FINANCIAL COMPANIES REGULATED OF MULTIPLE PURPOSE IN MEXICO; AN ALTERNATIVE MODEL TO MANAGE THE PROBABILITY OF DEFAULT OF REVOLVING LOANS

Trejo García, Jose Carlos; Martínez García, Miguel Ángel; Carrillo García, Mayra Daniela

FINANCIAL COMPANIES REGULATED OF MULTIPLE PURPOSE IN MEXICO; AN ALTERNATIVE MODEL TO MANAGE THE PROBABILITY OF DEFAULT OF REVOLVING LOANS

EMPRESAS FINANCIERAS REGULADAS EN OBJETOS MÚLTIPLES EN MÉXICO; UN MODELO ALTERNATIVO PARA GESTIONAR LA PROBABILIDAD DE INCUMPLIMIENTO EN CRÉDITOS REVOLVENTES

Jose Carlos Trejo García 1 / jtrejog@ipn.mx
Escuela Superior de Economía del Instituto Politécnico Nacional, Mèxico
Miguel Ángel Martínez García 2 / mmartinezga@ipn.mx
Escuela Superior de Economía del Instituto Politécnico Nacional, Mèxico
Mayra Daniela Carrillo García 3 / danielita_051191@hotmail.com
Escuela Superior de Economía del Instituto Politécnico Nacional., Mèxico

Abstract: This research suggests a particular probability of default model to Microfinances Regulated in Mexico (SOFOM ER), based on the use of own mixed variables (continuous and dichotomous) according to the credit behavior and the optimization by maximum likelihood. Financial indicators such as savings, assets and profits were used nationwide to reaffirm an implementation of new and more profitable model in the sector of Mexico, indicating a yield of about 4% with the optimized credit model. Within the limitations of the research lies the lack of robust information and free consultation, due to bank secrecy in Mexico.
Keywords: econometric methods, estimation models, optimization techniques, banks and credit.
Resumen: En esta investigación se propone un modelo de probabilidad de incumplimiento particular para las Microfinancieras reguladas en México (SOFOM E.R.), con base en el uso de variables mixtas (continuas y dicotómicas) propias del comportamiento del propio crédito y el uso de un modelo crediticio optimizado por máxima verosimilitud. Indicadores financieros como ahorro, activos y utilidades, fueron utilizados a nivel nacional para reafirmar una implementación de nuevo modelo inmediato y rentable en el sector de México, indicando un rendimiento de 4% aproximadamente con el modelo crediticio optimizado. Dentro de las limitaciones de la investigación radica la falta de información robusta y libre de consulta, debido al secreto bancario en México.
Palabras clave: métodos econométricos, modelos de estimación, técnicas de optimización, crédito y banca.

INTRODUCTION

This research shows the need to improve the credit risk management of Regulated Microfinance (IMFRs) in Mexico, also called Financial Companies Regulated of Multiple Purpose (SOFOMERS). This kind of improvement implies estimations of probability of default and expected losses of their consolidated revolving loans. This sector is mainly engaged in lending to mid-level people, in order to support this sector of the population.

Being the Expected Losses the principle issue in credit risk management, this research proposes a mathematical model to particular variables for SOFOMERS, estimating a Probability of Default as an alternative model versus the National Banking and Securities Commission (CNBV), which is specifically applicable to bigger institutions of credit in Mexico.

The estimation of credit risk over the last four decades has been one of the fastest growing areas in statistics and probability models. The techniques used for its management have been Credit Scoring and Behavioral Scoring; applications for the estimation of financial risk in relation to retail loans. In the Mexican retail credit market, the portfolio is divided into three types of pools; Revolving Loans, Personals Loans and Mortgage[1].

All loan applicants and clients consolidated of all those financial institutions in Mexico, have a qualification in the origination loan application and in the behavior of each loan.

There are two important objectives in techniques for the credit rating, one is the need to identify the risk of consumption, while the second one is to minimize the percentage of defaulted customers, and thereby banking and credit institutions optimize their portfolios for a good and better business.

In this case, it is necessary to develop new and better models based on historical customer information, helping in generating models to the origination and behavioral loans to enhance and develop greater credit business with consolidated clients (Mays, 2004).

According to CNBV [2] , the credit cards issued by SOFOMERS in Mexico represents almost 44% of revolving cards issued by banking institutions.

This research makes a contribution for adopting predictive models of expected losses, particularly for IMFRs, promoting the development and updating the rating techniques of credit losses in Mexico and considering the particular local and international regulation, also identifying positive impacts to loan loss provisions and helping to give a positive effect to assets, net income and profitability of SOFOMERS.

The especial contributions of this research is to consider the literature on credit risk management, contributing value added for IMFRS. The rest of the paper is organized as follows.

Reviewing the sector and regulation about SOFOMERS in Mexico, the Section 2 contributes to have the main elements of managing

credit risk. In Section 3, the structure and local regulation are really important to know them, in order to comply with legal rules about the business of revolving loans in Mexico. In Section 4, it is explained the methodology of the calculation of a particular Probability of Default, considering all mathematical concepts to develop a particular model for SOFOMERS. In this way, Section 5 is the explanation of dataset and methodological optimization used to choose the better variables according to the assumptions established in the previous section, while Section 6 presents projections of savings, profits and returns under the optimization developed with the model proposed in probability of default of SOFOMERS in Mexico, extending the analysis from previous sections. Finally, Section of Conclusions provides concluding remarks, reflecting high compliance in the approaches of this research.

BACKGROUND

Prior to the eighties of the twentieth century in some developing countries, governments were the agent to give credit to small farmers. Most of these were authorized and assigned to representatives of rural populations for administration, but many times never reached all farmers.

In other countries, only farmers requested the local moneylender loans, paying disproportionate interests. As capital requirements of a population were always insufficient and substantial losses (Robinson, 2000 pp. 15-34).

In the seventies, in Bangladesh, Dr. Mohammed Yunus (Nobel Peace Prize in 2006), encouraged the social and economic development with group lending scheme for landless, currently the institution is the Grameen Bank, with more than 7 million customers.

The concept of IMFRs was created in order to the sectors of the low-income population without access to common financial services in order to stimulate their productive activities.

The statistical advances and data used through analytical models allow to financial institutions to automate decisions on acceptance or rejection of credit applications, as well as the management of a loan portfolio (consolidated clients) on a cross-selling. Few years ago, the credit management was only on experience or with an executive expertise.

Nowadays, one of the most used models for evaluating credit is the Scoring Model which determines a score for a client requesting a loan or for a client with the possibility of defaulting on its partial payments.

The literature regarding the type of credit rating was developed by Rosenberg & Gleit (1994); the option pricing model with Merton (1974), with similar themes Hand and Vinciotti (2003); less important but equally in credit terms as Thomas Crook & Edelman (1992); Lewis (1992); while credit issues and a combination of Behavioral Scoring with Elizabeth Mays (1998).

Analytical models for credit risk management have been extensively studied and used by countless areas such as medicine, sociology and of

course by financial institutions for over 50 years, showing successful results (Thomas 2004).

Analytical Scoring Models can be defined as a set of quantitative methods and techniques used to predict the probability of default of each client, a kind of risk that a banking institution has of non-recovering the loans.

The Scoring Models are tools that use the classification of each applicant or customer consolidated by risk level, based on the historical information from credit applications and credit reports on payment.

Identifying those consolidated or potential customers which no generate a business entity, it is necessary to adopt guidelines on best treatment and provisions in case of default and / or migration from buckets of default to buckets as write-down, such is the case with the investigations of Hsia (1978), which describe the importance of scoring methods. In other way, researches of computed data are suggested by Reichert, Cho and Wagner (1983), Joanes (1993) Hand and Henley (1997); detailing the study of particular relationships between the distributions of good and bad customers with the ability to accept and reject people with similar features.

The history of the scoring models back when Fisher (1936) introduces the idea of discriminating different groups within a specific population. This idea was developed more fully by Durand (1941), applying financial context to discriminate between "Good" and "Bad" payers (Thomas, 2004).

In relation to the creation of Scoring Models, Bill Fair and Earl Isaac[3] in the late fifties began with the development of an analytical system for credit risk analysis. In the 60s with the creation of new financial instruments, such as: credit cards, enhanced scoring models and showed their real importance and usefulness.

According to Myers (1963) this type of model is superior as a predictor than any qualitative expert judgment. Another important milestone in this context was the development of Z-Score proposed by Altman (1968) which has been applied in many companies in the financial sector.

COMPLIANCE

Currently, the Mexican Financial System is regulated directly and indirectly by the Secretariat of Finance and Public Credit (SHCP), with Federal Executive dependence, through the relevant committees; National Banking and Securities Commission (CNBV), National Commission of Insurance and Bonds (CNSF), National Savings System Commission for Retirement (CONSAR) and the Central Bank or Banxico (Banco de Mexico).

Based on Banxico Statistical Reports in October 2014, the Banking Sector figures as the main lender Mexican society sector with a share of 88.5% at the end of June 2014, transforming the people deposits (savings) into source of public financing.

With the growing demand for credit in the last decade for different types of population, the Non-Banking Sector of complementary services (parallel), better known as ancillary credit business; represents 10.2% compared to Gross Domestic Product in Mexico. In addition, SOFOMERS as credit auxiliary, has reflected an annual growth of 11.6% in the last three years.

Moreover, information issued in November 2014 by the CNBV[4], at the end of June 2014 a major credit instruments in Mexico by SOFOMERS are Credit Card with approximately 11 million of plastic authorized by 7 regulated entities recognized[5], representing almost 44% of revolving loans issued by commercial banks.

All this implies the need for adopting predictive models of expected losses, particularly in SOFOMERS sector.

Based on these statistics, this research includes information only for IMFRs that provide loans through SOFOMERS, which are regulated by the CNBV.

In July 2006 the General Law of Organizations and Auxiliary Credit Activities was amended to create SOFOMERS that have the aim to provide credit for the purchase of various goods or to fund various activities without being limited to a sector or activity in particular.

The lending by SOFOMERS is not based on attracting funds from the public in deposit accounts or any other type of account that represents an obligation to repay the funds raised and the interest generated.

Therefore, the CNBV has the power to issue prudential regulation to limit excessive risk by SOFOMERS to avoid performing operations that could lead to lack of liquidity or solvency. Management of credit risk by estimating the expected loss (EL) is indicated in the CUIFE[6] and which mentions regulation CUB[7].

Although some financial institutions in Mexico develop their own scoring models based on guidelines established by the CNBV[8], most of the entities make provisions based on guidelines established by the own commission.

Vendors or consultants create scoring models based on the information and the specific parameters provided by the areas of Credit Risk Management. Therefore, these institutions must have the full information and ensure that the development of such models is consistent with the objectives of the risk appetite adopted by the Bank.

The international agreement on banking regulation and supervision called "Capital Accord"[9], requires to financial institutions of member countries, a review of its capital injections to cover the risks.

Those parameters require to have tools that allow them to establish measurement models (scorings and ratings), discrimination of clients based on their risk profile, monitoring systems and current risk models of exposure at default and loss given default of each loan.

Basel deal also requires to financial institutions not only to adapt their systems of calculating capital consumption, also to modify reporting systems (financial statements) and information analysis. Both elements, i.e. financial reporting and data analysis are keys to manage large databases,

capable of providing accurate information to quantify the risks of each operation, a real challenge for Banks.

So the key to analyzing the risk rating process, element using internal models (IRB[10]), accordance with the rules of Basel a credit scoring model allows to measure the probability of default. For this purpose, the CNBV's standard method enables to validate the current data and obtain results by an IRB method of behavioral scoring.

In general, the credit is an agreed compromise between a person and a financial institution, this institution predicts the debtor capacity payment (individual or company).

The credit allows to buy and to satisfy the needs of consumers, paying interests and the cost of access to financing. In Mexico, the CNBV mentions that all credit activity means placing monetary resources through operations of loan with discounts and considering a credit risk level based on collaterals, generating a receivable in favor and a risk of default[11].

Speaking very generally of Retail Credit, it can be described from the point of view of risk management, as a cycle in which every consumer is involved; Origination, Management and Recovery.

In the Retail Credit Portfolio are all direct loans, including liquidity that have no real guarantee, denominated in domestic currency, foreign, UDI (investment units) or VSM (times of minimum wage) and the interest generated to individuals, derivative from credit card operations, personal loans, credit for the Acquisition of Consumer Durables (known as ABCD), which includes the automotive credit and leasing transactions that are held with individuals[12].

Based on various investigations of Banxico, the credit risk is particularly the case when the contract cannot be fulfilled by the debtor to the creditor (grantor of credit).

Recently, besides the case of default, are incorporated events affecting the value of a loan, without necessarily means default from the debtor. Normally, the factors to be taken into account when is analyzed the credit risk are: the probability of default and / or migration in the credit quality of the obligor, the correlations between failures, concentration or portfolio segmentation, exposure to each debtor and the recovery rate in case of default by the debtor[13].

For the determination of provisions in Mexico for retail market and managed by each financial institution, it is taken from the regulation imposed by the CNBV through the CUB[14].

Factors such as Loss Given Default (LGD), the loss of the creditor in case of default of the debtor, after taking into account all the costs involved in recovery (recovery costs, court costs, etc.).

Regarding to the Exposure at Default (EaD) is the balance owed by the borrower at any given time in case of default.

While the credit risk factor analyzed in this research is the "Probability of Default" (PD), which is a measure of how likely it is that a borrower defaults with its contractual obligations. Its minimum value is zero,

indicating that it is impossible to default on its obligations, and its maximum value is one or one hundred percent when defaults at all.

Based on current regulations, the estimate should be based on credit and risk factors mentioned above. It is defined that the Provisions (P) or Expected Losses (EL) is the mean of the distribution of income, i.e., indicates how much can be the loss on average and is usually associated with the policy of loan loss provisions that the institution should have against credit risks.

With last considerations, the EL is calculated by considering the percentage of reserves (PD * LGD) multiplied by EaD[15].

METHODOLOGY

In practical terms, Scoring models allow a significant reduction in execution times of various financial processes for granting and monitoring credits, thus allowing a greater automation and drastically reducing the need for human intervention in the evaluation and assessment of credit risk.

The main users of those models are banks, financial institutions and insurance companies. Key features they have in common are the ability to manage the risk; the handle significant amounts of capital, small reductions in the risk of the portfolio and increases in profitability.

Benefits reported by the application of those models not only affect banks and financial institutions, also directly to all customers in the financial sector by reducing the erroneous discrimination of customers requesting a loan and provides a more objective analysis, being important to incorporate significant variables, concentrating on a single model of multiple factors that can predict the risk of an application or tracking payments.

The data collected and processed for analysis was according to the ideas of Weber (1999), in order to build a particular model of Probability of Default (PDC) for the product of revolving loans managed by SOFOMERS, all based on the local regulation; CUB, Article 92 and Section III.

The representative sample and used in this research was with 3,102[16] clients consolidated as "goods" and "bads" revolving loans managed by SOFOMERS. In this sense, the model developed predicts the probability that a customer does not perform two or more consecutive payments[17] over the next 12 months, calculating the probability of recovering.

The statistical technique used is the discriminant analysis and logistic regression as the CUB adopts for validation, proposing an alternative model to adopt better significance tests.

The necessary variables to work are quantitative and dichotomous, the dependent variable has these assumptions Y_i; 1 as bad loan or a Default ≥ 2 months late, and 0 or good loan if Comply < 2 months late.

The statistical technique most commonly used by the financial industry corresponds to the logistic regression (Thomas et al., 2002). This technique is less restrictive, being an alternative to the use of discriminant

analysis. In recent years, there have been a number of new techniques called Data Mining (Weber, 1999), which have also been used to construct Scoring models.

These techniques have the advantage of not having too many requirements and assumptions for the input variables, increasing its validity. Additionally, it has been extensively used for building Scoring Models for understanding complex patterns of a particular customer segment, having the ability to model nonlinear relationships between variables.

The linear probability models have a number of problems that have led to the search for alternative models with better estimations of dichotomous variables.

To prevent the estimated endogenous variable can be outside the range (0, 1), the available alternatives non-linear probability models, which the function specification used guarantees a result in estimated within the 0-1 range.

Since the use of a distribution function, this ensures that the estimation result is bounded between 0 and 1, in principle there are several possible alternatives, the most common is the logistic distribution function, which has led to the logit model, relating endogenous variable Yi with the explanatory variables Xi through a distribution function in order to study (Yi = 1).

Once estimated parameters, the interpretation of a logit model implies to note the sign, which indicates the movement direction that the probability increases with the movement of its explanatory variable, however, the amount of the parameter does not match the magnitude of variation the probability.

Through linearization of the model and based on the general equation of Logit Model, yi is defined as the probability of state or alternative 1:

$$E(Y_i) = Prob(Y_i = 1) = p_i = \frac{e^{\beta_0 + \beta_k X_k}}{1 + e^{\beta_0 + \beta_k X_k}} \qquad (1)$$

(1)

For obtaining the maximum likelihood of β estimators, it is derived L(β) for each βj parameters with j= 1, 2, ..., p and equate to zero. In terms of matrix

$$\frac{\partial L(\beta)}{\partial L(\beta_0)} = \sum_{i=1}^{n} y_i \, x_i - \sum_{i=1}^{n} x_i \left(\frac{e^{\beta_0 + \beta_k X_{ki}}}{1 + e^{\beta_0 + \beta_k X_{kj}}} \right)$$

$$
\begin{bmatrix} \dfrac{\partial L(\beta)}{\partial L(\beta_0)} \\ \dfrac{\partial L(\beta)}{\partial L(\beta_1)} \\ \vdots \\ \dfrac{\partial L(\beta)}{\partial L(\beta_j)} \\ \vdots \\ \dfrac{\partial L(\beta)}{\partial L(\beta_p)} \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^{n} y_i\,(1) \\ \sum_{i=1}^{n} y_i\, x_{i1} \\ \vdots \\ \sum_{i=1}^{n} y_i\, x_{ij} \\ \vdots \\ \sum_{i=1}^{n} y_i\, x_{ip} \end{bmatrix} + \begin{bmatrix} \sum_{i=1}^{n} (1)\left(\dfrac{e^{\beta_0+\beta_k X_{ki}}}{1+e^{\beta_0+\beta_k X_{kj}}}\right) \\ \sum_{i=1}^{n} x_{i1}\left(\dfrac{e^{\beta_0+\beta_k X_{ki}}}{1+e^{\beta_0+\beta_k X_{kj}}}\right) \\ \vdots \\ \sum_{i=1}^{n} x_{ij}\left(\dfrac{e^{\beta_0+\beta_k X_{ki}}}{1+e^{\beta_0+\beta_k X_{kj}}}\right) \\ \vdots \\ \sum_{i=1}^{n} x_{ip}\left(\dfrac{e^{\beta_0+\beta_k X_{ki}}}{1+e^{\beta_0+\beta_k X_{kj}}}\right) \end{bmatrix}
$$

Expressing each of these derivatives in a column vector

$$
\frac{\partial L(\beta)}{\partial L(\beta_0)} = \sum_{i=1}^{n} y_i\, x_i - \sum_{i=1}^{n} x_i \left(\frac{e^{\beta_0+\beta_k X_{ki}}}{1+e^{\beta_0+\beta_k X_{kj}}}\right)
$$

Resulting a logit model of probability

$$
PD = \left(\frac{e^{\beta_0+\beta_k X_{ki}}}{1+e^{\beta_0+\beta_k X_{kj}}}\right) = \left(\frac{1}{1+e^{-(\beta_0+\beta_k X_{kj})}}\right) \qquad (2)
$$

(2)

To avoid problems of collinearity, it is helpful to apply correlation tests to selected variables, thus avoiding redundant information (collinearity), which can distort the predictive ability of the estimated and discriminant function.

Moreover, the collinearity does not include highly significance to explanatory variables given by poor estimation of the parameters.

That is the calculation of

$$
R^2 = 1 - \frac{D(\beta)}{D(\beta_0)} \qquad (3)
$$

(3)

OPTIMIZATION

The PD regulatory model of revolving loans has 6 variables. The set obtained for this research includes 3,102 established customers and managed by the Mexican financial institutions as of June 2014.

The variables analyzed according to the regulatory methodology of CNBV[18] are: Aim (Y), Number of Defaults (ACT), History of Defaults (HIST), After Months Credit (ANT), Pay-Balance Ratio (%PAGO) and Balance Payable-Credit Limit Ratio (%USO).

While the PD model determined with the same sample set has 5 variables as a proposal for this research of revolving loans managed by SOFOMERS in Mexico at the end of June 2014.

The variables proposed and based on the CNBV[19] are: Aim (Y), Credit Limit (limite_credito), History of Defaults (HIST), Pay-Balance Ratio (%PAGO) and the Costumer Payment (PAGO1).

The 5 variables proposed in this research, and analyzed econometrically, show no collinearity because the levels of correlation among the independent variables are less than 41%; so the relation of limite credito, %PAGO y PAGO1 with the dependent variable (Y) is inverse and more than 50%.

The impact of HIST with Y is direct, positive and greater than 92%. This correlation test shows that the proposed model could be predictive.

A good model must satisfy two conditions, the first is to have a strong predictive capacity and the second is that the parameter estimation has high accuracy. An additional condition is that the model must be simple, i.e. containing a minimum of explanatory variables and satisfies the above two conditions.

Using the sample of 3,102 customers or loans and the proposed variables, a Logit model explained in the methodology is estimated by calculating the probability that a customer do not pay its loan.

So, the proposed model of Probability of Default, one particular for SOFOMERS sector would be:

$$PD = \left(\frac{1}{1+e^{-(-7.1423-0.099*limite\_credito+2.7542*Hist-3.1066*\%pago-0.0624*pago1)}} \right) \quad (4)$$

(4)

The estimation of the Logit Probability expressed in equation 4, requires 5 iterations to estimate the model. The LR chi2 function indicates that the coefficients are jointly significant in explaining the probability that 3,102 customers are in default, then it is correct to say that the statistical value of Prob> chi2 indicates that it could be rejected in 1% the hypothesis that all coefficients are zero.

Moreover, the Pseudo R2 statistic indicates that approximately 91% of the variation in the dependent variable is explained by the variation

of the independent variables of the model proposed for Mexican case of SOFOMERS.

The quality of model adjustment is 98.22%[20], the result of the ratio of correct provisions against number of observations, stating in general that the model predicts correctly the PD.

In this sense, it is necessary to mention that in the moment to perform the logistic regression for those variables currently applied by the regulation, it is obtained an error with the result of the regression, due to the high collinearity of the ACT variable with HIST one, indicating that the explanatory variables and currently used by the CNBV are not needed to construct the model, so the ACT variable must be excluded.

Using statistical criteria such as partial correlations could help get all possible variables and choose the best between them.

To determine whether a variable should be included in a model because it has a significant weight, the Wald test statistic it is applied. The test results from testing the null hypothesis $H_0: \beta_i = 0$, against the alternative $H_1: \beta_i \neq 0$.

Significance levels for the proposed model indicate that all p-values (values of significance) have high significance (p-value = 0.0000 <5%), although the intercept the research allows in future to include better variables.

While the score obtained, with the proposed variables, the logit transformation of Equation 16 is applied to obtain the probability of default calculated (PDC).

The interpretation of the coefficients is by measuring the variation of the estimated logit model for a unit change in the explanatory variable given. So, if HIST increases in one unit of default, the estimated logit increases in 2.75 units, suggesting a positive relationship to the PDC.

Otherwise, if limite de credito, %PAGO y PAGO1 is greater each time, so the PDC decreases, due to higher solvency and/or payment capacity. Thus, the credit theory is applied correctly in the econometric regression posed by logit estimation. The opportunities in the logit models are calculated using logistic regression or odds, i.e. by an antilogaritmical transformation of $\beta \# e^{\beta}$.

Figure 1 indicates generally a good distribution and minimal observations outside the normal behavior from other observations.
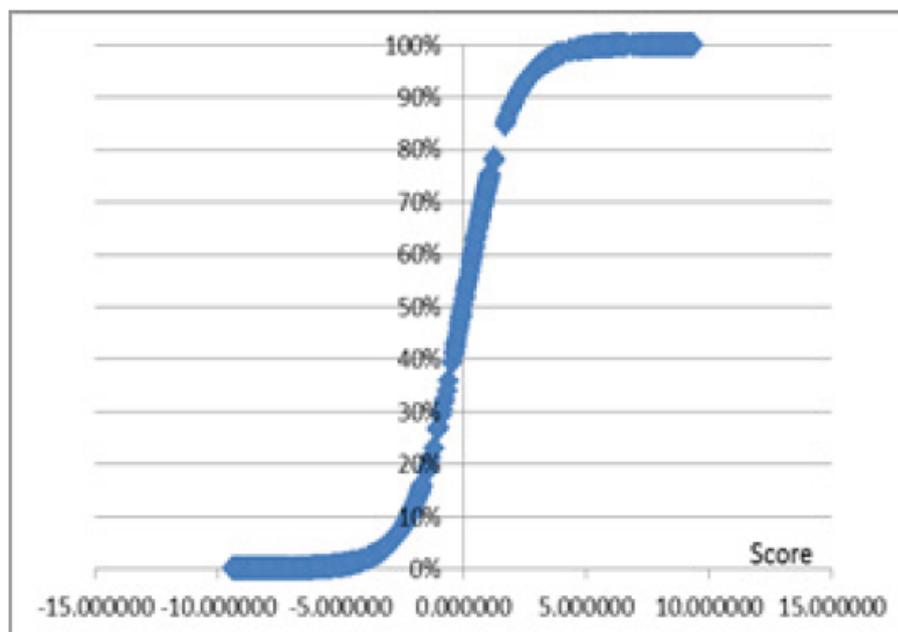
Figure 1

Score Logit Distribution and Probability of Default, Proposed Model for Revolving Loans

Source: Prepared with the SAS Enterprise Guide 7.1 package, based on logistic regression and logit distribution

The figure above shows an optimum distribution, because in most of the observations there are no elements outside the normal behavior.

The average PD obtained by the CNBV methodology is 42.7%, however it is interesting to validate the model proposed (PDC) in this research, if there is an optimum logit distribution and with attachment to the theorical behavior. The average PDC is 52%, i.e. 9.3% more than the PD of the current regulation (CNBV).

Also, the results obtained show that the PD according to the current methodology of CNBV is higher in A-1 and C-1 risks levels than proposed model, according to the category of bad customer (more than two months of defaults).

With the model proposed were obtained results of PDC higher in most of risk levels, due to the SOFOMERS sector as part of IMFRs is more risk than big banking institutions in México, indicating that the proposed model for SOFOMERS has no relation to the model of PD used by big banks.

So, with the proposed model any bad revolving loan managed in SOFOMER sector, it would be classified better and more realistic than the current model of PD issued by CNBV.

Thus, the PDC calculated with the proposed model implies greater coherence and better performance. Therefore, it is necessary to validate the discrimination level between good and poor populations' clients, using the model proposed in this research.

Table 1
Probability of Default Average, Good and Bad Revolving
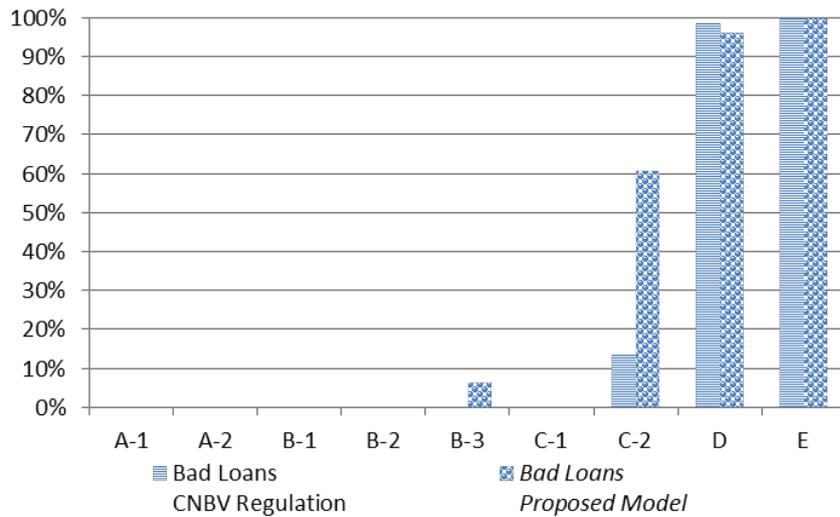Loans by Risk Levels, CNBV versus Proposed Model

Risk Level	Good Loans Current Regulation	Bad Loans Current Regulation	Good Loans Proposed Model	Bad Loans Proposed Model
A-1	1.65%	0.00%	0.15%	0.00%
A-2	4.00%	0.00%	4.08%	0.00%
B-1	5.60%	0.00%	5.88%	0.00%
B-2	7.17%	0.00%	7.55%	0.00%
B-3	8.84%	0.00%	9.40%	0.00%
C-1	12.60%	100.00%	11.34%	100.00%
C-2	22.68%	100.00%	25.96%	100.00%
D	46.66%	100.00%	57.95%	100.00%
E	0.00%	99.91%	0.00%	100.00%

Source: Prepared with the SAS Enterprise Guide 7.1 package,
CNBV PD (2014) and Proposed Model of PDC with Risk Levels

By population testing discrimination, this is a multivariate technique to study simultaneously the behavior of an independent group of variables with the intention to classify a number of cases previously defined and exclusive groups together.

Once calculated the PD y PDC variables, these are used to be multiplied with LGD[21] variable, obtaining in this way the loan loss provisions percentage.

Validating by current methodology (PD*LGD) against the proposed model (PDC*LGD), the Percentage and Risk Level in the revolving portfolio behavior shown in the CNBV[22] methodology, it is identified that the proposed model has better discrimination of good and bad loans in every risk level.

Figure 2

Discrimination of Bad Revolving Loans by Risk Level Current Model versus Proposed Model

Source: Prepared with the SAS Enterprise Guide 7.1 package, based on the
Percentage Distribution of Expected Losses and Risk Levels, CNBV (2014)

Therefore, it is shown with the proposed model the detection of bad loans is greater than the current model used by the CNBV, ensuring better and greater discrimination of bad or defaulting loans, as shown notably the B-3 and C-2 risk levels. While risk groups D and E, maintain their discriminatory trend between the two models.

Once the score was calculated with model proposed of PDC with the logit regression formula, it is necessary to know if calculated values of each loan used with the sample are well identified in the respective risk level group.

As greater is the difference of goods and bad loans, it will be greater the discriminatory power of the model proposed (PDC).

The Kolmogorov-Smirnov (KS) is a nonparametric test for goodness of fit, proving that two independent samples come from the same distribution. These differences are determined not only by the mean, also by the dispersion or the skew symmetry.

The test is based on the null and alternative hypotheses as follows:

H0: The distribution score for good and bad loans is equal.

H1: The distribution score for good and bad loans is different.

$$F_n(x) = \frac{1}{n} \sum_{1=1}^{n} I_{X_i \le x} \quad (5)$$

$$D_n = \sup_x |(F_n(x) - F(x)) \quad (6)$$

Where x the is the superior distance. In this way, it can be compared the discriminatory KS index between the separation level of good and bad loans with the Current or Regulatory Model (PD) versus the Proposed Model (PDC). In this way, it can be compared the discriminatory KS

index between the separation level of good and bad loans with the Current or Regulatory Model (PD) versus the Proposed Model (PDC).

Therefore, it is shown with the proposed model the detection of bad loans is greater than the current model used by the CNBV, ensuring better and greater discrimination of loans. Table 2 shows the discrimination power among the risk levels by the percentage of loan loss provisions (PDC * SP) is greater in C-1 (49.12%) with the proposed model, which allows to the SOFOMERS sector in Mexico the ability to identify better the good and bad loans, applying more effective collection policies to improve that risk level.

Table 2.
Population Separation Level by Good and Bad Revolving
Loans, Current Model (CNBV) versus Proposed Model

Risk Level	K-S Separation Current Methodology	K-S Separation Proposed Model
A-1	30.50%	46.80%
A-2	38.46%	44.42%
B-1	42.32%	45.05%
B-2	43.77%	44.96%
B-3	44.83%	49.50%
C-1	47.14%	**49.12%**
C-2	**49.02%**	46.15%
D	3.80%	3.80%
E	0.00%	0.00%

Source: Based on the methodology of Kolmogorov-Smirnov

It even allows the Retail Risk Unit of each SOFOMER, the possibility to propose an optimal level of cross-selling with costumers located in higher risks levels (betters) to C-1, representing greater chance of business for the benefit of a better and more profitable business in retail loans.

The model based on logistic regression consists of two basic elements: a set of weights related to a group of variables or attributes that characterize a set of customers and a threshold or cut-off. With the set of weights is possible to characterize the patterns that describe both populations (good and bad loans) and determine which of the input variables used are really important in terms of risk levels for a good prediction.

Regarding the cut-off, it determines the border between being a "bad" or "good" customer. In general, the Mexican banking industry has the cost of accepting a bad customer is in several times greater than rejecting a good customer.

As shown in Table 2, the Proposed Model shows a separation index of 49.12% in the risk level C-1, greater than the one determined with the Current Model (49.02%). This allows to determine the optimum risk appetite in the Mexican SOFOMERS sector, i.e. how is disposed the sector to accept to win in an assets level or Return on Assets (ROA[23]), according to the cost of allowing to have loans in risk level C-1 with the proposed model.

EVALUATION OF CURRENT MODEL VERSUS PROPOSED MODEL

The loan loss provisions are created considering the outstanding balance recorded on the last day of each. According to the information from the National Banking and Securities

Commission (CNBV[24]), at the end of June 2014, one of the main credit instruments in Mexico are Credit Cards. There are approximately 23 million of plastics authorized by 24 recognized financial institutions.

SOFOMERS have authorized approximately 11 million plastic, i.e., 44% from the revolving loans managed by big banking institutions in Mexico. This implies an annual growth over the past four years of 11% on average.

Nowadays, in Mexico there are three SOFOMERS which comprise approximately the 93% of the market in revolving loans; Banamex Cards, Santander Consumer and Ixe Cards.

The loans classified as write-off at the end of June 2014, were about US $ 8269.6 million of Mexican pesos (MM), i.e. approximately 5.14% of the total of SOFOMER's outstanding amount and accounted by CNBV.

In Mexico, approving a credit card to people with incomes between MXN $5,000 and $17,000 pesos, does not ensure a constant payment in all cases, having higher probability of default in periodic payments.

Following up to the current Mexican regulation in relation to calculate loan loss provisions for all credit cards, the CNBV's statistics show at the end of June 2014 a total amount of MXN $23,479.5 million.

The SOFOMER's loan loss provisions have been in annual growth over the last twelve months by 8%, based on increases of: the probability of default, the loss given default (loss severity) and growth in the monetary of the exposure at default.

This situation is increasing, affecting net income flow and financial assets of all regulated microfinance, reflecting a ROA in national level of 3.50%.

For the subject of revolving credit risks, the CNBV requires to rank the loan loss provisions in an accounting way; as the concept of Asset, in subaccounts named as Preventive Estimation of Credit Risk.

In order to obtain the Expected Losses (EL=PD* LGD*EaD) in credit risk management, in this research was considered the current regulation (CNBV) and the results obtained with the data information of 3,102 sample in revolving loans managed by SOFOMERS; as well as the particular estimation of expected losses with the proposed model (PDC) in this research. See Table 3.

Table 3.

Savings Comparison on Expected Losses of Revolving Loans in the Mexican SOFOMERS sector; CNBV versus Proposed Model at the end of June 2014

	Expected Losses (MXN)		Savings with the model proposed of PD_C	
	Current Methodology (CNBV)	Proposed Model	Differences (MXN)	% Saving
Sample: 3,102 loans	$10,830,996	$10,274,279	$556,717	3.52%
National Level: SOFOMERS in Mexico	$23,479,452,839	$22,653,768,296	$825,684,543	3.52%

Source: Based on the calculation of the Expected Losses Percentage and National Information with data as of June 2014, CNBV (2014)

In Table 3, the amount of EL obtained with the sample of 3,102 in revolving loans, according to the proposed model (PDc), it was MXN

$ 10.3 million versus the current methodology of MXN $ 10.8 million. So, the Mexican SOFOMERS sector could have had savings of MXN $ 556,700 (3.52%) at the end of June 2014. If this trend would have continued as at the system, the savings would have been MXN $ 825.7 million. This implies to analyze the effect of ROA at the end of June 2014, which was a 3.50% of return in the Mexican SOFOMERS sector.

Table 4.
Assets and EL Performance of Revolving Loans in the
Mexican SOFOMERS sector at the end of June 2014

	Total Assets CNBV Balance	Expected Losses CNBV Methodology	Expected Losses Model Proposed (PD$_C$)	ROA %
Sample: 3,102 costumers	$109,969,686	$10,830,996	$10,274,279	**3.50%**

Source: Based on the calculation of the Expected Losses Percentage
and National Information with data as of June 2014, CNBV (2014)

However, considering the savings of SOFOMERS in calculating EL, this implies the possibility of considering the Proposed Model (PDC); the Net Income reported by the CNBV at the end of June 2014 would have been MXN $ 7,607 million, increasing MXN $ 1,646.5 million. In the same way, the Balance Assets in nationwide would increase the ROA from 3.50% to 4.44%, representing higher return to the Mexican SOFOMERS sector at the moment of considering an updated model of Probability of Default. See Table 5.

Table 5.
Net Income and Assets Performance of Revolving Loans in
the Mexican SOFOMERS sector at the end of June 2014

	Net Income CNBV	Net Income with Proposed Model (PD$_C$)	Assets with Proposed Model (PD$_C$)	ROA % with Proposed Model
Sample: 3,102 costumers	$4,019,489	$5,084,972	$114,526,403	**4.44%**
National Level: SOFOMERS in Mexico	$5,961,432,009	$7,607,988,656	$171,340,873,335	**4.44%**

Source: Based on the calculation of the Expected Losses Percentage
and National Information with data as of June 2014, CNBV (2014)

CONCLUSIONS

When the financial risk is mentioned, it is always normal to think about portfolio management, financial instruments and various valuation techniques, such as Brownian motion and Black - Scholes (1972) formula, or bonds valuation by Merton technique (1974). While market risk is the potential loss and the most important financial risk; also credit risk is so important in the financial risk management in any institution, especially when the retail lending business is contributing over the half of the financial institutions proceeds.

This research was focused on the analysis in the business of retail credit risk for the Mexican SOFOMERS sector, especially revolving loans, commonly known as credit cards.

Mathematical ideas were reviewed in relation to logit regression and econometric theory, considering legal conditions related to Loan Loss Provisions issued by CNBV and Basel II. This research had the aim to propose a particular model of Behavioral Scoring with a special Probability of Default, using discriminant analysis and a database composed by mixed variables (continuous and dichotomous).

It can be confirmed with this research that the model proposed to SOFOMERS satisfies two conditions; it has a strong predictive ability and the estimation of the parameters have high accuracy. An additional condition is that the model is simple and contains a minimum of explanatory variables. However, it opens the door to future research to continuously and particular improvements in prediction and thereby improving profitability for Retail Banking and SOFOMERS in Mexico.

Thus, the validation of the profitability for revolving loans in SOFOMERS sector was obtaining the indicator of ROA. The savings obtained with the proposed model of probability of default (PDC) was approximately MXN $ 556.717 (3.52% compared to the current CNBV methodology). If this trend were continuous to national levels, a national average saving would have been MXN $ 825.7 MM.

It was also validated that the savings would increase the assets accounting in SOFOMERS sector by MXN $171,340.9 MM (versus MXN $ 170,515.2 MM by CNBV methodology), and it would increase net income to MXN $7607.9 MM (versus MXN $5961.4 MM by CNBV methodology), according to the data obtained at the end of June 2014. Thus, the ROA of 3.50% obtained with the PD of CNBV methodology would have pass to 4.44% with the proposed model (PDC).

Given the low publication of research related to Behavioral Scoring models to measure the behavior of credit risk for SOFOMERS in Mexico, it implies to CNBV to have ample field of research; especially to adopt a particular methodology on loan loss provisions of expected losses for SOFOMERS, as the possibility to develop more profitable and predictive models of credit risk in revolving loans.

References

Altman, Edward. (1968). "Financial Ratios, Discriminant Analysis and the Prediction of Corporate Bankruptcy," Journal of Finance, 23(4), 589-609.

Banxico, (2005), Definiciones Básicas de Riesgos. Revisado en: http://www.banxico.org.mx/sistema-financiero/material-educativo/intermedio/riesgos/%7BA5059B92-176D-0BB6-2958-7257E2799FAD%7D.pdf

Black, Fisher, y Scholes, Myron – Scholes (1972). "The Valuation of Option Contracts and a Test of Market Efficiency," Journal of Finance, 399-417.

CNBV, (2014). Normatividad Comisión Nacional Bancaria y de Valores. Revisado en: http://www.cnbv.gob.mx/Paginas/NORMATIVIDAD.aspx

Durand, David. (1941). Risk Elements in Consumer Instalment Financing, Studies in Consumer Instalment Financing, New York: National Bureau of Economic Research.

Fisher Ronald. (1936). The Use of Multiple Measurements in Taxonomic Problems, Annals of Eugenics, 7(2), 179-188.

Hand, David and Henley William (1997). "Statistical Classification Methods in Consumer Credit Scoring: A Review," Journal of the Royal Statistical Society: Series A (Statistics in Society), 160: 523-541.

Hand, David and Vinciotti, Veronica (2003). Local Versus Global Models for Classification Problems, American Statistician, 124-131.

Hsia, Daniel. (1978). Credit scoring and the equal credit opportunity act, Hast. Law. J., 30, 371-448.

Joanes, Derrick. (1993). Reject Inference applied to logistic regression for credit scoring, IMA J. Math. Appl. Bus. Industry, 5, 35-43.

Lewis, Edward (1992). An Introduction to Credit Scoring, California: Fair, Isaac and Co., Inc.

Mays, Elizabeth, (1998). "Credit Risk Modeling, Design and Application," Fitzroy Dearborn publishers, Chicago.

Mays, Elizabeth, (2004). Credit Scoring for Risk Managers, the Handbook for Lenders, Thomson South Western, Mason, Ohio.

Merton, Robert, (1974). "On the Pricing of corporate debt: the risk structure of interest rates," Journal of Finance 29, 449-470.

Myers, James and Forgy, Edward (1963). "The development of numerical credit evaluation systems". J. Am. Statist., 58, 799-806.

Reichert, Alan, Cho, Chien. and Wagner, George (1983). "An Examination of Conceptual Issues Involved in Developing Credit-Scoring Models," J. Bus. Econ. Stat., 1 (2), 101-14.

Robinson, Margarite, (2000). "La revolución microfinancieras, Finanzas sostenibles para los pobres," Banco Mundial, Washington DC E.E.U.U.

Rosenberg, Eric, and Gleit, Alan, (1994). "Quantitative Methods in Credit Management: A Survey," Operations Research, 42, 589-613.

Thomas, Lyn, Crook, Jonathan, & Edelman, David, (1992). Credit scoring and credit control, Oxford University Press, Oxford.

Thomas, Lyn, Edelman, David, & Crook, Jonathan, (2002). Credit scoring and its applications, SIAM Monographs on Mathematical Modeling and Computation. Philadelphia, PA: Siam.

Thomas, Lyn, Edelman, David, Crook, Jonathan, (2004), Readings in Credit Scoring, Oxford University Press, Oxford.

Weber, Richard, (1999). Applications of Fuzzy Logic for Creditworthiness Evaluation. Soft Computing in Financial Engineering, Physica-Verlag, Heidelberg, 388-401.

Notes

[1] Title II, Chapter V Loan Loss Provisions; documented in the General Prudential Provisions Applicable to Credit Institutions, or commonly known as the Unique Banking Dispositions issued by the National Banking and Securities Commission, published in the Official Gazette on December 2, 2005 and last amended on 31st July 2014.

[3] Fair – Isaac, Company.

[4] Section of Supervised Statistical Information in Banking Sector, with information as of June 2014, Loans for Credit Card http://www.cnbv.gob.mx/SECTORES-SUPERVISADOS/BANCA-MULTIPLE/Paginas/Información-Estadística.aspx.

[5] Financiera CrediScotia, Ixe Tarjetas, Santander Consumo, Sociedad Financiera Inbursa, Soriana, Tarjetas Banamex and Credito Familiar; based on the Statistical Reports on Credit Card Number by SOFOMER, CNBV November 2014.

[6] General Provisions Applicable to Credit Auxiliary Organizations,

[7] Exchange, Credit Unions, Financial Companies and SOFOMERS, Article 70, Section VII.

[8] General Prudential Provisions Applicable to Credit Institutions, or commonly known as the Unique Banking Dispositions issued by the National Banking and Securities Commission, published in the Official Gazette on December 2, 2005 and last amended on 31st July 2014.

[9] Unique Banking Dispositions published on December 2, 2005 and last amended on 31st July 2014, Appendix 15 Minimum Requirements for Internal Methodologies Authorization.

[10] Section III Credit Risk – The Internal Ratings-Based Approach (IRB), Overview (page 52). Basel II: International Convergence of Capital Measurement and Capital Standards: a Revised Framework, Comprehensive Version (BCBS) (June 2006 Revision), http://www.bis.org/publ/bcbs128.htm.

[11] Unique Banking Dispositions published on December 2, 2005 and last amended on 31st July 2014, Title I General Provisions Chapter I Definitions, Article 1, Section I.

[12] Unique Banking Dispositions published on December 2, 2005 and last amended on 31st July 2014, Definitions Chapter I, Article 1, Section XXIX.

[13] Banxico 2005, Basic Definitions of Risks. http://www.banxico.org.mx/sistema-financiero/material-educativo/intermedio/riesgos/%-7BA5059B92-176D-0BB6-2958-7257E2799FAD%7D.pdf

[14] Unique Banking Dispositions published on December 2, 2005 and last amended on 31st July 2014, Title Two Prudential provisions of Chapter I, Granting Credit and Chapter V Loan Qualification Method.

[15] Unique Banking Dispositions published on December 2, 2005 and last amended on 31st July 2014, Title Two Prudential provisions of Chapter I, Granting Credit and Chapter V Loan Qualification Method, First Section, Section B.

[16] The source database of some microfinance clients (SOFOMER s) at end-June 2014: http://archive.ics.uci.edu, Datasets Section.

[17] Unique Banking Dispositions published on December 2, 2005 and last amended on 31st July 2014, Annex 33 Series B Criteria for the items comprising the financial statements, B-6 Credit Portfolio.

[18] Unique Banking Dispositions published on December 2, 2005 and last amended on 31st July 2014, Title Two Prudential provisions of Chapter I, Granting Credit and Chapter V Loan Qualification Method, art. 92 section III.

[19] Unique Banking Dispositions published on December 2, 2005 and last amended on 31st July 2014, Title Two Prudential provisions of Chapter I, Granting Credit and Chapter V Loan Qualification Method, art. 92 sections I and III.

[20] By Stata 13 program, "estat class" command.

[21] Unique Banking Dispositions published on December 2, 2005 and last amended on 31st July 2014, Title Two Prudential provisions of Chapter I, Granting Credit and Chapter V Loan Qualification Method, art. 92 sections I and III. Conditions: ACT < 10, SP=75% and ACT ≥ 10, SP=100%.

[22] Unique Banking Dispositions published on December 2, 2005 and last amended on 31st July 2014, Title Two Prudential provisions of Chapter I, Granting Credit and Chapter V Loan Qualification Method, Section Five: The provisioning and classification by risk level, article 129.

[23] Return on Assets; which according to the CNBV is a result by the ratio of Net Income Flow between Total Assets.

[24] Supervised Statistical Information Section of Commercial Banks (November 2014), with information as of June 2014 Retail Loans for Credit Card Industry http://www.cnbv.gob.mx/SECTORES-SUPERVISADOS/ BANCA-MULTIPLE/Paginas/Información-Estadística.aspx.

[2] Section of Supervised Statistical Information in Banking Sector, with information as of June 2014, Loans for Credit Card http://www.cnbv.gob.mx/SECTORES-SUPERVISADOS/BANCA-MULTIPLE/Paginas/Información-Estadística.aspx

Author notes

1 Doctor en Ciencias Económicas, Profesor-Investigador y Coordinador de la EARF, Escuela Superior de Economía del Instituto Politécnico Nacional. Línea de investigación: Teoría Económica, correo: jtrejog@ipn.mx

2 Doctor en Ciencias Económicas, Profesor-Investigador y Jefatura de Posgrado SEPI, Escuela Superior de Economía del Instituto Politécnico Nacional, Línea de investigación: Teoría Económica, correo: mmartinezga@ipn.mx

3 Maestrante en Ciencias Económicas, Escuela Superior de Economía del Instituto Politécnico Nacional. Línea de investigación: Teoría Económica, correo: danielita_051191@hotmail.com

Additional information