Credit risk management at retail in Mexico: An econometric improvement in the selection of variables and changes in their characteristics

Administración del riesgo crediticio al menudeo en México: una mejora econométrica en la selección de variables y cambios en sus características

José Carlos Trejo García [*], Miguel Ángel Martínez García, Francisco Venegas Martínez

Instituto Politécnico Nacional, SEPI ESE IPN, Mexico

Abstract

The early prediction of bad debtors for revolving credits in Mexico is a relevant issue today. The credit behavior econometric model proposed considers the changes in the characteristics of the consolidated accredited and provides better results than those obtained with the methodology utilized by the CNBV on provision matters. The results obtained show that the possibility of replacing the current model, minimizing the expected loss and increasing the ROA per financial institution at a national level by 2.20%, complies with the methodological criteria and the statistical tests in accordance with the Compiled Banking Regulation and Basel II guidelines on credit risk issues.
© 2017 Universidad Nacional Autónoma de México, Facultad de Contaduría y Administración. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

JEL classification: G21; E51; C5; C80; C61
Keywords: Banking; Credit; Econometric models; Data estimation methodology; Optimization techniques

---

[*] Corresponding author.
  E-mail address: jtrejog@ipn.mx (J.C. Trejo García).
  Peer Review under the responsibility of Universidad Nacional Autónoma de México.

Resumen

La predicción temprana de malos deudores para créditos rotatorios en México es un asunto de relevancia actual. El modelo econométrico propuesto de comportamiento crediticio considera los cambios en las características de los acreditados consolidados y proporciona mejores resultados que los obtenidos con la metodología utilizada por la CNBV en materia de provisiones. Los resultados obtenidos muestran que la posibilidad de reemplazar el modelo vigente, minimizando la pérdida esperada y aumentando el ROA por entidad financiera a nivel nacional en un 2.20%, cumple con los criterios metodológicos y pruebas estadísticas de acuerdo a la Circular Única de Bancos y lineamientos de Basilea II en materia de riesgo crediticio.
© 2017 Universidad Nacional Autónoma de México, Facultad de Contaduría y Administración. Este es un artículo Open Access bajo la licencia CC BY-NC-ND (http://creativecommons.org/licenses/by-nc-nd/4.0/).

Códigos JEL: G21; E51; C5; C80; C61
Palabras clave: Banca; Crédito; Modelos econométricos; Metodología de estimación de datos; Técnicas de optimización.

Introduction

Credit risk management has been one of the areas with greater growth in recent decades. The scoring techniques most used for credit risk management have been Credit Scoring and Behavioral Scoring, as well as various tools for the estimation of financial risk in relation to loans or financing to the retail market. In the Mexican consumer credit market, the classification divides in three types of pools: revolving credits, personal credits, and mortgages.[1]

All the credit applicants, as well as the consolidated clients of financial institutions in Mexico, have a score both in the credit request and in the behavior of the credit's life. There are two important objectives in the credit scoring techniques: the need to identify the consumer credit risk, and to minimize the percentage of defaulting clients. Using the latter, the banking or credit institutions optimize their pools for a good and better business.

Therefore, it is required to develop new models based on the historical information of the clients that allow the generation of decision models in the granting of credits, and consider the behavior of the consolidated clients (Mays, 2004).

According to the information – at the close of June 2014, by the National Banking and Securities Commission (*Comisión Nacional Bancaria y de Valores* – CNBV)[2] – one of the main credit instruments in Mexico are credit cards, given that there are approximately 23 million cards authorized by 24 recognized financial institutions.

This research develops a credit scoring technique that takes into consideration both domestic and international regulations. Furthermore, the proposed model is optimal in the sense that it minimizes the expected loss in the credit card sector, with positive repercussions in determining

---

[1] Second Title Prudential Regulations, Chapter V Credit Pool Qualification; documented in the General Provisions Applicable to Credit Institutions, or commonly known as the Compiled Banking Regulation issued by the National Banking and Securities Commission, published in the Official Journal on December 02, 2005, with a most recent modification on July 31, 2014.

[2] Statistics Information Section of the Universal Bank Supervised Sector, with data to the closing close of the month of July, 2014, for Credit Card Consumer Credit http://www.cnbv.gob.mx/SECTORES-SUPERVISADOS/BANCA-MULTIPLE/Paginas/Información-Estadística.aspx.

the credit reserves, generating a positive effect on the assets, net results, and the profitability of the financial entities (Universal Bank).

The document is organized as follows: the second section presents the theoretical framework in terms of statistical and econometric progress in the credit risk management; the third section establishes the bases and regulatory criteria of revolving credits in Mexico; the fourth section presents the proposed methodology and provides the statistically significant evidence; the fifth section establishes the optimization model; the sixth section presents the comparison of the profitability between the proposed model and the CNBV; in the seventh section the obtained results are presented and discussed; finally, we present the conclusions.

Advancements in credit risk management

The advancements in credit risk management information technologies allow financial and credit institutions to automate the acceptance or rejection decisions of a credit application and the management of a credit pool (consolidated clients) in a cross-sell. Some years ago, said credit management was only carried out with the experience or perception of the executive. Nowadays, one of the most commonly used models for credit evaluation is the Scoring Model, which determines a score for clients that apply for credit, identifying those who have the possibility of defaulting with their payments. The literature regarding credit scoring is broad and it suffices to mention the models of: Rosenberg and Gleit (1994), Merton (1974), Hand and Jacka (1998), Thomas, Crook, and Edelman (1992), and Lewis (1992); whereas for Behavioral Scoring there is the work by Mays (1998).

Analytical scoring models

In specialized literature there is a broad set of quantitative methods and techniques to predict the probability of a client to default[3] and, therefore, for the granted credit to not be recovered by any financial institution. Scoring models are tools that utilize the classification of the applicants or consolidated clients by risk level based on the supply of client information on the credit applications and payment behavior. The model provides the sum of points for each quantifiable element, producing a rank-ordered scoring or scoring scale.

The identification of potential clients that could generate losses to the financial institutions requires seeking better guidelines for its treatment and reserve a minimum capital in case of default and/or the migration of acceptable buckets to buckets categorized as overdue. Refer to the investigations by Hsia (1978), Reichert, Cho, and Wagner (1983), Joanes (1993), and Hand and Henley (1997), which detail the study of particular relations between the distribution of good and bad clients.

The history of scoring methods dates back to Fisher (1936) with the idea of differentiating groups within a specific population. This idea was further developed by Durand (1941), applying a financial context to differentiate between a "Good" and a "Bad" payer. Likewise, Thomas, Edelman, and Crook (2002) analyze the advantages and limitations of the creation of scoring model vendors (Bill Fair and Earl Isaac[4]) which, at the end of the 1950s, began with the development of

---

[3] When the probability of default is mentioned, it is usual to think about the Merton (1974) model, which is based on the formula by Black and Scholes (1972). In this research, an alternative model shall be developed.

[4] Fair–Isaac Company.

Table 1
Proof of multicollinearity with variables utilized by the CNBV.

	y	*act*	*hist*	*ant*	*%payment*	*%use*
y	100%					
act	**87%**	100%				
hist	**97%**	**91%**	100%			
ant	**−74%**	**−62%**	−41%	100%		
%payment	**−93%**	−1%	−1%	1%	100%	
%use	46%	25%	44%	−43%	−1%	100%

Source: Own elaboration with Stata 13, implementation of correlation tests.
Bold values means a higer correlation than 50%.

a credit risk analysis system. In the 1960s, with the creation of credit cards, the Scoring models evidenced their importance and utility. According to Myers and Forgy (1963), this type of models is a superior indicator to any qualitative expert judgment. Another important aspect in this context was that of the Z-Score proposed by Altman (1968), which has been implemented in the financial sector.

Optimization in financial entities

In practical terms, Scoring models allow a significant reduction in the execution times of the different financial processes for the issuance and monitoring of a credit, enabling greater automation and reducing the need for human intervention in the evaluation and estimation of credit risk. The main users of these types of models are banks and financial institutions, such as insurance companies or retail chains (consumer credits). Among the main characteristics of the Scoring models is the possibility of managing and administering the risk. The reported benefits for the implementation of these models not only affect banks and financial institutions, but also directly affect all the clients of the financial sector, as the erroneous differentiation of clients that apply for any kind of credit is reduced and it provides a more objective analysis, concentrating multiple factors into a single model that can affect the risk of an application or the monitoring of payments.

Local and international regulatory framework

Although some financial institutions in Mexico develop their own Scoring models based on the guidelines established by the CNBV,[5] most of the entities constitute their reserves based on the guidelines established by said commission. The vendors or consultants in the creation of scoring models build them based on the specific information and parameters provided by the credit risk administration areas of the financial institution. Therefore, the financial institution must ensure that the development of said models is done in accordance with the objectives of the bank and under its risk tolerance (Table 1).

The international agreement on banking regulation and supervision, denominated "New Capital Agreement",[6] demands that the financial institutions of the member countries carry out a review

---

[5] Compiled Banking Regulation with its last modification on July 31, 2014, Annex 15 Minimum Requirements for the Authorization of Internal Methodologies.

[6] International agreement known as Basel II and approved in 2004.

of their capital allocated to cover risks. These parameters force them to have tools available that allow them to establish scoring and rating models with the purpose of differentiating between clients according to their risk profile and evaluating the exposure and severity of the credit risk. The Basel II agreement also obligates the financial entities to not only adapt their capital consumption calculation systems, but to also modify their reporting (financial statements) and information analysis systems. Both elements, that is the financial statements and the analysis of the same, are key aspects in Basel II, necessary to manage large databases capable of providing the information to quantify the risk of each operation, which entails a true challenge for the banks.

The key element to analyze the risk rating process through internal models (IRB[7]), according to the Basel II standard, is for the financial and credit institutions to have a credit scoring model that allows them to measure the probability of default for the credit granted. To such effect, the standard or regulatory method of the CNBV allows the validation with current data and the behavioral scoring IRB method for the Mexican case through a sample of national information.

Credit risk management in the retail market in Mexico

Generally, credit is a compromise made between a natural person or legal entity and a financial institution with the purpose of granting purchasing power to the debtor in advance. Credit makes it possible to make purchases in advance based on their payment capacity. In Mexico, the CNBV mentions that any credit activity signifies the allocation of resources, both their own and those of third-parties (savers), through operations of loans, discounts, assumptions, guarantees or credits in their broadest sense, as well as any banking operation that generates or allows a credit claim in favor of the financial institutions with a degree of default.[8]

The consumer credit portfolio is comprised of all direct credits, including liquid assets that do not have real estate guaranty, denominated in national or foreign currency, UDIs, or in VSM, as well as the interests these generate, granted to natural persons, derived from credit card operations, personal credits, credits for the Acquisition of Consumer Durable Goods (*Adquisición de Bienes de Consumo Duradero* – ABDC), which considers, among others, auto loans and finance lease operations celebrated with natural persons.[9]

Based on different Banxico researches, credit risk is the particular case in which the contract cannot be complied with by the debtor to the creditor (the granter of the credit). Recently, in addition to the case of default, events that affect the value of a credit, without it necessarily signifying the default of the debtor, have been incorporated. Ordinarily, the factors that must be considered when measuring credit risk are: the probabilities of default and/or of the migration of the credit quality of the debtor, the correlations between defaults, the concentration or segmentation of the portfolio, the exposure to each debtor, and the recovery rate in the case of default by the debtors.[10]

In Mexico, for the determination of provisions of the consumer credit portfolio managed by each financial institution, we use the regulation imposed by the CNBV through the Compiled

---

[7] 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.

[8] Compiled Banking Regulation, updated on July 31, 2014, First Title General Provisions of Chapter I Definitions, Article 1, Section I.

[9] Compiled Banking Regulation, updated on July 31, 2014, First Title General Provisions of Chapter 1 Definitions, Article 1 Section XXIX.

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

Banking Regulation (CBR)[11] as basis. Factors such as the Severity of the Loss Given Default, which is what is lost by the creditor in the case of default by the debtor considering all the costs implied in the recovery (retrieval costs, legal costs, etc.) regarding Exposure at Default, is the balance owed by the debtor at a given time in the case of default. The credit risk measurement in this research is the Probability of Default, which expresses how probable it is for a borrower to stop complying with their contractual obligations. The minimum value is zero, which would indicate that it is impossible for them to fail to comply with their obligations, and its maximum value is one or one-hundred percent, when it is certain they will fail to comply.

Based on current regulation, the estimation of the Probability of Default must be by credit. The Reserves or Expected Losses is the measurement of the distribution of profits and losses. That is, it indicates how much can be lost on average and is normally associated with the preventive reserves policy that the institution must have in the case of credit exposure. To this end, the Expected Losses are calculated taking into consideration the percentage of reserves (Probability of Default*Loss Given Default) by the Exposure at Default.[12]

Proposed methodology and mathematical optimization

The methodology utilized in this research for the construction of an optimal Behavioral Scoring model for the product of revolving credits managed by the Mexican financial institutions is based on the local regulatory guidelines of the CBR, which stipulates in article 92 Section III the different stages for the credit risk management of the revolving credits portfolio in Mexico.

According to the local regulation mentioned, it is necessary to detail each stage for the development of the proposal.

Stage 1. Data and criteria in the determination of the objective

First, a sample was obtained of clients classified in advanced as "good" and "bad" (according to Huang, Chen, & Wang, 2007). A sample of 43,323[13] accounts was considered in this manner, with the regulatory criteria and the event of default under the criteria that a client does not make two or more consecutive payments[14] within the next 12 months; this is an essential objective for this investigation.

Stage 2. Information compilation and process

The construction of the alternate Behavioral Scoring model (Weber, 1999) is carried out through logistical differentiation and regression, which is employed to validate with better levels of significance.

---

[11] Compiled Banking Regulation 2014, updated on July 31, 2014, Second Title Prudential Provisions of Chapter 1 Credit Issuance and Chapter V Credit Portfolio Qualification.

[12] Compiled Banking Regulation 2014, updated on July 31, 2014, Second Title Prudential Provisions of Chapter V Credit Portfolio Qualification, First Section, Part B.

[13] The extraction of the client database of some Mexican banks and of other topics related to the July 2014 closure, you can refer to the following page: http://archive.ics.uci.edu, on the Datasets section.

[14] Compiled Banking Regulation with last modification on July 31, 2014, Annex 33 Series B Criteria relative to concepts that form part of the financial statements, B-6 Credit Portfolio.

For this, the work is carried out using independent quantitative variables and the dependent dichotomous variable, under the condition:

$$\text{Dependent variable } Y_i = \begin{cases} 1 \text{ (bad); if Default} > 2 \text{ months behind} \\ 0 \text{ (good); if Default} \geq 2 \text{ months behind} \end{cases}$$

The statistical technique most used by the financial industry corresponds to logistic regression (Thomas, Edelman, & Crook, 2004). This technique is less restrictive, being an alternative to the differential analysis. It is important to note that in recent years a new series of techniques has appeared, called Datamining (Weber, 1999), which have also been used for the construction of Scoring models. These techniques have the advantage of not having too many requirements and assumptions for the input variables, increasing their validity. This technique has been heavily used for the construction of Scoring methods for the understanding of complex patterns of a determined sector of clients, having the capability to model nonlinear relations between the variables.

Stage 3. Probabilistic estimation

Linear probabilistic models pose a series of problems that have led to the search of other alternative models that allow more trustworthy estimations when there are variable dichotomies. To avoid the estimated endogenous variable to be out of range (0, 1), the available alternative is to use nonlinear probability models, where the specification function utilized guarantees a result in the estimation within the range (0, 1).

Given that the use of a distribution function guarantees that the result of the estimation is delimited between 0 and 1, the possible alternatives are several, with the logistic distribution function being the most usual, as it has given rise to the *Logit* model, linking the endogenous variable Y_i with the explicative variables X_i through a distribution function ($Y_i = 1$). Regarding the interpretation of the parameters estimated in a *Logit* model, the sign of the same indicates the direction in which the probability moves when the corresponding explicative variable increases.

Through the linearization of the model and parting from the general equation of the *Logit* Model, Y_i is defined as the probability of the state or alternative 1, departing from the following:

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

This satisfies $0 \leq y_i \leq 1$. Therefore:

$$p_i + p_i e^{\beta_0 + \beta_k X_k} = e^{\beta_0 + \beta_k X_k} \tag{2}$$

$$p_i = (1 - p_i)e^{\beta_0 + \beta_k X_k} = e^{\beta_0 + \beta_k X_k} \tag{3}$$

$$\frac{p_i}{(1 - p_i)} = e^{\beta_0 + \beta_k X_k} \tag{4}$$

The quotient between the probability that an incident occurs, or that option 1 is chosen, against the probability that the event in question does not occur, or that option 0 is chosen, is denominated the Odds Ratio. Its interpretation is the "advantage" or preference for option 1 against option 0, that is, the number of times it is more probable for default to occur against it not occurring.

$$\text{Odds Ratio} = \frac{p_i}{(1 - p_i)} = e^{\beta_0 + \beta_k X_k} \tag{5}$$

Taking natural logarithms of the Odds Ratio, the equation of the Logit model is linearized, observing the objective that the estimated values are within the range (0,1), obtaining the following expression:

$$Ln[\![\frac{p_i}{(1-p)}]\!] = Ln(e^{\beta_0 + \beta_k X_k}) = \beta_0 + \beta_k X_k \tag{6}$$

The new generated variable $Ln[\![\frac{p_i}{(1-p_i)}]\!]$ represents, in a logarithmic scale, the difference between the probabilities that default in revolving credits occur (Lee, Cheung, & Chen, 2005).

The Odds Ratio as it is built (quotient divided by probabilities) will always be greater or equal to 0. The field of variation of the ration goes from 0 to $+\infty$, and its interpretation is carried out if the value is equal, less than or greater than the unit: if the antilogarithm has a value of 1, it means that the probability of default is the same for it not occurring; if the ratio is less than 1, it indicates that the occurrence of the probability of default has less probability than the occurrence of compliance by the client; whereas if it is greater than 1, then the client's probability of default is greater than the probability of compliance.

$$\text{Odds Ratio Quotient} = \frac{p_i/(1-p)}{p_j/(1-p_j)} = \frac{e^{\beta_0 + \beta_k X_{ki}}}{1 + e^{\beta_0 + \beta_k X_{kj}}} = e^{\beta_k(X_{ii} - X_{jj})} \tag{7}$$

Stage 4. Estimation with maximum verisimilitude

For the estimation of the Logit Model searching to obtain the predictor parameters, the maximum verisimilitude (MV) method is utilized. Given that p_i takes two values; with probability of default p_i and its counterpart $1 - p_i$, it has a Bernoulli style distribution.

$$P(y_i) = p_i^{y_i}(1 - p_i)^{1-y_i} \ y_i = 0, 1 \tag{8}$$

The MV function for a random sample of n data (x_i, y_i) is calculated as follows:

$$P(y_1, \ldots, y_n) = \prod_{i=1}^{n} p_i^{y_i}(1 - p_i)^{1-y_i} \tag{9}$$

By implementing logarithms

$$Ln\,P(y) = \sum_{i=1}^{n} y_i Ln(p_i) + \sum_{i=1}^{n} (1 - y_i) Ln(1 - p_i), \tag{10}$$

The verisimilitude logarithmic function satisfies

$$Ln\,P(y) = \sum_{i=1}^{n} y_i Ln\left(\frac{p_i}{1-p_i}\right) + \sum_{i=1}^{n} Ln(1 - p_i) \tag{11}$$

$$Ln[\![\frac{p_i}{(1-p)}]\!] = Ln(e^{\beta_0 + \beta_k X_k}) = \beta_0 + \beta_k X_k \tag{12}$$

Now, the previous equation is substituted in the logarithmic verisimilitude function. From here, the verisimilitude function is obtained in logarithms in terms of the β parameters given by

$$L(\beta) = \sum_{i=1}^{n} y_i \beta_0 + \beta_k X_k - \sum_{i=1}^{n} Ln(1 + e^{\beta_0 + \beta_k X_k}) \tag{13}$$

To obtain the maximum verisimilitude β estimators, $L(\beta)$ is derived for each of the β_j parameters with $j = 1, 2,..., p$ and is equated to zero. In terms of matrices:

$$\frac{\partial L(\beta)}{\partial L(\beta_0)} = \sum_{i=1}^{n} y_i x_i - \sum_{i=1}^{n} x_i \left(\frac{e^{\beta_0 + \beta_k X_{ki}}}{1 + e^{\beta_0 + \beta_k X_{kj}}} \right) \tag{14}$$

$$\begin{bmatrix} \dfrac{\partial L(\beta)}{\partial L(\beta_0)} \\[2mm] \dfrac{\partial L(\beta)}{\partial L(\beta_1)} \\[2mm] \vdots \\[2mm] \dfrac{\partial L(\beta)}{\partial L(\beta_j)} \\[2mm] \vdots \\[2mm] \dfrac{\partial L(\beta)}{\partial L(\beta_p)} \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^{n} y_i (1) \\[2mm] \sum_{i=1}^{n} y_i x_{i1} \\[2mm] \vdots \\[2mm] \sum_{i=1}^{n} y_i x_{ij} \\[2mm] \vdots \\[2mm] \sum_{i=1}^{n} y_i x_{ip} \end{bmatrix} + \begin{bmatrix} \sum_{i=1}^{n} (1) \left(\dfrac{e^{\beta_0 + \beta_k X_{ki}}}{1 + e^{\beta_0 + \beta_k X_{kj}}} \right) \\[2mm] \sum_{i=1}^{n} x_{i1} \left(\dfrac{e^{\beta_0 + \beta_k X_{ki}}}{1 + e^{\beta_0 + \beta_k X_{kj}}} \right) \\[2mm] \sum_{i=1}^{n} x_{ij} \left(\dfrac{e^{\beta_0 + \beta_k X_{ki}}}{1 + e^{\beta_0 + \beta_k X_{kj}}} \right) \\[2mm] \sum_{i=1}^{n} x_{ip} \left(\dfrac{e^{\beta_0 + \beta_k X_{ki}}}{1 + e^{\beta_0 + \beta_k X_{kj}}} \right) \end{bmatrix}$$

If each of these derivations is expressed in a vector column

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

If the vector column is then equated to zero

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

The estimated β is the vector of parameters that fulfills the matrix system; p_i is calculated in terms of these estimators and with it an estimation is obtained for y_i, so that the probability estimation is:

$$\sum_{i=1}^{n} y_i x_{ij} = \sum_{i=1}^{n} x_i y_{estimated_i} \tag{15}$$

The linearization of Eq. (14) allows the modeling of PI through a multiple logistic regression:

$$PI = \left(\frac{e^{\beta_0 + \beta_k X_{ki}}}{1 + e^{\beta_0 + \beta_k X_{kj}}} \right) = \left(\frac{1}{1 + e^{-(\beta_0 + \beta_k X_{kj})}} \right) \tag{16}$$

Empirical results

To validate the proposed methodology, variables have been selected that are relevant in the determination of the probability of default proposed.

Determination and variable selection

To avoid multicollinearity problems, it is of great help when the selected variables have proof of correlation. To this end, the set of explicative and significant variables is selected, thus avoiding including little significant variables or variables with redundant information (collinearity), which could distort the predictive capability of the estimated differential function.

To measure the correlation, the following calculation is considered:

$$R^2 = 1 - \frac{D(\beta)}{D(\beta_0)} \tag{17}$$

With a database of 43,323 consolidated clients, administrated by the Mexican financial institutions at the close of August 2014, the 6 variables were considered in accordance with the regulatory methodology of the CNBV[15]: Compliance/Non-compliance Variable (*Y*), Number of Defaults (*ACT*), History of Default (*HIST*), Credit Months Transpired (*ANT*), Payment-Balance Relation (*%PAYMENT*), and Balance Payable Relation – Credit Limit (*%USE*).

Whereas with the proposed model and in relation with the database of the same 43,323 clients, only 5 significant variables were identified: Compliance Variable (*Y*), Credit Months Transpired (*ANT*), Credit Limit (*credit_limit*), History of Default (*HIST*), and Payment-Balance Relation (*%PAYMENT*).

Variable *Y* for both models of study (current and proposed) was recoded from the criteria mentioned in the methodology: binary variable that adopts a value of 1 (does not comply) and a value of 0 (complies).

As can be observed, multicollinearity problems are notorious when using all the variables based on the current methodology of the CNBV, mainly *HIST* and *ANT* with a correlation above 60%. Furthermore, the lack of impact with *Y* on behalf of *%USE* can be observed, which could cause less prediction with the model currently used by the Mexican regulation.

Whereas the absence of multicollinearity, as well as the high impact of the independent variables (*ANT*, *Credit Limit*, *HIST*, and *%PAYMENT*) on the *Y* variable, do influence in a significant manner and this could reflect major predictability in measuring the level of non-compliance.

In this manner, the variables proposed in this investigation show an absence of multicollinearity, given that the correlation levels between endogenous independent variables are less than 47%, whereas the relation of *ANT*, *credit limit*, and *HIST* with the dependent variable (*Y*) is greater than 56%. The impact of *%PAYMENT* with *Y* is greater than 92%. This correlation test indicates that the proposed model could be predictive and more stable.

Regression and logistical significance

In accordance with the inference techniques on logistical regression (Joanes, 1993), a good model must satisfy two conditions. The first is that it has a strong predictive capability and the second, that the estimation of the parameters has a high accuracy. An additional condition is for the model to be as simple as possible, meaning that it includes the minimum number of explicative variables and that it meets the two previous conditions. Utilizing the sample of 43,323 clients and the proposed variables in Table 2, a *Logit* model is estimated as explained

---

[15] Compiled Banking Regulation 2014, last updated July 31, 2014, Second Title Prudential Provisions of Chapter I Issuance of Credits and Chapter V Credit Portfolio Qualification, article 92, section III.

Table 2
Absence of multicollinearity with proposed variables.

	y	ant	credit limit	hist	%payment
y	100%				
ant	**−74%**	100%			
credit limit	**−57%**	48%	100%		
hist	**97%**	−41%	−37%	100%	
%payment	**−93%**	1%	1%	−1%	100%

Source: Own elaboration with Stata 13, implementation of correlation tests.
Bold values means a higer correlation than 50%.

in the methodology, calculating in turn the probability that a client pays their loan of revolving credits.

In the estimation of the *Logit* Probability, 9 iterations were necessary to estimate the model. The *LR chi2* function indicates that the coefficients are jointly significant to explain the probability that the 43,323 clients are in non-compliance, thus the statistical value $Prob > chi^2$ indicates that the hypothesis that all the coefficients are equal to zero can be rejected at a 1%. Furthermore, the *Pseudo R2* statistic indicates that approximately 97% of the variance of the dependent variable is explained by the variation of the independent variables of the model proposed for Mexico.

The quality adjustment of the model is of 99.38%,[16] resulting from the reason of correct provisions against the number of observations; in general, this confirms that the model correctly predicts the observations.

Significance of the logistic regression proposed

It is necessary to emphasize that at the time of carrying out the logistic regression for those variables contemplated by the current regulation, an error presented itself in when obtaining the results. This was due to the high multicollinearity of the *ACT* variable with *HIST*. Said explicative variables currently used by the CNBV are not at all necessary to continue with the model, and therefore, the *ACT* variable must be excluded.

By using statistical criteria such as that of partial correlations, it is possible to obtain all the potential variables and choose the best among them. In this regard, to determine if a variable must be included in a model by a significant weight, the Wald test is implemented. The test is a contrast of the null hypothesis $H_0 : \beta_i = 0$, against the alternative $H_1 : \beta_i \neq 0$.

The estimation shown in Table 3 allows the calculation of the score function from Eq. (4). The levels of significance for the model proposed indicate that all the significant values—*p-values*—have high significance, given that they are lower than 5%. To obtain the score with the proposed variables, the *logit* transformation of Eq. (16) is implemented to obtain the calculated probability of non-compliance (PI_C).

The interpretation of the coefficients is done through the measurement of the variation of the estimated *logit* model for a unitary variation of the given explicative variable. Thus, if *ANT* increases in a default, the estimated *logit* increases by 0.01 units, which suggests a positive relation regarding the PI_C. It is the same for the *Credit Limit* and *HIST*, the PI_C increases at a greater limit as well as at a greater increase of historic defaults. On the contrary, if *%PAYMENT* is greater, then

[16] Through the Stata 13 program, the command "*estat class*" is implemented.

Table 3
Estimation of the *logit* regression of the proposed model (PI_C).

Y	Coef.	Std. Err.	z	P > z	[95% Conf.	Interval]
ant	0.0129560	0.0006613	19.59	0.0000	0.0142522	0.0116598
credit limit	0.0000161	2.47E−06	6.51	0.0000	0.0000112	0.0000209
hist	2.74826	0.0796168	34.52	0.0000	2.592214	2.904306
%payment	−7.687266	0.7029944	−10.94	0.0000	−9.06511	−6.309423
_cons	−7.022915	0.316211	−22.21	0.0000	−7.642677	−6.403153

Source: Own elaboration with Stata 13 and based on the sample of 43,323 clients.
Observations = 13,000; LR chi^2(4) = 53,200; Prob > chi^2 = 0.0000; Pseudo R^2 or McFadden test = 0.9626; Log likelihood = −722.0952.

Table 4
Estimation of the logistic regression of the proposed model (PI_C).

Y	Coef. *Odds*	Std. Err.	Z	P > z	[95% Conf.	Interval]
ant	1.013040	0.051042	19.59	0.0000	1.114391	0.911689
credit limit	1.000016	2.47E−06	6.51	0.0000	1.000011	1.000021
Hist	15.61544	1.243252	34.52	0.0000	13.35932	18.25258
%payment	0.0004586	0.0003224	−10.94	0.0000	0.0001156	0.0018191
_cons	0.0008912	0.0002818	−22.21	0.0000	0.0004795	0.0016563

Source: Own elaboration with Stata 13, in accordance with the sample of 43,323 clients.

the PI_C decreases by 7.6 units. According to the aforementioned – the econometric model posed through the *logit* estimation – the revolving credits risk management is adequate for the credit risk management.

The opportunities in the *logit* models are calculated through logistic regression or odds, that is, through the anti-logarithmic transformation of $\beta \gg e^\beta$ (Table 4).

These results show compliance with the theory posed, given that if *odd* < 1 it means that the occurrence of the PI_C tends to be a negative relation (less opportunity). Whereas if *odd* > 1 it means that the occurrence of the PI_C tends to be a positive relation (greater opportunity).

Therefore, *ANT*, *credit limit*, and *HIST* variables indicate that if the defaults and the credit limit increase, then the *PI* increases as well and approximately by 17 times, contrary to the *%Payment* as the coefficient is less than the unit.

Logit distribution

The *logit* distribution obtained with the current methodology of the CNBV[17] is as shown in the following figures.

Fig. 1 indicates that there is no optimal distribution, given that there are negative observations outside the normal behavior regarding the other observations.

The average *PI* of the methodology of the CNBV is of 43.4%. Therefore, it is necessary to validate through the model proposed (PI_C) in this investigation if there is an optimal *logit* distribution that adheres to the theoretical behavior.

---

[17] Compiled Banking Regulation 2014, last update on July 31, 2014, Second Title Prudential Provisions of Chapter I Issuance of Credits and Chapter V Credit Portfolio Qualification, article 92, sections I and III. ACT ≥ 4, PI = 100% and ACT < 4, the calculation of the PI is done through the model posed in article 92 section III.
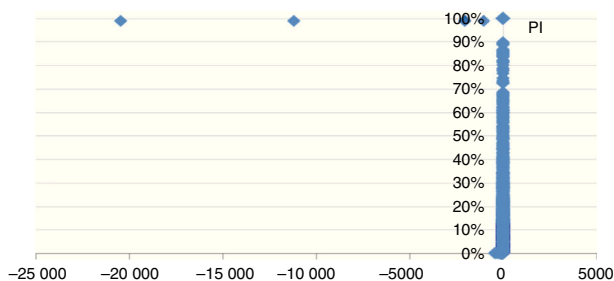
Figure 1. Logistic score distribution versus probability of default current methodology of the CNBV.
Source: Own elaboration with Stata packaging 13, based on the logistic regression and the *logit* distribution.
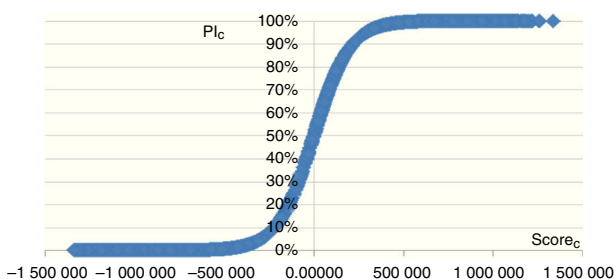

Figure 2. Logistic score distribution versus probability of default through the model proposed.
Source: Own elaboration with Stata 13, in accordance with the logistic regression and the logit distribution.

Fig. 2 indicates that there is in fact one optimal distribution, given that there are no observations outside the normal behavior regarding most of the observations. The PI_C average is of 42.6%, that is, less than the PI of the current regulation (43.4%).

Distribution analysis of the percentage of reserves in Mexico

Through Demographic Differentiation tests, the differential analysis consists on a multivariate technique that allows the simultaneous study of the behavior of a group of independent variables with the intent of classifying a series of cases into previously defined and mutually exclusive groups (Fisher, 1936).

Once the aforementioned PI and PI_C have been calculated, these are used to take them with the SP,[18] thus obtaining the reserve percentage. By validating through the current methodology ($PI*SP$) against the proposed model ($PIC*SP$) in the behaviors shown of the Reserves Percentage and Degrees of Risk indicated in the methodology of the CNBV,[19] it is identified that the proposed model has a better differentiation of good and bad clients on each degree of risk (Fig. 3).

---

[18] Compiled Banking Regulation 2014, last update on July 31, 2014, Second Title Prudential Provisions of Chapter I Issuance of Credits and Chapter V Credit Portfolio Qualification, article 92, section III. ACT < 10, SP = 75% and ACT ≥ 10, SP = 100%.

[19] Compiled Banking Regulation 2014, last update on July 2014, Second Title Prudential Provisions of Chapter I Issuance of Credits and Chapter V Credit Portfolio Qualification, Section Five of the constitution of reserves and its classification by degree of risk, article 129.
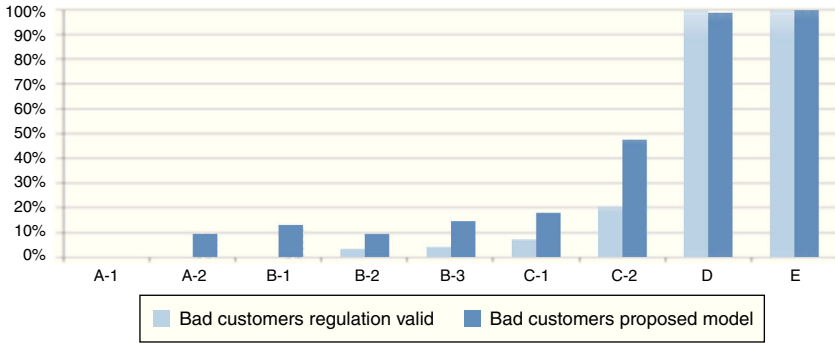
Figure 3. Degree of differentiation of bad clients per degree of risk current model versus proposed model.
Source: Own elaboration with Stata 13, based on the distribution of the Reserves Percentage between the Degrees of Risk, CNBV (2014).

Consequently, it is shown that with the proposed model the detection of bad clients is greater than with the current model utilized by the CNBV, which guarantees a better and greater differentiation of bad or non-compliant clients, as it is notably shown for the A-2 to C-2 degree of risk. Whereas the risk groups A-1, D and E maintain their differential tendency between both models.

K–S test

Once the score has been calculated with the formula estimated through the *logit* regression, the plan is to determine if these values calculated on the sample do a good job of identifying to what group they belong. The greater the difference in the scores of the groups, the greater the differential capability of the model used. The Kolmogorov Smirnov (*K–S*) indicator is a non-parametric test for the goodness of fit, to prove that two independent samples derive from the same distribution of a continuous random variable. These differences are determined not only through the medians, but also through dispersion, symmetry or obliquity. The test is built on the null and alternative hypotheses as follows:

H_0 *(:)*. The distribution of the score for the good and bad accounts is equal.

H_1 *(:)*. The distribution of the score for the good and bad accounts is not equal.

The test statistic is

$$F_n(x) = \frac{1}{n}\sum_{1=1}^{n} I_{X_i \leq x} \tag{18}$$

With two statistical tails given by

$$D_n = \sup_x |(F_n(x) - F(x)) \tag{19}$$

where $\sup_x$ is the greater distance of the group.

In this manner, the differential *K–S* index between the degree of separation of good and bad clients with the Current Model versus the Proposed Model can be compared.

In the previous table it can be observed that with the proposed model, the differential power between degrees of risk per reserves percentage (PI_C*SP) is greater in C-1, which allows for the Consumer Bank in Mexico to have the possibility to better detect the good and bad clients, as well as implement more effective collection policies to be able to improve in this area.

It even provides the Retail Market Risks area of the financial institutions the power to propose an optimal cross-selling level with clients in higher (better) degrees of risk to C-1, this represents a greater possibility of business in favor of profitability in the business of retail market credits.

Selection of the cut-off point

Regarding the level of differentiation presented in the previous section, mainly in Table 5, it is possible to identify the degrees of risk that the banking sector could accept in the administration of the portfolio, based on the level of probability of default given the maximum distribution separation of good and bad clients.

The model based on logistic regression is comprised by a set of weights related to a group of variables or attributes, characterizing a group of clients, and thus allowing the determination of a cut-off point (threshold).

Regarding the cut-off point, this determines the limit between being a "Bad" or a "Good" client. In general, it is understood in the banking industry that the cost of accepting a bad client is many times greater than that of rejecting a good client (Costa, Boj, & Fortiana, 2012).

Regarding the set of weights, it is possible to characterize the patterns that describe both populations (good and bad clients) and to determine which of the input variables used are really significant in terms of a good prediction by segments for the administration of the portfolio of revolving credits.

According to the above, the criteria for the selection of the cut-off point in the proposed logistical model are based on distances. In the application of the behavior scoring technique, the quality adjustment criterion was carried out through the K–S differentiation technique, considering that the threshold indicates the greatest separation distance between the good and bad clients, as is referenced in Eq. (19).

As can be observed in Table 6, the proposed model shows a separation index of 56.68% in the C-1 segment or degree of risk, greater than the 56.38% determined with the current regulation of the CNBV in the C-2 degree of risk.

Table 5

Level of Population Separation by good and bad clients CNBV versus Proposed Model.

Class	Current K–S separation methodology	Proposed K–S separation model
A-1	53.07%	56.22%
A-2	54.35%	56.34%
B-1	54.94%	56.39%
B-2	55.37%	56.45%
B-3	55.74%	56.51%
C-1	56.06%	56.68%
C-2	56.38%	56.66%
D	25.62%	13.40%
E	0.00%	0.00%

Source: Own elaboration in accordance with the Kolmogorov–Smirnov methodology.

Table 6
Determination of the Cut-off point with K–S and its relation with the probability of default: CNBV versus Proposed Model.

Class	Current methodology			Proposed model		
	K–S separation	Probability of default for		K–S separation	Probability of default for	
		Good clients	Bad clients		Good clients	Bad clients
A-1	53.07%	1.12%	100.00%	56.22%	0.03%	100.00%
A-2	54.35%	3.89%	100.00%	56.34%	4.02%	100.00%
B-1	54.94%	5.77%	100.00%	56.39%	5.72%	100.00%
B-2	55.37%	7.18%	100.00%	56.45%	7.14%	100.00%
B-3	55.74%	8.90%	100.00%	56.51%	8.94%	100.00%
C-1	56.06%	12.05%	100.00%	56.68%	12.92%	100.00%
C-2	56.38%	21.67%	100.00%	56.66%	21.99%	100.00%
D	25.62%	44.54%	100.00%	13.40%	60.05%	100.00%
E	0.00%	100.00%	100.00%	0.00%	-	100.00%

Source: Own elaboration based on the Kolmogorov–Smirnov methodology.

With the maximization of the K–S coefficient, a cut-off point of a C-1 risk level was determined for the model proposed, with a probability of default for good clients of 12.92%. Said probability is greater than the 12.05% obtained with the current model of the CNBV for the same level of risk.

According to the table above, it can be observed that it is costlier to continue managing the portfolio with a C-2 risk level using the current model than the one managed with the proposed model with a C-1 risk level. In this manner, the PI_C calculated with the proposed model implies greater coherence and, in a certain manner, better behavior, given that in the A-1 to B-2 risk levels (levels with less risk) a lower PI_C is reflected in the proposed model. Whereas in the B-3 to D risk levels (levels with greater risk), the proposed model tends to coherently consider a greater PI_C. In the case of risk level E, it was not possible to consider any sort of credit in the proposed model, given that it is the worst risk level.

Therefore, the optimal threshold in which the Consumer Bank in Mexico can allow a maximum risk appetite with the administration of credit cards is the C-1 cut-off point with the proposed model. This allows analyzing how willing the banking in Mexico would be in accepting a gain with a level of assets (ROA), with the cost of allowing clients that are in the *C-1* degree of risk or lower (worse) with the proposed model.

Credit assets and profitability of retail banking in Mexico

In Mexico, the consumer preventive credit reserves are determined considering the outstanding balance recorded on the last day of the month and the rating obtained. In accordance with the information at the close of June 2014 by the National Banking and Securities Commission (CNBV),[20] one of the main credit instruments in Mexico are credit cards as there are approximately 23 million cards authorized by 24 recognized financial entities, which have had a minimum annual growth of only 0.14% on average.

---

[20] Statistic Information Section of the Universal Bank Supervised Sector (August 22, 2014), with data up to the close of the month of June 2014 for Consumer Credits for Credit Cards http://www.cnbv.gob.mx/SECTORES-SUPERVISADOS/BANCA-MULTIPLE/Paginas/Información-Estadística.aspx.

Table 7
CNBV reserves comparative and proposed model, with level of savings in the consumer banking in Mexico at the close of June 2014.

	Reserves (MXN)		Savings with the reserves proposed	
	CNBV current methodology	Proposed model	Difference (MXN)	% Savings
Sample of 43,323 clients	$266,165,772.98	$255,052,060.18	$11,113,712.80	4.18%
Information of consumer banking in Mexico	$33,422,191,923.94	$32,026,653,203.41	$1,395,538,720.53	4.18%

Source: Own elaboration in accordance with the calculation of the Reserves Percentage and National Information with data up to the close of June 2014, CNBV (2014).

It is worth mentioning that approximately 54% of issued cards are being managed by the financial entities of BBVA Bancomer and Banamex, while the entities of BanCoppel, Santander Consumo, Banorte-IXE, and HSBC cover approximately 35%. The past-due portfolio of said revolving credits is equivalent to the close of June 2014, for an amount of MXN $90,593 million Mexican pesos (MM), that is, approximately 34% of the total amount authorized by the financial entities authorized by the CNBV. Thus, in this sense, it is a clear example of the problem it represents for the financial credit institutions in Mexico to approve a loan to people that do not guarantee a payment and not reflecting their non-compliance in the periodic payments.

Following up on the current Mexican regulation for the calculation of credit card reserves (to cope with the possible non-compliances), the same statistics of the CNBV reflect that at the close of June 2014 the reserves at a national level added up to MXN $33,422 MM which, in the same manner as the concentration of credit cards by Bancomer and Banamex, added up to approximately 51% of the reserves. Santander Consumo, Tarjetas Banorte-IXE, HSBC, and Scotiabank amounted to approximately 42% together.

Said provisions have seen an annual growth of approximately 12% during the last twelve months, as the growing consumption has been reflected by a greater probability of default, an increase in the severity of the loss, and in the monetary exposure to default. This is a situation that affects the net yields of the cash flows and the financial assets of all entities, reflecting a yield on assets (*ROA*[21]) of 2.18% at a national level, with American Express having registered a greater yield of 6.6% and Banco Bicentenario with the lowest yield with a negative 82.6%.

Result in credit provisions and yield

For the treatment of the estimations for revolving credit risks, the CNBV classifies, at book value, the reserves in the General Balance under the concept of Assets, and in Preventive Estimations subaccounts for Credit Risk, with a classification of Consumer Credits (Credit Card).

A comparative is shown in Table 7 considering the case when the reserves are obtained through the product of *PI*SP*EI* and based on the current regulation, in addition to the results obtained with the information of the 43,323 clients of the sample.

---

[21] *Return on Assets*: that in accordance with the CNBC it is comprised by the result of the Net Flow divided by the Total Assets.

Table 8
Assets and profitability indicators in consumer banking in Mexico at the close of June2014.

	Total assets balance CNBV	Current methodology provisions CNBV	Provisions of the proposed model	ROA % indicators CNBV
Sample 43,323 clients	$55,257,733,112	$266,165,773	$255,052,060	2.18%
Information of consumer banking in Mexico	$6,938,662,851,615	$33,422,191,924	$32,026,653,203	2.18%

Source: Own elaboration based on the calculation of the Reserves Percentage and National information, with data up to the close of June 2014, CNBV (2014).

Table 9
Assets and profitability indicators in consumer banking in Mexico with the proposed model (PI_C) at the close of June 2014.

	Net result CNBV	Net result with proposed model	Total assets with proposed model	ROA % with proposed model
Sample 43,323 clients	$1,204,850,112	$1,215,963,825	$55,268,846,825	2.20%
Information of consumer banking in Mexico	$151,291,923,163	$152,687,461,883	$6,940,058,390,336	2.20%

Source: Own elaboration in accordance with the calculation of the Reserves Percentage and National Information, with data up to the close of June 2014, CNBV (2014).

As shown in Table 7, under the proposals made and with the calculation of the reserves of the sample of 43,323 clients per MXN $255 MM versus the current methodology which added up to MXN $266 MM, a savings of $11 MM (4.18%) could be obtained at the close of June 2014. If this tendency continues at the Consumer Banking level in Mexico, the savings could be of $1395 MM. This implies analyzing the Return on Assets effect that was implied with the Current Methodology, thus the ROA registered a 2.18% (Table 8).

By considering the Consumer Banking savings that imply the use of the Proposed Model (PI_C), the Net Result reported by the CNBV at the close of June 2014 of MXN $151,291 MM would increase by MXN $1395 MM, that is, MXN $152,687 MM.

Similarly, the Balance Sheet Assets at a national level would increase. Therefore, the ROA would go from 2.18% to 2.20%, which represents greater profitability for Consumer Banking in Mexico when considering an updated PI_C model (see Table 9).

Conclusions

Although the market risk is the most significant potential loss in financial institutions, credit risk does not stop being important for its administration in any financial institution, especially when the retail market credit business contributes with more than half of the banking income.

According to the above, the analysis approach in this investigation was the retail market credit risk, especially the one associated with revolving credits, commonly known as Credit Cards. The legal conditions were reviewed with the Provisions of the CNBV and Basel II to propose a new

Behavioral Scoring classification model, utilizing the linear differential analysis with a database comprised of mixed and continuous variables, and a dichotomy (good or bad clients).

To determine which variables ought to be excluded, different tests were carried out for variable selection hypotheses; such is the case of the non-multicollinearity validation. Taking into consideration the methodology of the CNBV in terms of segmentation of degrees of risk and the results obtained, it was shown that with the credit management proposed, the "cost" of administrating credits at B-3 to D levels of risk is greater than the cost of maintaining clients at the A-1 to B-1 levels of risk.

Furthermore, the optimal cut-off point for the management of the portfolio in the proposed model is the C-1 risk level, in which it is identified before the maximum separation of good and bad clients (K–S = 56.68%) with a drastic increase in the PI_C of 12.92%. Both results are greater and better indicators than the current model.

Likewise, it can be confirmed with this investigation that the proposed model satisfies two conditions: the first is that it has a strong predictive capability (99.38%), and the second that the estimation of the parameters have a high dependent variable variability relation regarding the independent variable at a 97%. However, this opens the door for future investigations for the continuous improvement of the prediction and with it, for the improvement of profitability for Retail Banking in Mexico.

The validation of the degree of profitability for consumer banking in Mexico through the indicator known as Returns on Assets (ROA) showed rather important results. Firstly, the savings obtained with the proposed model for probability of default (PI_C) for the analyzed sample was of approximately MXN \$11 MM (4.8% regarding the current methodology of the CNBV). If this tendency is constant, banking would have had approximate savings of MXN \$1395 MM at a national level. Secondly, it was also possible to validate that the savings would increase the countable record of Assets to MXN \$55,268 MM (versus MXN \$55,257 MM with the CNBV model) and the Net Result would increase to MXN \$1215 MM (versus MXN \$1204 MM with the CNBV model) at the close of June 2014; with both cases being aided by the decrease of expected losses. With this, the ROA would go from the 2.18% obtained with the PI model of the CNBV, to a possible ROA of 2.20% with the proposed PI_C mode.

The creation of reserves and the punishment of credits are significant operations that greatly impact the level of capitalization, to the point that credit provisions are the second most significant expense after wages and salaries; therefore, these must be reported immediately from the Finance or Risks area the moment they originate. In this manner, the CEO and the authorized bodies (Risks Committee and Board of Directors) of the financial institution have knowledge of the financial effect and thus are able to make important decisions regarding the credit limits and the risk appetite that the credit institution is willing to tolerate with regard to their levels of profitability.

With scarce research related to Behavioral Scoring models to measure the credit risk behavior of Consumer Banking in Mexico, it is understood that the banking sector itself and the CNBV still have a broad margin of discussion in this regard. Particularly to update the current methodology in the CBR regarding retail provisions.

References

Altman, E. I. (1968). Financial ratios, discriminant analysis and the prediction of corporate bankruptcy. *Journal of Finance*, *23*(4), 589–609. http://dx.doi.org/10.1111/j.1540-6261.1968.tb00843.x

Black, F., & Scholes, M. (1972). The valuation of option contracts and a test of market efficiency. *Journal of Finance*, *27*(2), 399–417. http://dx.doi.org/10.1111/j.1540-6261.1972.tb00969.x

CNBV. (2014). *Normatividad Comisión Nacional Bancaria y de Valores, Banca Múltiple, Circular Única de Bancos*. Available in: http://www.cnbv.gob.mx/Paginas/NORMATIVIDAD.aspx [Consulted in 20/12/2016]

Costa, T., Boj, E., & Fortiana, J. (2012). Bondad de ajuste y elección del punto de corte en regresión logística basada endistancias. Aplicación al problema de Credit Scoring. *Anales del Instituto de Actuarios Españoles*, *18*, 19–40. http://dx.doi.org/10.1007/s11009-014-9415-6

Durand, D. (1941). *Risk Elements in Consumer Instalment Financing, Studies in Consumer Instalment Financing*. NewYork, NY: National Bureau of Economic Research.

Fisher, R. A. (1936). The use of multiple measurements in taxonomic problems. *Annals of Eugenics*, *7*(2), 179–188. http://dx.doi.org/10.1111/j.1469-1809.1936.tb02137.x

Hand, D. J., & Jacka, S. D. (1998). Consumer credit and statistics. In D. J. Hand, & S. D. Jacka (Eds.), *Statistics in finance* (pp. 69–81). London: Arnold.

Hand, D. J., & Henley, W. E. (1997). Statistical classification methods in consumer credit scoring: A review. *Journal of the Royal Statistical Society: Series A*, *160*(Part 3), 523–541. http://dx.doi.org/10.1111/j.1467-985X.1997.00078.x

Hsia, D. C. (1978). Credit scoring and the equal credit opportunity act. *The Hasting Law Journal*, *30*, 371–448.

Huang, C.-L., Chen, M.-C., & Wang, C.-J. (2007). Credit scoring with a data mining approach based on support vectormachines. *Expert Systems with Applications*, *33*(4), 847–856. http://dx.doi.org/10.1016/j.eswa.2006.07.007

Joanes, D. N. (1993). Reject inference applied to logistic regression for credit scoring. *IMA Journal of Mathematics Applied in Business and Industry*, *5*(1), 35–43. http://dx.doi.org/10.1093/imaman/5.1.35

Lee, M. K. O., Cheung, C. M. K., & Chen, Z. (2005). Acceptance of internet-based learning medium: The role of extrinsic and intrinsic motivation. *Information & Management*, *42*(8), 1095–1104. http://dx.doi.org/10.1016/j.im.2003.10.007

Lewis, E. M. (1992). *An introduction to credit scoring*. pp. 86. San Rafael, CA: Fair, Isaac and Co., Inc.

Mays, E. (1998). *Credit risk modeling, design and application*. Chicago: Fitzroy Dearborn Publishers.

Mays, E. (2004). *Credit scoring for risk managers. The handbook for lenders*. Mason, OH: Thomson South Western.

Merton, R. C. (1974). On the pricing of corporate debt: The risk structure of interest rates. *Journal of Finance*, *29*, 449–470. http://dx.doi.org/10.1111/j.1540-6261.1974.tb03058.x

Myers, J. H., & Forgy, E. W. (1963). The development of numerical credit evaluation systems. *Journal of the American Statistical Association*, *58*, 799–806. http://dx.doi.org/10.2307/2282727

Reichert, A. K., Cho, C., & Wagner, G. M. (1983). An examination of conceptual issues involved in developing credit scoring models. *Journal of Business Economics and Statistics*, *1*(2), 101–114. http://dx.doi.org/10.2307/1391851

Rosenberg, E., & Gleit, A. (1994). Quantitative methods in credit management: A survey. *Operations Research*, *42*, 589–613. http://dx.doi.org/10.1287/opre.42.4.589

Thomas, L. C., Crook, J. N., & Edelman, D. B. (1992). *Credit scoring and credit control*. Oxford: Oxford University Press.

Thomas, L. C., Edelman, D. B., & Crook, J. N. (2002). *Credit scoring and its applications. SIAM monographs on mathematical modeling and computation*. Philadelphia, PA: Siam.

Thomas, L. C., Edelman, D. B., & Crook, J. N. (2004). *Readings in credit scoring*. Oxford: Oxford University Press. http://dx.doi.org/10.1137/1.9780898718317

Weber, R. (1999). Applications of fuzzy logic for credit worthiness evaluation. In R. R. Ribeiro, H.-J. Zimmermann, R. R. Yager, & J. Kacprzyk (Eds.), *Soft computing in financial engineering* (pp. 388–401). Heidelberg: Physica-Verlag.