International Journal of Economics and Financial Issues

ISSN: 2146-4138

available at http: www.econjournals.com

International Journal of Economics and Financial Issues, 2016, 6(S3) 161-169.

Special Issue for "Asia International Conference (AIC 2015), 5-6 December 2015, Universiti Teknologi Malaysia, Kuala Lumpur, Malaysia"

Corporate Default Prediction with Industry Effects: Evidence from Emerging Markets

Maryam Mirzaei[1]*, Suresh Ramakrishnan[2], Mahmoud Bekri[3]

[1]Department of Finance & Accounting, Universiti Teknologi Malaysia, Johor, Malaysia, [2]Department of Finance & Accounting, Universiti Teknologi Malaysia, Johor, Malaysia, [3]School of Economics and Business Engineering, Karlsruhe Institute of Technology (KIT), Karlsruhe, Germany. *Email: mirzaei.maryam83@yahoo.com

ABSTRACT

The accurate prediction of corporate bankruptcy for the firms in different industries is of a great concern to investors and creditors. Firm-specific data accompany with industry and macroeconomic factors offer a potentially large number of candidate predictors of corporate default. We employ a predictor selection procedure based on non-parametric regression and classification tree method (CART) and test its performance within a standard logistic regression model. Overall entire analyses indicate that the orientation between firm-level determinants and the probability of default is affected by each industry's characteristics. As well, our selection method represents an efficient way of introducing non-linear effects of predictor variables on the default probability.

Keywords: Default Prediction Modeling, Industry Effects, Emerging Markets
JEL Classification: E00

1. INTRODUCTION

Prediction of corporate default is of a great concern to investors/creditors, borrowing firms, and governments. As a result of the collapse of global financial crisis 2008, which started in USA with sub-prime lending crises (Soludo, 2009), many voices have called for a revolution of existing default warning systems to detect or prevent default problems in real time. Many studies evident that firms follow leverage targets (Graham and Harvey, 2001; Fama and French, 2002; Flannery and Rangan, 2006). It can be said that any wrong decision about capital structure may lead the firm to financial distress and eventually to bankruptcy (Eriotis, et al., 2007). Thus, the impact of external factors should be considered. Debt financing plays an important role in emerging market finance. Although both developed markets and emerging markets have experienced an increase in leverage over the past decades, the increase has been more pronounced in emerging markets (Mitton, 2008). This study investigates the probability of default across industries for Iranian listed corporations. We also examine the impact of industry environment on the probability of default.

The substantial body of literature examining the relevance of firms-specific determinants to probability of default has implication for improving the default prediction models (e.g. Chudson, 1945; Jackendoff, 1962; Meyer and Pifer, 1970; Deakin, 1972; Courtis, 1978). While some studies are concerned with the relation between macroeconomic determinants and probability of default (Goudie and Meeks, 1991; Vassalou and Yuhang, 2004; Liou and Smith, 2007; Reisz and Perlich, 2007; Bharath and Shumway, 2008). Overall, little is known about the impact of industry environment on the probability of default, particularly in emerging markets.

Kale et al. (1991) argue that industry characteristics are important in estimating the risk of business. A single information can affect the industries differently across whole market. Therefore, industries tend to have different issues and challenges, which could differently influence the default probability of firms. For that reason, industry plays significant role in explaining the national market volatility (Nishat, 2001). According to Nishat (2001), financial reforms and changing industry specific policies may cause the change in industries risk level. Furthermore, Grinold et al. (1989) argue that some industries are internally more volatile than others. Thus, level of riskiness may vary due to the industries' structure. Moreover, the major industries are more likely to raise financing. Therefore, exploring how the specific environment of each industry could differently affect the probability of default of firms is an essential issue.

The limited literature examining the impact of industry on the firm's probability of default does not perceive the ambiguous theoretical implications. According to Kayo and Kimura (2011) firms tend to have similar properties that operate within a particular industry. A number of empirical evidence highlights the impact of industry on probability of default. In this regard, Opler and Titman (1994) and Maksimovic and Phillips (1997) support the significance of industry effects on probability of default. Likewise, Berkovitch and Israel (1998) model the decision to reveal default as a strategic variable. They show that the fraction of firms that go default is higher for firms operating in mature industries than firms in growth industries. According to Acharya and Srinivasan (2003), industry conditions at the time of default affect the recovery rate. In general, there are variations of behavior across industries that may indirectly affect the probability of default. This impact, indirectly provides insights about the nature of industry and its impact on the firm capability of repayment. Though, few past studies have used dummy variables to characterize the industry. Such techniques do not provide a clear depiction which shows the existent effect of a particular industry (Kayo and Kimura, 2011). Moreover, these studies do not take into account institutional settings within developing and emerging countries.

In conjunction with repeated debt crises over the last three decades, the risks of excessive reliance on foreign currency borrowings have been exposed by developing countries. To avoid these risks, there is a strong case for governments to raise long-term resources from the domestic debt market. The government has been recognized as a key player in the initial stage of bond market development in its role as an issuer, regulator and promoter in developing markets. Iran, which is the second largest economy in the Middle East and North Africa region, suffered from internal and external economic shocks. In the context of Iran, default prediction literature mainly tapped to firm-level characteristics and macro-economic factors. Whereas, the risky situation due to government policies influenced the Iranian firms differently across industries. For instance, the cut-off subsidies have a dramatic effect on specific industries. Therefore, this spat warrants the need to examine the impact of industry characteristics on probability of default among Iranian listed firms.

Due to these criteria, Iran's recent past, presents a rare - arguably a laboratory like - case for the study in hand for various reasons, (1) Firstly, on the basis of its national income and market infrastructure development, Iran is a developing economy (World Bank, 2010), (2) Secondly, Iran is a country where 50% of the economy is centrally planned. The Iranian government policy affects macroeconomic factors, i.e., unemployment rate, the inflation rate, and the rate of economic growth. Eventually, the Tehran Stock Exchange (TSE) is an emerging institution in the early stages of its development, which is different from stock exchanges in developed markets (Bagherpour Velashani, 2007), (3) thirdly, Iran is a bank oriented economy, which has less developed stock and bond market. The global financial crises directly affected the banking sector of the world. Consequently, due to high dependency on banking sector, the industries of Iran have learned a great lesson, (4) finally, Iran's economy heavily relies on oil revenues. This revenue has been used to implement a range of policies. In addition, Iranian firms faced a various internal and external economic shocks: High inflation rate, international sanctions and change in oil price. Despite the significance of industries behavior in explaining the firms' probability of default, this area remained untapped in Iran. Therefore, to capture the real impact of industry, the current study warrants the need to investigate the industry characteristics effects on the probability of default of Iranian firms across industries.

2. LITERATURE AND HYPOTHESIS DEVELOPMENT

The literature identifies the industry effects on probability of default but the argument seems to stand under-explored in the context of emerging markets. A number of studies tend to ignore the significance of industry in their model specification. In addition, researchers face problem in establishing the industry-specific variables due to data limitations. Therefore, most of the past studies removed the industry effect by including an industries dummy (Chava and Jarrow, 2004). However, prior studies support the importance of industry effects on probability of default. Opler and Titman (1994) find that the adverse consequences of leverage on bankruptcy are more pronounced in concentrated industries. In related research, Lang and Stultz (1992) reveal competitive intra-industry effects of bankruptcy announcements. According to Acharya and Srinivasan (2003), industry conditions at the time of default affect the recovery rate. The argument between firm-specific and industry effect is indecisive across emerging markets. However, there exist enormous institutional differences. As well, the effect of sectoral behavior on probability of default determinants may differ across different markets. Despite that, the unique behavior of each industry varies between countries due to different financial settings. It is affirmed that, the lack of developed bond markets is often one of the reasons for the intensity of the financial crisis across developing countries. As the financial system in most emerging and developing economies is centered on banks, an important aspect of the development of bond markets is the impact on the banking system. Since the Asian financial crisis of 1997-98, attention has increasingly focused on the relative roles of the banking sector and of the capital market in developing economies. In many instances, the domestic bond market, where it exists, is generally under-developed, in both breadth and depth, compared to the banking system and the equity market. It has been argued that, over-reliance on bank lending for debt financing exposes an economy to the risk of a failure in the banking system.

According to Noravesh et al. (2007), Iranian industries are heavily dependent on bank financing, which has been shown to affect the investment decisions of firms. Consequently, during uncertain macro environment, banks are resistant to advance long-term loans to the private industries, and often retreat to short-term lending. Based on significance of industries in the performance of firms, it reveals the importance to investigate the industry effects on probability of default of firms in emerging economies. To detail the characteristics of each industry, following the Kayo and Kimura's (2011) approach, which justify the characteristics influencing leverage, and in order to capture the more realistic effect of industry on default prediction, this study employs munificence, dynamism and firm's concentration of an industry. Consistent with Kayo and Kimura (2011), firms operating in industries with abundance of resources amount to have greater opportunities. Hence, these firms tend to be more profitable as compared to firms nested in low munificent environment. They reported negative relationship between long-term leverage and munificence. It is demonstrated that firms are normally very dependent on their internal funds as their main source of financing but as the internal resources reserves, they will rely on debt financing and, subsequently, external equity as a last resort for financing. Many studies evident that firms follow leverage targets (Graham and Harvey, 2001; Fama and French, 2002; Flannery and Rangan, 2006). It can be said that any wrong decision about capital structure may lead the firm to financial distress and eventually to bankruptcy (Eriotis et al., 2007). In purview of default prediction inferences, this inverse association further corroborates negative relationship between munificence and probability of default.

Hypothesis 1: There is a significant and negative relationship between the probability of default and munificence.

Based on empirical work dealing with risk prisma, when the environment becomes more unstable, all the firms nested within the industry are more exposed to business risk. As a result, high dynamism reduces the employment of leverage because unstable environment creates uncertainty in future income of firms. From theoretical perspectives, agency cost theory reveals negative relationship between dynamic environment and leverage. In relation to this theory, firms during dynamic environment tend to use equity financing to reduce the transaction cost arising from increased risk. It could also attributable to the reason that debt becomes more expensive due to uncertain outcomes. On the other hand, stable environment warrants the need for debt financing, as the level of dynamism decreases. In other word, the firms operating in similar industry are exposed to systematic risk (business risk), when the environment under which they are operating becomes unstable. These arguments further dominate the debate on default risk because highly leveraged firms during dynamic environment tend to be more prone to bankruptcy as compared to less leveraged firms. Consequently, this study expects dynamism to exhibit a positive relationship to default probability.

Hypothesis 2: There is a significant and positive relationship between the probability of default and dynamism.

There is general consensus that competitive markets, trigger a strong disciplinary power and lead inefficient companies out of the market. The relationship between market concentration and default is assessed in two orientations. On the one hand, competition is the force that contributes to mortality. The competition increases with the number of players in the market enhancing the level of mortality. Another theoretical orientation emphasizes the fact that the low level of market concentration facilitates greater survival of companies, as a result there are more solvent. In conjunction with above argument, as the industries are subject to different level of growth and risk, hence, the concentration level may be different for various industries. The study therefore, expects negative relationship between probability of default and concentration across industries.

Hypothesis 3: There is a significant and negative relationship between the probability of default and Herfindahl-Hirschman (HH) index.

3. METHODOLOGY AND DATA

3.1. Sample and Data

In accordance with the arguments mentioned above, the society of study contains non-financial firms of Iran, which are listed on TSE for the period of 7-year from 2005 to 2011. The study relies on secondary data, which is mainly retrieved from the publications of TSE: Iran's largest stock exchange, which first opened in 1967. Our sample consists of all listed Iranian corporations with debt financing and available data from four industries, namely, chemical products, cement, food and sugar. To investigate the industry effects on probability of default, the firms are selected from four different industries, due to the following reasons: (1) Firms are selected from the industries, with higher contribution in Iran's economy, (2) the number of firms in the selected industry is acceptable to investigate the industry effect, and (3) the data availability. After eliminating companies with missing and outlier data, the final number of observations is 289, including: 151 Non-default and 138 default.

3.2. Variable Construction

3.2.1. Dependent variable

Default is the dependent variable and it is a dummy variable. If the default variable gets a value of 1 in some year, it means that it has failed in payments on that given year, or it has gone into bankruptcy on that given year. If the firm has gone to bankruptcy it doesn't have any more information after that year. Based on the background of Iranian listed companies, the criteria whether the listed company is specially treated (ST) by the Teheran Stock Exchange categorizes companies into two classes based on their financial condition: (1) Normal and (2) distressed. Distressed companies are referred to ST companies and are classified as such if their accumulated losses are more than 50% of stockholder equity (Iran Business Law Article 141).

3.2.2. Independent variables

Most researchers selected financial ratios as predictor variables based on their popularity and predictive ability in the previous bankruptcy research studies (Altman, 1968; Beaver, 1966). It is evoked by researches that there exists significant relationship between default prediction and firm specific variables. The incentive is the belief that all factors affecting the corporate default or success are in financial statements (Lincoln, 1984). The selected ratios allow for a very comprehensive financial analysis of the companies including profitability, liquidity, leverage and activity. The proxies used for each financial ratio are summarized in Table 1.

According to the importance of economic conditions and its effects on bankruptcy, researchers attempted to develop default prediction models based on macroeconomic factors (El Hennawy and Morris, 1983; Taffler, 1984; Goudie and Meeks, 1991; Liou and Smith, 2007). There are various macroeconomic factors, which have been considered by different researchers. As growth in gross national product is taken to be an overall pointer of a nation's economic health, its effect on default is examined by Taffler (1999) and Bunn (2003). On the basis of empirical literature, in this study we employ three macroeconomic indicators, including: Gross domestic product (GDP), interest rate and inflation.

Prior studies indicate that industry effect is an important component in determining the probability of default. To detail the characteristics of each industry, following the Kayo and Kimura's (2011) approach, which justify the characteristics influencing leverage, and in order to capture the more realistic effect of industry on default prediction, we employ munificence, dynamism and firm's concentration of an industry. According to Dess and Beard (1984), the ability of an environment to preserve a constant expansion is called munificence. Consistent with Kayo and Kimura (2011), firms operating in industries with abundance of resources amount to have greater opportunities. Hence, these firms tend to be more profitable as compared to firms nested in low munificent environment. According to Opler and Titman (1994), there are significant differences across industries in terms of the effects of environmental characteristics on firms. In the viewpoint of the dynamism, Simerly and Li (2000) further argued that across industries the environmental dynamism is different. Thus, it differently affects the similar activities across industries. Based on empirical work dealing with risk prisma, when the environment becomes more unstable, all the firms nested within the industry are more exposed to business risk. On the basis of industry concentration, it can be divided into high and low concentrated industries. Generally, in terms of their characteristics, both types of industries vary (Almazan and Molina, 2005). In other words, low concentration industries (competitive industries) are exposed to high risk and high volatility in profitability; therefore, they use less amount of leverage. In contrast, more concentrated industries utilize greater level of leverage, because they are more profitable, more stable and less exposed to risk. The formulation of industry variables are summarized in Table 1.

3.3. Variable Selection

Selection of predictor variables is an important step in all bankruptcy prediction studies. To date no unified theory has been generally accepted. Most of the previous studies used a brute empirical approach of initial choice of variables (based also on expert knowledge) followed by a step-wise procedure to select the variables in the final logit or discriminant model. Such a procedure is not statistically rigorous. Different sequencing and/or initial ordering of variables need not result in a unique selection. As an attempt to overcome this deficiency some authors started using data mining techniques (Cho et al., 2010). These are also better suited to capture potential non-linearities in the relations between financial distress and predictor variables. We use decision tree in selection of bankruptcy predictors and their subsequent use in prediction models. This approach is compared to more conventional methods of variable selection for the logistic regression model. Both approaches are described in detail in the next two subsections.

3.4. Decision Tree

Decision trees are the most popular and powerful techniques for classification and prediction. The foremost cause behind their recognition is their simplicity and transparency, and consequently relative improvement in terms of interpret ability. Decision trees allocate data to predefined classification groups. For instance, in terms of business default prediction, this technique assigns each firm to a failed or non-failed group. Decision tree is a non-parametric and introductory technique, which is capable to learn from examples by a procedure of simplification. Generally, decision trees are binary trees, which include a set of branches (paths from roots to leaf nodes), leaf nodes (objects classes) and

Table 1: Formulation of independent variables

Type	Ratio	Empirical studies
Firm-specific determinants		
Liquidity	CA/CL; CA/CL; WC/TA	Shin and Lee (2002), Altman (1968)
Profitability	NP/TA; NP/CA; NP/L; EBIT/TA	Gunsel and Cukur (2007), Zhao et al. (2008)
Leverage	CL/TA; L/E; LTL/E	Beaver (1966)
Activity	R/L; S/CA	Zhao et al. (2008), Ravi and Pramodh (2008)
Industry determinants		
Municence	Regression time against the sales of sector over the period of study	Kayo and Kimura (2011)
Dynamism	Standard error of municent slope co-efficient divided by the mean value of sale	Kayo and Kimura (2011)
HH index	Sum of squares of percentage of market shares held by the RMS within a sector	Kayo and Kimura (2011)
Macroeconomic determinants		
GDP	GDP per capita	Bunn and Redwood (2003)
Interest rate	Commercial banks lending interest rate	Jia and Wilson (2002), Liou and Smith (2007)
Inflation	CPI	Jia and Wilson (2002), Liou and Smith (2007)

CPI: Consumer price index, GDP: Gross domestic product, NP: Notes payable, CA: Current assets, CL: Current liabilities, WC: Working capital, TA: Total assets nodes (decision rules), which classifies objects according to their attributes (Dimitras et al., 1996). Therefore, the decision tree takes the form of top-down term structure, which divides the data to generate leaves. Under the structure, one target class is central, and each record flows through the tree along a path is determined by a series of tests, until it obtains a terminal node (Quinlan, 1986).

The basic algorithm for decision tree is a greedy algorithm that constructs decision tree in a top-down recursive divide and conquer manner. The information gain measure is used to select the test attribute at each node in the tree. Such a measure is referred to as an attribute selection measure or a measure of a goodness of split. The attribute with the highest information gain is chosen as the test attribute for the current node. This attribute minimizes the information needed to classify the samples in the resulting partitions and reflects the least randomness or impurity in these partitions. Let, S be a set consisting of s data samples. Suppose the class label attribute has m distinct values defining m distinct classes, C_i (for $I = 1,...., m$). Let, s_i be the number of samples of S in class C_i. The expected information needed to classify a given sample is given by:

$$I(s_1, s_2, ..., s_m) = -\sum_{i=1}^{m} p_i \log_2(p_i)$$

Where, p_i is the probability that an arbitrary sample belongs to class C_i and estimated by s_i/s. Note that a log function to the base 2 is used since the information is encoded in bits. Let, attribute A have v distinct values,

$$\{a_1, a_2,..,a_v\}$$

Attribute A can be used to partition S into v subsets,

$$\{S_1, S_2,..,S_v\}$$

Where, S_j contains those samples in S that have value a_j of A. If A was selected as the test attribute (i.e., the best attribute for spliting) then, the subsets would correspond to the branches grown from the node containing the set S. Let s_{ij} be the number of samples of class C_i in a subset s_{ij}. The entropy or expected information based on the partitioning into subsets by A, is given by:

$$E(A) = \sum_{j=1}^{v} \frac{s_{1j} + ... + s_{mj}}{s} I(s_{1j} + ... + s_{mj})$$

The term,

$$\frac{s_{1j} + ... + s_{mj}}{s}$$

Acts as the weight of the j^{th} subset and is the number of samples in the subset divided by the total number of samples in S. The smaller the entropy value is, the greater is the purity of the subset partitions. Note that for a given subset S_j,

$$I(s_1, s_2, ..., s_m) = -\sum_{i=1}^{m} p_i \log_2(p_i)$$

Where,

$$p_{ij} = \frac{s_{ij}}{|s_j|}$$

And is the probability that a sample in S_j belongs to class C_i.

We use CART classification tree in selection of bankruptcy predictors and their subsequent use in prediction models. CART builds classification and regression trees for predicting continuous dependent variables (regression) and categorical predictor variables (classification). Our reasons for selecting CART from the family of artificial intelligence methods are similar to those of Li et al. (2010). From a practical point of view, one important advantage of decision trees in bankruptcy prediction is the ability to generate easily understandable decision rules. This feature is not shared by many artificial intelligence approaches.

3.5. Logistic Regression Model

Our basic model of the probability of bankruptcy is the logistic regression. The logistic regression model has been extensively applied in the literature (Chen, 2011; Min and Jeong, 2009). There are several reasons to employ this model. First, the logistic regression has been widely used. Second, it is relatively easy to understand and readily available in virtually all software packages. Finally, logistic regression has resulted to be a fairly robust and reliable tool for forecasting financial distress. We use the following equation to test our hypotheses:

$$\text{Probability of default } (x_1,..,x_i) = \frac{1}{1 + e^{-(\beta_0 + \beta_1 x_1 + ... + \beta_t x_t)}}$$

Where, x_1 to x_i are independent variables such as firm-specific variables and β_1 to β_i are coefficients, which are estimated by the model. The logistic regression, also uses the form:

$$P(x) = \frac{e^{\beta_0 + \beta_1 X}}{1 + e^{\beta_0 + \beta_1 X}}$$

It is easy to see no matter what values β_0, β_1 or X take, $p(x)$ will have values between 0 and 1. This model can be explained as the probability of default based on firm's given characteristics. In this model, maximum probability function is applied. Hence, the weights are employed to make best use of the probability of default for the identified failed companies and the probability of non-default for non-failed companies. Thus, based on this technique, using a broken-off point, a firm is classified as failed or non-failed. Logistic regression is also able to verify the significance of individual variables in the model (Allison, 2001; Hosmer and Lemeshow, 2000). Logit models are not the only bankruptcy prediction models we consider. While our paper focuses on the use of classification trees in selection of bankruptcy predictors for standard parametric models like logit, it is also straightforward to use the classification tree for bankruptcy prediction.

4. RESULTS AND DISCUSSION

4.1. Results of Variable Selection

In this section we present the results of variable selection using CART classification tree. The estimated classification tree on the

Table 2: Classification results

Model	LR					CART				
	(1)	(2)	(3)	(4)	(5)	(1)	(2)	(3)	(4)	(5)
Accuracy	77.5	80.6	78.4	85.1	81.9	78.2	81.94	83.08	86.48	79.16
RMSE	0.48	0.35	0.24	0.35	0.37	0.43	0.36	0.20	0.33	0.4
ROC	0.78	0.89	0.9	0.89	0.86	0.83	0.88	0.92	0.93	0.87

The overall accuracy, RMSE and ROC for LR and CART classification tree, (1) Overall data, (2) Chemical industry, (3) Cement industry, (4) Food industry, (5) Sugar industry. LR: Logistic regression, RMSE: Root mean squared error, ROC: Receiver operating characteristic

matched sample is presented in Figure 1. Given the sample size it resulted optimal to estimate a simple tree with nodes determined by thresholds on five variables. The first node is determined by variable interest rate. In the left branch of the tree (if the value of interest rate is more than 13.5), the tree contains a refinement based on variable cash/current liabilities which determines liquidity. Based on the same variable a threshold is found also on the right branch of the tree, which is further refined by using the information on net profit/liabilities which shows the firm's profitability. After threshold based on the profitability, there are two further branches based on the activity measure, which is defined as sales over the current assets, and dynamism, that measures the environment instability.

4.2. Logistic Regression and Hypotheses Testing

This study demonstrates the importance of including industry effects in probability of default models. To do so, three different kinds of measures are used to capture the industry effects, including: (1) Munificence, (2) dynamism, and (3) HH index. Examining the sign and size of the coefficients, the intercepts indicate that the firm's default probabilities differ across industries. Based on the results from stepwise logistic regression shown in Table 2, when the firm goes default ($Y = 1$), for variables with negative coefficients the probability of firm's default decreases, and inversely variables with positive coefficients cause an increase in the firm's probability of default. The estimates of the logistic regression model are presented in Table 3.

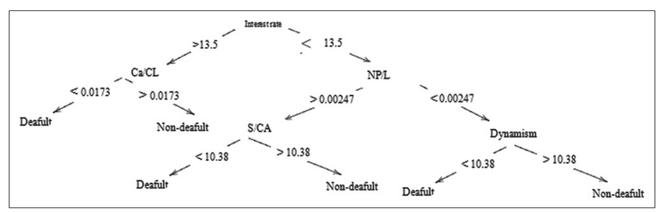
The estimated probability of default model confirms the study expectations, regarding the impact of the individual indicators on firm's default probability. Accordingly, the negative sign of profitability coefficient indicates that the Iranian firms with higher profitability tend to be more able to repay their obligations. As expected, liquidity is negatively related to the probability of default. In other words, the firm has ability to repay its obligations without incurring too much cost, in case of higher liquidity. Based on the results in terms of macroeconomic variables, the most noteworthy indicator is interest rate, which has strong influence on the firm's probability of default. It signifies that higher interest rate tends to increase the probability of default. Interest rate persistently remains the most important factor to probability of default among the studied industries. The size of the coefficient further explains the impact of interest rate on probability of default. This is consistent with the findings of Kritzer (1985), Liou and Smith (2007). Emerging and developing markets tend to have volatile business cycles and experience economic crises more frequently than developed economies. In particular, emerging and developing markets economies face volatile and highly interest rates (Neumeyer and Fabrizio, 2005).

Table 3: Estimated LR model with step-wise selection procedure and CART classification tree

Type	(1)	(2)	(3)	(4)	(5)	(6)
CA/CL	−0.35**	−0.75**	−4.3**		−0.75**	−1.22**
NP/TA	−0.11*	−0.45**	−2.28**	−15.44**	0.42**	−1.93**
LTL/E			0.04*	−0.01*	0.04*	0.03*
S/CA	−0.01*	−0.04*			−0.04*	−0.01*
WC/TA			−1.06**	−0.34**		
GDP	−0.01*		−0.43**	−0.29*		
Inflation			0.07		−0.17*	
Interest rate	0.43**		0.52**	0.37**	0.47*	1.33**
Municence	−0.03*					−0.13*
Dynamism	0.01*	0.08*	−0.1*	0.02*		0.08*

*Donates statistical significance at %5, **Donates statistical significance at %1 (1) LR Overall data, (2) DT Overall data, (3) LR Chemical industry, (4) LR Cement industry, (5) LR Food industry, (6) LR Sugar industry.

Figure 1: Estimated optimized tree on the full sample



In analyzing Iran's capital market, paying attention to this issue is necessary, as about 20% of the value of Iran's capital market is assigned to banking industry. Changes in interest rate lead to further changes in demands of shares of banks, listed on the stock exchange. The latter changes cause an increase or a decrease of the price of the shares of banks. This negatively influences the indices of TSEs. Likewise, GDP per capita is significantly related to the firm probability of default. As a whole, GDP per capita negatively affects the firm probability of default across industries. It also implies that increasing GDP per capita tends to decline the probability of default, which is consistent with Kritzer (1985), and Bunn and Redwood (2003). Looking into overall data, the influence of inflation on probability of default is not denoting. Although, the trend of inflation shows a great increase during the sample period, but, this variable is insignificantly related to the probability of default. An explanation may be given by the government policies, which employ the oil income in order to cover the inflation effects on the industrial sector. Notwithstanding, Iran's economic growth has been hampered by double-digit rates of inflation. However, the Iranian government policy affects the macroeconomic factors, including: The unemployment rate, the inflation rate and the rate of economic growth. The full sample results reported in Table 3 support Hypothesis 1 that affirms the existence of significant relationship between the probability of default and munificence. The results of split sample tests for industries depict that munificence preserved negatively indicative association with probability of default; however, the substance of this variable remained less significant across other industries, particularly chemical, cement and food industries. These findings corroborate that except for sugar industry, the level of industry's growth tend to be an unimportant indicator of default prediction in other Iranian industries.

Hypothesis 2 predicts that the relation between probability of default and dynamism is positively significant. Among industry indicators, the study reports dynamism as most important factor which has compelling impact on the probability of default, particularly across chemical products, cement and sugar industry. Based on overall sample, looking into environmental instability across industries, the negative sign of dynamism coefficient is inconsistent with the study expectations, which predicts the positive relationship between dynamism and probability of default. The possible explanation in the notion of Iranian firms could be that they tend to reduce the usage of debt as the environment of industries become more dynamic. These inferences further extend inverse relationship between dynamism and probability of default. Another argument apparent to this fact may be that with the increased level of industrial dynamism, Iranian firms opt to issue equity financing, which operate under dynamic environment. As a whole, it is clearly observed that two industries including sugar and cement are the most affected industries. The results also depict that the firms operating under sugar industry are more affected by the level of industry's growth. During last decade, a number of companies operating under sugar industry faced problem to repay their obligation and went default. An explanation for this issue could be the specific environment and government policies which have more focused on the major industries such as oil, chemical product and food industries.

Based on overall sample, the impact of HH index on probability of default is insignificant. Thus, this immaterial inference could attributable to the reason, that developing industries are closely held, where few families and corporate groups controlling the ownership structure of corporations. This is why industry concentration maintained insignificant relationship with the probability of default. This is in contradiction to developed economies having low level of industry concentration, and industries are widely held. This result eventually rejects the hypothesis 3, which indicates a significant relationship between HH index and the probability of default. In general, the significant relationship between firm-specific, industry and macroeconomic indicators, and firm's probability of default across industries, evident that nature of each industry tends to affect the mechanism between probability of default and its determinants. Moreover, the size of magnitude of each significant variable varies across industries. This reveals the effect of each variable in the process of probability of default determinants. In other words, the different sizes of coefficients indicate the different degrees of economic importance of each indicator on probability of default. These analyses, further point out a substantial divergence between the overall sample and the industries' results. The noteworthiness of each variable on probability of default is noticeable across sectors.

Root mean squared error, and receiver operating characteristic metrics were used to evaluate each algorithm for financial prediction performance by applying logistic regression and decision tree methods to the overall and sectorwise dataset. The result of performance measure is presented in Table 2. As it can be seen, classification trees result to be best suited for predicting bankruptcy cases for different industries. In addition, the models performance varies across industries. This might be explained due to the different nature and environment of industries which affects the determinants of probability of default, indirectly.

5. CONCLUSION

The corporate default prediction undergoes an enormous augment in the number of researches. However, several open issues remained less explored. For decades the ultimate goal of the most of the prediction models was (and it still is) to increase the prediction power of the models. Mostly, this leads to some complicated techniques. A variety of factors ascertain the quality of an econometric study, including: The model, the algorithm and the data. The robust approaches endeavor to initiate reliable and accurate models, by considering noteworthy determinants and well-founded techniques.

In this study, the idea of increasing the performance of default prediction models is investigated. Therefore, this study extensively explores the explanatory variables distributions and their potential influence on the prediction power of the models. To do so, this study highlights the sensitivity of the determinants of the probability of default, by employing the industry variables. Under the research framework, this study confirms the industry effects on the firm's probability of default. Although, it is ascertained that the firm-level variables are the most predominate determinants, due to their explanatory power. Nevertheless, the overall entire analyses indicate that the orientation between firm-level determinants and the probability of default is affected by each industry's characteristics.

Though, the employment of the industry-based analysis, the study points out the indirect impact of industry characteristics on the determinants of the probability of default. The findings depict that the indirect impact is clearly indictable, due to the changing of the sign and magnitude of the determinants across industries. In summary, the conclusions of the empirical analysis in this study highlight that by applying proper explanatory variables and techniques, it is possible to increase the performance of the default prediction models. The study of default prediction provides an early warning signal and detects areas of weaknesses. Accurate default prediction usually leads to many benefits such as cost reduction in credit analysis, better monitoring, and an increased debt collection rate.

REFERENCES

Acharya, V.B.S., Srinivasan, A. (2003), Understanding the Recovery Rates of Defaulted Securities, Working Paper. London Business School.

Allison, P.D. (2001), Logistic Regression Using the SAS System: Theory and Application. Cary, NC: SAS Publishing, BBU Press.

Almazan, A., Molina, C. (2005), Intra-industry capital structure dispersion. Journal of Economics and Management Strategy, 14(2), 263-279.

Altman, E.I. (1968), Financial ratios, discriminant analysis and the prediction of corporate Bankruptcy. Journal of Finance, 23(4), 589-609.

Bagherpour Velashani, M.A. (2007), Auditor Switches and Emerging Markets: The Case of Iran. Canberra: Australian National University.

Beaver, W.H. (1966), Financial ratios as predictors of failure. Journal of Accounting Research, 4(3), 71-111.

Beaver, W., McNichols, M., Rhie, J. (2005), Have financial statements become less informative? Evidence from the ability of financial ratios to predict bankruptcy. Review of Accounting Studies 10, 93-122.

Berkovitch, E., Israel, R. (1998), The bankruptcy decision and debt contract renegotiations. European Finance Review, 2, 1-27.

Bharath, S.T., Shumway, T. (2008), Forecasting default with the Merton distance to default model. Review of Financial Studies, 21(3), 1339-1369.

Bunn, P., Redwood, V. (2003), Company Accounts Based Modelling of Business Failures and the Implications for Financial Stability, Bank of England Working Paper, No. 210.

Bunn, P., Redwood, V. (2003), Company Accounts Based Modelling of Business Failures and the Implications for Financial Stability, Bank of England Working Paper, No. 210.

Chava, S., Jarrow, R. (2004), Bankruptcy prediction with industry effects. Review of Finance, 8, 537-569.

Chen, S., Hrdle, W.K., Moro, R.A. (2011), Modeling default risk with support vector machines. Quantitative Finance, 11(1), 135-154.

Cho, S., Wang, J., Kuo, C.C.J., Cha, B.H. (2010), Block-Based Image Steganalysis for a Multi - Classiffier. Multimedia and Expo (ICME), 2010 IEEE International Conference on. p1457-1462.

Chudson, W. (1945), The Pattern of Corporate Financial Structure. New York: National Bureau of Economic Research.

Courtis, J.K. (1978), Modelling a financial ratios categoric framework. Journal of Business Finance and Accounting, 5, 371-386.

Deakin, E.B. (1972), A discriminant analysis of predictors of business failure. Journal of Accounting Research, 10(1), 167-179.

Dess, G., Beard, W. (1984), Dimensions of organizational task environments. Administrative Science Quarterly, 29(1), 52-73.

Dimitras, A., Slowinski, R., Susmaga, R., Zopounidis, C. (1999), Business failure prediction using rough sets. European Journal of Operational Research, 114, 263-280.

El Hennawy, R.H.A., Morris, R.C. (1983), The significance of base year in developing failure prediction models. Journal of Business Finance and Accounting, 10(2), 209-223.

Eriotis, N., Vasiliou, D., Neokosmidi, Z.V. (2007), How firm characteristics affect capital structure: An empirical study. Managerial Finance, 33(5), 321-331.

Flannery, M.J., Rangan, K.P. (2006), Partial adjustment toward target capital structures. Journal of Financial Economics, 79(3), 469-506.

Fama, E.F., French, K.R. (2002), Testing trade-off and pecking order predictions about dividends and Debt. Review of Financial Studies, 15(1), 1-33.

Grinold, R., Andrew, R., Dan, S. (1989), Global factors: Fact or fiction? Journal of Portfolio Management, 16(1), 79-89.

Graham, J.R., Harvey, C.R. (2001), The theory and practice of corporate finance: Evidence from the field. Journal of Financial Economics, 60, 187-243.

Goudie, A.W., Meeks, G. (1991), The exchange rate and company failure in a macro-micro model of the UK company sector. Economic Journal, 101(406), 444-457.

Gunsel, N., Cukur, S. (2007), The effects of macroeconomic factors on the London stock returns: A sectoral approach. International Research Journal of Finance and Economics, 10, 140-150.

Hosmer, D.W., Lemeshow, S. (2000), Applied Logistic Regression. New York: Wiley.

Jackendoff, N. (1962), A Study of Published Industry Financial and Operating Ratios. Philadelphia, PA: Temple University, Bureau of Economic and Business Research.

Jagtiani, J., Allen, L., Peristiani, S., Saunders, A. (2003), The role of bank advisors in mergers and acquisitions. Journal of Money, Credit, and Banking, 36, 197-224.

Jia, L., Wilson, N. (2002), Corporate failure rates and the impact of the 1986 insolvency act: An econometric analysis, Managerial Finance, 28 (6), 61-71.

Kale, J.R., Noe, T.H., Ramirez, G.G. (1991), The effect of business risk on corporate capital structure: Theory and evidence. Journal of Finance, 46, 1693-1716.

Kayo, E.K., Kimura, H. (2011), Hierarchical determinants of capital structure. Journal of Banking and Finance, 35, 358-371.

Kritzer, H.M, 1985. The shadow of panitives: An unsuccessful effort to bringit into view. In: Arlen, J. 2013. Research Handbook on the Economics of Torts.

Lang, L., Stultz, R. (1992), Contagion and competitive intra-industry effects of bankruptcy announcements: An empirical analysis. Journal of Financial Economics, 32, 45-60.

Li, H., Sun, J., Wu, J. (2010), Predicting business failure using classification and regression tree: An empirical comparison with popular classical statistical methods and top classification mining methods. Expert Systems with Applications, 37, 5895-5904.

Lincoln, M. (1984), An empirical study of the usefulness of accounting ratios to describe levels of insolvency risk. Journal of Banking and Finance, 8(2), 321-340.

Liou, D.K., Smith, M. (2007), Macroeconomic variables and financial distress. Journal of Accounting, Business and Management, 14, 17-31.

Maksimovic, V., Phillips, G. (1997), Asset Efficiency and Reallocation Decisions of Bankrupt Firms, Working Paper, University of Maryland.

Meyer, P.A., Pifer, H.W. (1970), Prediction of bank failures. Journal of Finance, 25(4), 853-868.

Mitton, T. (2008), Why have debt ratios increased for firms in emerging markets? European Financial Management, 14, 127-151.

Min, J.H., Jeong, C. (2009), A binary classification method for bankruptcy prediction. Expert Systems with Applications, 36, 5256-5263.

Mohmad Isa, H. (2004), Determination of Capital Structure and Prediction of Corporate Financial Distress in Malaysia. Ph.D. Dissertation, Universiti Putra Malaysia.

Nishat, M. (2001), Industry risk premia in Pakistan. Pakistan Development Review, 40(4), 929-952.

Neumeyer, P.A., Fabrizio, P. (2005), Business cycles in emerging economies: The role of interest rates. Journal of Monetary Economics, 52(2), 345-380.

Noravesh, I., Dianati Dilami, Z., Bazaz, M.S. (2007), The impact of culture on accounting: Does grays model apply to Iran? Review of Accounting and Finance, 6(3), 254 272.

Opler, T., Titman, S.H. (1994), Financial distress and corporate performance. Journal of Finance, 3, 1015-1040.

Quinlan, J.R. (1986), Induction of decision trees. Machine Learning, 1, 81-106.

Ravi Kumar, P., Ravi, V. (2008), Bankruptcy prediction in banks and firms via statistical and intelligent techniques - A review. European Journal of Operational Research, 180(1), 1-28.

Reisz, A.S., Perlich, C. (2007), A market-based framework for bankruptcy prediction. Journal of Financial Stability, 3(2), 85-131.

Shin, K.S., Lee, Y.J. (2002), A genetic algorithm application in bankruptcy prediction modeling. Expert system with Applications, 23(3), 321-328.

Simerly, R., Li, M., (2000), Environmental dynamism, capital structure and performance: A theoretical integration and an empirical test. Strategic Management Journal, 21(1), 31-49.

Soludo, C. (2009), Global Financial and Economic Crisis: How Vulnerable is Nigeria. Abuja: Central Bank of Nigeria.

Tew, Y.H., Enylina, N. (2005), Predicting corporate financial distress using logistic regression: Malaysian evidence. Working Paper. Universiti Teknologi Mara.

Taffler, R.J. (1984), Empirical models for the monitoring of UK corporations. Journal of Banking and Finance, 8(2), 199-227.

Taffler, R.J. (1999), Rational Asset Pricing and Bankruptcy Risk: A Z-Score Perspective, Helsinki European Finance Association's 26th Annual Meeting.

Vassalou, M., Yuhang, X. (2004), Default risk in equity returns. Journal of Finance, 59(2), 831-868.

World Bank. (2010), World Bank Development Indicator the International Bank for Reconstruction and Development. Washington, D.C: World Bank.

Zhao, S., Nossal, K., Kokic, P., Elliston, L. (2008), Productivity growth: Australian broadacre and dairy industries. Australian Commodities, 15(1), 236-242.