



Finance & Banking Studies

IJFBS ISSN: 2147-4486

Contents available at www.ssbfnets.com/ojs
DOI:

Risk Pricing in Emerging Economies: Credit Scoring and Private Banking in Iran

Yiannis Anagnostopoulos

Kingston Business School, Department of Accounting, Finance and Informatics, Lodon, UK

Abstract

Iran's banking industry as a developing country is comparatively very new to risk management practices. An inevitable predictive implication of this rapid growth is the growing concerns with regard to credit risk management which is the motivation of conducting this research. The paper focuses on the credit scoring aspect of credit risk management using both logit and probit regression approaches. Real data on corporate customers are available for conducting this research which is also a contribution to this area for all other developing countries. Our questions focus on how future customers can be classified in terms of credibility, which models and methods are more effective in better capturing risks. Findings suggest that probit approaches are more effective in capturing the significance of variables and goodness-of-fitness tests. Seven variables of the Ohlson O-Score model are used: CL_CA, INTWO, OENEG, TA_TL, SIZE, WCAP_TA, and ROA; two were found to be statistically significant in logit (ROA, TL_TA) and three were statistically significant in probit (ROA, TL_TA, SIZE). Also, CL_CA, ROA, and WCAP_TA were the three variables with an unexpected correlation to the probability of default. The prediction power with the cut-off point is set equal to 26% and 56.91% for defaulted customers in both logit and probit models. However, logit achieved 54.85% correct estimation of defaulted assets, 0.37% more than what probit estimated.

Key Words: *Credit Scoring, Probit, Logit Models, Risk Management, Risk Pricing, Credit Scoring*

JEL classification: G2

Introduction

In light of the 2008 financial crisis, risk management is becoming more important than ever. According to the Basel Committee, the largest source of serious banking problems is credit risk (Kalapodas & Thomson, 2006). All while at the same time, there has been a steady increase rate of defaults and bankruptcies which made credit risk an important part of Financial Institutions' (FIs) risk management. It is the degree of value fluctuations in debt instruments and derivatives due to changes in the underlying credit quality of borrowers and counterparties. The loss distribution of a credit risk model is composed of two components: a weighting vector to characterise holding of credits and the multivariate distribution of the credit losses. A bank's credit portfolio at time t is given by $P_t = w'A_t$ where w is the weight of A which is an $[N \times 1]$ vector of a bank's credit portfolio and the credit risk model characterises ΔP_{t+1} (Lopez & Saidenberg, 1999).

Lang et al. (2010) emphasised that corporate governance problems and weak credit risk management caused the financial crisis in 2008 which shows how the world's economy can suffer because of this type of risk. Measuring credit risk is important not only because of the loss which may occur from inappropriate lending but also because of miscalculation in the associated credit risk premium which poses a potential danger of losing a good customer. In another study, done by Fatemi et al. (2006), the fact that an increase in variety of the types of counterparties (from individuals to sovereign governments) and the forms of obligations (e.g. auto loans and complex derivative transactions) has made credit risk management a high priority task in FIs. Furthermore, the first step of building a strong and stable financial system is to minimise the non-performing loans (Sanjeev, 2007) and this is directly related to sound credit assessment and monitoring. Berger, Allen & De Young (1997) studies have shown that bad loans makes banks increase spending on monitoring, working out or writing of the loans, and become more diligent in administering their outstanding loans. In fact, bad credit risk management is not only manifested in excess expenditures but also subpar monitoring and underwriting practices which eventually pave the road to non-performing loans.

To that effect, market risk management is not a considerable risk aspect for banks mainly due to the nature of Islamic banking and because interest rates are being set by the Central Bank of Iran (CBI) and they rarely change. Also equity investments weigh only 2%-3% of the portfolio of total assets of Iranian banks as is the same case with bond portfolios.

Although credit risk management has started to be looked at seriously, there are still a lot of concerns with regards to this matter. For example, members of the so called Majlis (Iran's parliament) disclosed....(as of September 2011), banks have \$27 billion unpaid, outstanding loans, which they cannot collect, while a delegate from our committee to the Financial Corruption Panel has reported that \$8 billion worth of loans have been given to 90 people. Through a rough calculation, we find that this is about \$90 million per person (PressTVa, 2011).. Furthermore, the authorities of Iran's General Inspections Organization announced that Iranian banks have around USD 38 billion of delinquent loans while they are only capitalised at USD 20 billion. The organisation added that the 2010 average for late debts in Iran's state owned banks is over 15% compared to the 3%-5% global standard figures. Sanjeev (2007)

brought up some global figures for the ratio of bad debts to total loans which was 2.5% for the United States, 12.5% for India, 25% in China, and 45% in Thailand.

Extending credit by Iranian banks which is expected to practice Islamic banking is of certain distinctive characteristics. Iran's three largest banks according to Financial Times (2010) are state owned. Credit risk management is a vast area for research. It involves four stages: identification, measurement, management and control of credit risk (Kalapodas & Thomson, 2006). Judgmental approaches are still widely used in Iran while economic growth starves for credit for further development and as a result loan customers are increasing. Credit Scoring is important in a way that a major credit institution cannot evaluate each customer individually and base credit decision on subjective judgment (Greene, 1998). Credit institutions have almost fifty years of history in scoring their customers (Tasche, 2006). There is a strong need for modern credit scoring models to help make appropriate credit decisions. What is more, using credit scoring techniques improves the accuracy of credit assessment, reduces lending costs and allows banks expand

their credit facilities to marginal borrowers that would otherwise not be able to receive credit (Berger, Frame & Miller, 2002). To do that, statistical models are applied to historical data for each customer (financial, behavioural, or economical) to score them (Greene, 1998).

Rating and scoring are rather different concepts. Historically rating has been used to refer to pricing of bonds issued by large corporate while scoring is mostly associated to retail credit rationing. Furthermore, ratings are often being developed by experienced practitioners while scoring is being conferred on experts in statistics. Nowadays, both rating and scoring are used to determine the Probability of Default (PD) in order to calculate price of a loan and make credit decisions. According to Basel II and III regulation, there are two approaches for measuring credit risk, the standardised methods and the Internal Rating Based (IRB) approach. The lack of availability of data for most customers as a standard approach is the bottleneck of the standardised methods and has led credit institutions towards the IRB approach. Using their in-house credit rating or scoring models, PDs are being calculated which is later being used for minimum capital requirement calculations (Tasche, 2006). What variables are important in a credit scoring model, how they can be adjusted if an adjustment is needed, how future customers can be classified in terms of their credibility and which credit model should be used are the main questions.

We utilise both logit and probit regressions which are applied to the data provided by our case study private Iranian bank using Ohlson's (1980) proposed variables. Two outcomes are expected, either a default or a non-default customer (dichotomous regression analysis). Then, the results of the logit and the probit regressions are being compared with each other. Also, a function is developed to predict the probability of default of future customers for that bank based on the coefficients that the two regression analyses estimate. As this research is using confidential data, any information which can be used to reveal the identity of that bank is not being mentioned in this paper.

To our knowledge, there is no research available which analyses the results of a credit scoring model using highly confidential real time data in Iran from the review of literature to date. Furthermore, contextual research in this field and across the whole set of developing economies is very limited due to the nature of 'closed' economies and the highly confidential nature of the data such research requires.

This paper contributes to the topic of credit scoring to help credit institutions especially those in Iran make credit decisions. Our research is laid out on four further sections. Section 2 is a review of previous related research conducted in this area. Section 3 is dedicated to the methodology and data. Section 4 is the analysis followed by section 5, the conclusion. Appendix I is the sample output of regression analysis and Appendix II is the histogram of disturbances distributions.

Literature Review

Sanjeev (2007), studies a series of factors which affect loans negatively within an Indian context concluding that external factors have a higher influence compared to the internal factors (out of which economic downturn and wilful default are the most influential) followed by poor credit scoring skills of managers which has a significant role in default on loans. This study signifies the importance of good credit scoring skills on loan quality. However one possible limitation of this study is its small sample size.

In order to obtain a good predictive model, taking into account the availability and the quality of data, method selection and variable selection are very important. The Federal Reserve System Task Force on Internal Credit Models (1998, 2009) and Basel Committee on Banking Supervision (Basel III, 2009) highlighted two issues with regard to internal credit rating models: maintaining consistency of a chosen IRB approach and controlling the quality of inputs to that model. Kalapodas and Thomson (2005) evaluated different credit risk management methodologies. They studied the judgmental method, credit scoring, and portfolio models. A good risk management practice is to rely collectively on all approaches. When credit risk is assessed judgmentally, the risk manager will use information which may potentially threaten counterparty's commitment to the Financial Institutions (FIs). Three main areas to look at in this regard are: the guarantees of the counterparty, its capacity and its characteristics which are defined as the borrower's willingness to repay (Hempe, Simonson & Coleman, 1994).

The capacity of a borrower is being measured using financial performance and ratios extracted from financial statements. The rationale for this approach is also examined in the literature. Koch & McDonald (2000) conclude that lending to highly leveraged companies is hazardous. Bass (1991) and Rice and Coyle (1999) also established the same conclusion with lending to companies over dependent on external financing and companies with a series of negative cash flows. Also the external factors and the borrower's sensitivity to such factors are measured. Factors generally measured are the country risk, economic recession and exchange rate risk. The credit risk then materialises when the value of the collateral decreases in the market or the guarantor's financial stability decreases.

There are a number of sources available and being used when assessing credit risk of loan customers including agency reports, online resources, or reference books (Bass, 1991) and accounting figures like balance sheet and profit and loss accounts (Koch & McDonald, 2000). These data are susceptible to creative accounting practices as well as hiding data that may negatively affect the decision. In this regard, the judgmental method outperforms the statistical models. In general, credit risk assessment is a pattern relation task, where human beings are more efficient than computers (Hoch, 2001). Furthermore, the judgmental method establishes early relation between the borrower and the lender. On the other hand, perhaps the most important issue with judgmental method is that decisions are made subjectively. Libby's study (1975) was the first to examine the impact of human expectations on subjective selections. They developed a relationship between what a person chooses to see and his expectations. Therefore, with a large number of applicants, credit managers focus only on what catches their eyes and not always what is there. This is also being confirmed by Plous (1993) in his study.

The other method for credit risk management is credit scoring. Credit scoring models use data being provided by the borrower to the FI to predict the Probability of Default (PD). The first thing is to extract data from the borrower. Data can be either categorical (i.e. gender, own home or rent etc.) or continuous (i.e. income) which can be grouped into intervals to be treated as categorical (Hand, 2001). Once the predictor variables are identified they are used to measure a borrower's creditworthiness.

The values are called the attributes and the output of the model should be the score (Hand & Henly, 1997). The results of scores are a continuous or a discrete number per customer. These numbers should be mapped to grades provided by Basel II for calculating the capital adequacy requirements and the original numbers being kept in the institution's discretion (Tosche, 2006). Expert systems mixed with statistical models, often in different combinations, are called hybrid models. These are often the statistical models and their results can be ruled out by expert decisions (Tosche, 2006).

Methods used in credit scoring are generally either data mining or regression methods. Subbanarasimha, Arinze, and Anadarajan (2000) and Gorr, Nagin, and Szczypula (1994) both concluded that linear regression models outperforms data mining techniques while Shuhui, Wunsch, Hair, and Giesselmann (2001) proved it to be the other way around. However, these studies were based on their sample sizes and variables which make their studies tailor made to their data in hand. Kim (2008) studied decision tree analysis, artificial neural networks, and linear regression. They used 60 simulated examples to compare the performance of these methods based on the number of independent variables, the number of classes of the independent variables, and sample size using SAS Enterprise Miner. They found out that for continuous independent variables, a statistical technique like linear regression is superior to data mining techniques like neural networks or decision trees. They also found that if variables are continuous and categorical, statistical methods still outperform the mining techniques when there is one variable only and otherwise it is the neural network which outperforms others for two or more variables.

Credit institutions predominantly use statistically based score variables when it comes to scoring. One possible problem is the population drift explained by Hand & Henley (1997) who claims that as the population evolves in time, distributions change. Two main reasons are that the applications are closely related to the economic pressure and a change in competitive environment. This issue can degrade the performance of a credit decision model. Such models should be dynamically adjusted to this issue. Choosing the inputs used in a model is as important as choosing the model itself. Matousek & Stewart (2009) conducted a study of 681 international banks' Fitch credit ratings and how each one is reflecting

their financial performance using probit models. Their research concluded that banks with larger assets, greater capitalisation and higher return on asset have higher ratings. It is a common approach in choosing the explanatory variables of a model based on some hypotheses such as checking whether a customer's liquidity position is important in forecasting the probability of default which is what may be expected. In fact, given some financial figures, thousands of ratios may be generated. But risk management is half art and half knowledge and to choose which ratio to use is critical in conducting a good research. Beaver (1966) for example, analysed a list of ratios individually for predicting bankruptcy and found cash flow to total debt signals firm's failure well ahead.

Altman (1968), in the same line of argument, extended Beaver's approach by introducing a discriminant function to combine ratios using multivariate analysis (Wang & Campbell, 2010). Beaver's univariate approach which put emphasis on individual signals of impending cases made Altman develop his study in the field (Altman, 1968). In his famous Z-score model used five ratios: working capital to total assets which he found to be the most valuable liquidity ratio, retained earnings to total assets which is a measure of cumulative profitability over time considering the size of the firm, operating income after depreciation to total assets as a measure of true productivity of a firm's assets, sales to total assets to measure managerial performance in sales generating ability of the firm's assets, and the market value of equity to total liabilities to proxy for insolvency. They found that such variables outperformed that of Beaver's (Wang & Campbell, 2010). Multiple Discriminant Analysis (MDA) in the selection of variables is followed by a big number of researchers.

Ohlson (1980) criticised Altman's work because of the statistical restrictions the MDA model has. Their concerns were the normality assumptions of predictors' distribution. Alternatively, in a dichotomous scenario (failed or not failed), the variance-covariance matrix of the predictors should be the same for both outcomes. In their study, the logit regression has been utilised on these variables for predicting bankruptcy.

Specifically, to proxy for size the logarithm of total assets to price-level index is used, whereas for liquidity the current liabilities to current assets ratio is utilised. For leverage; total liabilities to total assets, working capital to total assets for the same reason as Altman. For profitability measurement net income to total asset (ROA), OENEG which is equal to one if total liabilities exceeds total assets and zero otherwise to proxy for financial risk, CHIN which is equal to $(NI_t - |NI_t|)/(|NI_t| + |NI_t - 1|)$ used for past income performance as well as INTWO which is one if Net Income (NI) is negative and zero otherwise, and finally FUTL which is the fund provided by operations divided by total liabilities. Matousek & Stewart (2009) chose equity to total assets ratio used to proxy for measuring overall risk. For liquidity measurement they used liquid assets to total assets. To proxy for the size of the bank, the natural logarithm of total assets is used.

Furthermore for performance measurement they used the NIM (Net Interest Margin), Net Operating Income (NOI) normalised for total assets, and operating expenses to total operating income are being used.

Research and Methodology

Part one explains the logic of the model chosen, part two explains data availability and part three discusses the variables chosen as explanatory variables of the model.

Logit versus probit

In a regression, one needs to assess $E(Y_i | X_{1i}, X_{2i}, \dots, X_{ki})$ where Y is a dichotomous (a binary) response variable and X 's are regressors. There are three ways to approach this: the Linear Probability Model (LPM), Logit and Probit. A simple regression model is composed of a response variable (Y_i), explanatory variable (X_i), disturbances (U_i), the coefficient (β_2), and an intercept (β_1): $Y_i = \beta_1 + (\beta_2 \times X_i) + U_i$. An LPM model is a simple linear regression where the response variable is dichotomous. LPM is widely used because of its simplicity (West, 2000) but there are some problems with this model: non normality of disturbances, heteroscedastic variances of the disturbances, questionable value of R^2 as a measure of goodness of fit, and non-fulfilment of $0 \leq E(Y_i | X) \leq 1$.

The non-normality of disturbances is assumed in the Ordinary Least Square (OLS) model but in LPM it is not tenable to assume so. The reason behind this is that disturbances like the response variable are

dichotomous and hence they follow the Bernoulli distribution rather than normal distribution. To see this more clearly, when $Y_i = 1$, $U_i = 1 - \beta_1 - (\beta_2 \times X_i)$ and when $Y_i = 0$, $U_i = -\beta_1 - (\beta_2 \times X_i)$. But this may not be the worst problem with LPM since when the objective is point estimation, normality is not a problem. Also with large sample size, LPM will follow OLS under the normality assumption. The other problem is the heteroscedastic variances of the disturbances. This is because $var(U_i) = P_i \times (1 - P_i)$. And since the P_i depends on X_i , and var_{U_i} depends on P_i , then var_{U_i} depends on X_i . This problem can affect the efficiency of the model since the variance is not at minimum. But this is not a serious problem too since it can be solved using Weighted Least Squares (WLS) and divide the regression equation by it. In the literature, WLS is equal to: $WLS = \sqrt{P_i \times (1 - P_i)}$. Since P_i is unknown, we should calculate \hat{Y}_i which is the estimated of $E(Y_i | X)$ using an OLS approach and use that to obtain \hat{Y}_i .

Another issue is that R^2 as a measure of the goodness of fit is much lower than one in dichotomous response models. This is because Y values either lie alongside the X axis or along the line corresponding to one (as shown in figure 1). Therefore no LPM can fit such a scatter well which leads to question R^2 as a measure of the goodness of fitness.

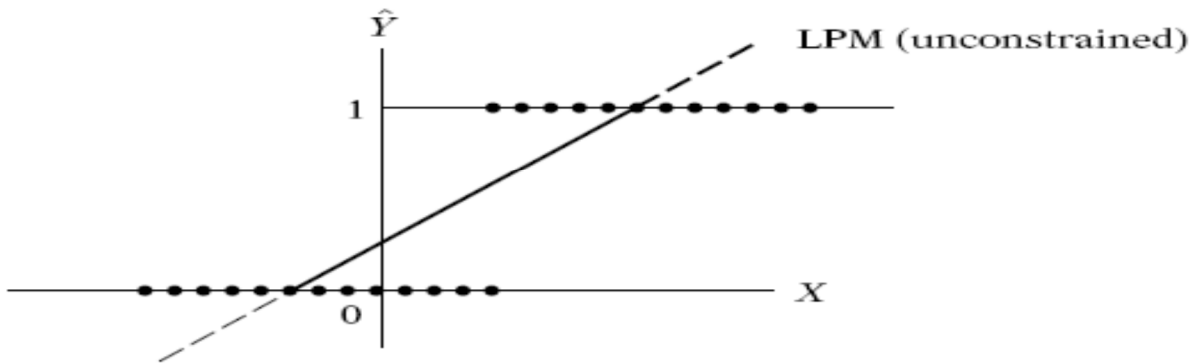
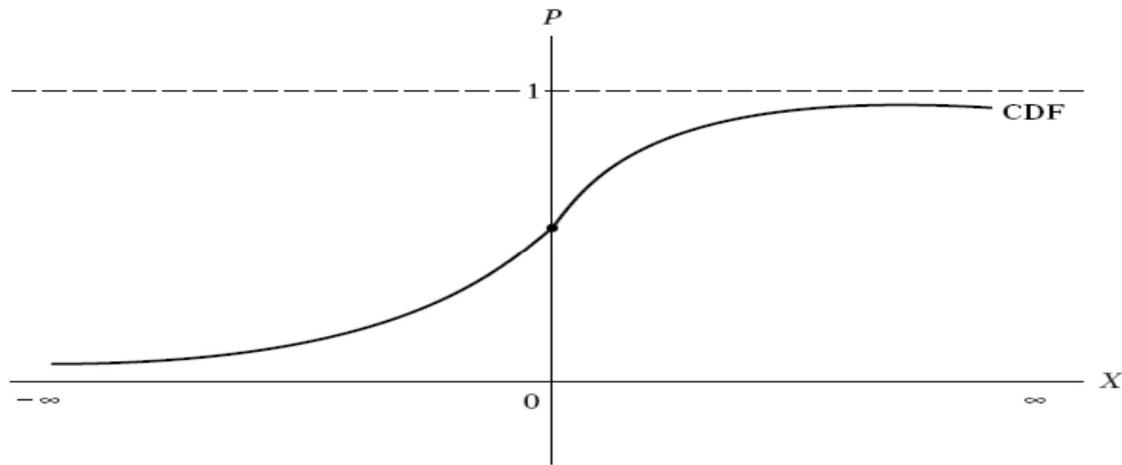


Figure 1: Unconstrained LPM Goodness of Fitness

The other problem is that sometimes an LPM approach does not give us results between 0 and 1. This is a problem since in the end LPM is used to estimate probabilities. Practitioners take two approaches to solve this real problem. They either use OLS to estimate \hat{Y}_i and see if its value is less than zero, they convert it to zero and if it is more than one, they put one instead or they use logit or probit models to have a guarantee that the estimated probabilities lie between zero and one. This is the main reason why logit and probit models were being developed. LPM assumes that P_i and X 's have a linear relationship which makes this model unattractive. An attractive model is to produce predictions between zero and one as well as be non linear. That means to get to zero and one at a slower rate, just like the Cumulative Distribution Function curve (CDF) shown in figure 2.



A cumulative distribution function (CDF).

Figure 2: Cumulative Distribution Function

The figure above is the starting point of logit and probit models. This non linearity comes from a difference in $E(Y_i | X)$ when compared to LPM: $P_i = E(Y_i | X) = \frac{e^z}{1+e^z}$ where z is the logistic distribution function and is equal to $Z_i = \beta_1 + \beta_2 \times X_i$. However since the nonlinearity is between P_i and X_i as well as β 's, OLS procedure cannot be carried on. Therefore it should be liberalised. As a result the odds ratio is being introduced and set equal to: $\frac{P_i}{1-P_i}$ and now the logit L can be written as:

$L_i = \ln\left(\frac{P_i}{1-P_i}\right) = Z_i = \beta_1 + \beta_2 \times X_i$. This way L is linear with X but the actual probabilities are not. If for a dichotomous variable where the probability is either zero or one, the logit cannot be estimated since $\ln\left(\frac{1}{0}\right)$ or $\ln\left(\frac{0}{1}\right)$ are meaningless. This is where Maximum Likelihood methods (ML) are being used.

In logit a general CDF is assumed and everything is being built up using that. Probit (or often called normit) method assumes that the CDF is following a normal distribution pattern. The same concepts explained in the logit model are then applied to the probit model. The choice between these three methods is then easy to make. The LPM model is not being chosen in this research because of its major flaws in the model explained earlier. So the choice is down to either logit or probit. The fact is that these two models produce very similar results. Figure 3 shows the outcome of logit versus the probit model:

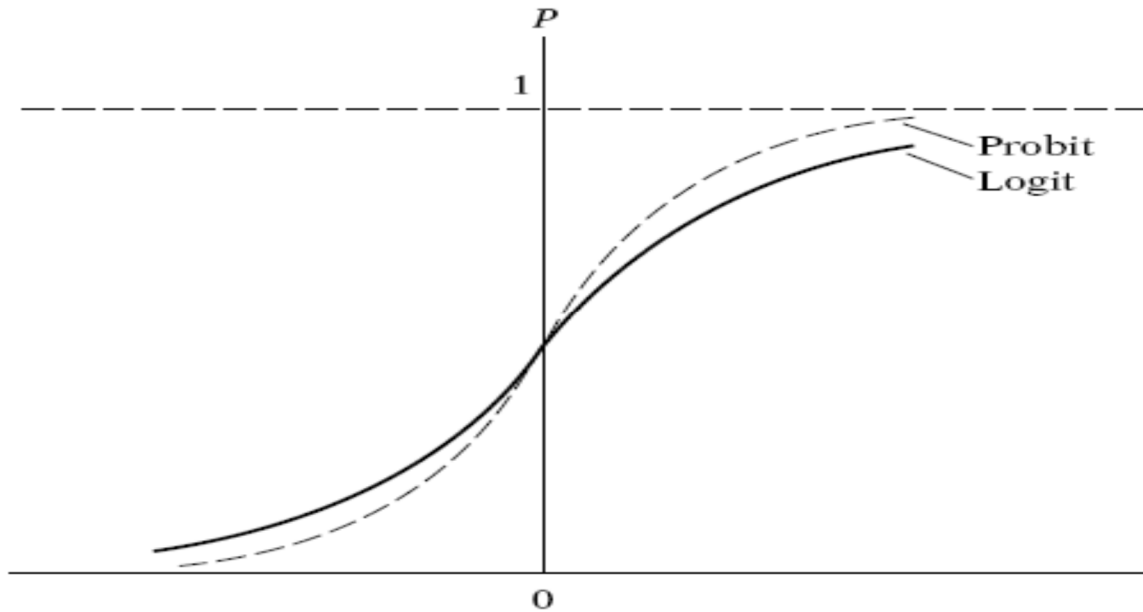


Figure 3: Logit Versus Probit, Gujarati

The only reason why logit is being regularly chosen between the two is its mathematical simplicity. However both logit and probit models are being used in this research so that Figure 3 is also tested empirically. We utilize ML method, estimates coefficients for the two regression models.

The likelihood function, assuming $f_i(Y_i)$ is the probability that Y_i is one or zero, is given by: $f(Y_1, Y_2, \dots, Y_n) = \prod_1^n p_i^{Y_i} (1 - p_i)^{1-Y_i}$. For this equation to be easily used, a log likelihood function is used which takes a natural logarithm of the both sides of the likelihood equation. It is therefore written as: $LLF = \sum_1^n \left[Y_i \ln \left(\frac{p_i}{1-p_i} \right) \right] + \sum_1^n \ln(1 - p_i)$. Using the odds ratio definition and replacing $\ln \frac{p_i}{1-p_i}$ by $\beta_1 + \beta_2 \times X_i$, the LLF is converted into a function of β_1 and β_2 and X 's are known. Then the ML method partially differentiates the unknowns and sets them equal to zero to find its maximum.

Variables Description

Variables used by Ohlson in his O-score model are thoroughly explained in part two. Since he was the first to implement logit regression in forecasting future and others built their models based on his study, his work is very important in credit scoring. Therefore the same variables are used in this research to compare his results to that of the Case Bank. The seven variables chosen for Case Bank are therefore listed in the Table 1 below:

Table 1: Variables' Description

Variable Symbol	Variable Description	Purpose
CL_CA	Current Liability to Current Asset	To Proxy For Liquidity
INTWO	$\forall NI > 0$ INTWO = 1, 0 Otherwise	To Proxy For Financial Performance
OENEG	If Total Liabilities > Total Assets = 1, 0 otherwise	To Proxy For Financial Risk
ROA	Return on Asset	To Proxy For Profitability
SIZE	$\ln(\text{Total Assets})$	To Proxy For The Size of The Firm
TL_TA	Total Liabilities to Total Assets	To Proxy For Leverage
WCAP_TA	Working Capital to Total Assets	To Proxy For Liquidity

CHIN and FUTL are the two variables Ohlson used but there were no data available to make them. Therefore they are inevitably excluded from this research. The next section therefore, is dedicated to expand on availability of data.

Data

Before a regression test is being carried on and the results are explained, it is vital to know more about the data being used. Representing data characteristics and statistics can be looked at from different perspectives. Therefore *Available Data* explains what is available and used in this research, *Data Statistics* reports on some descriptive statistics about data and *Co-linearity Between Variables* explains the co-linearity of variables used in this regression analysis. The direct implications of these statistics on the regression models are explained in part 5, the conclusion.

Available Data

In order to be able to carry on with the regression analysis, first a number of variables are selected. To choose the variables, first a summary of what is in hand in terms of data seems to be necessary. Data provided by Bank A were a mixture of consolidated and nonconsolidated accounting information on companies who were credit customers of that bank. There were a total of 726 customers out of which only 35% were consolidated accounting information. However, two of the 726 customers were being dropped out of the sample because of bad values. Data available for those customers go back to 2002 until 2012 out of which only 25.97% defaulted. A big number of financial information is extracted for each customer which is included all those variables used in this research.

However, what is more important in regression analysis is what the statistics show about the available data and ratios used in the equations. This information is included in the following section.

Data Statistics

Regarding the collected data, a descriptive statistics table on data is presented in Table 2.

Table 2: Descriptive Statistics for Variables

	CL_CA	INTWO	OENEG	ROA	SIZE	TL_TA	WCAP_TA
Mean	1.140454	0.121547	0.093923	0.060593	11.07925	0.823046	0.074199
Median	0.926210	0.000000	0.000000	0.041306	11.04001	0.859655	0.060972
Maximum	52.84785	1.000000	1.000000	1.320388	18.50021	5.000000	0.988954
Minimum	0.000577	0.000000	0.000000	-2.044118	4.418841	0.014838	-4.029412
Std. Dev.	2.286650	0.326988	0.291923	0.151324	2.244174	0.292469	0.376364
Skewness	17.39147	2.316385	2.784011	-3.231247	0.013785	4.722101	-2.222505
Kurtosis	372.1053	6.365638	8.750717	66.76132	3.114212	66.10610	24.75899
Jarque-Bera	4146364.	989.1670	1932.888	123902.6	0.416437	122825.8	14878.56
Probability	0.000000	0.000000	0.000000	0.000000	0.812030	0.000000	0.000000
Sum	825.6889	88.00000	68.00000	43.86909	8021.376	595.8855	53.72015
Sum Sq. Dev.	3780.399	77.30387	61.61326	16.55590	3641.257	61.84424	102.4129
Observations	724	724	724	724	724	724	724

Each of these data are represented using statistical distributions, with mean, median, maximum, and minimum to illustrate the range of data for each variable and to gain a better understanding of what values on average each variable may take. But perhaps the most important attribute to see here is the standard deviation. This shows how dispersed data are around their mean. The current ratio is the most widely scattered among all with the *SIZE* being next. On the other hand, the *ROA* is the least dispersed among others followed by *OENEG*. Skewness is the other attribute shown above which shows how asymmetrical the distribution is around its mean. The current ratio has the highest positive skewness which speaks for

having extreme values above its mean and ROA has the highest negative skewness which indicates having outliers below the mean value.

A perfect symmetrical distribution has zero skewness. Another attribute is the Kurtosis values shown above which indicates how flat or sharp is the distribution compared with the normal distribution which has Kurtosis equal to three. The highest among these attributes is current ratio which means this distribution has flatter tails and more extreme values. On the other hand *SIZE* has the smallest Kurtosis value which shows that this distribution has flatter middles and less extreme values. It is also important to mention the descriptive statistics of the residuals of the logit and probit regressions.

This is mainly because of the assumptions made in regression models (e.g. the linear probability model) that the disturbances are normally distributed because of statistical inference explained more in section 4.1.3. Residuals are the difference between the real observations and the estimated ones. Table 3 reports the descriptive analysis of the residuals after regression is applied. To test for their normality, the Jarque-Bera (JB) test is carried on.

Table 3: Descriptive Statistics for Variables

	Logit	Probit
Mean	3.82E-12	0.000277
Median	-0.231195	-0.230461
Maximum	0.853123	0.859304
Minimum	-0.739636	-0.738449
Std. Dev.	0.435524	0.435460
Skewness	1.066346	1.064882
Kurtosis	2.230596	2.230043
Jarque-Bera	155.0674	154.7167
Probability	0.000000	0.000000
Sum	2.77E-09	0.200758
Sum Sq. Dev.	137.1394	137.0992
Observations	724	724

JB is a large sample test based on the OLS residuals. JB test uses the sample size (n), the skewness (S) and the kurtosis (K) of the residuals and is computed as following: $JB = n \left[\frac{S^2}{6} + \frac{(K-3)^2}{24} \right]$. The test uses the null hypothesis that the residual are normally distributed. It also uses the chi-square distribution with two degrees of freedom and a computed p values appear in Table 3. Therefore, since the computed p values for JB test is exceptionally low, the null hypothesis can be rejected that the residuals are normally distributed. The histograms of the residuals are included in Appendix 2.

Co-linearity Between Variables

One of the important factors in choosing what explanatory variables to use is to choose those variables with minimum co-linearity among them (ideally no co-linear variables produce the best conclusive results). X_2 and X_3 , in a two variable regression model are considered collinear if the relation $\lambda_2 X_{2i} + \lambda_3 X_{3i} = 0$ given that λ_2 and $\lambda_3 \neq 0$. If any two variables are perfectly correlated to each other, the individual effect of each one of the variables on the dependant variable is not identifiable. So if we have a two variable regression model like the example above, if X_2 and X_3 are perfectly correlated to each other, the regression essentially has only one variable instead of two. That is why ideally, explanatory variables should be perfectly uncorrelated to each other to produce the best model.

Our analysis provides the partial correlation between any two variables in a matrix with the diagonal equal to all ones. This matrix is shown in Table 4

Table 4: Co-linearity Between Variables.

	CL_CA	INTWO	OENEG	ROA	TL_TA	WCAP_TA	SIZE
CL_CA	1.000000	0.071576	0.045451	-0.104965	0.100274	-0.350273	0.011123
INTWO	0.071576	1.000000	0.430851	-0.441743	0.278605	-0.242405	-0.089868
OENEG	0.045451	0.430851	1.000000	-0.316249	0.464109	-0.295238	-0.224595
ROA	-0.104965	-0.441743	-0.316249	1.000000	-0.623970	0.414925	0.113819
TL_TA	0.100274	0.278605	0.464109	-0.623970	1.000000	-0.603356	-0.266814
WCAP_TA	-0.350273	-0.242405	-0.295238	0.414925	-0.603356	1.000000	0.044054
SIZE	0.011123	-0.089868	-0.224595	0.113819	-0.266814	0.044054	1.000000

The highest correlation among variables, as shown above, is between *ROA* and the *TL_TA*. Knowing that the perfect correlation is indicated by one, variables used in this research are fairly uncorrelated to each other. The negative sign indicates that the rise in one variable causes the fall in the other. The least correlation is between the current ratio and the size.

Analysis and Results

This section represents the logit and probit analysis of the data into two main sections. This section investigates results on each specific predicted variable's coefficient while all others remain the same and following section includes results on the overall model concerning all coefficients together.

Regression Results - The Coefficients

The regression results produced are shown in Appendix I. 724 observations were regressed using Maximum Likelihood (ML) logit and probit prediction model. To start with regression, initial values are given by the software. These values are denoted by $C(X)$, $x=0..8$. Therefore $C(1)$ is the intercept, $C(2)$ is the *CL_CA* ratio, $C(3)$ is *INTWO* and so on. In the next part of the table, variable specific output is shown. Following parts include (i) presentation and implications of estimated coefficients, (ii) reports the standard error of coefficients, and (iii) contains inference issues and significance of predicted coefficients. Finally a comparison of the results to that of Ohlson's is included.

The Estimated Coefficients

Regression analysis calculates the multipliers of each input to the system. Future predictions of if a customer is likely to default or not is made using the calculated coefficients of the regression analysis. The estimated coefficients for logit regression are reported in Table 5. These coefficients are slightly different in probit model and are represented in Table 6.

Table 5: Logit Estimated Coefficients

Variable	Coefficient
C	-1.190285
CL_CA	-0.029693
INTWO	0.342383
OENEG	-0.109633
ROA	1.316707
SIZE	-0.065823
TL_TA	0.935493
WCAP_TA	0.081307

Table 6: Probit Estimated Coefficients

Variable	Coefficient
C	-0.725657
CL_CA	-0.018286

INTWO	0.203919
OENEG	-0.067942
ROA	0.790104
SIZE	-0.040052
TL_TA	0.566473
WCAP_TA	0.042740

Coefficients are the regressed β 's calculated using logit regression. To interpret these coefficients take *INTWO* of logit regression as an example. This shows that for a unit increase in *INTWO*, the weighted log of odds ($\ln(\frac{p}{1-p})$) will go up by the coefficient (i.e. by 0.342383) provided all other variables remain the same. The negative correlation indicates that for explanatory variables like the current ratio, a unit increase will decrease the log of odds by the coefficient associated to that variable (i.e. by 0.029693).

Standard Error of the Estimated Coefficients

The standard errors of the estimated coefficients are important when the precision of data is being analysed. The standard error column in the output serves for the same purpose. The real question is to what extent one can be confident in estimators to be accurate or reliable. Are they going to change going from one sample to the other sample given a population? If the standard errors are small, then the coefficients are likely to be precise. Among these, *ROA* has the highest standard error for both logit and probit and hence the highest degree of uncertainty in estimated values for this variable. On the other hand, *SIZE* has the smallest standard error both in logit and probit making its coefficient the most precise among others. All the results in terms of standard error are shown in the Table 7 below:

Table 7: The Standard Error of Logit and Probit Regressions

Variable	SE (logit)	SE (probit)
C	0.717991	0.423175
CL_CA	0.068252	0.038804
INTWO	0.304574	0.181824
OENEG	0.344316	0.204740
ROA	0.760941	0.457451
SIZE	0.040805	0.024252
TL_TA	0.485581	0.283696
WCAP_TA	0.323694	0.191239

The standard errors are used to test the inference null hypothesis which is explained in the following section.

Inference

If the purpose of the research was point estimation only, since the OLS model does not make assumptions on the normality of disturbances will suffice. But in order to both estimate and consider inference, an assumption is made about disturbances following a normal distribution. To have a judgment on population parameters in a random process the related hypothesis is tested.

The null hypothesis is set to $H_0: E \left(\hat{\beta}_i \right) = 0$ with the alternative hypothesis set to $H_1: E \left(\hat{\beta}_i \right) \neq 0$.

Perhaps the most important implication of testing the null hypothesis is to understand the significance of each individual predicted coefficient in affecting the probability of default while all others remain constant. To do that, a test of null hypothesis is carried on with a set nominal significance level. Based on a common practice in Iran, this level is set to 10%. A t-ratio test is carried on for this purpose and is defined as:

$$t - stat = \frac{\hat{\beta} - E(\hat{\beta})}{SE(\hat{\beta})}. \text{ Therefore the t-stat ratio will reduce to: } t - stat = \frac{\hat{\beta} - 0}{SE(\hat{\beta})} \text{ assuming that } E(\hat{\beta})$$

equals zero. If the t calculated is greater than a critical value ($t_{\alpha/2}$) or less than a negative critical value ($-t_{\alpha/2}$), then the null hypothesis is rejected and that coefficient is considered to be statistically significance. Otherwise H_0 cannot be rejected and the corresponding coefficient is not statistically significant. Having 724 number of observations, and since it is greater than 300, based on t-student distribution, the degree of freedom denoted by ϑ is considered to be ∞ . Hence, for corresponding 90% significance interval, $t_{1-0.90/2} = t_{0.05}$ is equal to 1.6449 (retrieved from critical values of Student's t-distribution for different probability levels table). To find out which coefficient is significant a series of calculations are done in tables 8 and 9 for the logit and the probit models respectively:

Table 8: The Significance of the Estimated Coefficients for Logit Regression

Variable	Coefficient	SE	T-Ratio	Compared With	Is It Significant?
C	-1.190285	0.717991	-1.657799	-1.6449	<input checked="" type="checkbox"/>
CL_CA	-0.029693	0.068252	-0.435048	-1.6449	<input type="checkbox"/>
INTWO	0.342383	0.304574	1.124137	1.6449	<input type="checkbox"/>
OENEG	-0.109633	0.344316	-0.318410	-1.6449	<input type="checkbox"/>
ROA	1.316707	0.760941	1.730366	1.6449	<input checked="" type="checkbox"/>
SIZE	-0.065823	0.040805	-1.613089	-1.6449	<input type="checkbox"/>
TL_TA	0.935493	0.485581	1.926545	1.6449	<input checked="" type="checkbox"/>
WCAP_TA	0.081307	0.323694	0.251184	1.6449	<input type="checkbox"/>

Table 9: The Significance of the Estimated Coefficients for Probit Regression

Variable	Coefficient	SE	T-Ratio	Compared With	Is It Significant?
C	-0.725657	0.423175	-1.714791	-1.6449	<input checked="" type="checkbox"/>
CL_CA	-0.018286	0.038804	-0.471227	-1.6449	<input type="checkbox"/>
INTWO	0.203919	0.181824	1.121518	1.6449	<input type="checkbox"/>
OENEG	-0.067942	0.204740	-0.331847	-1.6449	<input type="checkbox"/>
ROA	0.790104	0.457451	1.727187	1.6449	<input checked="" type="checkbox"/>
SIZE	-0.040052	0.024252	-1.651484	-1.6449	<input checked="" type="checkbox"/>
TL_TA	0.566473	0.283696	1.996758	1.6449	<input checked="" type="checkbox"/>
WCAP_TA	0.042740	0.191239	0.223491	1.6449	<input type="checkbox"/>

Comparing the results for logit and probit, it is easy to note that in probit one more coefficient is statistically important which is the *SIZE*.

Comparison of Results to O-Score Model:

Since the model is using Ohlson variables, it is now possible to compare this result to his. In his work he found that the size of the company, the measure of financial structure, the measure of performance, and the measure of current liquidity are significant (Ohlson, 1980). However, based on his variables, only the financial performance (*ROA*) and the financial structure (*TL_TA*) were significant. He also predicted the coefficients' signs which show if the relationship with the explained variable is direct or indirect. He predicted using common sense and literature review that *TL_TA*, *CL_CA*, and *INTWO* should be positively correlated while *SIZE*, *WCAP_TA*, and *ROA* are negatively correlated. He also left *OENEG* indeterminate. A summary of Olson's results and our results are shown in 10. Olson made three distinctive assumptions to build his o-score model. Therefore, three results are produced: one for each assumption. The first model predicts the bankruptcy (which was his dependant variable) within one year. The second model predicts

bankruptcy in two years, given that the company did not fail within the subsequent year. The last model (model 3), predicts the bankruptcy within one or two years (Ohlson, 1980).

Table 10: Comparison of Results

Variable	Expected	Ohlson			Case Bank (logit)	Case Bank (probit)
		Model I	Model II	Model III		
C		-	+	+	-	-
CL_CA	+	+	-	+	-	-
INTWO	+	+	-	-	+	+
OENEG		-	-	-	-	-
ROA	-	-	-	-	+	+
SIZE	-	-	-	-	-	-
TL_TA	+	+	+	+	+	+
WCAP_TA	-	-	-	-	+	+

The first model, which predicts bankruptcy within one year seems to be a better model compared to the other two models in terms of goodness-of-fit statistics. This is why the base of comparison between Ohlson and our Case Bank is his first model. The ratios *CL_CA*, *ROA*, and *WCAP_TA* are different between the results when data are regressed both by logit and by probit models.

Regression Results – The Scoring Model

In the previous section, each individual coefficient is being considered provided all other coefficients remain constant. In this section a number of various discussions are being carried on for the whole credit scoring system considering all variables together. In this section, the significance of the model is discussed, the goodness-of-fit of the model is explained and finally, the power of the model is investigated.

The Significance of the Overall Model

To understand the overall significance of the estimated parameters, Likelihood Ratio (LR) and inference (e.g. hypothesis testing) are being looked at. The LR follows the Chi square (χ^2) distribution. The χ^2 table shows that for a degree of freedom greater than 100, $\forall k > 100, \sqrt{2\chi^2} - \sqrt{2k-1} = Z$ converges to the standard normal distribution. Therefore, Z distribution is tested to measure the significance of the whole model. So the critical value is equal to $Z = 1.2816$. This number should be compared with the LR statistic to indicate the significance of the model.

If LR is greater than the critical value, then the null hypothesis explained in section 4.1.3 is rejected and therefore the model is statistically significant otherwise the null hypothesis cannot be rejected and the model is statistically insignificant. The LR ratio of the logit model is 11.76143 and is greater than the critical value. Therefore this model is statistically significant. This value is 12.07791 for the probit regression which is a better result compared to that of logit's.

The Goodness-of-Fitness Tests

The following two sections explain the McFadden R^2 and the H-L models for the goodness-of-fitness tests. The goodness-of-fitness is a measure of how well the regression line includes discrete observations of the model. It tries to explain to what extent the model can estimate the dependant variable given the estimated coefficients and the explanatory variables.

The Goodness-of-Fitness of the Model – McFadden R^2

This analysis is based on Maximum Likelihood (ML) for estimation. As also explained in section 3.1, the objective is not to minimise the Residual Sum of Squares (RSS) but to maximise the Log Likelihood Function (LLF). The problem with R^2 is the interpretation of the results. As an example, a dichotomous

variable like the Probability of Default (PD) variable we are considering is in the end converted from a continuous set of numbers between [0,1]. Those who are closer to one (normally $\hat{P}_i \geq 0.5$) are being granted a loan and those who are close to zero are not. But when the observations of historical PD is not evenly divided between PD = 0 and PD = 1, then the number 0.5 will change accordingly and one may have to score 0.8 and above to be considered a PD = 1. In general, if less observations are a non-default (i.e. PD = 0), the model may have to accept a greater score compared to default (i.e. PD = 1). Therefore a measure of pseudo- R^2 is being introduced and is defined as: $Pseudo - R^2 = 1 - \frac{LLF}{LLF_0}$ where LLF_0 is the ML of the restricted model where the model is considered to have only an intercept, i.e. the slope parameters are set equal to zero and the LLF is the usual maximised value of the log likelihood function of the logit model.

Because the likelihood function, which is a joint probability and its value is between zero and one, the LLF is a negative number. The improved R^2 model will cause LLF become less negative which leads the pseudo R^2 to rise. The pseudo R^2 is being referred to as McFadden R^2 and is an improved model for ordinary goodness-of-fitness measure R^2 . Case Bank's logit model has the McFadden R^2 value of 0.014183 which means only 1.4% of the variations in PD is explained by this model (i.e. almost none of the movements of the explanatory variables attribute to the movements of the PD). This value is not much different when the model is built on the probit approach and is equal to 0.014564. However, in our dataset, 74.03% of observations are non-default customers which makes the sample less biased towards defaulted customers. Also disturbances are not normally distributed as discussed in Data Statistics. Therefore, the R^2 model is not the best method for measuring the goodness-of-fitness of the model.

The Goodness-of-Fitness of the Model – Hosmer-Lemeshow $H - L$

This model is based on Hosmer-Lemeshow (1980) and is a measure of goodness-of-fitness of models. The idea is to compare the fitted expected values to the actual values in randomised groups of a given size (the default size is 10). The test is carried on the basis of the predicted probability that PD = 1. Obviously, the bigger this number is, the worse a model fits data. Andrew test is another measure for the goodness-of-fitness which is a more general form of Hosmer-Lemeshow's model. This model will take selected series (independent variables) to take into consideration. A proposed consideration of this model is when one set of group shows the model fits while the other set of group shows exactly the opposite.

This can happen when subjects are widely different in values of the covariates (Hosmer, Hosmer, Cessie, and Lemeshow, 1997). The .quantile of risk. is defined as the high and low value of the depicted probability for each decile. Then for each regressand's outcome (PD = 1 or PD = 0) the actual value is compared to the expected value. The total number of observations is reported in each decile which sums up to 724. Finally the H-L value is the correspondent Hosmer-Lemeshow values. These numbers are shown for both the logit and the probit models in Table 11 and Table 12 respectively:

Table 11: Logit Model

Group Number	Quantile of Risk		Regressand=0		Regressand=1		Total Observations	H-L Value
	Low	High	Actual	Expected	Actual	Expected		
1	0.0458	0.1988	64	59.3812	8	12.6188	72	2.04989
2	0.1991	0.2179	56	56.9131	16	15.0869	72	0.06991
3	0.2180	0.2313	55	56.5822	18	16.4178	73	0.19673
4	0.2314	0.2437	56	54.8684	16	17.1316	72	0.09808
5	0.2438	0.2574	53	54.6773	20	18.3227	73	0.20499
6	0.2574	0.2691	58	53.0375	14	18.9625	72	1.76299
7	0.2692	0.2814	44	52.1791	28	19.8209	72	4.65719
8	0.2815	0.2959	51	51.9281	22	21.0719	73	0.05746
9	0.2960	0.3180	49	49.9874	23	22.0126	72	0.06380
10	0.3193	0.7396	50	46.4457	23	26.5543	73	0.74776

In table 11, the biggest H-L value corresponds to the 7th decile (H-L = 4.65719) and the lowest to the 8th decile (H-L = 0.05746). The total number of H-L values is 9.90880 with the probability of Chi-square value with eight degree of freedom equal to 0.2715. This means that according to this test, the model is statistically significant at 27.15% which is more than the 10% significance level. However if the same analysis is carried out by the probit model, the Chi-square's probability with eight degree of freedom reduces down to 0.1151 which is much closer to that of logit. A detailed Hosmer-Lemeshow test on Bank A's probit model is shown below:

Table 12: Probit Model

Group Number	Quantile of Risk		Regressand=0		Regressand=1		Total Obrv.	-L Value
	Low	High	Actual	Expected	Actual	Expected		
1	0.0315	0.1973	64	59.6475	8	12.3525	72	1.851
2	0.1977	0.2170	56	57.0069	1	14.9931	72	0.085
3	0.2172	0.2305	55	56.6272	18	16.3728	73	0.208
4	0.2308	0.2440	56	54.8851	16	17.1149	72	0.095
5	0.2440	0.2571	52	54.6668	21	18.3332	73	0.518
6	0.2571	0.2697	60	53.0049	12	18.9951	72	3.499
7	0.2698	0.2823	43	52.1281	29	19.8719	72	5.791
8	0.2824	0.2969	51	51.8645	22	21.1355	73	0.050
9	0.2972	0.3186	49	49.9317	23	22.0683	72	0.057
10	0.3198	0.7384	50	46.4381	23	26.5619	73	0.751

The Prediction Evaluation

The method used is to prompt the user for a cut-off value as an input and then produces a 2 × 2 table of correct and incorrect classification based on that cut-off value. The first table classifies each observation based on having a prediction probability above or below the cut-off value. In this table the probability of $P(\text{dependant variable}=1) \leq \text{cut-off value}$ is plugged into the equation and its probability is estimated. Then the number of correct estimations where the estimated dependant variable = 0 for all actual probabilities \leq the specified cut-off value is recorded. This way the table classifies the correct and incorrect prediction percentages. The important thing to note here is to input the correct cut-off value. Since the proportion of the default to non-default customers observed is 25.97, then as expected the cut-off value produces its best estimates at this threshold.

The Figure 4 below shows the movement of the correct percentages for when the dependant variable is estimated zero and one as opposed to increase in cut-off values.

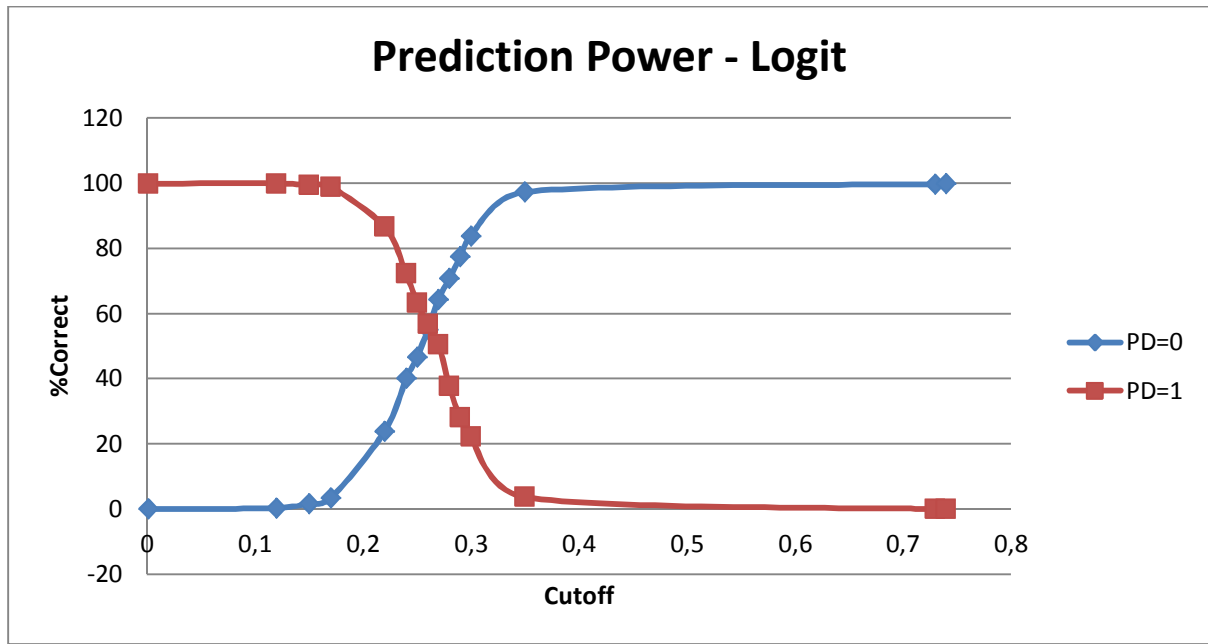


Figure 4: Correct Predictions versus the Cut off, Logit Model

In the probit model, this chart also has a threshold of 25.97. The Figure 5 shows how the model predicts for PD = 0 and PD = 1.

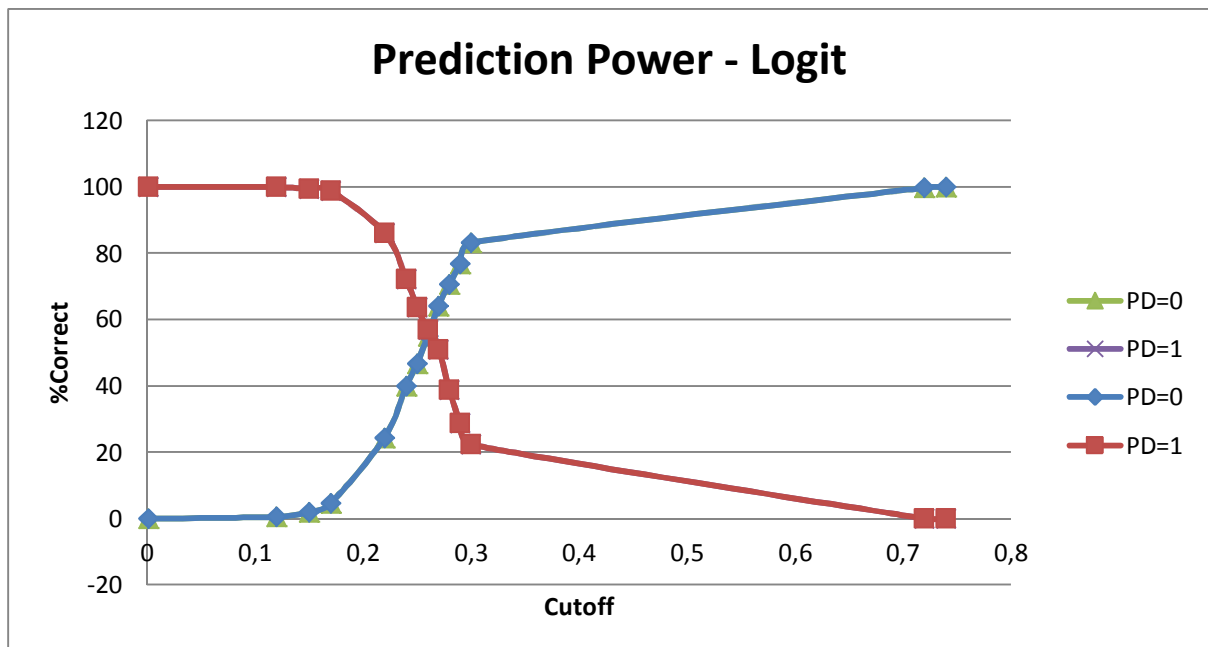


Figure 5: Correct Predictions versus the Cut off, Probit Model

In that threshold both logit and probit scored 56.91% correct estimation for $PD = 1$. However, logit slightly outperformed probit when $PD = 0$. While probit achieved a 54.48% correct estimation, logit did achieve a similar 54.85%.

Conclusion

In the light of the 2008 financial crisis, some bankers may argue if the risk management practices and theories behind them are any good. Some may even question the abilities of the regulatory authorities in capturing risky banks and lost their faith in organizations like the Financial Services Authorities (FSA). In the end, Iran with its current risk management practices was being hit by this crisis far less than countries like the United Kingdom or the United States. So why would a bank want to implement what has being fatally failed in the Europe and the USA?

The answer to this question is related to investigate why those banks have failed. Was that because models were predicting wrong numbers or was that because of banks being greedy, not taking risk management practices seriously, or if they were being over reliant on what their systems were predicting. The answer to this question is not the purpose of this research and is out of its scope but implementing credit scoring systems is an inevitable part of growing banking industry in Iran. Of course there is a human judgment attached to all risk management systems and being over reliant on just the output numbers without having a correct understanding of them can cause serious damage, so bad that some bankers may call switching to credit scoring models a mistake. The established results sends missing signals when being compared to the Ohlson's. Only two of Ohlson's variables were statistically significant, *ROA* and *TL_TA*, using the same approach Ohlson took in his research (i.e. logit regression). *WCAP_TA* which is a measure of liquidity is positively correlated to the probability of default. That means if we can fully trust the quality of data, customers of Case Bank will score a higher probability of default value if they tie up their liquidity to their day to day business. Also *ROA*, which is the Net Income (NI) divided by total assets, is positively correlated to the probability of default. That means, surprisingly, the more NI a company earns, the more its probability of default is. Therefore Bank A should be more careful about following up their bad loans since customers do not pay back loans using their income. The *CL_CA* figure is also negatively correlated to the probability of default. It can be explained by the fact that customers are not fully utilising their Weighted Average Cost of Capital (WACC) to benefit from leverage's cost reduction benefits. That is why customers with a higher level of current liability (normalised for current asset) are less likely to default.

Another issue to think about with this model relates to the data set in hand. Based on available data, since there are far less observations for defaulted loans, the model can be biased towards customers' probability of default. Therefore the prediction power of the model is not satisfying. Also the quality of data is in question. Only 35% of customers provided consolidated accounting figures. To look more deeply into data, 83.46% of customers are acting responsibly with regard to paying back their loans among consolidated accounting figures while this number reduces down to 69.01% among non-consolidated accounting figures.

There are also some aspects to be looked at when we compare the logit and the probit models. Based on Case Bank's data set, probit regression outperformed logit in two ways. First in probit approach, the *SIZE* variable is an extra addition to the number of statistically significant coefficients compared to logit regression coefficients. Therefore, Ohlson's variables work better for probit regression approach. According to Hosmer-Lemeshow goodness-of-fitness test, probit regression fits data up to the industry level confidence interval (10%) while logit regression fits the data for twice that confidence level. Therefore logit provides a better fit to data.

References

- Altman, E. I. (1968). Financial ratios, discriminant analysis and the prediction of corporate bankruptcy., *Journal of Finance*, vol. 23, pp.589–609.
- Basle Committee on Banking Supervision (1999), *Credit Risk Modelling: Current Practices and Applications*. Available from: www.bis.org/publ/bcbs49.htm [Accessed: August 18, 2010].
- Bass, R. M. V. (1991) *Credit Management*. UK: Stanley Thornes (Publishers) Ltd.

- Beaver, W.H. (1966). Financial Ratios as Predictors of Failure. *Empirical Research in Accounting*, pp. 71-111.
- Berger, A., Frame, W.S. & Miller, N.H. (2002). Credit Scoring and the Availability, Price, and Risk of Small Business Credit..
- Berger, Allen, N. & De Young, R. (1997). Problem Loans and Cost Efficiency in Commercial Banks., *Journal of Banking and Finance*, vol. 21, no., pp. 849-870.
- Etemadi, H., Rostamy, A.A.A. and Dehkordi, H. F. (2009), 'A genetic programming model for bankruptcy prediction: Empirical evidence from Iran', *Expert Systems with Applications*, Vol. 36, Issue 2, Part 2, March, pp. 3199-3207
- Fatemi, A. & Fooladi, I. (2006). Credit risk management: a survey of practices., *Managerial Finance*, vol. 32, no. 3, pp. 227-233.
- Federal Reserve System Task Force On Internal Credit Risk Models (1998), *Credit Risk Models at Major U.S. Banking Institutions: Current State of the Art and Implications for Assessments of Capital Adequacy*. Available from: <http://www.federalreserve.gov/boarddocs/press/General/1998/19980529/study.pdf> [Accessed: August 18, 2010].
- Financial Times (2008), *InfoViewer: Iran's banks struggle with credit shortages*. Available from: http://us.ft.com/ftgateway/superpage.ft?news_id=fto110420081809010198&page=2 [Accessed: August 8, 2010].
- Gorr, W.L., Nagin, D. & Szczypula, J. (1994). The Comparative Study of Artificial Neural Network and Statistical Models for Predicting Student Grade Point Averages ., *International Journal of Forecasting* , vol. 10, no. pp. 17-34.
- Greene, W., & N. (1998). Sample selection in credit-scoring models., *Japan and the World Economy*, vol. 10, pp. 299-316.
- Gujarati & (2004). Qualitative Response Regression Models.. In *Basic Econometrics*, 4th ed, The McGraw-Hill Companies, New York, pp.580-635
- Hand, D.J. & Henley, W.E. (1997). Statistical classification methods in consumer credit scoring: a review., *Journal of the Royal Statistical Society, Series A (statistics in Society)*, vol. 160, no. 3, pp. 523-541.
- Hand, D.J. (2001). Modeling consumer credit risk., *IMA Journal of Management Mathematics*, vol. 12, no. pp. 139-155.
- Hemple, G. H., Simonson, D. G. & Coleman, A. B. (1994) *Bank Management: Text and Cases*. New York, USA: John Wiley and Sons.
- Hlawatsch, S. & Reichling, P. (2010). A Framework for Loss Given Default Validation of Retail portfolios., *The Journal of Risk Model Validation*, vol. 4, no. 1, pp. 23-48.
- Hoch, S. J. (2001) *Combining models with intuition to improve decisions. Wharton on Making Decisions*. New York, USA: John Wiley and Sons, Inc.
- Hosmer, D.W. & Lemeshow, S. (1980). Goodness of Fit Tests for the Multiple Logistic Regression Model., *Communications in Statistics - Theory and Methods*, vol. 9, no. pp. 1043-1069.
- Hosmer, D.W., Hosmer, T., Cessie, S.L. & Lemeshow, S. (1997). A Comparison of Goodness-of-Fit Tests for the Logistic Regression Model., *Statistics in Medicine*, vol. 16, no. pp. 965-980
- Kalapodas, E. & Thomson, M.E. (2006). Credit risk assessment: a challenge for financial institutions., *IMA Journal of Management Mathematics*, vol. 17, no., pp. 25-46.

- Kim, Y.S. (2008). Comparison of The Decision Tree, Artificial Neural Network, and Linear Regression Methods Based on The Number and Types of Independent Variables and sample size., *Expert Systems with Applications*, vol. 34, no. pp. 1227–1234.
- Koch, T. W. & McDonald, S. S. (2000) *Bank Management*. Harcourt College Publishers.
- Lang, W.W. & Jagtiani, J.A. (2010). The Mortgage and Financial Crises: The Role of Credit Risk Management and Corporate Governance..
- Liby, R. (1975). Accounting Ratios and The Prediction of Failure: Some Behavioral Evidence., *Journal of Accounting Research*, vol. 13, no. 1, pp. 150-161.
- Lopez, J.A., Saidenberg, M.R. (1999). Evaluating Credit Risk Models., Jun 30.
- Makiyan, S. Z. (2003), 'Role of rate of return on loans in the Islamic banking system of Iran', *Managerial Finance*, Vol. 29, Issue 7
- Mashayekh, B. and Bazaz, M. S. Corporate (2008), 'Governance and Firm Performance in Iran', *Journal of Contemporary Accounting & Economics*, Vol. 4, Issue 2, pp. 156-172
- Matousek, R. & Stewart, C. (2009). A Note on Ratings of International Banks., *The Journal of Financial Regulation and Compliance*, vol. 17, no. 2, pp. 146-155.
- Ohlson, J.A. & (1980). Financial Ratios and the Probabilistic Prediction of Bankruptcy., *Journal of Accounting Research*, vol. 18, no. 1, pp. 109-131.
- Plous, S. (1993) *The Psychology of Judgement and Decision Making*. New York: McGraw-Hill
- Prasad, K. N, Banouei, A. A. and Swaminathan A. M. (1990), 'Economic consequences of non-optimal inventory holdings in India and Iran an integrated linear programming and input-output approach', *Engineering Costs and Production Economics*, Vol. 19, Issues 1-3, May, pp. 73-79
- PressTv (2009a), *In Iran, 90 people owe \$8 billion to banks*. Available from: www.presstv.ir/detail.aspx?id=104317§ionid=351020102 [Accessed: August 8, 2010].
- Presstv (2009b), *USD 38 billion delinquent loans to Iranian banks*. Available from: <http://www.presstv.com/detail.aspx?id=108239§ionid=351020102> [Accessed: August 8, 2010].
- Rice, T. & Coyle, B. (1999) *Framework for Credit Risk Management*. UK: Financial World Publishing.
- Sadr, K. (1999). The Role of Musharakah Financing in the Agricultural Bank of Iran., *Arab Law Quarterly*, vol. 14, no. 3, pp. 245-256.
- Sanjeeve, G.M. (2007). Bankers' Perceptions on Causes of Bad Loans in Banks., *Journal of Management Research*, vol. 7, no. 1, pp. 40-46.
- Saunders, A. & Cornett, M.M. (2003). Credit Risk: Individual Loan Risk.. In *Financial Institutions Management: A Risk Management Approach*, McGraw-Hill/Irwin, New York, pp. 259-302
- Shuhui, L., Wunsch, D.D., Hair, E.O. & Giesselmann, M.G. (2001). Comparative Analysis of Regression and Artificial Neural Network Models for Wind Turbine Power Curve Estimation ., *Journal of Solar Energy Engineering* , vol. 123, no. pp. 327–332.
- Subbanarasimha, P., Arinze, B. & Anadarajan, M. (2000). The Predictive Accuracy Of Artificial Neural Networks and Multiple Regression in The Case of Skewed Data: Exploration of Some Issues., *Expert Systems With Applications* , vol. 19, no. pp. 117–123.
- Taktak, N. B., Zouari, S. B. S., and Boudriga, A., (2010) 'Do Islamic banks use loan loss provisions to smooth their results?', *Journal of Islamic Accounting and Business Research*, Vol. 1, Issue 2
- Tasche, D. & (2006). Validation of internal rating systems and PD estimates..

Usmani, M.T. & (1999).The Concept of Musharakah and Its Application as an Islamic Method of Financing., *Arab Law Quarterly*, vol. 14, no. 3, pp. 203-220.

Valadkhani, A. and Nameni, M. (2011), 'How can Iran's black market exchange rate be managed?', *Journal of Economic Studies*, Vol. 38, Issue 2

Wang, Y. & Campbell, M. (2010).Financial Ratios and The Prediction of Bankruptcy: The Ohlson Model Applied To Chinese Publicly Traded Companies., *ASBBS*, vol. 17, no. 1, pp. 334-338.

Waring, A., Glendon, I.A. & (1998).Managing Risk., *International Thomson Business Press*.

Appendix:

The Following two tables show the sample output of running logit and probit regressions

Dependent Variable: PD				
Method: ML - Binary Logit (Quadratic hill climbing)				
Date: 08/23/10 Time: 10:48				
Sample: 1 724				
Included observations: 724				
Estimation settings: tol= 0.00010				
Initial Values: C(1)=-1.11252, C(2)=-0.01009, C(3)=0.26530, C(4)=-0.06473, C(5)=1.03409, C(6)=-0.04697, C(7)=0.67886, C(8)=0.06178				
Convergence achieved after 5 iterations				
Covariance matrix computed using second derivatives				
Variable	Coefficient	Std. Error	z-Statistic	Prob.
C	-1.190285	0.717991	-1.657799	0.0974
CL_CA	-0.029693	0.068252	-0.435048	0.6635
INTWO	0.342383	0.304574	1.124137	0.2610
OENEG	-0.109633	0.344316	-0.318410	0.7502
ROA	1.316707	0.760941	1.730366	0.0836
SIZE	-0.065823	0.040805	-1.613089	0.1067
TL_TA	0.935493	0.485581	1.926545	0.0540
WCAP_TA	0.081307	0.323694	0.251184	0.8017
McFadden R-squared	0.014183	Mean dependent var		0.259669
S.D. dependent var	0.438756	S.E. of regression		0.437648
Akaike info criterion	1.151274	Sum squared resid		137.1394
Schwarz criterion	1.201935	Log likelihood		-408.7613
Hannan-Quinn criter.	1.170827	Deviance		817.5225
Restr. deviance	829.2839	Restr. log likelihood		-414.6420
LR statistic	11.76143	Avg. log likelihood		-0.564587
Prob(LR statistic)	0.108682			
Obs with Dep=0	536	Total obs		724
Obs with Dep=1	188			

App I.1: Logit Regression Output

Dependent Variable: PD			
Method: ML - Binary Probit (Quadratic hill climbing)			
Date: 08/27/10 Time: 02:59			
Sample: 1 724			
Included observations: 724			
Estimation settings: tol= 0.00010			
Initial Values: C(1)=-0.69532, C(2)=-0.00630, C(3)=0.16581, C(4)= -0.04046, C(5)=0.64630, C(6)=-0.02935, C(7)=0.42429, C(8)=0.03861			
Convergence achieved after 4 iterations			
Covariance matrix computed using second derivatives			

App I.2: Probit Regression Output

Variable	Coefficient	Std. Error	z-Statistic	Prob.
C	-0.725657	0.423175	-1.714791	0.0864
CL_CA	-0.018286	0.038804	-0.471227	0.6375
INTWO	0.203919	0.181824	1.121518	0.2621
OENEG	-0.067942	0.204740	-0.331847	0.7400
ROA	0.790104	0.457451	1.727187	0.0841
SIZE	-0.040052	0.024252	-1.651484	0.0986
TL_TA	0.566473	0.283696	1.996758	0.0459
WCAP_TA	0.042740	0.191239	0.223491	0.8232
McFadden R-squared	0.014564	Mean dependent var		0.259669
S.D. dependent var	0.438756	S.E. of regression		0.437584
Akaike info criterion	1.150837	Sum squared resid		137.0992
Schwarz criterion	1.201498	Log likelihood		-408.6030
Hannan-Quinn criter.	1.170390	Deviance		817.2060
Restr. deviance	829.2839	Restr. log likelihood		-414.6420
LR statistic	12.07791	Avg. log likelihood		-0.564369
Prob(LR statistic)	0.098026			
Obs with Dep=0	536	Total obs		724
Obs with Dep=1	188			

The histogram of the residuals for both logit and probit models are shown. Jarque-Bera (JB) test implies that residuals are not normally distributed for Case Bank. In the following two histograms, this can be seen better for both logit and probit residuals respectively.