INVESTIGATING THE EFFECTIVENESS OF CREDIT SCORING MODELS IN ENHENCING CREDIT RISK MANAGENT OF MICROFINANCE INSTITUTIONS IN ZIMBABWE

WALTER SIBANDA*
WASHINGTON CHIWANZA*
WALTER GACHIRA*

*Dept. of Financial Engineering, Harare Institute of Technology, P. O. Box BE 277, Zimbabwe

Abstract

The current high levels of non-performing loans in MFI's in Zimbabwe call for a search of effective tools for risk management. Effective credit risk management is essential for the long term sustainability of microfinance Institutions. The study examines the credit assessment methods currently in use in Zimbabwe and their effectiveness. It concludes that the current credit assessment methods are qualitative in nature base on the 5Cs approach. The weakness is that there is more subjectivity in decision making. The study then examines the effectiveness of two statistical credit scoring methods Logistic Regression Model and Probit Model in predicting credit default. Secondary data was obtained from Merit Financial Services Loan Management System. Individual loan data consisting of 1055 personal loans disbursed from1 January up to 31 December 2015 was used the study concluded that Probit Regression had a slightly higher accuracy rate that lied within the Tasche (2005) range. With an accuracy rate of 75.02%, Probit regression model has the ability to ensure effective and sound credit risk estimates that can enable Microfinance Institutions to enhance credit risk management. The model should be used in conjunction with existing qualitative methods and not as a standalone tool.

Keywords: predictive power, Credit Risk Management, Portfolio at Risk: Non Performing Loans, Cumulative Distribution Function, Probability Distribution Function (PD)

1. Introduction

1.1 Back ground to the study

Micro-Finance Institutions in Zimbabwe play a pivotal role in the provision of services to the financially excluded population, particularly the poor and the informal sector (Fedhila, 2012). The Zimbabwean Microfinance Sector grew massively during the period before 2003. After the year 2003, the number of licensed MFI's declined from more than 1600 to less than 150 in 2013 (ZAMFI , 2013) 2015 to 147 as at 30 June 2015 serving

224,300 clients through 495 branches throughout the country. According to the MPS (2015) total loans for the sector increased from $157 million as at 31 December 2014 to $163.5 million as at 31 March 2015.

In the MPS (2015) report the growth of the microfinance sector continues to be hampered by funding challenges largely attributed to the general market illiquidity and high cost of funding escalating within the Zimbabwean Financial Sector.

In offering financial services to the poor and promoting financial inclusion in Zimbabwe, Microfinance Institutions are exposed to a spectrum of risks ranging from credit risk, liquidity risk, market risks and operational risk. Credit risk is the risk resulting from a trading partner not fulfilling his or her obligation as per the contract on due date or any time thereafter. This risk can greatly jeopardise the smooth functioning of an Institution. Credit risk therefore is one of the greatest concerns to most Microfinance Companies in Zimbabwe.

Dr Mangudya in the 2015 Monetary Policy attributed poor performance to poor lending technologies. The slow adoption by Microfinance Institutions of adequate lending technologies to ensure sound and effective risk management has resulted in huge defaults thus affecting their liquidity, capital structures and operations.

The study seeks to find the effectiveness of credit scoring models in enhancing credit risk management of Microfinance Companies in Zimbabwe. Effective credit risk management is the process of managing an institution’s activities which create credit risk exposures, in a manner that significantly reduces the likelihood that such exposures will impact negatively on a microfinance institution’s earnings and capital such as loan defaults that lead to huge non-performing loans. M La Torre 2003 defined credit risk management as an identification measurement and the control of credit risks (expected and unexpected changes) in order to price the investment correctly and to reduce losses exacerbated by changes in future events or outcomes.

Credit scoring is a technique that assists Lenders/Creditors to decide whether to grant credit to applicants who apply to them. Baklouti 2008; 2014 indicated that the process of effective credit assessment plays an important role in the financial decision making in MFI’s as it enables faster credit approval and diminishes the possible risks associated with customer’s repayment defaults. Credit scoring models help financial institutions to assign credit applicants to one of two groups: a ‘good credit’ group that is likely to repay the financial obligation or a ‘bad credit’ group that should be denied credit because of a high likelihood of defaulting on the financial obligation.

Traditionally, MFIs have used subjective qualitative scoring–the use of defined parameters such as experience in the business, net margin of the business, profitability and disposable income–to analyse businesses and credit risk. These parameters are defined using industry standards, institutional experience and stated lending policies. Mark Schreiner (2000) added that Scoring by most Microfinances relied inherently on qualitative information. However, this technique seems to be inefficient, inconsistent and above all non-uniform because of subjectivity in choice of risk weights and scores.

According to the ZAMFI (2015) monthly report there were numerous reports and complaints of weak credit assessments systems leading to client over-indebtedness, inadequate explanation of terms and conditions of loans and over-deductions, leading to excess loan repayments. This largely burdens the borrower leading to huge loan delinquencies. Delinquency levels in the Zimbabwean Microfinance Sector have remained high as reflected by the level of Portfolio at Risk (PaR 30 days) at 13.31% as at June 2015, remains above the prudential regional benchmark of 5%.The Portfolio at risk of MFI's in Zimbabwe after a decrease from 25.52% in December 2012 to 16.03% in December 2013 has been increasing from 11.29% in December 2014 to 12.05% in March 2015. With the increasing number of non-performing loans within the economy great emphasis need to be narrowed down to efficient credit assessment systems that help in screening clients, tracking and monitoring loan portfolio performance within MFI's. Microfinance institutions need to strictly vet their clients.

1.2 Objectives of the study

- Evaluate the effectiveness of credit scoring models in enhancing credit risk management of microfinance institutions.
- To identify the current credit assessment techniques being used by MFI's in Zimbabwe
- To determine how credit scoring models enhance Credit Risk Management of MIF's
- To determine the most appropriate credit scoring model that is efficient in predicting loan defaults.

1.3 Research Hypothesis

H_0: The Statistical Credit Scoring Models accurately predict client's probability of default(PD)

H_1: The Statistical Credit Scoring Models do not accurately predicts client's PD

H_0: Credit scoring models enhances credit risk management of MFI's

H_1: Credit Scoring models do not enhance credit risk management of MFI's

2. LITERATURE REVIEW

2.1 Basel 2 requirements

Basel 2 emphasised two main approaches for organisations to adopt in order to manage their credit risks (Joris Van Gool (2000));

1. the standardized approach.
2. internal rating based approach.

The former approach measures credit risk in a standardized manner, executed by external credit rating agencies. The latter approach allows and encourages financial institutions to develop their own internal measures for the assessment of credit risk.

2.2 Credit Scoring Models

Lee et al., 2002; and Abdou, 2009; indicated that there are three approaches to credit scoring models used by Financial Institutions namely judgmental, statistical and non-statistical techniques.

2.2.1 Judgemental Credit Scoring Models

These models are qualitative in nature. They are rules-based models. Dynamic Business information Journal (2012) pg 2 states that judgmental scoring is the easiest to implement because it uses credit policies and decision process, the number of rules are easily set, and the grading scale is simple. Therefore, it is easier to understand. This is why it is the most widely used model by most Microfinance Institutions in developing countries. Every loan request is analysed individually by a credit analyst. The success of a judgmental technique purely depends on the experience and expert knowledge of credit analyst. Micro-lending program rely heavily on credit officers visits to applicants' businesses rather than just on business documents. Credit analysts use current as well as past experience to evaluate a client and hence grant a loan.

However Hand (1998) argued that judgmental approach lacks quantification of credit risk. More so loan officers, like other people, may seriously have cognitive bias in processing information that affect his judgments and beliefs, which leads to behavioural bias which affects objective evaluation hence there is need to incorporate statistical techniques to enhance the accuracy of the decision to be made by the loan officer pertaining to a certain client.

2.2.2 Statistical Credit Scoring Models

According to Thomas, (2000),Abdou and Aktan(2009), there are three main statistical scoring models namely Linear discriminate analysis, logistic regression and probit regression model.

2.2.2.1 Logistic Regression Model

Thomas (2000) and Schreinier (2000) outlined that explanatory variables in a Logistic Regression model which can be in the form of lenders characteristics, loan characteristics, and borrower's characteristics whereby theses variables are linearly combined to determine the score of the applicant. Logistic regression uses dependent characteristics of the borrower to estimate the likelihood of that applicant to default. The main problem with logistic regression is that there is no universally accepted approach to choose the candidate explanatory variables for a credit scoring model. A literature review indicates that most authors refer to expert advice and (or) prior studies to explain their choices. Hence one Logistic Model cannot be universally applicable to several microfinance institutions. In order to increase the accuracy of a logistic model studies discussed above it is necessary to put weights on different categories of clients based on past information. For any decision, one assesses the circumstances and determines a weight of evidence.

2.2.2.2 Probit Model

Abdou et al(2007) developed a Probit model for an Egyptian bank, in his study on the applicability of credit scoring to Egyptian banks. Probit is a technique that finds coefficient values, such that this is a probability of a unit value of a binary coefficient. Under a probit model, a linear combination of the independent variables is transformed into its cumulative probability value from a normal distribution. The method requires finding value for the coefficients in this linear combination, such that this cumulative probability equals the actual probability that the binary outcome is one. In his study Probit Regression model had a

86.92% accuracy rate. Results from this research showed that Probit Regression can predict accurate and reliable credit risk estimates.

2.3 Empirical results

There is no unanimous agreement among the authors on the best credit scoring models. Each model has its own particular strengths and weaknesses.

Hand (1997) concluded that what is best depends on the objective of the classification, the data structure and the characteristics used. In particular, combinations of different approaches are often used as they might generate the best results in particular circumstances.

Banks use statistical tools extensively to monitor their credit portfolio. Statistical Credit Scoring Models involve the use of past loan information to determine conditions necessary to grant or reject an application. Statistical Models rely on the credit history of those debtors who are accepted and granted credit by the banks. In developed countries, lenders often rely on statistical credit scoring models to predict risks based on the performance of past loans with characteristics similar to current loans to make informed decisions.

In a study undertaken in Bolivia, the profit of Bank increased by $2 million the year after implementing credit scoring (Kuhn and Olsen, 2008). ACCION (in Bolivia, Ecuador, and Peru) and Women's World Banking (in Colombia and the Dominican Republic) are also examples of MFIs that have adopted the credit scoring in their granting decision. Two WSBI members, Banco BCSC in Colombia and Banco Estado in Chile, have incorporated credit scoring systems and have found that scoring system is very important tool for making faster and more accurate credit decisions.

Alexandria Business Associations (ABA), a Microfinance Service Provider in Egypt started full implementation of credit scoring in August 2007. El Tabaa (2007), the Executive Director of (ABA) said: 'In the first eight months following the implementation of the credit scoring model, ABA registered productivity in monthly disbursements of almost 15 percent over the previous year and an increase in amount of loans disbursed of 30 percent over the previous year. Some other interesting statistics include a slight lowering in administrative cost per loan disbursed. Repayment rates on credit approved through scoring is almost 99.5%. Since the average correct classification rate became an important criterion/tool in evaluating the classification capability of the scoring models, Adbou (2007) concluded that logit had a high classification rate of 90.54% and 75.38% respectively it was important to compare the different models'. In majority of the studies it was deduced that Logit Model has a higher ability to predict loan defaults compared to Probit model. Credit scoring is vulnerable to several statistical limits and that it cannot model risks such as unwillingness to repay and inability due to natural catastrophes, even though those risks often significantly influence default Schreiner, (2003)

.

3.0 RESEARCH METHODOLOGY

3.1 Research Design

3.1.1 Primary and secondary data

Primary data was collected through questionnaires. This information was used to establish the credit assessment techniques currently being used by Zimbabwean MFI's as well as indicating the main causes of loan defaults for individual loans. Sampling was judgmental. MFI's in Harare were selected for this study due to time constraints and costliness of reaching out to all MFI's in other towns and cities around Zimbabwe. The research sample was 20 MFI's located in Harare. Secondary data was obtained from Merit Financial Services Loan Management System. Individual loan data was needed for the evaluation of effectiveness of models. Dataset consisted of 1055 personal loans disbursed from 1 January 2015 up to 31 December 2015.

Based on the responses from questioners Probit and Logistic Regression models were constructed in Stata (Statgraphics Centurion) using individual loan information Database of Merit Financial Services. Due to the unavailability of sensitive data such as non-performing loans, Clients Loan repayments behaviour from other Microfinance the researchers constructed the models using information from Loan Management System of Merit Financial Services. Likelihood ratio tests for both models was undertaken in order to determine the significance of the variables chosen for the model. The likelihood ratio tests enables the determination of the models viability and validity. Furthermore, Chi-Square tests were conducted so as to determine whether the data fits with the model. Deviance were calculated to determine the most efficient model in predicting defaults. Test to determine which model has the highest predictive power was conducted. The models predictive powers and accuracy rates will be obtained from the Stata software. These models are important because they constitute a key role in Credit Scoring models which use past data to determine whether the client will default or not.

Table 1.1: Explanatory Variables

Variable Characteristic	Description
X_1	Clients Age
X_2	Clients Employer
X_3	Loan Purpose
X_4	Other Debts
X_5	Deduction Method
X_6	Loan Amount
X_7	Loan Tenure (Years)
X_8	Gender
X_9	Marital Status
X_{10}	Clients Experience
X_{11}	Net Income
X_{12}	Accommodation

Variables in Table 1.1 constituted the independent variables part of the two models. Test for Multicollinearity before the variables were fitted into the models was done. Multicollinearity is a situation whereby the independent variables are strongly related i.e. having a high positive correlation, which affects the stability of the model.

Correlation Matrix of the Independent Variables was constructed to determine the correlation amongst the independent variables. Where there was a high positive correlation amongst the variable i.e higher than 50% those variables were dropped.

3.1.2 Logistic Regression Model

Logit (P) ~ logistic distribution: The Applicants Credit Score linearly combines all the clients Attributes X

$$\text{Logit (P)} = \beta_0 + \beta_{1i} X_1 + \beta_{2i} X_2 + \beta_{3i} X_3 + \beta_{4i} X_4 + \beta_{5i} X_5 + \beta_{6i} X_6 + \beta_{7i} X_7 + \beta_{8i} X_8 + \beta_{9i} X_9 + \beta_{10i} X_{10} + \beta_{11i} X_{11} + \beta_{12i} X_{12}$$

Vector Format

$$\beta' = \begin{Bmatrix} \beta_1 \\ \beta_2 \\ . \\ . \\ \beta_i \end{Bmatrix} \qquad X' = \begin{Bmatrix} X_3 \\ X_2 \\ . \\ . \\ X_i \end{Bmatrix}$$

$$\text{Logit (P)} = \beta' X$$

$$\text{Probability of Default} = \frac{1}{1 + e^{-\text{logit (P)}}}$$

$$0 < PD < 1$$

Where: X denote the attributes of an individual *i*. (their number is n)

β coefficients of each attribute

3.1.3Probit Model

The probit model Probit (P) is the CDF of the standard normal distribution

$$\text{Probit (P)} = \beta_0 + \beta_{1i} X_1 + \beta_{2i} X_2 + \beta_{3i} X_3 + \beta_{4i} X_4 + \beta_{5i} X_5 + \beta_{6i} X_6 + \beta_{7i} X_7 + \beta_{8i} X_8 + \beta_{9i} X_9 + \beta_{10i} X_{10} + \beta_{11i} X_{11} + \beta_{12i} X_{12}$$

Vector Format

$$\beta' = \begin{Bmatrix} \beta_1 \\ \beta_2 \\ . \\ . \\ \beta_i \end{Bmatrix} \qquad X' = \begin{Bmatrix} X_3 \\ X_2 \\ . \\ . \\ X_i \end{Bmatrix}$$

$$\text{Probit (P)} = \beta' X$$

$$0 < PD < 1$$

$$\text{Probability of Default} = \Phi(\beta' X)$$

Where

Φ = Value from the cumulative normal distribution

$X_{1i} \ldots X_{Ki}$ = Clients Attributes

$\beta_{1i} + \ldots + \beta_{Ki}$ = Coefficients

The Pearson Chi-Square test was used to determine whether the logit and probit model functions adequately fits the observed data thus providing a measure of reliability of the models. After the models were constructed the predictive powers of the models was tested using the accuracy rate and their R – Statistics. The accuracy rate was obtained from the values which the models have correctly predicted the probabilities of default as true.

4. DATA ANALYSIS AND REPRESENTATION

4.1 Analysis of Questionnaires

Analysis of the qualitative and quantitative information obtained from questionnaires was undertaken. Of the 20 MFIs given questionnaires 18 MFI's responded. The research questionnaire had 85% response rate. From the response 63.64% related loan defaults to the inefficiencies in credit analysis. This indicates that efficiency in credit analysis can ensure a robust foundation for efficient credit risk management. Inefficiencies in credit analysis can range from poor credit scoring systems, poor management and corporate governance etc. Other causes were mainly attributed to the liquidity challenges in the economy which affects the client's ability to service their loans. The research findings from questionnaires indicated that most MFI's are using the traditional 5C's to manage their risk. From those who responded 48% indicated that expert knowledge from credit analyst and loan officers provides an efficient way to manage their risk management because they are the ones who work with the people. However there is need for improvements in order to base their decision using something or a system that is quantitative in nature. From the responses it can be deduced that 85% of MFI's in the sample have not yet adopted the use of credit bureaus in Zimbabwe as indicated by 5% of the population who have used information for their loan underwriting processes.

Almost half of the responds (47%) pointed out the need for improvements and innovation of the current risk management systems being used by MFI's in Zimbabwe. The other 47% of the sample indicated that their risk management tools were effective. Statistical credit scoring models do have a place in MFIs in Zimbabwe. This is signified by 60% of responds agreeing that statistical credit scoring models can be an efficient tool in enhancing credit risk management of MFI's in Zimbabwe. Also 13% indicated that they strongly agree that credit scoring if implemented can reduce loan defaults and 27% were indifferent. From these findings it can be presumed that Statistical credit scoring can improve the prediction of risk estimates of a significant number of MFIs thus providing an efficient way of managing loan defaults.

In order to build the proposed two models Logit and Probit personal loan Dataset of Merit Financial Services was used. The Individual Dataset consists of 1055 personal loans disbursed from 1 January 2015 up to 31 December 2015. The models was estimated using statistical software STATGRAPHICS Centurion.

4.2 Basis for default

If the outstanding instalment due for the month is greater than 0 client is considered to have defaulted that particular month hence binary code 0. If instalment paid is greater than the principal and interest due for that month binary code is 1.

Table 1.2: Explanatory Variable

Variable	Code	Comment
Income	Income	
Age	Age	Age ranges from 21 to 60
Marital Status	Mar_Stat	0 Not Married 1 Married
Loan Amount	Ln_Amt	
Other Debts	Oth_Dbts	0 Debts 1 No Debts
House Owned or rented	Rent_Owned	0 Rents 1 Onwed
Duration	Dur	In years up to 2 years
Loan Officer	LO_Exp	0 less than 5 years 1 more than 5 years
Clients Experience	Clie_Expr	0 Experience< 5 years and 1 Experience >5years
Deduction Method	Ded_Meth	O Stop or Debit Order and 1 Employer Deduction
Purpose	Purp	0 Consumption 1 Business
Coporate Guarantee	Cor_Guar	0 No Copoarte Guarantee 1 Coporate Guarantee
Sex	Sex	0 Male 1 Female
Occupation	Occ	0 Blue Color 1 White Colour
Company	Co	0 Private 1 SSB
Loan Quality	Default	0 Default/Bad 1 Not Default/Good

In Table 1.2 each customer is linked to 12 independent variables.

Clients Occupation is directly linked to the Income hence it is dropped to avoid problems of Multicollinearity. These 12 variables were selected on the basis of how each characteristic will likely affect a client’s ability to service their loans.

4.3 Multicollinearity Test

Correlation matrix was used to test for correlation between the twelve independent variables.

Table 1.3 Correlation matrix for coefficient estimates

	CONSTANT	Age	Company	Purpose	Other Debts	Dedctn Mthd	Loan Amnt	Sex	Marital Status	Experience	Net Income	Accomodation	Period
CONSTANT	1	-0.5546	-0.104	0.0096	-0.1839	-0.1273	0.0393	-0.1852	-0.7643	-0.1726	-0.1015	-0.1075	-0.1519
Age	-0.5546	1	-0.0552	0.0141	0.0229	0.0142	-0.0245	0.0421	0.0187	0.0331	0.0421	0.0206	-0.0248
Company	-0.104	-0.0552	1	0.0415	0.1904	0.3786	0.0459	-0.0214	-0.016	0.0044	0.0865	0.0579	0.0482
Purpose	0.0096	0.0141	0.0415	1	-0.1664	0.0356	-0.1623	0.0221	-0.0715	0.1085	0.0849	-0.1171	0.0413
Other Debts	-0.1839	0.0229	0.1904	-0.1664	1	0.0868	**-0.0832**	0.0219	0.0846	0.0909	**-0.1023**	-0.0198	-0.0661
Dedctn Mthd	-0.1273	0.0142	0.3786	0.0356	0.0868	1	0.0802	0.1136	-0.0534	-0.03	0.0781	-0.0497	-0.1336
Loan Amnt	0.0393	-0.0245	0.0459	-0.1623	-0.0832	0.0802	1	-0.0088	-0.0125	-0.0067	0.3876	-0.0445	-0.2582
Period	-0.1519	-0.0248	0.0482	0.0413	-0.0661	-0.1336	-0.2582	-0.0539	0.0198	0.0363	0.1015	0.0065	1
Sex	-0.1852	0.0421	-0.0214	0.0221	0.0219	0.1136	-0.0088	1	0.1384	0.0554	0.065	-0.1539	-0.0539
Marital Status	-0.7643	0.0187	-0.016	-0.0715	0.0846	-0.0534	-0.0125	0.1384	1	-0.0215	-0.0462	0.1218	0.0198
Experience	-0.1726	0.0331	0.0044	0.1085	0.0909	-0.03	-0.0067	0.0554	-0.0215	1	-0.0587	-0.1534	0.0363
Net Monthly Income	-0.1015	0.0421	0.0865	0.0849	-0.1023	0.0781	**0.3876**	0.065	-0.0462	-0.0587	1	-0.0127	0.1015
Accomodation	-0.1075	0.0206	0.0579	-0.1171	-0.0198	-0.0497	-0.0445	-0.1539	0.1218	-0.1534	-0.0127	1	0.0065

Source: Results StataGraphics Centurion

Table 1.3 shows the correlation matrix for coefficient variables results. There is no correlation with absolute figures greater that 50% amongst the predictor variables ruling out the issue of Multicollinearity. However, Net Income and Loan Amount recorded a highest correlation of 38.6%. That means there is a direct relationship between the Loan Amount and clients net Income. This correlation maybe high because loan amount to be grated to clients is dependant mainly on Net Income. More so there is a negative correlation between Other Debts variable and the Loan Amount the higher your other obligations the lower your loan. These variables had a correlation of -0.0832. Other Debts variable also has a negative correlation to the net income i.e. debts affects the overall clients net income.

4.4. Logistic Regression Model

The logistic regression model is constructed using the above mentioned 12 variables. STATGRAPHICS Centurion was used for the model construction.

Table 1.4 : Estimated Regression Model

Logistic Regression - Default

Estimated Regression Model (Maximum Likelihood)

		Standard	*Estimated*
Parameter	*Estimate*	*Error*	*Odds Ratio*
CONSTANT	0.0393825	0.644398	
Age	0.00588602	0.00827906	1.0059
Company	-0.340809	0.224852	0.711195
Purpose	0.232815	0.264446	1.26215
Other Debts	-0.564673	0.140872	0.568546
Dedctn Mthd	0.769979	0.149465	2.15972
Loan Amnt	0.00017574	0.0000992139	1.00018
Period	-1.7284	0.375584	0.177568
Sex	0.434878	0.200414	1.54477
Marital Status	0.485897	0.493108	1.62563
Experience	-0.0276202	0.0118282	0.972758
Net Monthly Income	0.0000298787	0.000102098	1.00003
Accomodation	0.0801044	0.173007	1.0834

Source: Results Stata Graphics Centurion

The above table 1.4 indicates all the variables used for the model were statistically significant at 95% confidence level as indicated by the Standard Error Figure which is approximately equal to 0. The Standard Error measures the degree of error variance between the estimated parameters

4.4.1 Estimation of Probability of Default

$$PD = \frac{1}{1 + e^{-\mathrm{logit}(x)}}$$

were :

Logit (P)= 0.0393825 + 0.00588602*Age - 0.340809*Company + 0.232815*Purpose - 0.564673*Other_Debts + 0.769979*Dedctn_Mthd + 0.00017574*Loan_Amnt - 1.7284*Period + 0.434878*Sex + 0.485897*Marital Status - 0.0276202*Experience + 0.0000298787*Net _Monthly_Income + 0.0801044*Accomodation

4.4.2 Likelihood Ratio Test

The Likelihood ratio test tests whether the Client's Characteristics have an Effect on then PD. It test how reliable are the coefficients in explaining the dependant variable.

Table 1.5 : Analysis of Deviance

Analysis of Deviance

Source	*Deviance*	*Df*	*P-Value*
Model	87.0313	12	0.0034
Residual	1340.74	1042	0.0609
Total (corr.)	1427.77	1054	

Percentage of deviance explained by model = 6.09561
Adjusted percentage = 4.27459

Source: Statgraphics Centurion Results

Since the P-Value in the analysis of deviance Table 1.5 is less than 5%. This means that there is 95% confidence that the coefficient of the independent variables do have an effect on the dependant variable. Therefore clients characteristics such as age , gender, employment status, experience, marital status, loan amount, accommodation, deduction method and other debts do have an effect on the clients likelihood to default.

4.4.3 Chi-Square Goodness of Fit Test

The Pearson Chi-Square test is used to determine whether the logistic regression function adequately fits the observed data.

Table 1.6: Goodness of fit test

Logistic Regression - Default

Chi-Square Goodness of Fit Test

	Logit		*TRUE*	*TRUE*	*FALSE*	*FALSE*
Class	*Interval*	*n*	*Observed*	*Expected*	*Observed*	*Expected*
1	less than -0.119894	211	89.0	81.9549	122.0	129.045
2	-0.119894 to 0.177504	211	88.0	106.661	123.0	104.339
3	0.177504 to 0.584126	211	135.0	126.446	76.0	84.5544
4	0.584126 to 1.00093	211	149.0	145.265	62.0	65.7355
5	1.00093 or greater	211	162.0	162.674	49.0	48.3264
Total		1055	623.0		432.0	

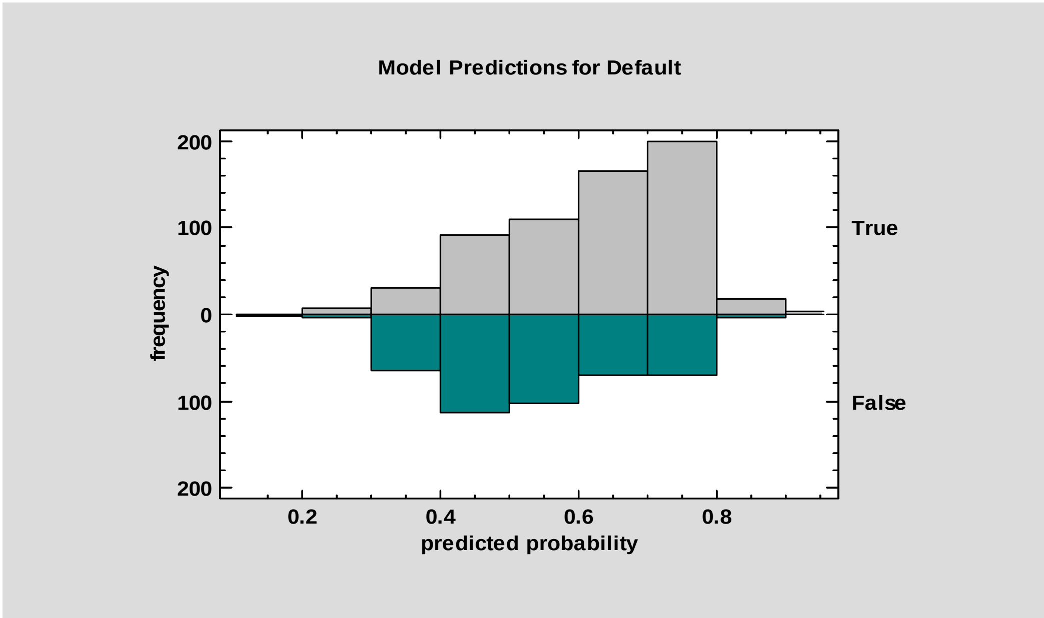
Chi-square = 9.35757 with 3 d.f. P-value = 0.054895

Source: Statagraphics Centurion Results

Table 1.6 shows goodness of fit. Chi-square = 9.35757 with 3 degrees of freedom P-value = 0.054895 this test determines whether the logistic function adequately fits the observed data. Because the P-value >0.05, we conclude that there is a perfect fit between the observed and expected frequencies at 95% Confidence Interval.

4.4.4 Model prediction capability

Fig 1.1: Model Prediction for Default



Source: StataGraphics Centurion

As depicted in the fig 1.1 there are many cases of a high probability prediction of good credit, which were confirmed as true. The model had 74.83% classification rate. Which means that it managed to predict more than 70% of the loans which were said to have defaulted. Prediction capability for the Model describes the relationship between different cut-off points and the per cent correctly classified.

4.4.5 Prediction Performance

Table 1.7 : Predictive Power of the Logit Model

Logistic Regression - Default

Prediction Performance - Percent Correct

Cutoff	TRUE	FALSE	Total
0.0	100.00	0.00	59.05
0.05	100.00	0.00	59.05
0.1	100.00	0.00	59.05
0.15	100.00	0.00	59.05
0.2	100.00	0.23	59.15
0.25	99.68	0.23	58.96
0.3	98.88	1.16	58.86
0.35	96.47	6.48	59.62
0.4	93.90	16.20	62.09
0.45	88.28	24.54	62.18
0.5	79.29	42.59	64.27
0.55	79.66	57.87	74.83
0.6	61.80	66.44	63.70
0.65	47.99	75.23	59.15
0.7	35.31	82.87	54.79
0.75	19.58	91.90	49.19
0.8	3.37	99.31	42.65
0.85	0.80	99.77	41.33
0.9	0.48	100.00	41.23
0.95	0.32	100.00	41.14
1.0	0.00	100.00	40.95

Source: StataGraphics Centurion results

Table 1.7 shows a summary of the prediction capability of the fitted model. First, the model is used to predict the response using the information in each row of the data file. If the predicted value is larger than the cut-off, the response predicted to be TRUE. If the predicted value is less than or equal to the cut-off, the response is predicted to be FALSE. Table 1.7 shows the percent of the observed data correctly predicted at various cut-off points. For example, using a cut-off equal to 0.55, 79.6629% of all TRUE responses were correctly predicted, while 57.8704% of all FALSE responses were correctly predicted, for a total of 74.8341%. Therefore the total Predictive Power of the Logistic regression model is 74.83%.

4.5. Probit Regression Model

The probit model **F(x'β)** is the CDF of the standard normal distribution

$$\mathbf{F(X'\beta) = \Phi X'B = \Phi\,(\beta_{1i}\,X_{1i} + \beta_{\,2i}\,X_{2i} + \ldots. + \beta_{Ki}\,X_{\,Ki})}$$

Table 1.8: Estimated Regression Model (Maximum Likelihood)

Probit Analysis - Default

Estimated Regression Model (Maximum Likelihood)

		Standard
Parameter	*Estimate*	*Error*
CONSTANT	0.0277337	0.396904
Age	0.00345142	0.0050671
Company	-0.207262	0.139358
Purpose	0.143778	0.162899
Other Debts	-0.34308	0.0865352
Dedctn Mthd	0.472313	0.0920749
Loan Amnt	0.000101688	0.000057349
Period	-1.05499	0.229209
Sex	0.27111	0.122613
Marital Status	0.298749	0.304919
Experience	-0.0170229	0.00727101
Net Monthly Income	0.0000220978	0.0000626116
Accomodation	0.0519837	0.105839

Source: StataGraphics Centurion Results

Table 1.8 shows the results of fitting a probit regression model to describe the relationship between Default and 12 independent variable(s).

The equation of the fitted model is

Default = normal (Probit (X)

Where

Probit (P) = 0.0277337 + 0.00345142*Age - 0.207262*Company + 0.143778*Purpose –

0.34308*Other Debts + 0.472313*DedctnMthd + 0.000101688*Loan Amnt - 1.05499*Period + 0.27111*Sex + 0.298749*Marital Status - 0.0170229*Experience + 0.0000220978*Net Monthly Income + 0.0519837*Accomodation

4.5.1. Likelihood Ratio Test

It test whether the variables have significance to the outcome

Table 1. 9 Analysis of Deviance

Probit Analysis - Default

Analysis of Deviance

Source	*Deviance*	*Df*	*P-Value*
Model	86.8621	12	0.0091
Residual	1340.91	1042	0.0433
Total (corr.)	1427.77	1054	

Percentage of deviance explained by model = 6.08376
Adjusted percentage = 4.26274

Source: StataGraphics Centurion Results

Table 1.9 shows results of analysis of deviance. The p value of the model is less than 5% hence we conclude that at 95% confidence interval the estimated coefficients do have an impact on the dependant variable .

4.5.2. Chi-Square Goodness of Fit Test

Table 1.10: Goodness of Fit Test

Probit Analysis - Default

Chi-Square Goodness of Fit Test

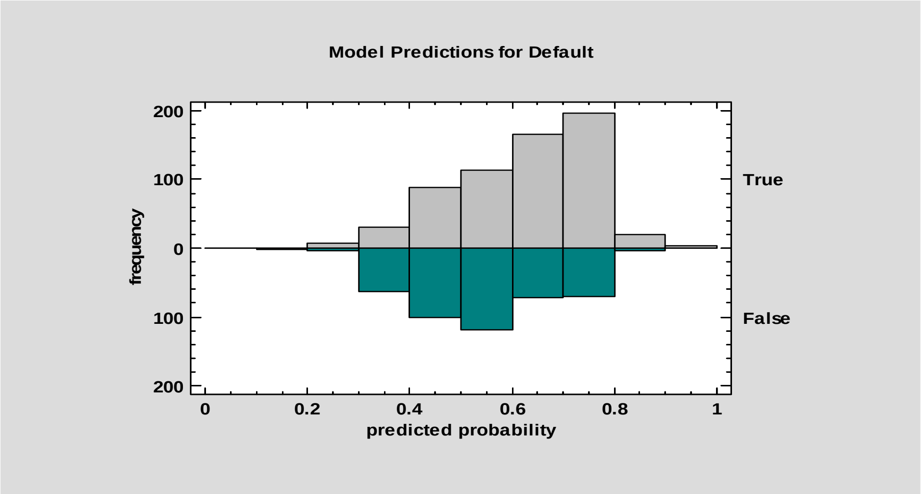
Class	*Probit Interval*	*n*	*TRUE Observed*	*TRUE Expected*	*FALSE Observed*	*FALSE Expected*
1	less than -0.0712663	211	89.0	82.4013	122.0	128.599
2	-0.0712663 to 0.106795	211	89.0	106.839	122.0	104.161
3	0.106795 to 0.362664	211	136.0	126.274	75.0	84.726
4	0.362664 to 0.614659	211	147.0	144.973	64.0	66.0267
5	0.614659 or greater	211	162.0	162.809	49.0	48.1905
Total		1055	623.0		432.0	

Chi-square = 8.87469 with 3 d.f. P-value = 0.0510031

Table 1.10 shows results of goodness of fit test. This test determines whether the probit model adequately fits the observed data. Because the P-value >0.05, we conclude that the fitted model adequately fits the data at 95% confidence level.

4.5.3 Model prediction capability

Fig 1.2: Model Prediction Capability



Source: Statgraphics Centurion Results

As shown in Figure 1.2 there are many cases of a high probability prediction of good credit, which were confirmed as true. Where the prediction probability exceeds 0.60 there are a few false results, i.e. bad credits, for probabilities of good credits less than 0.60 there are more false results.

4.5.3. Predictive Power of the Probit Model

Table 1.12: Classification Rate

Probit Analysis - Default

Prediction Performance - Percent Correct

Cutoff	*TRUE*	*FALSE*	*Total*
0.0	100.00	0.00	59.05
0.05	100.00	0.00	59.05
0.1	100.00	0.00	59.05
0.15	100.00	0.23	59.15
0.2	100.00	0.23	59.15
0.25	99.68	0.23	58.96
0.3	98.88	1.16	58.86
0.35	96.47	6.25	59.53
0.4	94.06	15.74	61.99
0.45	88.44	24.54	62.27
0.5	79.94	39.12	63.22
0.55	79.82	58.10	75.02
0.6	61.80	66.44	63.70
0.65	47.51	75.46	58.96
0.7	35.15	83.10	54.79
0.75	19.26	91.90	49.00
0.8	3.53	99.31	42.75
0.85	0.80	99.77	41.33
0.9	0.48	100.00	41.23
0.95	0.32	100.00	41.14
1.0	0.00	100.00	40.95

Table 1.11 shows a summary of the prediction capability of the fitted model. First, the model is used to predict response using information in each row of the data file. If the

predicted value is larger than the cut-off, the response is predicted to be TRUE. If the predicted value is less than or equal to the cut-off, the response is predicted to be FALSE. Table 1.11 shows the percent of the observed data correctly predicted at various cut-off values. For example, using a cut-off equal to 0.55, 69.8234% of all TRUE responses were correctly predicted, while 58.1019% of all FALSE responses were correctly predicted, for a total of 75.0237%.

4.6. Comparison of the models

Probit and Logit models were compared using their predictive power, accuracy rates and percentage of deviance from the observed constant model. On this criterion a model with the highest predictive power, highest accuracy rate as well as a lower percentage of deviance will be the one that predicts credit risks more.

4.5.1. Predictive Powers

From the classification parameters of the two models it can be deduced that correct risk estimates i.e. probability of default parameters have the ability to determine the ability of the client in defaulting by more than 75% on average according to Tasche's (2005) range.

Table 1.12: Average correct classification rates for Probit and regression Models

Cut-off	0.35	0.4	0.45	0.5	***0.55***	0.6	0.65	0.7	0.75
Probit	59.53	61.99	62.27	63.22	***75.02***	63.70	58.96	54.79	49.00
Logit	59.62	62.09	62.18	64.27	***74.83***	63.70	59.15	54.79	49.19

From Table 1.12 above it can be summarised that probit regression models has a slightly higher predictive ability compared to logit model. Thus the higher the classification rate the more the ability of credit analyst and loan officers are able to detect credit risk. Difference in the accuracy rate of (75.02%) Probit regression model and (74.83%) Logit model may not be statistically significant.

4.5.2. Percentage of Deviance

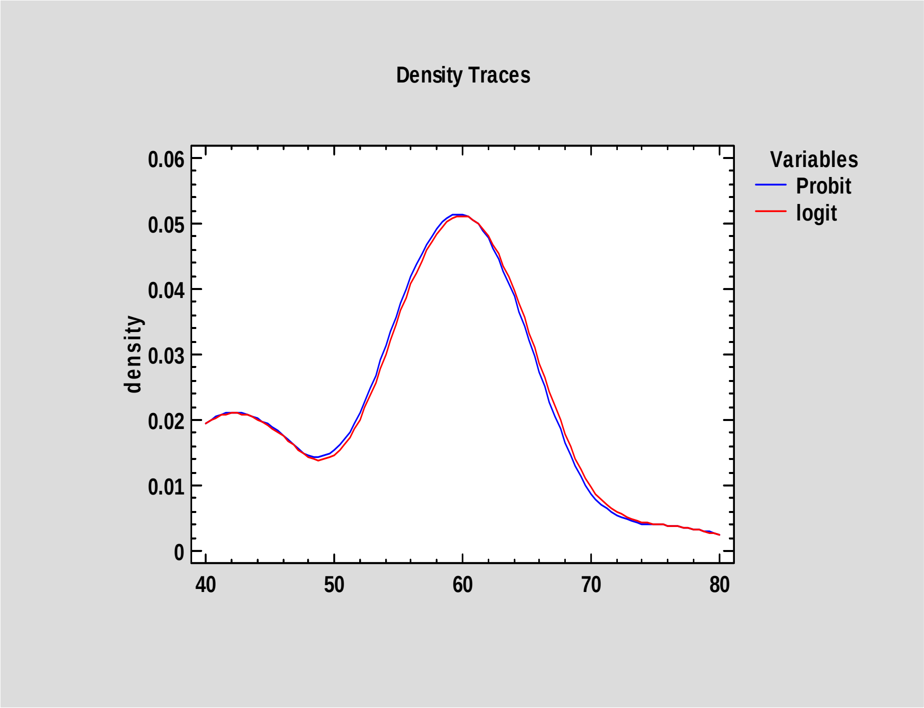
The reliability tests of the models are obtained from the deviance of the fitted model from the observed constant only model. It indicates which model is deviating more from the observed constant only model.

Table 1.13: Percentage of Deviance

	Probit	Logit
Percentage of deviance	6.08376	6.09561

Table 1.13 depicts the percentage of deviance of the two models. From the table it shows that Probit Model had a slightly lower percentage of deviance, 6.08376% as compared to logistic regression, and 6.09561% which indicates that probit regression is slightly more reliable than logit model.

Fig1.3: Distribution Functions



Source: Statgraphics Centurion Results

Fig 1.3 the logit and probit models give similar results and the main difference between logit and probit model is that probit has slightly flatter tails.

5.0 INTERPRITATION

From the responses of the questionnaire it can be established that credit scoring models can enhance credit risk management of Microfinances in Zimbabwe. However it must be noted that other respondents (27%) were not clear on the effectiveness of credit scoring when used alone. Hence there is need to use it in combined with other scoring techniques such as judgmental credit scoring to back their decisions to accept or reject a loan.

Test to determine the goodness of fit for the models were undertaken and the results confirmed that the dataset used adequately fitted the logit and probit models at 95% confidence interval. The results of the research are therefore reliable.

The main hypothesis of the research was to determine whether credit scoring models accurately predicts probability of default at 95% confidence level. The benchmark of an effective credit scoring model is measured by Tasche's (2005) range **[0.75 - 0.9].** This range indicates that a scoring model with a predictive power lying within the range is effective and efficient in accurately predicting credit risk.

From the results Probit and Logit Models had 75.02% and 74.83% prediction powers respectively. Probit Regression had a slightly higher accuracy rate that lied within the Tasche (2005) range. The higher the classifications rate the more the ability of credit analyst and loan officers are able to detect credit risk thus ensuring a sound credit risk management tool.

With the accuracy rate of 75.02% Probit regression model has the ability to ensure effective and sound credit risk estimates that can enable Microfinance's to enhance their credit risk management. This is in line with studies by Scheiner (2000),Vogelsang (2013) who concluded that credit scoring models have the ability to ensure effective and sound credit risk estimates thereby enabling MFI's to enhance their credit risk management tools.

. From the hypothesis credit scoring do have a place in the lending criteria of lenders in Zimbabwe. An accuracy rate of 75,02% shows that credit scoring models can effectively predict likelihood to default.

6.0 CONCLUSION

Credit scoring models do enhance credit risk management of MFI's in Zimbabwe as indicated by the 75.02% and 74.83% accuracy rate of the Probit and Logit models respectively. Probit regression model is slightly more accurate at (75.02% accuracy) in predicting defaults than the logit model. However, the Probit model's prediction power lies at the lower end of Tasche's range (0:75 - 0:9). According to Tasche (2005) it is difficult to entirely replace judgmental credit scoring with an automated credit scoring system. J.V Gool at al (2009) points out that credit scoring should become a refinement tool to judgemental credit scoring which has proven to be stable, easy to use, and also to have a certain discriminatory power. Either of the two models could be used.

References

1. Abdou, H., Masry, A. E., &Pointon, J. (2007). On the Applicability of Credit Scoring Models in Eqyptian Banks. Banks and Bank Systems, 2 (1), 4-20.
2. Baklouti Ibtissem1, AbdelfettahBouri 2013 credit risk management in microfinance: the
3. Chen, M., & Huang, S. (2003). Credit Scoring and Rejected Instances Reassigning Through Evolutionary Computation Techniques. Expert Systems with Applications, 24.
   conceptual framework
4. Dynamic Business information Journal (2012) pg 2
5. Hand, D.J. (1998). Reject inference in credit operations. In Mays, E. (Eds.), Credit Risk Modeling: Design and Application, New York, Glenlake Pub. Co.
6. L. Thomas, J. Crook and D. Edelman (eds), Credit Scoring and Credit Control. Oxford: Oxford University Press.
7. Lee, T. C., Chiu, (2002). Credit Scoring Using the Hybrid Neural Discriminant Technique .Expert Systems with Applications, 23, 245-254.
8. Schreiner, M. (2003). Scoring: The Next Breakthrough in Micronance. CGAP Occasional
9. Thomas, L. (1992). Methods for credit scoring applied to slow payers.
10. Van Gool, J., Verbeke, W., Sercu, P., Baesens, B. (2011). Credit scoring for microfinance: is it worth it?. International Journal of Finance & Economics, 17(2),
11. Vogelgesang, U. (2003). Microfinance in times of crisis: the effects of competition, rising indebtness, and economic crisis on repayment behavior.