Kappa statistic calculated as K = [(p_o

– p_e)/ (1 – p_e)]. Kappa value of 0.55 was observed, which indicates moderate performance.

For calculating the statistics, we use a classification threshold that optimizes our model's accuracy to predict borrower defaults, using the attributes selected. Although, deciding a threshold based on test data introduces a slight look-ahead bias, this approach is suitable for two reasons. First, our performance statistics - F statistic and Kappa statistic are not sensitive to the threshold used while modeling. Second, banks prefer to adjust their classification models based on estimated delinquency rate to decide upon the acceptance threshold.

We analyzed the results starting with a simplified version of the regression specifications. The results of our logistic regression model are reported in table II. The observed signs of the coefficients are comparable to the expected behavior of the attributes, except for EBIDTA margin (which is an insignificant variable in default prediction). Consistent with the expected relationships, our results reveal that firms with higher Z-scores have lesser tendencies of defaults, and thereby emerge as a strong contender in predicting default (p-value of 0.011). Importantly, coefficient of -2.410 supports this assertion. In addition, the results also indicate a fair prediction power of an important liquidity parameter - CFO (coefficient -1.178 and p-value 0.058), thus indicating that defaulting firms generally face liquidity problems. These findings corroborate the results of related studies [Hamer, 1983 and Gerantonis et al. (2009)]. Results also suggest that the ability of attributes like ICR, TOL/ Adj. TNW and EBIDTA margin is rather limited in predicting defaults by the borrowing firms.

Table II. Regression specifications of logistic regression model

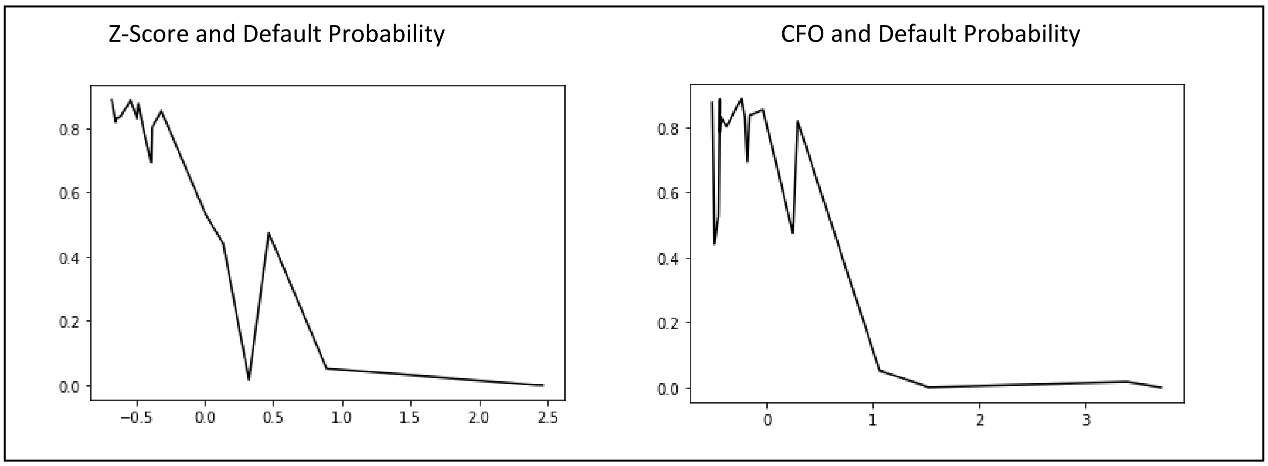
Predictor	Coefficient	SE	z-value	p-value	Expected sign of predictor coefficient
CFO	-1.178	0.621	-1.898	0.058**	-
ICR	-0.167	0.945	-0.177	0.860	-
TOL/ Adj. TNW	0.203	0.529	0.384	0.701	+
Z Score	-2.410	0.945	-2.549	0.011*	-
EBIDTA Margin	0.945	0.783	1.207	0.860	-
Pseudo R^2	0.371				

Note: SE denotes the standard error. ** Significant at 10% level, * Significant at 5% level.

We addressed a central issue, i.e. to authenticate the extent to which Z-Score model can be integrated in credit underwriting process, so that banks can save costs arising out of defaults. We posited that probability of default increases with downward movement in Z-Score and CFO. Figure 1 illustrates the relationship between Z-Scores and CFO levels with default probability of firms. While no model can unambiguously discriminate between defaulters and non-defaulters, the relevance of attributes can be established nonetheless. The subsequent discussion complements our assertion regarding the efficacy of Z-Score in predicting defaults.

As mentioned earlier, cut-off for considering default = 1 was 0.60. In order to understand the outcomes of considering different cut-off points, we used the receiver operating characteristic (ROC). The curve plots the proportions of actual defaults classified as defaults (sensitivity) against proportions of no-defaults, classified in turn, as defaults (1-specificity), at all values of cut-off point. Notably, AUC in the given case is 0.79 (refer figure 2). Higher AUC (nearing 1) indicates better model performance in differentiating both the positive and negative classes. For further illustrating the scope of differences in the results of logistic regression, decision tree and random forest algorithms, we examined the accuracy of the model variables in predicting defaults for all three techniques. Since, the decision trees are supervised machine learning algorithms, which split the data according to certain parameters, the results of decision tree and random forest algorithms offer intuitive information.

Figure 1: Relationship of Z-Score and CFO with firm default



Note: CFO values are the standardized values. Y-axis depicts the probability of default while Z-Score and CFO have been depicted on X-axis.

To further examine the predictive ability of the selected variables, we used supervised learning algorithms in the form of random forests. Simply put, it's a collection of multiple unrelated decision trees. Random forest model was estimated with 50 random decision trees. Notably, it is better to use random forest algorithms over other machine learning algorithms, as random forest algorithms fix well with errors in class populations (Zhu et al., 2019). Additionally, we used Gini Index (G) as a selection metric for splitting attributes in the decision tree. For a default, X_i possible levels as $L_1; L_2; \ldots\ldots; L_n$. Thus, G at internal node of the tree is calculated as:

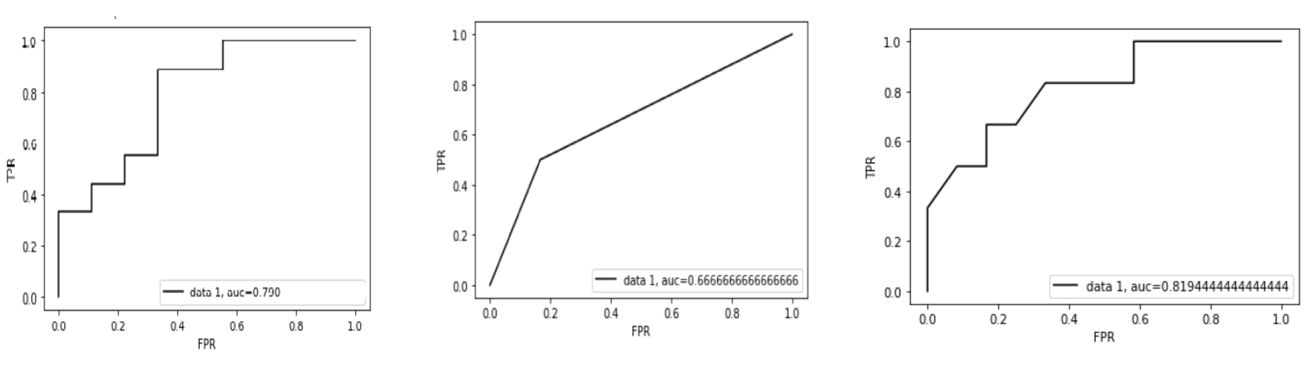
$$G(Xi) = \sum_{n=1}^{N} Pr(Xi = Ln)(1 - Pr(Xi = Ln)) \qquad (4)$$

$$= 1 - \sum_{n=1}^{N} Pr(Xi = Ln)^2 \qquad (5)$$

Table III. Evaluation metrics - techniques of comparison

Classifier	AUC	Accuracy	f1-score		Precision		Recall	
			0	1	0	1	0	1
Logistic Regression	0.79	79%	0.75	0.80	0.86	0.73	0.67	0.89
Decision Tree	0.67	67%	0.80	0.55	0.77	0.60	0.83	0.50
Random Forest	0.82	78%	0.85	0.60	0.79	0.75	0.92	0.50

Figure 2: ROC curves (logistic regression, decision tree and random forest)



Logistic regression | Decision tree | Random forest

Note: The ROC plots the True Positive Results (TPR) referred as sensitivity against the False Positive Results (FPR) referred as specificity at various threshold values and helps to effectively differentiate signals from noise. The area under the curve (AUC) determines the ability of a classifier to differentiate between classes.

We report the results of accuracy and AUC for different methods in table III. While the logistic regression and random forest were comparable in their performance, as they generated accuracy rates of 79% and 78% respectively, the same in case of decision tree was noted to be relatively less (67%). Clearly, logistic regression and random forest seemed to outperform the decision tree approach. Figure 2 depicts the ROC curves of the three methods. Proximity of the ROC curve to the top left corner of the graph indicates higher recall rate of the model. A point on the ROC curve closest to the top left corner indicates the best threshold with minimal classification errors with lowest number of false positives and false negatives. Graphs in figure 2 support the assertion that random forest and logistic regression outperform the decision tree. However, the results collectively indicate substantial predictive power of the variables used in the study.

CONCLUSION

We employed a dataset consisting of anonymous bank information from large public sector banks in India to comprehensively evaluate the ability of Altman's Z-Score to assess financial health, and thereby predict delinquency status of bank borrowers. We reported evidence that Z-Score significantly predicts loan defaults, and is thereby a superior metric in comparison to other financial parameters used by banks like - interest coverage ratio, liabilities to net worth, EBIDTA margin and even cash flow from operations. It would be beneficial to lenders to incorporate Z-Score in their formal credit underwriting structures. This is important, as loan defaulters bring losses that are often higher than profits, which banks earn from non-defaulters. We also noted that the ill effects of deterioration in cash flow from operations would be passed on towards making firms default. Thus, firm default was found to be consistent with cash flow reactions. Our study contributes beyond Hamer (1983) and Gerantonis et al. (2009) who only studied Z-Score reactions in predicting defaults.

Our results also posit that decision tree and random forest algorithms outperform the logistic regression model in predicting default. However, our study does have its own some limitations. First, our model tested the default prediction for manufacturing firms that were listed. Importantly, due to differences in the asset base and capital structures, we had to exclude trading and service firms from our analysis. Secondly, we had to resort to a smaller data set due to the confidentiality of default information of banks' clients. However, we believe that the results of our study cannot be undermined, as we offer a firm theoretical base to incorporate Z-Score in banks' credit analysis structure.

REFERENCES

Altman, E. I. (1968). "Financial ratios, discriminant analysis and the prediction of corporate bankruptcy", *The Journal of Finance*, Vol. 23 No. 4, pp. 589-609.

Altman Edward, I., Haldeman Robert, G., and Narayanan, P. (1977). "Zeta analysis: A new model to identify bankruptcy risk of corporations", *Journal of Banking and Finance*, Vol. 10, pp. 29-54.

Altman, E. I., and Saunders, A. (1997). "Credit risk measurement: Developments over the last 20 years", *Journal of Banking & Finance*, Vol. 21 No. 11-12, pp. 1721-1742.

Altman, E. I., Esentato, M., and Sabato, G. (2020). "Assessing the credit worthiness of Italian SMEs and mini-bond issuers", *Global Finance Journal*, Vol. 43, 100450.

Bao, W., Lianju, N., and Yue, K. (2019). "Integration of unsupervised and supervised machine learning algorithms for credit risk assessment", *Expert Systems with Applications*, Vol. 128, pp. 301-315.

Butaru, F., Chen, Q., Clark, B., Das, S., Lo, A. W., and Siddique, A. (2016). "Risk and risk management in the credit card industry". *Journal of Banking & Finance*, Vol. 72, pp. 218-239.

Beaver, W. H. (1966). "Financial ratios as predictors of failure", *Journal of Accounting Research*, pp. 71-111.

Betts, J., and Belhoul, D. (1987). "The effectiveness of incorporating stability measures in company failure models", *Journal of Business Finance & Accounting*, Vol. 14 No. 3, pp. 323-334.

Blum, M. (1974). "Failing company discriminant analysis", *Journal of accounting research*, pp. 1-25.

Bobyl, V. V. (2014). "Use of neuro-fuzzy scoring model in evaluation of credit risk of borrower", *Finances and Credit*, Vol. 36, pp. 18-25.

Breiman, L., and Cutler, A. (2004). "Random Forests", URL: http://www. stat. berkeley. edu/users/breiman. *RandomForests/cc_papers. htm*.

Brewer III, E., and Koppenhaver, G. D. (1992). "The impact of standby letters of credit on bank risk: A note", *Journal of Banking & Finance*, Vol. 16 No. 6, pp. 1037-1046.

Criminisi, A., Shotton, J., and Konukoglu, E. (2012). "Decision forests: A unified framework for classification, regression, density estimation, manifold learning and semi-

supervised learning", *Foundations and trends® in computer graphics and vision*, Vol. 7 No. 2-3, pp. 81-227.
Crouhy, M., Galai, D., and Mark, R. (2000). "A comparative analysis of current credit risk models", *Journal of Banking & Finance*, Vol. 24 No. 1-2, pp. 59-117.
Deakin, E. B. (1972). "A discriminant analysis of predictors of business failure", *Journal of Accounting Research*, pp. 167-179.
Deakin, E.B. (1977). "Business failure prediction: an empirical analysis". In E. Altman and A. Sametz (eds.). Financial Crisis: Institutions and Markets in a Fragile Environment. New York: John Wiley, pp. 117-138.
de Salceda Ruiz, L. M. G. (1997). A credit risk assessment model for dairy farm borrowers in Mexico. University of Illinois at Urbana-Champaign.
Dietterich, T. G. (2000). "An experimental comparison of three methods for constructing ensembles of decision trees: Bagging, boosting, and randomization", *Machine Learning*, Vol. 40 No. 2, pp. 139-157.
Ericsson, J., and Renault, O. (2006). "Liquidity and credit risk", *The Journal of Finance*, Vol. 61 No. 5, pp. 2219-2250.
Gerantonis, N., Vergos, K., and Christopoulos, A. G. (2009). "Can Altman Z-score models predict business failures in Greece?", *Research Journal of International Studies*, Vol. 12 No. 10, pp. 21-28.
Gordy, M. B. (2000). "A comparative anatomy of credit risk models", *Journal of Banking & Finance*, Vol. 24 No. 1-2, pp. 119-149.
Haberman, S. J. (1978). *Analysis of Qualitative Data* Vol. 1, New York: Academic Press.
Hamer, M. M. (1983). "Failure prediction: sensitivity of classification accuracy to alternative statistical methods and variable sets", *Journal of Accounting and Public Policy*, Vol. 2 No. 4, pp. 289-307.
Hastie, T., Tibshirani, R., and Friedman, J. (2009). The Elements of Statistical Learning. SSS.
Jagtiani, J., Saunders, A., and Udell, G. (1995). "The effect of bank capital requirements on bank off-balance sheet financial innovations", *Journal of Banking & Finance*, Vol. 19 No. 3-4, pp. 647-658.
Jiménez, G., and Saurina, J. (2006). "Credit cycles, credit risk, and prudential regulation", *International Journal of Central Banking* Vol. 2 No. 2, pp. 65-98.
Kaminskaya, T. E., and Ivanchenko, Y. S. (2011). "Methodical Aspects of Credit Risk Assessment", *Bulletin of PNU*, Vol. 23 No. 4.
Kaveri, V. S. (1980). *Financial ratios as predictors of borrowers' health: with reference to small-scale industries in India*. New Delhi: Sultan Chand.
Liu, Y., Zhou, Q., Zhao, X., and Wang, Y. (2018). "Can listing information indicate borrower credit risk in online peer-to-peer lending?", *Emerging Markets Finance and Trade*, Vol. 54 No. 13, pp. 2982-2994.
Moscato, V., Picariello, A., and Sperlí, G. (2021). "A benchmark of machine learning approaches for credit score prediction", *Expert Systems with Applications*, Vol. 165, 113986.
Nicolo, A., and Pelizzon, L. (2006). "Credit derivatives: Capital requirements and strategic contracting", In *EFMA 2004 Basel Meetings*.
Ohlson, J. A. (1980). "Financial ratios and the probabilistic prediction of bankruptcy", *Journal of Accounting Research*, pp. 109-131.

Pai, G. R., Annapoorani, R., and Pai, G. A. V. (2006). "A principle component analysis - multiple discriminant analysis model for prediction of corporate failure in India", *Finance India*, Vol. 20 No. 4, pp. 1413.

Pantalone, C. C., and Platt, M. B. (1987). "Predicting failure of savings & loan associations", *Real Estate Economics*, Vol. 15 No. 2, pp. 46-64.

Piesse, J., and Wood, D. (1992). "Issues in assessing MDA models of corporate failure: A research note", *The British Accounting Review*, Vol. 24 No. 1, pp. 33-42.

Platt, H. D., and Platt, M. B. (1991). "A note on the use of industry-relative ratios in bankruptcy prediction", *Journal of Banking & Finance*, Vol. 15 No. 6, pp. 1183-1194.

Saunders, A. (1994). Financial institutions management: a modern perspective.

Schlesselman, J. J. (1982). "*Case-control studies: design, conduct, analysis",* Vol. 2, Oxford University Press.

Srivastava, S. S. (1986). "*Management and monitoring of industrial sickness",* Concept Publishing Company.

Suresh, P., and Paul, J. (2014). "*Management of banking and financial services",* Pearson Education India.

Taffler, R. J. (1982). "Forecasting company failure in the UK using discriminant analysis and financial ratio data", *Journal of the Royal Statistical Society: Series A (General)*, Vol. 145 No. 3, pp. 342-358.

Taffler, R. J. (1984). "Empirical models for the monitoring of UK corporations", *Journal of Banking & Finance*, Vol. 8 No. 2, pp. 199-227.

Williamson, O. (1963). "Managerial discretion and business behaviour", *American Economic Review,* Vol. 53 No. 5, pp. 1032-57.

Wilson, R. L., and Sharda, R. (1994). "Bankruptcy prediction using neural networks", *Decision Support Systems*, Vol. 11 No. 5, pp. 545-557.

Yadav, R. A. (1986). "*Financial ratios and the prediction of corporate failure",* Concept Publishing Company.

Zhu, L., Qiu, D., Ergu, D., Ying, C., and Liu, K. (2019). "A study on predicting loan default based on the random forest algorithm", *Procedia Computer Science*, Vol. 162, pp. 503-513.