

A MODEL FOR THE GRANTING OF CREDITS AND RISK ESTIMATION IN THE AGRICULTURAL SECTOR

Dr. Javier Chavez Ferreiro

Instituto Tecnológico de Morelia, México

Dr. German Narvaez Vasquez

Instituto Tecnológico y de Estudios Superiores de Monterrey, México

MIF. Emmanuel Soriano Basilio

Fiidecomisos Instituidos en Relación con la Agricultura, México

Abstract

This study aims to propose a statistical model for the granting of credits in the agricultural sector that will generate an appropriate credit risk management. (Case of the Trust Funds to Agriculture, FIRA). The methodology used was the discrimination technique, commonly known as credit scoring. The score tables allow to determine default probabilities of guaranteed loans, these probabilities and the contingent balance are inputs to estimate credit risk through three models commonly used in practice: CyRCE, Montecarlo and Credit Risk+. The results led to the conclusion that the parameters determined through a scoreboard based on logistic regression and estimation of credit risk with the Monte Carlo model, allow having a balance between revenue and expenditure.

Keywords: Model, credit risk, agricultural sector

Introduction

The Financial institutions through lending become a pillar for economic growth of a country or region. The strength of the capital structure of these institutions contributes to a stable financial system that will help mitigate economic and financial impacts. To achieve robustness, governments must issue rules of supervision and capital requirements to reduce the risk of insolvency. This led to the authorities in several countries to form the Basel Committee on Banking Supervision which issued the document "International Convergence of Capital Measurement and Capital Standards", also called "The agreement BIS 1988" (later known as Basel I).

According Sandica (2010), the basis for a better credit risk assessment is to define the probability of default of a counterparty who has an obligation to comply with the financial institution. In this regard various techniques have been developed, including the development of a scoring model (credit scoring). The pioneer of this model is 1941 Durand when applied discriminant analysis proposed by Fisher (1936) to classify borrowers into good and bad based on their characteristics (Sandica, 2010, Rencher, A. (2002). Scoring models allow approval of a loan so that the probability of default can be inferred. Therefore, a scoring model the financial intermediary is able to extend credit to a desired level of risk.

The present work aims to develop a comprehensive tool for managing credit risk, generally applicable to financial institutions and in particular for the Trust Funds to Agriculture (FIRA). To achieve this objective, in the first part refers to the problematic and background. In the second part presents the theoretical framework which underpins the methodology, in particular develops the method of scoring tables, also known as credit scoring, which is a tool to classify borrowers into categories of behavior: good and bad.

Both estimates score tables as obtained through credit risk models are the basis for the estimation of an insurance premium based on the proposed model. In the third part describes the methodological process which is based on the application of credits guaranteed by FIRA between 2006 and 2011 in the models referred to the theoretical framework. Then in the following section presents the results obtained, which shows the operation of the model from the data. Finally we present the conclusions, highlighting among others the lack of knowledge of the probability of default of a borrower fails to generate policies for credit risk management that conserve the assets of an institution and increase profits.

Literature Riview

Scoring model (Credit Scoring): The scoring model is a set of decision-making models and their underlying techniques that help lenders in consumer credit. These techniques they decide who will get the credits, and which operational strategies will improve the gain of the lenders. The techniques of the scoring model (credit scoring) evaluate the risk of lending to a particular consumer. (Thomas, Edelman & Crook, 2002).

Lenders must make two types of decisions: a) give the credit to a new applicant and b) how to manage the existing applicants, including whether they should increase the credit limit. In both cases, the scoring model is a predictor of risk and behavior based on samples of behavior in the past loans, whose consumers would be expected to be similar to those that will be evaluated with the score model. The scoring model for the first decision is

known as Score Origination and the second is a behavioral Score. In particular, this paper develops a Score of Origination.

Discrimination model: While statistical methods were the first to be used to construct the score tables, these have been modified. The most common statistical method nowadays is logistic regression (logit model) that takes as input the discriminant analysis. Other methods are linear regression and classification tree (model tree) which separates the set of applicants in a number of subgroups based on their attributes and then classifies the subgroups in satisfactory or unsatisfactory. (Thomas, Edelman & Crook, 2002).

In the process of granting credit the lender must choose between two options: grant or reject the credit. The scoring model tries to help the election. Therefore, the model classifies those credited in one of the following two groups: good or bad. Good behavior would be acceptable to the lender while bad means that the lender would have liked to reject it.

Discriminant analysis: The two groups classified by the lender should be good and bad. Fisher suggested that a measure M of separation of populations with common sampling variance is:

$$M = \frac{\text{Distance between sample means of two groups}}{(\text{sampling variance in each group})^{1/2}} \tag{1}$$

If m_G , m_B and S are the mean of the sample of the group of good, the average of the groups and the poor common variance respectively, the above equation is redefined is:

$$M = w^T \frac{m_G - m_B}{(w^T S w)^{1/2}} \tag{2}$$

Differentiating the equation 2 with respect to w , the value of M is maximized when:

$$\frac{m_G - m_B}{(w^T S w)^{1/2}} - \frac{(w(m_G - m_B))^T (S w^T)}{(w^T S w)^{3/2}} = 0$$

$$\text{So } (m_G - m_B)(w^T S w) = (S w^T)(w(m_G - m_B))^T \tag{3}$$

Intuitively the score table aims to separate the groups to search for an equation that maximizes the separation between the averages.

Value of information: The Value of information allows to estimate the predictive value of a feature based on its attributes. (Siddiqi, 2006). The equation to estimate is given by:

$$VI = \sum_{i=1}^a \ln \frac{f(x_i)_g}{f(x_i)_b} \tag{4}$$

Where $f(x_i)_g$ = relative frequency of good in the attribute, i ; and, $f(x_i)_b$ = frequency of evil in the attribute, i ; a = number of attributes that divide the property.

Linear regression: Linear regression models attempt to explain the value of a dependent variable Y based on the value of one or more explanatory variables X. X variables can be quantitative or qualitative, being the first those that take a numeric value and qualitative, also known as dummy variables, are those involving the existence or absence of a quality or attribute, such as being male or female. The method of quantifying the attributes consists in constructing dummy variables that assume the value of 1 if the attribute is present and 0 if it's not (Gujarati, 2006). The relationship between the dependent variable and the independent variable is given by the following equation:

$$Y = w_0 + w_{1,1}X_{1,1} + w_{1,2}X_{1,2} + \dots + w_{1,m-1}X_{1,m-1} + \dots + w_{n,1}X_{n,1} + \dots + w_{n,m-1}X_{n,m-1} \quad (5)$$

If the qualitative variable (characteristics) i (i = 1,..., n) has m categories (attributes), simply insert m-1 dummy variables. This is extended to each of the n characteristics to be used to predict the value of Y.

The dependent variable Y can take two values: 1 if the loan is classified as bad and 0 if not. With equation (5) the purpose is to find the values of the w to explain in a better way the behavior of the credit, however the results can be estimated to be greater than 1 in spite of the above, the linear regression results differ from the logistic regression only queues (Greene, 2011), which is the most widely used regression models to score. Therefore, the linear regression results are equally valid for a scoring model. (Thomas, Edelman, & Crook, 2002).

Equation (5) can be solved by Ordinary Least Squares, method which estimates the values of w's. The solution is given in matrix form by (Gujarati & Porter, 2010):

$$W = (X'X)^{-1} X'Y \quad (6)$$

Where X is a matrix containing 0 and 1 indicating the qualitative characteristics of each variable in each of the appropriations, X' is the transposed matrix of X, and Y is a vector containing 0 and 1 based on the behavior of each credit.

Logistic regression : Logistic regression is a discriminant function that attempts to find the best combination of features that explain the probability of default. If p_i is the probability that the applicant i in the sample has failed, meaning that Y_i is 1, it is desirable to find a vector w such that (Gujarati & Porter, 2010):

$$\log\left(\frac{p_i}{1-p_i}\right) = w_0 + w_1x_1 + \dots + w_px_p = wx^T \quad (7)$$

Taking exponential of both sides of the equation (7) we obtain:

$$p_i = \frac{e^{w'x}}{1+e^{w'x}} = F(x'_i w) \quad (8)$$

Unlike linear regression, logistic regression allows the variable pi only takes values between 0 and 1, which is the expected range for a variable that considers a probability (in this case of default). The components of the vector w are found through the maximum likelihood method, by finding the vector that maximizes the log-likelihood function which is (Novales, 1993).

$$L = \prod_{Y_i=1} F(x_i'w) \prod_{Y_i=0} [1 - F(x_i'w)] \tag{9}$$

Weight of evidence: Both linear regression and logistic regression model can be performed on dummy variables or based on WOE for its acronym in English (Weight of Evidence). The $WOE_{i,k}$ of the attribute i which belongs to a k characteristic is estimated by:

$$WOE_{i,k} = \sum_{i=1}^a \ln \left(\frac{f(x_i)_g}{f(x_i)_b} \right) * 100 \tag{10}$$

Where $f(x_i)_g$ = relative frequency of good in the attribute, i; and, $f(x_i)_b$ = frequency of evil in the attribute, i; a= number of attributes that divide the property.

Score board: In general the relationship between odds (division of good to bad) and the score can be presented as a linear transformation: (Siddiqi, 2006):

$$\text{Score} = \text{offset} + \text{factor} * \ln(\text{odds}) \tag{11}$$

The scoreboard is developed by using odds specified to a score "points to double odds" (PDO) is also specified. The factor and offset can be calculated by solving the following equations:

$$\text{Score} = \text{offset} + \text{factor} * \ln(\text{odds})$$

$$\text{Score} + \text{pdo} = \text{offset} + \text{factor} * \ln(2 * \text{odds})$$

Solving for the different variables we obtain:

$$\text{Pdo} = \text{factor} * \ln(2) \tag{12}$$

$$\text{Factor} = \text{pdo} / \ln(2) \tag{13}$$

$$\text{Offset} = \text{score} - \text{factor} * \ln(\text{odds}) \tag{14}$$

Unless the scoreboard is performed using as input WOE, equation (11) can be modified as:

$$\begin{aligned} \text{score} &= \ln(\text{odds}) * \text{factor} + \text{offset} = - \left(\sum_{j,i=1}^{k,n} (\text{woe}_j * w_i) + a \right) * \text{factor} + \\ \text{offset} &= - \left(\sum_{j,i=1}^{k,n} \left(\text{woe}_j * w_i + \frac{a}{n} \right) \right) * \text{factor} + \text{offset} = \sum_{j,i=1}^{k,n} \left(- \left(\text{woe}_j * \right. \right. \\ & \left. \left. w_i + \frac{a}{n} \right) * \text{factor} + \frac{\text{offset}}{n} \right) \end{aligned} \tag{15}$$

Where: woe = weight of evidence for each attribute grouped, w = regression coefficient for each feature, a = the intercept estimated by logistic regression, n = number of features, k = number of groups (attributes) in each feature. With equation (15) is the score of each attribute. The addition of the scores for each attribute of credit equals the total score given to credit.

Validation scoreboard: The score board must show that effectively separates a proper way to loans classified as bad for those classified as good, giving a lower score to the former. (Thomas, Edelman & Crook, 2002) For this, we have developed a test statistics would, being those commonly used to carry out validation scoring models include: i) *Ginni index*, ii) *Divergence Index*, iii) *K-S (Kolmogorov-Smirnov)*, iv) *Value of Information*, v) *Stability Index of the Population*, vi) *T Test*.

Model of credit risk

The credit risk defined by the "General provisions applicable to credit institutions" issued by the National Banking and Securities Commission (CNBV, 2010) as a potential loss by the lack of payment of an accredited or counterpart. Credit risk derives from the possibility that the borrower may default. (Hull, 2007) In particular for this paper the credit risk will be defined as the possibility that FEAGA pays a guarantee to the financial intermediary for breach of the obligations of a borrower.

Expected loss: As this is an expected loss, financial institutions must take as given their declining profits under such loss. Through the generation of reserves for credit risks the loss is recognized. As detailed below, this loss is the basis for the charge of the insurance premium. The unexpected loss, which is the difference between credit risk and the expected loss is covered by the capital of the Institution and must also be considered as a component of the insurance premium.

Model CyRCE: The result of granting a loan can be classified into two scenarios: good (fulfilled) or bad (failed). The result of credit i classified as bad is based on a default probability p_i for a time horizon HT. This probability is associated to the characteristics of the borrower and it is estimated from the previously described scoring models. For simplicity, it is assumed independence between the borrowers, which means that if borrower fails does not imply that others will too. Once described the above, it is necessary to find the distribution of the portfolio loss where each loan has a balance of default S_i and a default probability p_i . The distribution can be estimated by a normal probability distribution, which requires obtaining a common default probability p_π . This probability is estimated by:

$$p_\pi = \frac{\sum_{i=1}^n p_i S_i}{S} \quad (16)$$

$$S = \sum_{i=1}^n S_i \quad (17)$$

Since each credit has a binomial behavior, that is to say, payment or non-payment, the risk of market is estimated through:

$$VaR_\alpha = p_\pi S + z_\alpha \sqrt{V} = p_\pi S + z_\alpha \sqrt{p_\pi(1 - p_\pi) \sum_{i=1}^n S_i^2} \quad (18)$$

The final equation CyRCE model shows that the value at risk is a function of the probability of default weighted index portfolio and portfolio concentration.

$$VaR_\alpha = S(p_\pi + z_\alpha \sqrt{p_\pi(1 - p_\pi)IHH}) \tag{19}$$

Credit Risk+Model: Unlike CyRCE model, the Credit Risk+ model models two random processes that any credit portfolio present: the number of default process and the process of the amount of losses. The Poisson distribution models discrete events in a continuous space or time. That is, the events of default (discrete) in a given period of time ht . In the portfolio you want to model the number of non-compliance from where it is impossible to predict the exact number or the right time. However, you can associate a probability to the number of defaults in a given period of time .

Being $p_n = Pr[N = n]$ the probability that exactly n failures occur. Using the Poisson distribution, p_n is given by (Gutierrez & Elizondo, 2002).

$$p_n = \frac{e^{-\mu} \mu^n}{n!} \tag{20}$$

Where μ represents the expected number of defaults in the portfolio during the period.

Montecarlo model: There is a portfolio made up of n borrowers, each with a balance, independent of each other with a default probability p_i for a time horizon ht . Since the default is a random variable, the result of each credit can be modeled through an indicator function that takes the following values: (Márquez, 2009)

$$I_i = \begin{cases} 1 & \text{si } u_i \leq p_i \\ 0 & \text{eoc} \end{cases} \tag{21}$$

Where I_i takes the value 1 if the borrower fails and 0 otherwise (EOC), and u_i is a random variable uniformly distributed in the range 0 – 1 $U \sim [0,1]$ for $i = 1,2,\dots,n$. With the given the above, the estimated loss LGD_w to the stage portfolio w is provided by:

$$LGD_w = \sum_{i=1}^n I_i * s_i \tag{22}$$

W scenarios are simulated that allow recreating the loss distribution LGD

$$LGD = \begin{pmatrix} LGD_1 \\ LGD_2 \\ \vdots \\ LGD_w \end{pmatrix} \tag{23}$$

The W scenarios are sorted from highest to lowest loss, being the VaR a α confidence level, the Y observation fulfills: (Hull, 2007)

$$Y = W * (1 - \alpha) \tag{24}$$

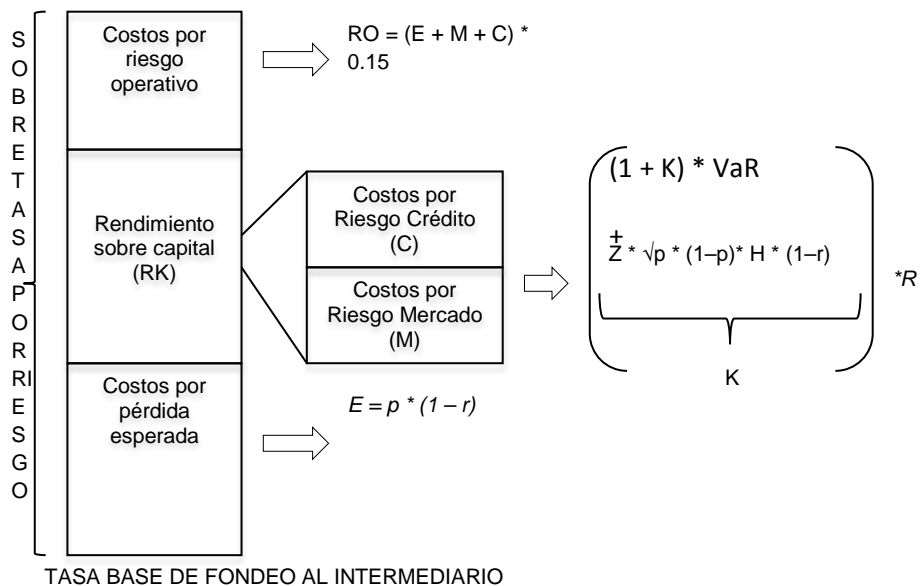
Back testing: The aim of back testing is to determine the predictive capacity of risk models. The above is done by factoring in a certain time horizon the number of times that the losses by credit risk exceed the VaR (violations). The suggested methodology was developed by Kupiec, which sets the number of maximum and minimum violations according to the following equation:

$$-2 \ln \left[(1-p)^{m \cdot e} p^{m \cdot (1-e)} \right] + 2 \ln \left[(1-e/m)^{m \cdot e} (e/m)^{m \cdot (1-e)} \right] \quad (25)$$

Model for estimating insurance premiums

The insurance contract is the agreement by which one of the parts, the insurer -FEGA-, undertakes to indemnify damage through an insurance sum of money to the other part, holder -Financial Intermediary-, in exchange for the payment of a price known as premium. In Figure 1. outlines the general model for estimating premium risks. The general concept is built on: (i) the basis of a funding rate to borrowers, (ii) a first component that integrates the cost per expected loss of credit risk, (iii) a second component comprising the return on capital that supports the operation, iv) a final component which includes consideration of the possible standard operating costs of the operation. (Bousoño, C., Heras, A., & Tolmos, P., 2008).

Figure 1: Components of the surcharge for risks. (Also called risk premium).



This figure shows the components of the model for risk surcharge developed by FIRA, where $E = p \cdot (1 - r)$ is the expected cost for credit risk. $K = z \cdot \sqrt{p \cdot (1 - p) \cdot H \cdot (1 - r)}$, is the capital to support the risks assumed by the institution which is composed of the capital allocated to cover unexpected

losses and allocated to cover market risk. $L = (I + K) * VaR$, is the cost of market risk (M), where VaR is market risk is the unit of currency in which the transaction is funded in percentage terms. $RK = (L+K) * R = M+C$, is the Return on capital is comprised of market risks and credit risks. $RO = (E+C+M) * 0.15$, operational risk is assumed by the institution for its operations. $ST = E+M+C+ RO$, is the surcharge or risk premium would be given by the sum of the cost for taking credit risk margin for operational risk and market risk.

Methodology

This research is a descriptive correlational design, to determine the most appropriate model in the granting of credits in the agricultural sector that could be the basis for credit risk management. To conduct this study, a sample of 110.278 guaranteed loans of natural people was taken from FEFA between January 2006 and May 2011. Each credit has 69 variables related to the profile of the borrower. The same sample was used for retrospective testing of the estimated VaR of the guaranteed portfolio.

The theoretical basis lies in its first part, in models of discrimination by an analysis of the different scoring models, being the most important the linear regression model and Logit. After making a comparison between the models of discrimination, the results were applied to the estimation models of credit risk, CyRCE , Credit Risk + and Montecarlo. Subsequently, the discrimination model results were applied to a parametric model for estimating risk premiums in the agricultural sector. In general terms this research wishes to present the relationship between the independent variables (model discrimination, credit risk, risk premiums and profit) with the dependent variable (heritage conservation and increased profits).

Results and Discussions

FIRA is constituted by four trusts called: FEFA, FONDO, FEFA and FOPESCA. To fulfill its mission, FIRA supports productive projects related to the rural sector mainly through bank financial intermediaries (banks) and non-bank financial intermediaries (IFNB) as SOFOLES, SOFOMES and Credit Unions.

Financial intermediaries are offered mainly two products: 1) Funding to credit: FIRA provides resources for financial intermediaries so that they are granted to final borrowers, through the FEFA, FUND and FOPESCA trusts. 2) Warranty Service: FIRA, through the FEFA trust guarantees to financial intermediaries the payment of the ultimate borrowers. The intermediaries hire the payment of a percentage of the unpaid balance when the accredited does not fulfill its obligations; this is called the warranty service. (FEFA charges an insurance premium to provide this service).

The problem arises in the guarantee payment procedure, where the following points have been identified: i) it does not assess the probability of default of the accredited based on their characteristics, ii) there is no default probabilities to estimate credit risk (expected and unexpected loss), iii) The guarantee fee is charged depending on the financial intermediary, however, the event that triggers the payment is the accreditor's breach, iv) Lack of knowledge of the probability failure by a borrower does not allow management policies generate credit risk that conserve the assets of an institution and increase its profits.

Generation developed sample: This research was supported with an initial database of 110.278 guaranteed loans, of which 43 were eliminated due to lack of information, being then 110.235 credits for analysis. The guaranteed loans presented 69 different features, applying to each proof of the value of information, as described above to determine their usefulness for the development of the scoring model. The warranty period of the credits is from 2006 to 2011 as shown in the following Table 1.

Table 1: Credit distribution opening year and total sample

Opening date credit		Total sample		
Year	No. Credit	Classification	No. Credit	Percent
2006	5,104	Good	101,751	92.30%
2007	5,504			
2008	13,975	Bad	666	0.60%
2009	28,794			
2010	41,612	Indeterminate	7,818	7.09%
2011	15,246			
Total	110,235		110,235	100.00%

This table shows the total appropriations requested in the 2006-2011 period without any discrimination and classification according to the purpose of the model, this is to determine the characteristics of those accredited financial institution would not have wanted to give them a loan as it is classify them as: Poor, credit that guarantee payment submitted and 90 days did not recover the total. Well, that credit is not granted payment guarantee. Determined, those payments were guaranteed 90 days but had recovered the full amount.

To develop the score table suggests Siddiqi, (2006), that the initial sample is divided as follows: i) 80% of the total credits used as a sign of development, ii) 20% of the total credits used as a test sample. To develop the scoreboard (scorecard) indeterminate credits were eliminated, leaving only the ones defined as good or bad, with these changes the total sample was reduced to 102.417 loans, this in order to give more power to the model.

Of these credits were removed 2 for presenting incomplete information, so the final sample is 102.415 credits. In the following Table 2 shows the selection process.

Table 2: Classification of receivables from development sample and test sample

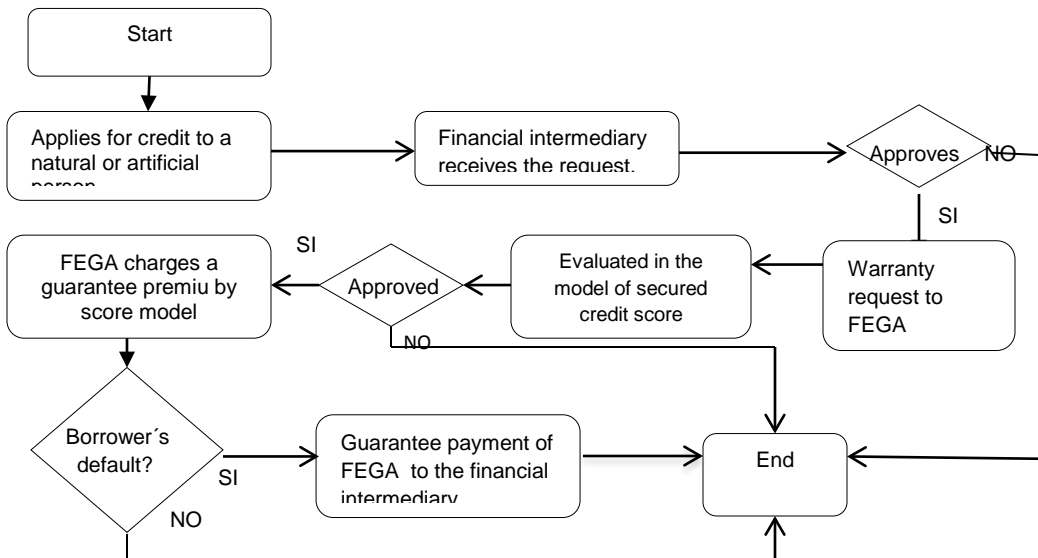
Development sample			Test sample		
Classification	No. Credit	Percent	Classification	No. Credit	Percent
Good	80,651	99.31%	Good	21,098	99.51%
Bad	562	0.69%	Bad	104	0.49%
Total	81,213	100.00%	Total	21,202	100.00%

*This table shows the proposed Siddiqi (2006) applied in our research, is that the initial sample was divided into: 80% of total credits used as "development sample" ($81.213 \text{ credits} = 102.415 * 0.80$) and 20% of the total credits used as "test sample", ($21,202 * 0.20 \text{ credits} = 102.415$). (Total = $102.415 + 81.213 + 21,202 \text{ credits}$).*

Implementation and comparison of models: The loans granted by financial intermediaries in the agricultural sector, with or without funding from FIRA, can be guaranteed by FEGA on payment of an insurance premium against default by the borrower. The payment of the premium is a function of the financial intermediary, leaving aside the credit characteristics of the borrower when the breach of the latter is the one that triggers the payment of the guarantee of FIRA to the IFNB.

The guarantee fee is charged by FEGA based payments and recoveries made by each financial intermediary, ie the price of the premium is based on each broker. No builds on the proven end, who is the trigger of the insurance payment. In addition, credit risk is not associated with the probability of default by the borrower end, in addition to just having a model for risk estimation. Therefore, it becomes necessary to develop a scoring model for lending in the agricultural sector The Figure 2. shows schematically the proposed model to generate a risk management system to solve the credit problems identified in the payment guarantee procedure that is based on a score model for the granting of loans.

Figure 2: Procedure for the guarantee payment considering a score model



This figure shows the model proposed by generating a system of credit risk management to solve the problems identified in the guarantee payment procedure that is based scoring model for lending, the results serve to estimate the risk of collateral portfolio credit risk premia and balance to be collected in the agricultural sector.

Scoring models, both the linear regression and the logistic regression, presents very similar indicators. This is consistent with the described above, (similar results in the center and differences at the ends); but the default probabilities estimated with the logistic regression model in the ends (highest and lowest score) are slightly higher. This is reflected in the estimation models of credit risk. Under the three different approaches-CyRCE, Credit Risk + and Montecarlo- the credit risk based on logistic regression model is superior to the linear regression model. In the case of credit risk estimated with logistic regression, the risk is always greater than the losses observed by the granting of guarantees, but with linear regression the estimated credit risk is less than the loss observed in the models. CyRCE and Credit Risk +, which is indicated by a 1 in the range of Kupiec indicator in the table above.

The premium insurance for risk parameters estimated with logistic regression is higher than the one estimated with linear regression. In both cases the premium is close to a balance, but only under the logistic regression model and Monte Carlo model we have a premium risk that covers all losses really observed. The logistic regression model is, for the target population in study, the most suitable model for the granting of credits in the agricultural sector that allows an adequate credit risk management by

using the Monte Carlo model and having a premium risk of balance, according to the models previously proposed. The following Table 3 below shows the results of applying various models designed to solve the problems initially. This information will allow us to make some conclusions.

Table 3: Variables results

Independent variable	Real Definition	Dimension	Indicators	Linear regression	Logistic regression
Discrimination model	Combination model of variables that best separate two groups depending on their characteristics.	Linear regression model	Ginni Index	0.37	NA
			K-S test	60%	NA
			Divergence Index	2.91	NA
			Information value	2.68	NA
		Logic model	Ginni Index	NA	0.40
			K-S test	NA	62%
			Divergence Index	NA	2.73
			Information value	NA	2.49
Credit risk estimation	Credit risk is the potential loss product of the failure of the counterpart in an operation that includes a commitment to pay.	CyRCE model	Value at risk	1.50%	1.65%
			Kupiec index	1	0
		Credit Risk + Montecarlo model	Value at risk	1.35%	1.49%
			Kupiec index	1	1
Insurance premium estimation	Amount of money paid by the insured to the insurance company in order to have insurance coverage contracted.	Parametric model	Premiums related to credit	0.61%	0.67%
			Estimated recovery relationship and value at risk	0.95	1.05
			Retrospective Test	1	0
Risk benefit ratio	Valuation assessment that relates earnings on invested capital or the value of production resources used and the profit generet	Risk benefit FEGA	Amounts of guarantee payment	1.56%	
			Risk benefit ratio	0.95	1.05

This table shows how the logistic regression model compared to the linear regression model is the most suitable model for lending in the agricultural sector, allowing adequate credit risk management. The combination of the logistic regression model, Monte Carlo model and the model for estimating risk premiums enables compliance with the solution of the problem.

Conclusion

The lack of knowledge of the probability of default of a borrower does not allow to generate policies for management of credit risk that conserve the heritage of an institution as well as to increase its profits, mainly due to the fact that: i) the default probability of the borrower based on their characteristics is not assessed ii) there are no default probabilities to estimate credit risk (expected and unexpected loss), iii) The guarantee premium is charged depending on the financial intermediary; however, the event that triggers the payment is the failure of the borrower.

With this background, it is proposed to create a system of credit risk management that solves the problems found in the warranty payment procedure that is based on a scoring model for the granting of credits, so the results serve to estimate the risk of collateral portfolio credit risk premiums and balance to be collected in the agricultural sector. Under these circumstances, the combination of the logistic regression model, Monte Carlo model and the model for estimating premium risks allow to reach the expectations raised in the general and specific hypothesis.

References:

- Bousoño, C., Heras, A., & Tolmos, P. (2008). *Risk factors and calculation of premiums by learning techniques. Insurance Institute of Science*, 180.
- CNBV. (2010). *General Provisions for Credit Institutions*. Mexico City, Mexico.
- Fisher, R.A. (1936). "The Use of Multiple Measurements in Taxonomic Problems". *Annals of Eugenics* 7 (2): 179–188
- Greene, W. (2011). *Econometric Analysis*. USA.
- Gujarati, D. (2006). *Principios de econometría*. España: McGraw Hill.
- Gujarati, D., & Porter, D. (2010). *Econometría*. México: McGraw Hill.
- Gutiérrez, J., & Elizondo, A. (2002). *Riesgo de Crédito: el enfoque actuarial*. Ciudad de México.
- Hull, J. C. (2007). *Risk management and financial institutions*. United States of America: Pearson Prentice Hall.
- Márquez, J. (2009). *Una nueva visión del riesgo de crédito*. Distrito Federal, México: Limusa.
- Novales, A. (1993). *Econometría*. Madrid: Mc Graw Hill.

- Rencher, A. (2002). *Methods of Multivariate Analysis*. Estados Unidos de América: Wiley Press.
- Sandica, A. M. (2010). Credit Scoring Modelling: A micro-macro approach. *Academy of economic studies. Doctoral school of finance-dofin* , 110.
- Siddiqi, N. (2006). *Credit Risk Scorecards: Developing and implementing intelligent Credit Scoring*. New Jersey: John Wiley & Sons, Inc.
- Thomas, L., Edelman, D., & Crook, J. (2002). *Credit Scoring and its applications*. United States of America: SIAM.