



Disponible en www.sciencedirect.com

www.cya.unam.mx/index.php/cya

Contaduría y Administración 62 (2017) 377–398



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

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

José Carlos Trejo García*, Miguel Ángel Martínez García
y Francisco Venegas Martínez

Instituto Politécnico Nacional, México

Recibido el 9 de diciembre de 2014; aceptado el 10 de agosto de 2015

Disponible en Internet el 8 de febrero de 2017

Resumen

La predicción temprana de malos deudores para créditos revolventes 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/>).

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

* Autor para correspondencia.

Correo electrónico: jtrejog@ipn.mx (J.C. Trejo García).

La revisión por pares es responsabilidad de la Universidad Nacional Autónoma de México.

<http://dx.doi.org/10.1016/j.cya.2017.01.003>

0186-1042/© 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/>).

Abstract

The early prediction of bad debtors in revolving loans in Mexico is a relevant current issue. The proposed econometric model of behavioral scoring considers the changes in the characteristics of consolidated clients and produces better results than those obtained with methodology used by the CNBV on provisions. The obtained results show the possibility of replacing the current model, minimizing the expected loss and increasing the ROA of Mexican financial institutions in 2.20% complying the methodology and statistical testing criteria according to the Unique Banking Dispositions and the guidelines of Basel II on credit risk. © 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/>).

Keywords: Banks; Credit; Econometric modeling; Data estimation methodology; Optimization techniques

JEL classification: G21; E51; C5; C80; C61

Introducción

La administración del riesgo crediticio durante las últimas décadas ha sido una de las áreas con mayor crecimiento. Las técnicas de calificación más utilizadas para la administración del riesgo crediticio han sido el *credit scoring* (otorgamiento de crédito) y el *behavioral scoring* (comportamiento crediticio), así como varias herramientas para la estimación del riesgo financiero con relación a los préstamos o financiamientos al menudeo. En el mercado mexicano crediticio al consumo se tiene la siguiente clasificación en tres tipos de *pools* (carteras): créditos revolventes, créditos personales y créditos a la vivienda¹.

Todos los solicitantes de crédito, así como los clientes consolidados de instituciones financieras en México, cuentan con una calificación tanto en la solicitud del crédito como en el comportamiento de la vida del crédito. Existen dos objetivos importantes en las técnicas para la calificación crediticia. Uno de ellos es la necesidad de identificar el riesgo de consumo, mientras que el otro es minimizar el porcentaje de clientes incumplidos, con lo cual las instituciones bancarias o crediticias optimizan sus carteras para un buen y mejor negocio.

Por lo anterior, se requiere desarrollar nuevos modelos con base en la información histórica de los clientes que permitan generar modelos de decisión en el otorgamiento y considerar el comportamiento con los clientes consolidados (Mays, 2004).

De acuerdo con la información, al cierre de junio 2014, por parte de la Comisión Nacional Bancaria y de Valores (CNBV)², uno de los principales instrumentos de crédito en México son las tarjetas de crédito, ya que existen aproximadamente 23 millones de plásticos autorizados por 24 entidades financieras reconocidas.

Esta investigación desarrolla una técnica de calificación crediticia considerando la regulación nacional e internacional. Además, el modelo propuesto es óptimo en el sentido de que minimiza la

¹ Título Segundo Disposiciones Prudenciales, Capítulo V Calificación de Cartera Crediticia; documentado en las Disposiciones de Carácter General Aplicables a las Instituciones de Crédito, o comúnmente conocidas como la Circular Única de Bancos que emite la Comisión Nacional Bancaria y de Valores, publicada en el Diario Oficial el 2 de diciembre 2005 y con última modificación el 31 de julio 2014.

² Sección Información Estadística del Sector Supervisado Banca Múltiple, con datos al cierre del mes de julio 2014 para Créditos al Consumo para Tarjeta de Crédito (<http://www.cnbv.gob.mx/SECTORES-SUPERVISADOS/BANCA-MULTIPLE/Paginas/Información-Estadística.aspx>).

pérdida esperada en el sector de tarjetas de crédito, con repercusiones positivas en la determinación de reservas crediticias, generando un efecto positivo en los activos, en los resultados netos y en la rentabilidad de las entidades financieras (Banca Múltiple).

El documento se organiza de la siguiente forma: en la segunda sección se presenta el marco teórico en términos de avances estadísticos y econométricos en la administración del riesgo crediticio; en la tercera sección se establecen las bases y criterios regulatorios de créditos revolventes en México; en la cuarta sección se presenta la metodología propuesta y se proporcionan las pruebas de significación estadística; en la quinta sección se establece el modelo de optimización; en la sexta sección se compara la rentabilidad entre el modelo propuesto y el de CNBV; en la séptima sección se presentan y discuten los resultados obtenidos; por último, se presentan las conclusiones.

Avances en la administración de riesgos crediticios

Los avances en las tecnologías de información en la administración del riesgo crediticio permiten que las entidades financieras crediticias automaticen las decisiones sobre aceptación o rechazo de una solicitud de crédito y la administración de una cartera crediticia (clientes consolidados) en un *cross-sell* (venta cruzada). Hace unos años, dicha administración crediticia se realizaba solo con la experiencia o percepción del ejecutivo. Ahora, uno de los modelos más usados para la evaluación de créditos es el *scoring model* (modelo de calificación), el cual determina un *score* (puntaje) para clientes que solicitan un crédito identificando a aquellos que tienen la posibilidad de incumplir con sus pagos. La literatura, con respecto de la calificación crediticia, es amplia, y basta mencionar los modelos de Rosenberg y Gleit (1994), Merton (1974), Hand y Jacka (1998), Thomas, Crook y Edelman (1992) y Lewis (1992), mientras que para calificaciones de comportamiento o *behavioral scoring* se cuenta con el trabajo de Mays (1998).

Modelos analíticos de scoring

Existen en la literatura especializada un conjunto amplio de métodos y técnicas cuantitativas para predecir la probabilidad de que un cliente falle o incumpla³ y, por consiguiente, no se recupere el crédito otorgado por alguna institución financiera. Los modelos *scoring* son herramientas que utilizan la clasificación de los solicitantes o clientes consolidados por nivel de riesgo con base en el suministro de la información de clientes en las solicitudes de crédito y comportamiento de pagos. El modelo proporciona la suma de puntos para cada componente cuantificable, produciendo una puntuación de rango ordenado o escala de puntuación.

La identificación de clientes potenciales que pueden generar pérdidas a la entidad financiera requiere buscar mejores lineamientos para su tratamiento y reservar un capital mínimo en caso de incumplimiento y/o migración de *buckets* (impagos) tolerables a *buckets* catalogados como vencidos. Véanse al respecto las investigaciones de Hsia (1978), Reichert, Cho y Wagner (1993), Joanes (1993) y Hand y Henley (1997), que detallan el estudio de relaciones particulares entre las distribuciones de clientes buenos y malos.

La historia de los modelos de *scoring* se remonta a Fisher (1936), con la idea de discriminar diferentes grupos dentro de una población específica. Esta idea fue desarrollada más ampliamente

³ Cuando la probabilidad de incumplimiento es mencionada, es usual pensar en el modelo de Merton (1974), el cual se basa en la fórmula de Black y Scholes (1972). En esta investigación se desarrollará un modelo alternativo.

por Durand (1941) aplicando contexto financiero para discriminar entre un «buen» y un «mal» pagador. De igual manera, Thomas, Edelman y Crook (2002) analizan las ventajas y limitaciones de la creación de proveedores de modelos *scoring* (*vendors* «Bill Fair y Earl Isaac»⁴), los cuales a finales de los cincuenta empezaron con el desarrollo de un sistema de análisis de riesgo de créditos. En los años sesenta, con la creación de las tarjetas de crédito, los modelos de *scoring* mostraron su importancia y utilidad. Según Myers y Forgy (1963), este tipo de modelos es superior como predictor que cualquier juicio experto cualitativo. Otro aspecto importante en este contexto fue el del Z-Score propuesto por Altman (1968), que ha sido aplicado en el sector financiero.

Optimización en entidades financieras

En términos prácticos, los modelos de *scoring* permiten una reducción significativa en los tiempos de ejecución de los distintos procesos financieros para el otorgamiento y seguimiento de un crédito, permitiendo con ello una mayor automatización y reduciendo la necesidad de la intervención humana en la evaluación y estimación del riesgo crediticio. Los principales usuarios de este tipo de modelos son los bancos y las instituciones financieras, así como las compañías de seguro o las cadenas de *retail* (créditos al consumo). Entre las principales características del *scoring* están la posibilidad de gestionar y administrar el riesgo. Los beneficios reportados por la aplicación de estos modelos no solo afectan a los bancos e instituciones financieras, sino también directamente a todos los clientes del sector financiero, pues reduce la discriminación errónea de clientes que solicitan algún crédito y provee un análisis más objetivo, concentrando en un solo modelo múltiples factores que pueden afectar el riesgo de una solicitud o un seguimiento de pagos.

Marco regulatorio local e internacional

Aunque algunas instituciones financieras en México desarrollan sus propios modelos de *scoring* con base en los lineamientos establecidos por la CNBV⁵, la mayoría de entidades constituyen sus reservas con base en los lineamientos establecidos por dicha comisión. Los vendedores o consultores en la creación de modelos *scoring* los construyen basándose en la información y en los parámetros específicos proporcionados por las áreas de administración de riesgo crediticio de la entidad financiera. Por lo tanto, la entidad financiera debe asegurarse de que el desarrollo de dichos modelos sea acorde a los objetivos del banco y con respecto de su tolerancia al riesgo.

El acuerdo internacional sobre regulación y supervisión bancaria denominado «Nuevo Acuerdo de Capital»⁶ exige a las entidades financieras de los países miembros una revisión de sus dotaciones de capital para cubrir los riesgos. Estos parámetros les obligan a disponer de herramientas que les permitan establecer modelos de medición (*scorings* y *ratings*) con el propósito de discriminar entre clientes según su perfil de riesgo y evaluar la exposición y la severidad en el riesgo de crédito. El acuerdo de Basilea II también obliga a las entidades financieras no solo a adaptar sus sistemas de cálculo del consumo de capital, sino también a modificar los sistemas de *reporting* (informes financieros) y de análisis de la información. Ambos elementos —es decir, los informes financieros y el análisis de los mismos— son aspectos clave en Basilea II necesarios para gestionar

⁴ Fair-Isaac Company.

⁵ Circular Única de Bancos con última modificación al 31 de julio 2014, Anexo 15 Requisitos Mínimos para la Autorización Metodologías Internas.

⁶ Acuerdo internacional y conocido como Basilea II y aprobado en 2004.

grandes bases de datos, capaces de ofrecer la información para cuantificar los riesgos de cada operación, lo que supone un verdadero reto para los bancos.

El elemento clave para analizar el proceso de calificación de riesgos mediante modelos internos (IRB)⁷, según la normativa de Basilea II, es que las entidades financieras crediticias dispongan de un modelo de *credit scoring* que les permita medir la probabilidad de incumplimiento del crédito que se otorga. Para tal efecto, el método estándar o regulatorio de la CNBV permite la validación con datos actuales y el método IRB de *behavioral scoring* para el caso mexicano a través de una muestra de información nacional.

Administración del riesgo crediticio al menudeo en México

De manera general, el crédito es un compromiso pactado entre una persona física o moral y una entidad financiera a fin de otorgarle capacidad de compra por adelantado al deudor. El crédito permite realizar compras por anticipado con base en su capacidad de pago. En México, la CNBV menciona que toda actividad crediticia significa la colocación de los recursos tanto propios como captados de terceros (ahorradores) mediante operaciones de préstamo, descuento, asunción, garantías o créditos en su más amplio sentido, así como cualquier operación bancaria que genere o permita un derecho de crédito a favor de las instituciones financieras con un grado de incumplimiento⁸.

La cartera de crédito al consumo se conforma de todos aquellos créditos directos, incluyendo los de liquidez que no cuenten con garantía de inmuebles, denominados en moneda nacional, extranjera, en UDIs o en VSM, así como los intereses que generen, otorgados a personas físicas, derivados de operaciones de tarjeta de crédito, de créditos personales, de créditos para la Adquisición de Bienes de Consumo Duradero (conocidos como ABCD), que considera entre otros al crédito automotriz y las operaciones de arrendamiento financiero que sean celebradas con personas físicas⁹.

Con base en diversas investigaciones de Banxico, el riesgo de crédito es el caso particular cuando el contrato no puede ser cumplido por el deudor hacia el acreedor (otorgante del crédito). Recientemente, además del caso de incumplimiento, se han incorporado eventos que afectan el valor de un crédito, sin que necesariamente signifique incumplimiento del deudor. Normalmente, los factores que se deben tomar en cuenta al medir riesgo de crédito son las probabilidades de incumplimiento y/o de migración en la calidad crediticia del deudor, las correlaciones entre incumplimientos, la concentración o segmentación de la cartera, la exposición a cada deudor y la tasa de recuperación en caso de incumplimiento de los deudores¹⁰.

Para la determinación de provisiones en México de la cartera crediticia al consumo y administradas por cada entidad financiera se tiene como base la regulación impuesta por la CNBV

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 Circular Única de Bancos, actualización 31 de julio 2014, Título Primero Disposiciones Generales del Capítulo I Definiciones, Artículo 1, Fracción I.

⁹ Circular Única de Bancos, actualización 31 de julio 2014, Título Primero Disposiciones Generales del Capítulo I Definiciones, Artículo 1, Fracción XXIX.

¹⁰ Banxico 2005, Definiciones Básicas de Riesgos (<http://www.banxico.org.mx/sistema-financiero/material-educativo/intermedio/riesgos/%7BA5059B92-176D-0BB6-2958-7257E2799FAD%7D.pdf>).

mediante la Circular Única de Bancos (CUB)¹¹, con factores como la Severidad de la Pérdida (SP o *Loss Given Default*), que es lo que pierde el acreedor en caso de incumplimiento del deudor una vez tomados en cuenta todos los costos implicados en la recuperación (costos de recobro, costos judiciales, etc.). Con respecto a la Exposición al Incumplimiento (EI o *Exposure at Default*), es el saldo que debe el deudor en un momento dado en caso de incumplimiento. La medida de riesgo crediticio en esta investigación es la Probabilidad de Incumplimiento (PI o *Probability of Default*), que expresa qué tan probable es que un acreditado deje de cumplir con sus obligaciones contractuales. El mínimo valor es cero, lo cual indicaría que es imposible que incumpla con sus obligaciones, y su máximo valor es uno o cien por ciento, cuando es seguro que incumpla.

Con base en la regulación vigente, la estimación de la PI debe ser por crédito. Las Reservas o Pérdidas Esperadas (R o PE) es la media de la distribución de pérdidas y ganancias. Es decir, indica cuánto se puede perder en promedio y normalmente está asociada a la política de reservas preventivas que la institución debe tener contra riesgos crediticios. Para ello, se calcula la PE considerando el porcentaje de reservas (PI*SP) por la EI¹².

Metodología propuesta y optimización matemática

La metodología que utiliza esta investigación para la construcción de un modelo óptimo de *behavioral scoring* para el producto de créditos revolventes administrados por las entidades financieras mexicanas se basa en los lineamientos regulatorios locales de la CUB, la cual estipula en el artículo 92 Fracción III las distintas etapas para administrar el riesgo crediticio del portafolio de créditos revolventes en México.

De acuerdo con la regulación local mencionada, es necesario detallar cada etapa para el desarrollo de la propuesta.

Etapas 1. Datos y criterio en la determinación del objetivo

Primero se obtuvo una muestra de clientes clasificados a priori como «buenos» y «malos» (de acuerdo con Huang, Chen y Wang, 2007). De esta manera, se consideró una muestra de 43,323¹³ cuentas con los criterios regulatorios y el evento de incumplimiento bajo el criterio de que un cliente no realice dos o más pagos consecutivos¹⁴ en los próximos 12 meses, objetivo esencial de la presente investigación.

Etapas 2. Compilación y proceso de información

La construcción del modelo alternativo de *behavioral scoring* (Weber, 1999) se realiza a través de la discriminación y la regresión logística, la cual se emplea para validación con mejores niveles de significación.

¹¹ Circular Única de Bancos 2014, actualización 31 de julio 2014, Título Segundo Disposiciones Prudenciales del Capítulo I Otorgamiento de Créditos y Capítulo V Calificación de Cartera Crediticia.

¹² Circular Única de Bancos 2014, actualización 31 de julio 2014, Título Segundo Disposiciones Prudenciales del Capítulo V Calificación de Cartera Crediticia, Sección Primera, Apartado B.

¹³ La extracción de la base de datos de clientes de algunos bancos mexicanos y de otros temas relacionados al cierre de julio 2014 se puede recurrir a la siguiente página: <http://archive.ics.uci.edu>, sección de *Datasets*.

¹⁴ Circular Única de Bancos con última modificación al 31 de julio de 2014, Anexo 33 Serie B Criterios relativos a los conceptos que integran los estados financieros, B-6 Cartera de Crédito.

Para ello, se trabajará con variables cuantitativas independientes y la variable dicotómica dependiente, bajo la condición:

$$\text{Variable dependiente } Y_i = \begin{cases} 1 \text{ (malo)} & ; \text{ si Impago} > 2 \text{ meses de atraso} \\ 0 \text{ (bueno)} & ; \text{ si Impago} \leq 2 \text{ meses de atraso} \end{cases}$$

La técnica estadística más utilizada por la industria financiera corresponde a la regresión logística (Thomas, Edelman y Crook, 2004). Esta técnica es menos restrictiva, siendo una alternativa del análisis discriminante. Es importante mencionar que en los últimos años han aparecido una serie de técnicas nuevas llamadas «minería de datos» (Weber, 1999), las cuales han sido utilizadas también para la construcción de modelos de *scoring*. Estas técnicas presentan la ventaja de no tener demasiados requerimientos y supuestos para las variables de entrada, aumentando su validez. Esta técnica ha sido intensamente utilizada para la construcción de modelos de *scoring* para la comprensión de patrones complejos de un determinado segmento de clientes, teniendo la capacidad de modelar relaciones no lineales entre las variables.

Etapa 3. Estimación probabilística

Los modelos probabilísticos lineales plantean una serie de problemas que han llevado a la búsqueda de otros modelos alternativos que permitan estimaciones más confiables cuando se tienen variables dicotómicas. Para evitar que la variable endógena estimada pueda encontrarse fuera del rango (0, 1), las alternativas disponibles son utilizar modelos de probabilidad no lineales, donde la función de especificación utilizada garantiza un resultado en la estimación comprendido en el rango (0, 1).

Dado que el uso de una función de distribución garantiza que el resultado de la estimación esté acotado entre 0 y 1, en principio las posibles alternativas son varias, siendo las más habituales la función de distribución logística, que ha dado lugar al modelo *logit*, relacionando la variable endógena Y_i con las variables explicativas X_i a través de una función de distribución ($Y_i = 1$). En cuanto a la interpretación de los parámetros estimados en un modelo *logit*, el signo de los mismos indica la dirección en que se mueve la probabilidad cuando aumenta la variable explicativa correspondiente.

A través de la linealización del modelo y partiendo de la ecuación general del modelo *logit*, se define Y_i como la probabilidad del estado o la alternativa 1, partiendo de lo siguiente:

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

Esto satisface $0 \leq y_i \leq 1$. Por lo tanto:

$$p_i + p_i e^{\beta_0 + \beta_k X_k} = e^{\beta_0 + \beta_k X_k} \quad (2)$$

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

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

Al cociente entre la probabilidad de que ocurra un hecho, o de que se elija la opción 1, frente a la probabilidad de que no suceda el evento en cuestión, o de que se elija la opción 0, se le denomina

odds ratio. Su interpretación es la «ventaja» o preferencia de la opción 1 frente al 0, es decir, el número de veces que es más probable que ocurra el incumplimiento frente al que no ocurra.

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

Tomando logaritmos naturales del *odds ratio* se linealiza la ecuación del modelo *logit*, respetando el objetivo de que los valores estimados se encuentren dentro del rango (0, 1), obteniéndose la expresión:

$$\text{Ln} \left[\frac{P_i}{(1 - p_i)} \right] = \text{Ln}(e^{\beta_0 + \beta_k X_k}) = \beta_0 + \beta_k X_k \quad (6)$$

La nueva variable $\text{Ln} \left[\frac{P_i}{(1 - p_i)} \right]$ generada representa en una escala logarítmica la diferencia entre las probabilidades de que ocurra incumplimiento en créditos revolventes (Lee, Cheung y Chen, 2005).

El *odds ratio* tal y como está construido (cociente entre probabilidades) siempre será mayor o igual que 0. El campo de variación del ratio va desde 0 hasta $+\infty$, y su interpretación se realiza en función de que el valor sea igual, menor o superior a la unidad: si el antilogaritmo toma el valor 1 significa que la probabilidad de incumplimiento es la misma de que no ocurra; si el ratio es menor que 1 indica que la ocurrencia de la probabilidad de incumplimiento tiene menor posibilidad que la ocurrencia del cumplimiento del cliente, mientras que si es mayor que 1, entonces la probabilidad de incumplimiento del clientes es mayor que el cumplimiento.

$$\text{Cociente Odds Ratio} = \frac{\frac{P_i}{(1-p)}}{\frac{P_j}{(1-p_j)}} = \frac{e^{\beta_0 + \beta_k X_{ki}}}{1 + e^{\beta_0 + \beta_k X_{kj}}} = e^{\beta_k (X_{ii} - X_{jj})} \quad (7)$$

Etapa 4. Estimación con máxima verosimilitud

Para la estimación del modelo *logit* en la búsqueda de obtener los parámetros predictores, se utiliza el método de máxima verosimilitud (MV). Como p_i toma dos valores —con probabilidad de incumplimiento p_i y su contraparte $1 - p_i$ —, tiene una distribución de estilo Bernoulli.

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

La función de MV para una muestra aleatoria de n datos (x_i, y_i) se calcula como:

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

Al aplicar logaritmos:

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

la función logarítmica de verosimilitud satisface:

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

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

Ahora la ecuación anterior se sustituye en la ecuación de la función logarítmica de verosimilitud. De aquí se obtiene la función de verosimilitud en logaritmos en términos de los parámetros β dada por:

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

Para la obtención de los estimadores β de máxima verosimilitud, se deriva $L(\beta)$ respecto de cada uno de los parámetros β_j con $j = 1, 2, \dots, p$ y se iguala a cero. En términos de matrices:

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

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

Si se expresa cada una de estas derivadas en un vector columna:

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

Si ahora se iguala a cero el vector columna:

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

La β estimada es el vector de parámetros que cumple el sistema matricial, se calcula p_i en términos de esos estimadores y con ello se obtiene una estimación para y_i , tal que la estimación de probabilidad sea:

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

Linealizando la ecuación 14, permite modelar la PI mediante una regresión logística múltiple:

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

Resultados empíricos

Para validar la metodología propuesta se han seleccionado variables relevantes en la determinación de la probabilidad de incumplimiento propuesto.

Determinación y selección de variables

Para evitar problemas de multicolinealidad es de gran ayuda que las variables seleccionadas cuenten con pruebas de correlación. Para ello se selecciona el conjunto de variables explicativas y significativas, evitando entonces el incluir variables poco significativas o con información redundante (colinealidad), lo que puede distorsionar la capacidad predictiva de la función discriminante estimada.

Para medir la correlación, se considera el siguiente cálculo:

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

Con una base de datos de 43,323 clientes consolidados y administrados por las instituciones financieras mexicanas al cierre de agosto 2014, se consideraron las 6 variables acorde a la metodología regulatoria de la CNBV¹⁵: variable de cumplimiento/incumplimiento (Y), número de impagos (ACT), historial de impagos (HIST), meses transcurridos de crédito (ANT), relación pago-saldo (%PAGO) y relación saldo a pagar – límite de crédito (%USO).

Para el modelo propuesto, y en relación con la base de datos de los mismos 43,323 clientes, se identificaron solo 5 variables significativas: variable de cumplimiento (Y), meses transcurridos de crédito (ANT), límite de crédito (límite_crédito), historial de impagos (HIST) y relación pago-saldo (%PAGO).

La variable Y para ambos modelos de estudio (vigente y propuesto) fue recodificada a partir de los criterios mencionados en la metodología: variable binaria que adopta el valor 1 (incumple) y el valor 0 (cumple) (tabla 1).

Como se observa, los problemas de multicolinealidad son notorios al momento de utilizar todas las variables con base en la metodología vigente de la CNBV, principalmente HIST y ANT, con una correlación superior al 60%. Además, se observa la falta de impacto con Y por parte de %USO, lo que puede provocar menor predicción con el modelo actualmente utilizado por la regulación mexicana.

La ausencia de multicolinealidad, así como el alto impacto de las variables independientes (ANT, límite_crédito, HIST y %PAGO) sobre la variable Y, sí influyen significativamente, y esto puede reflejar mayor predictibilidad para medir el nivel de incumplimiento.

De esta manera, las variables propuestas en esta investigación muestran ausencia de multicolinealidad, ya que los niveles de correlación entre variables independientes endógenas son menores al 47%, mientras que la relación de ANT, límite_crédito e HIST con la variable dependiente (Y)

¹⁵ Circular Única de Bancos 2014, actualización 31 de julio 2014, Título Segundo Disposiciones Prudenciales del Capítulo I Otorgamiento de Créditos y Capítulo V Calificación de Cartera Crediticia, art. 92, fracción III.

Tabla 1
Pruebas de multicolinealidad con las variables utilizadas por CNBV

	Y	ACT	HIST	ANT	%PAGO	%USO
Y	100%					
ACT	81%	100%				
HIST	93%	89%	100%			
ANT	–71%	–59%	–40%	100%		
%PAGO	–95%	–1%	–1%	1%	100%	
%USO	40%	23%	40%	–44%	–1%	100%

Fuente: elaboración propia con la paquetería Stata 13, aplicación de pruebas de correlación.

Tabla 2
Ausencia de multicolinealidad con las variables propuestas

	Y	ANT	límite_crédito	HIST	%PAGO
Y	100%				
ANT	–71%	100%			
límite_crédito	–52%	46%	100%		
HIST	93%	–40%	–32%	100%	
%PAGO	–95%	1%	1%	–1%	100%

Fuente: elaboración propia con la paquetería Stata 13, aplicación de pruebas de correlación.

es superior al 56%. El impacto de %PAGO con Y es mayor al 92%. Esta prueba de correlación indica que el modelo propuesto podría ser predictivo y más estable.

Regresión y significatividad logística

De acuerdo con las técnicas de inferencia en regresión logística (Joanes, 1993), un buen modelo debe satisfacer dos condiciones: la primera es que tenga una fuerte capacidad predictora y la segunda es que la estimación de los parámetros tenga una alta precisión. Una condición adicional es que el modelo sea lo más sencillo posible, esto es, que contenga el mínimo de variables explicativas y que satisfaga las dos condiciones anteriores. Empleando la muestra de 43,323 clientes y las variables propuestas en la tabla 2, se estima un modelo *logit* explicado en la metodología, calculando a su vez la probabilidad de que un cliente pague su préstamo de créditos revolventes.

En la estimación de la probabilidad *logit* fueron necesarias 9 iteraciones para estimar el modelo. La función de *LR chi²* indica que los coeficientes son conjuntamente significativos para explicar la probabilidad de que los 43,323 clientes se encuentren en incumplimiento; así, el valor estadístico de *Prob > chi²* indica que se puede rechazar en un 1% la hipótesis de que todos los coeficientes sean igual a cero. Además, el estadístico de *Pseudo R²* indica que aproximadamente el 97% de la variación de la variable dependiente es explicado por la variación de las variables independientes del modelo propuesto para México.

La calidad de ajuste del modelo es de 99.38%¹⁶, resultado de la razón de provisiones correctas contra número de observaciones; esto confirma en general que el modelo predice las observaciones correctamente.

¹⁶ Mediante el programa Stata 13, se aplicó el comando «*stat clas*».

Tabla 3

Estimación de la regresión *logit* del modelo propuesto (PI_C)

Y	Coef.	P>z	[95% Conf. Interval]	
ANT	0.012	0.0000	0.014	0.0116
SALDO	0.00002	0.0000	0	0
HIST	2.062	0.0000	2.059	2.201
%PAGO	-7.293	0.0000	-8.5698	-6.0548
_CONS	-3.04523	0.0000	-5.3252	-4.112

Observaciones = 13,000; LR $\chi^2(4) = 53,200$; Prob > $\chi^2 = 0.0000$; Pseudo R2 o McFadden test = 0.9626; Log likelihood = -722.0952.

Fuente: elaboración propia con la paquetería Stata 13 y con base en la muestra de 43,323 clientes.

Significación de la regresión logística propuesta

Es necesario destacar que al momento de realizar de nuevo la regresión logística para aquellas variables que actualmente la regulación contempla se obtuvo un error en la obtención de resultados, debido a la alta multicolinealidad de la variable ACT con HIST. Por lo tanto, dichas variables explicativas y utilizadas actualmente por la CNBV no son necesarias del todo para continuar con el modelo; de esta manera la variable ACT debe ser excluida.

Al utilizar criterios estadísticos como el de correlaciones parciales es posible obtener todas las variables potenciales y escoger entre ellas las mejores. Al respecto, para determinar si una variable debe ser incluida en un modelo por un peso significativo, se aplica la prueba de estadístico de *Wald*. La prueba resulta de contrastar la hipótesis nula $H_0 : \beta_i = 0$ contra la alternativa $H_1 : \beta_i \neq 0$.

La estimación mostrada en la [tabla 3](#) permite calcular la función *score* a partir de la ecuación (4). Los niveles de significación para el modelo propuesto indican que todos los valores de significación *p-values* cuentan con alta significación, puesto que son menores al 5%. Para obtener el score con las variables propuestas se aplica la transformación *logit* de la ecuación (16) para obtener la probabilidad de incumplimiento calculada (PI_C).

La interpretación de los coeficientes es mediante la medición de la variación del modelo *logit* estimado para una variación unitaria de la variable explicativa dada. Así, si ANT aumenta en un impago, el *logit* estimado aumenta en 0.01 unidades, lo que sugiere una relación positiva con respecto a la PI_C . El mismo caso es para el límite_crédito e HIST: a mayor límite como a mayor aumento de impagos históricos, la PI_C aumenta. En caso contrario, si %PAGO es mayor, entonces la PI_C disminuye en 7.6 unidades. De acuerdo con lo anterior, el modelo econométrico planteado mediante la estimación *logit* permite mejorar la administración del riesgo de créditos revolventes.

Las oportunidades en los modelos *logit* son calculadas mediante regresión logística u *odds*, es decir, mediante la transformación antilogarítmica de $\beta \gg e^\beta$ ([tabla 4](#)).

Estos resultados demuestran cumplimiento a la teoría planteada, ya que si $odd < 1$ significa que la ocurrencia de la PI_C tiende a ser una relación negativa (menor oportunidad), mientras que si $odd > 1$ significa que la ocurrencia de la PI_C tiende a ser una relación positiva (mayor oportunidad).

Por lo tanto, las variables ANT, límite_crédito e HIST indican que si aumentan los impagos y el límite de crédito, entonces la PI aumenta de manera conjunta y aproximada por 17 veces, caso contrario a %PAGO, puesto que el coeficiente es menor que la unidad.

Tabla 4
Estimación de la regresión logística del modelo propuesto (PI_C)

Y	Coef. Odds		[95% Conf. Interval]
ANT	1.01207229	1.01	1.1555
SALDO	1.00002	0.9111	1.0002
HIST	7.86167745	6.9588	8.1516
%PAGO	0.00068028	0.0001	0.0019
_CONS	0.04758537	0.0311	0.0544

Fuente: elaboración propia con la paquetería Stata 13, de acuerdo con la muestra de 43,323 clientes.

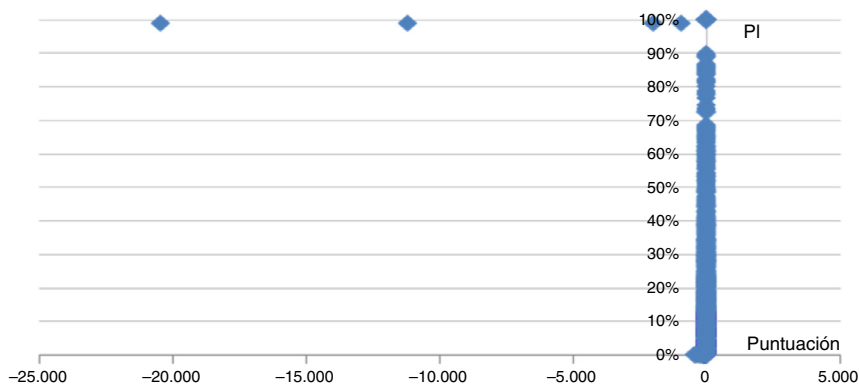


Figura 1. Distribución logística *score* versus probabilidad de incumplimiento con la metodología vigente de la CNBV. Fuente: elaboración propia con la paquetería Stata 13, de acuerdo con la regresión logística y a la distribución *logit*.

Distribución *logit*

La distribución *logit* obtenida con la metodología vigente de la CNBV¹⁷ es como se muestra en la figura 1.

La figura 1 indica que no existe una distribución óptima, puesto que hay observaciones negativas fuera del comportamiento normal respecto a las demás observaciones.

La PI promedio de la metodología de la CNBV es del 43.4%; por lo tanto, es necesario validar mediante el modelo propuesto (PI_C) en esta investigación si existe una distribución *logit* óptima y con apego al comportamiento teórico.

La figura 2 indica que sí existe una distribución óptima, puesto que no hay observaciones fuera del comportamiento normal respecto a la mayoría observaciones. El promedio de la PI_C es del 42.6%, es decir, menor que la PI de la regulación vigente (43.4%).

Análisis de distribución del porcentaje de reservas en México

Mediante pruebas de discriminación poblacional, el análisis discriminante consiste en una técnica multivariante que permite estudiar simultáneamente el comportamiento de un grupo de

¹⁷ Circular Única de Bancos 2014, actualización 31 de julio 2014, Título Segundo Disposiciones Prudenciales del Capítulo I Otorgamiento de Créditos y Capítulo V Calificación de Cartera Crediticia, art. 92, fracciones I y III. $ACT \geq 4$, $PI = 100\%$ y $ACT < 4$, el cálculo de la PI es mediante el modelo planteado en el art 92 fracción III.

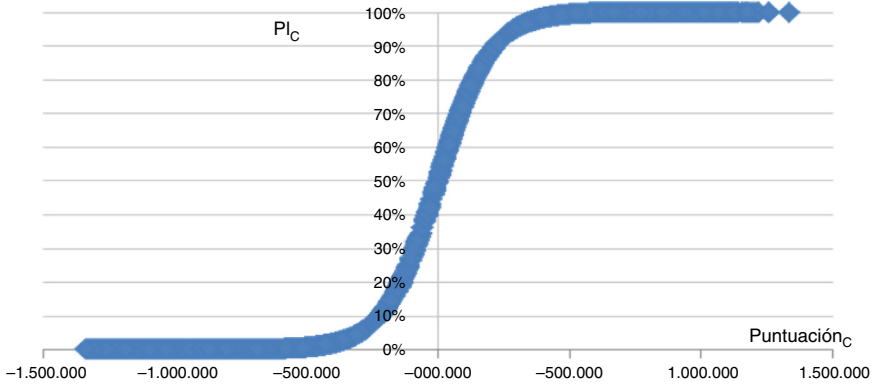


Figura 2. Distribución logística score versus probabilidad de incumplimiento mediante el modelo propuesto. Fuente: elaboración propia con la paquetería Stata 13, de acuerdo con la regresión logística y a la distribución *logit*.

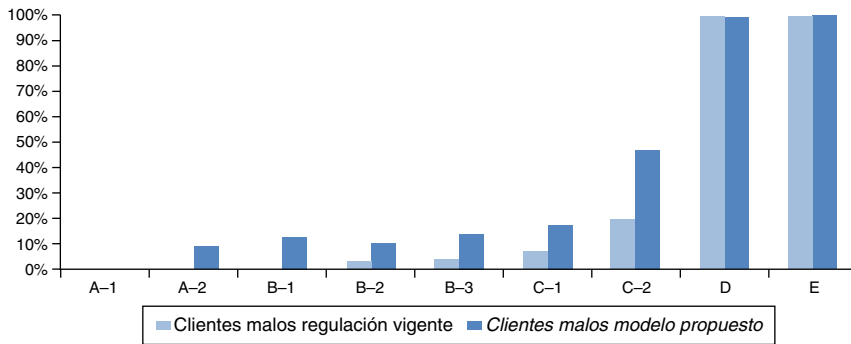


Figura 3. Grado de discriminación de clientes malos por grado de riesgo según el modelo vigente versus el modelo propuesto. Fuente: elaboración propia con la paquetería Stata 13, de acuerdo con la distribución del porcentaje de reservas entre los grados de riesgos (CNBV, 2014).

variables independientes con la intención de clasificar una serie de casos en grupos previamente definidos y excluyentes entre sí (Fisher, 1936).

Una vez calculadas las PI y PI_C mencionadas, estas son utilizadas para llevarlas en producto con la SP¹⁸, obteniendo así el porcentaje de reserva. Al validar mediante la metodología actual ($PI*SP$) contra el modelo propuesto (PI_C*SP), en los comportamientos mostrados del porcentaje de reservas y grados de riesgo señalados en la metodología de la CNBV¹⁹ se identifica que el modelo propuesto cuenta con mejor discriminación de clientes buenos y malos en cada grado de riesgo (fig. 3).

¹⁸ Circular Única de Bancos 2014, actualización 31 de julio 2014, Título Segundo Disposiciones Prudenciales del Capítulo I Otorgamiento de Créditos y Capítulo V Calificación de Cartera Crediticia, art. 92, fracción III. $ACT < 10$, $SP = 75\%$ y $ACT \geq 10$, $SP = 100\%$.

¹⁹ Circular Única de Bancos 2014, actualización 31 de julio 2014, Título Segundo Disposiciones Prudenciales del Capítulo I Otorgamiento de Créditos y Capítulo V Calificación de Cartera Crediticia, Sección Quinta De la constitución de reservas y su clasificación por grado de riesgo, art. 129.

En consecuencia, se muestra que con el modelo propuesto la detección de clientes malos es mayor que el modelo vigente utilizado por la CNBV, lo que garantiza una mejor y mayor discriminación de clientes malos o incumplidos, tal y como se muestra notablemente del grado de riesgo A-2 al C-2, mientras que los grupos de riesgo A-1, D y E mantienen su tendencia discriminatoria entre ambos modelos.

Prueba de Kolmogorov-Smirnov

Una vez que se ha calculado el *score* con la fórmula estimada por medio de la regresión *logit*, se pretende determinar si estos valores calculados en la muestra identifican bien a qué grupo pertenecen. Mientras mayor sea la diferencia de los puntajes de los grupos, mayor será la capacidad discriminante del modelo usado. El indicador de Kolmogorov-Smirnov (K-S) es una prueba no paramétrica para la bondad de ajuste para probar que dos muestras independientes provienen de la misma distribución de una variable aleatoria continua. Estas diferencias están determinadas no solo por las medias, sino también por la dispersión, simetría o la oblicuidad. La prueba se construye sobre las hipótesis nula y alternativa como sigue:

H₀. La distribución del *score* para las cuentas buenas y las cuentas malas es igual.

H₁. La distribución del *score* para las cuentas buenas y las cuentas malas no es igual.

El estadístico de prueba es:

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

Con dos colas el estadístico dado:

$$D_n = \sup_x | (F_n(x) - F(x)) | \quad (19)$$

donde \sup_x es la distancia superior del conjunto.

De esta manera se puede comparar el índice discriminatorio de K-S entre el grado de separación de clientes buenos y malos con el modelo vigente versus el modelo propuesto.

En la [tabla 5](#) se observa que con el modelo propuesto el poder de discriminación entre grados de riesgos por porcentaje de reservas (PI_C*SP) es mayor en C-1, lo que permite a la banca de consumo en México la posibilidad de detectar mejor a los clientes buenos y malos, así como aplicar políticas de cobranza más efectivas para poder mejorar ese segmento.

Incluso permite al área de riesgos al menudeo de las instituciones financieras el poder proponer un óptimo nivel de *cross-selling* (venta cruzada) con clientes que se encuentren en grados de riesgos superiores (mejores) a C-1, lo que representa mayor posibilidad de negocio en pro de la rentabilidad en el negocio de créditos al menudeo.

Elección del cut-off o punto de corte

En relación con el nivel de discriminación presentado en la sección anterior, principalmente en la [tabla 5](#), es posible identificar los grados de riesgo que la banca pudo aceptar en administrar la

Tabla 5

Nivel de separación de población por clientes buenos y malos CNBV vs modelo propuesto

Clase	Separación K-S Metodología vigente	Separación K-S Modelo propuesto
A-1	51.03%	53.11%
A-2	52.25%	54.24%
B-1	53.23%	55.10%
B-2	54.11%	55.15%
B-3	55.01%	55.99%
C-1	55.91%	56.33%
C-2	56.01%	56.59%
D	35.12%	27.00%
E	0.00%	0.00%

Fuente: elaboración propia de acuerdo con la metodología de Kolmogorov-Smirnov.

cartera, esto con base en el nivel de probabilidad de incumplimiento dada la máxima separación de distribución de clientes buenos y malos.

El modelo basado en la regresión logística está constituido por un conjunto de pesos relacionados a un grupo de variables o atributos, caracterizando a un conjunto de clientes, y con ello permitiendo la determinación un punto de corte o *cut-off* (umbral).

Con respecto del *cut-off*, este determina la frontera entre ser un «mal» o «buen» cliente. En general, en la industria bancaria se tiene que el costo de aceptar un cliente malo es varias veces mayor que el de rechazar a un cliente bueno (Costa, Boj y Fortiana, 2012).

Con respecto al conjunto de pesos, es posible caracterizar los patrones que describen ambas poblaciones (clientes buenos y malos) y determinar cuál o cuáles de las variables de entrada utilizadas son realmente importantes en términos de una buena predicción por segmentos para la administración de la cartera de créditos revolventes.

De acuerdo a lo anterior, el criterio para la elección del punto de corte, en el modelo logístico propuesto, es basado en distancias. En la aplicación de la técnica de *behavior scoring*, el criterio de calidad de ajuste se realizó mediante la técnica de discriminación de K-S, considerando que la mayor distancia en separación entre las distribuciones de clientes buenos y malos indica el umbral, tal y como se referencia en la ecuación (19).

Como se observa en la tabla 6, el modelo propuesto muestra un índice de separación del 56.68% en el segmento o grado de riesgo C-1, mayor que el determinado con la regulación vigente de la CNBV en el grado de riesgo C-2 por 56.38%.

Con la maximización del coeficiente K-S se determinó un punto de corte de nivel de riesgo C-1 para el modelo propuesto, con una probabilidad de incumplimiento para clientes buenos por 12.92%. Dicha probabilidad es mayor que la obtenida con el modelo vigente de la CNBV por 12.05% en el mismo nivel de riesgo.

De acuerdo con la tabla 6, se muestra que es más costoso seguir administrando la cartera con nivel de riesgo C-2 con el modelo vigente, que la administrada con el modelo propuesto con un nivel de riesgo C-1. De esta manera, la PI_C calculada con el modelo propuesto implica mayor coherencia y en cierta manera mejor comportamiento, ya que en niveles de riesgo de A-1 a B-2 (niveles con menor riesgo) se refleja un menor PI_C en el modelo propuesto, mientras que en los niveles de riesgo de B-3 a D (niveles con mayor riesgo) el modelo propuesto tiende a considerar

Tabla 6
 Determinación del *cut-off* con K-S y su relación con la probabilidad de incumplimiento CNBV versus modelo propuesto

Clase	Metodología vigente			Modelo propuesto		
	Separación K-S	Probabilidad de incumplimiento clientes buenos	Probabilidad de incumplimiento clientes malos	Separación K-S	Probabilidad de incumplimiento clientes buenos	Probabilidad de incumplimiento clientes malos
A-1	51.03%	1.12%	100.00%	53.11%	0.03%	100.00%
A-2	52.25%	3.89%	100.00%	54.24%	4.02%	100.00%
B-1	53.23%	5.77%	100.00%	55.10%	5.72%	100.00%
B-2	54.11%	7.18%	100.00%	55.15%	7.14%	100.00%
B-3	55.01%	8.90%	100.00%	55.99%	8.94%	100.00%
C-1	55.91%	12.05%	100.00%	56.33%	12.92%	100.00%
C-2	56.01%	21.67%	100.00%	56.59%	21.99%	100.00%
D	35.12%	44.54%	100.00%	27.00%	60.05%	100.00%
E	0.00%	100.00%	100.00%	0.00%	-	100.00%

Fuente: elaboración propia con base en la metodología de Kolmogorov-Smirnov.

coherentemente una mayor PI_C. En el caso del nivel de riesgo E, en el modelo propuesto no fue posible considerar a ningún crédito, ya que es el peor nivel de riesgo.

Por lo tanto, el umbral óptimo en que la banca de consumo en México puede permitir un máximo apetito de riesgo con la administración de tarjetas de crédito es en el punto de corte C-1 con el modelo propuesto. Esto permite analizar cuánto estaría dispuesta la banca en México en aceptar ganar con un nivel de activos (ROA), bajo el costo de permitir contar con clientes que se encuentren en el grado de riesgo C-1 o inferior (peor) con el modelo propuesto.

Activos crediticios y rentabilidad de la banca al menudeo en México

En México, las reservas preventivas de crédito al consumo son determinadas considerando el saldo insoluto registrado en el último día del mes y la calificación obtenida. De acuerdo a la información al cierre de junio 2014 por parte de la Comisión Nacional Bancaria y de Valores (CNBV)²⁰, uno de los principales instrumentos de crédito en México son las tarjetas de crédito, puesto que existen aproximadamente 23 millones de plásticos autorizados por 24 entidades financieras reconocidas, las cuales han tenido un crecimiento mínimo anual de tan solo el 0.14% en promedio.

Es notable mencionar que aproximadamente el 54% de tarjetas emitidas están siendo administradas por las entidades financieras de BBVA Bancomer y Banamex, mientras que las entidades de BanCoppel, Santander Consumo, Banorte-IXE y HSBC abarcan el 35% aproximadamente. La cartera vencida en dichos créditos revolventes equivale, al cierre de junio 2014, por un monto de MXN \$90,593 millones de pesos mexicanos (MM), es decir, aproximadamente un 34% del total del monto autorizado por las entidades financieras autorizadas por la CNBV, por lo que, en este sentido, es un claro ejemplo del problema que es para las instituciones financieras crediticias en México aprobar un préstamo a personas que no garantizan un pago y con ello no reflejen incumplimiento en los pagos periódicos.

Dando seguimiento a la regulación mexicana vigente para el cálculo de reservas de tarjetas de crédito (para hacer frente a los incumplimientos posibles), las mismas estadísticas de la CNBV reflejan que al cierre de junio 2014 las reservas sumaron a nivel nacional MXN \$33,422 MM, que de igual manera, en la concentración de tarjetas de crédito por Bancomer y Banamex, suman aproximadamente el 51% de reservas. Santander Consumo, Tarjetas Banorte-IXE, HSBC y Scotiabank suman en conjunto el 42% aproximadamente.

Dichas provisiones han estado en crecimiento anual durante los últimos 12 meses por 12% aproximadamente, ya que el consumo en aumento se ha reflejado por mayor probabilidad de incumplimiento, un aumento en la severidad de la pérdida y un crecimiento en la exposición monetaria al incumplimiento, situación que afecta a los rendimientos netos de flujo bancario y a los activos financieros de todas las entidades, reflejando un rendimiento sobre activos (ROA)²¹ por 2.18% a nivel nacional, siendo American Express el que registró mayor rendimiento por 6.6% y Banco Bicentenario con el menor rendimiento por 82.6% negativo.

²⁰ Sección Información Estadística del Sector Supervisado Banca Múltiple (22 de agosto 2014), con datos al cierre del mes de junio 2014 para Créditos al Consumo para Tarjeta de Crédito (<http://www.cnbv.gob.mx/SECTORES-SUPERVISADOS/BANCA-MULTIPLE/Paginas/Información-Estadística.aspx>).

²¹ *Return on Assets*, que de acuerdo a la CNBV se conforma por la razón del Resultado Neto Flujo entre el Activo Total.

Tabla 7

Comparativo de Reservas CNBV y Modelo Propuesto, con Nivel de Ahorro en la Banca de Consumo en México al cierre de Junio 2014

	Reservas (MXN)		Ahorro con reservas propuestas	
	CNBV Metodología vigente	Modelo propuesto	Diferencia (MXN)	% Ahorro
Muestra 13,000 clientes	\$79,849,731.89	\$76,515,618.05	\$3,334,113.84	4.18%
Información de la banca de consumo en México	\$33,422,191,923.94	\$32,026,653,203.41	\$1,395,538,720.53	4.18%

Fuente: elaboración propia de acuerdo con el cálculo del Porcentaje de Reservas e Información Nacional con datos al cierre de junio 2014 (CNBV, 2014).

Resultado en provisiones crediticias y rendimiento

Para el tratamiento de las estimaciones para riesgos crediticios revolventes, la CNBV clasifica contablemente a las reservas en el Balance General bajo el concepto Activo y en subcuentas de Estimaciones Preventivas para Riesgo Crediticio, con clasificación de Créditos al Consumo (tarjeta de crédito).

Si se obtienen las reservas mediante el producto de $PI \cdot SP \cdot EI$ y con base en la regulación vigente, además de los resultados obtenidos con la información de los 43,323 clientes de la muestra, se muestra un comparativo en la [tabla 7](#).

Como se muestra en el [tabla 7](#), bajo las propuestas realizadas y con el cálculo de reservas de la muestra de 43,323 clientes por MXN \$255 MM, versus la metodología vigente que sumó MXN \$266 MM, se podría obtener un ahorro al cierre de junio 2014 por MXN \$11 MM (4.18%). Si esta tendencia se siguiera a nivel banca de consumo en México, el ahorro podría ser de \$1,395 MM. Lo anterior implica analizar el efecto Rendimiento sobre Activos que implicó con la metodología vigente, por lo que el ROA registró un 2.18% ([tabla 8](#)).

Al considerar los ahorros de la banca de consumo que implica el utilizar el modelo propuesto (PI_C), el Resultado Neto reportado por la CNBV al cierre de junio 2014 por MXN \$151,291 MM aumentaría MXN \$1,395 MM, es decir, MXN \$152,687 MM.

De igual manera, los Activos del Balance a nivel nacional aumentarían, por lo tanto el ROA pasaría del 2.18% al 2.20%, lo cual representa mayor rentabilidad a la banca de consumo en México al considerar un modelo actualizado PI_C ([tabla 9](#)).

Tabla 8

Indicadores de Activos y Rentabilidad en la banca de consumo en México al cierre de junio 2014

	Total de Activos Balance CNBV	Provisiones metodología vigente CNBV	Provisiones modelo propuesto	ROA % Indicadores CNBV
Muestra 13,000 Clientes	\$16,577,319,933	\$7,984,9732	\$7,651,5618	2.18%
Información de la banca de consumo en México	\$69,386,628,51,615	\$33,42,21,91,924	\$32,02,66,53,203	2.18%

Fuente: elaboración propia con base en el cálculo del Porcentaje de Reservas e Información Nacional con datos al cierre de junio 2014 (CNBV, 2014).

Tabla 9

Indicadores de Activos y Rentabilidad en la banca de consumo en México con el modelo propuesto (PI_C) al cierre de junio 2014

	Resultado Neto CNBV	Resultado Neto con modelo propuesto	Total de Activos con modelo propuesto	ROA % con modelo propuesto
Muestra 43,323 Clientes	\$36,14,55,034	\$36,47,89,147	\$16,58,06,54,047	2.20%
Información de la banca de consumo en México	\$1,51,29,19,23,163	\$1,52,68,74,61,883	\$69,40,05,83,90,336	2.20%

Fuente: elaboración propia de acuerdo con el cálculo del Porcentaje de Reservas e Información Nacional, con datos al cierre de junio 2014 (CNBV, 2014).

Conclusiones

Aunque el riesgo mercado es la pérdida potencial más significativo en las instituciones financieras, el riesgo crediticio no deja de ser importante para su administración en cualquier entidad financiera, sobre todo cuando el negocio de créditos al menudeo es el que aporta más de la mitad de los ingresos a la banca.

De acuerdo con lo anterior, el enfoque de análisis en esta investigación fue el riesgo crediticio al menudeo, especialmente el asociado con los créditos revolventes y comúnmente llamados tarjetas de crédito. Se revisaron las condiciones legales con las disposiciones de la CNBV y Basilea II con la finalidad de proponer un nuevo modelo de clasificación *behavioral scoring* empleando el análisis discriminante lineal con una base de datos compuesta por variables mixtas continuas y una dicótoma (clientes buenos o malos).

Para determinar qué variables tienen que ser excluidas se realizaron diversas pruebas de hipótesis de selección de variables; tal es el caso de la validación de no multicolinealidad. Considerando la metodología de la CNBV en materia de segmentación de grados de riesgo y los resultados obtenidos, se demostró que con el modelo propuesto el «costo» de administrar créditos en niveles de riesgo de B-3 a D es mayor que el costo de mantener clientes en niveles de riesgo de A-1 a B-1.

Aunado a lo anterior, el punto de corte óptimo para la administración de la cartera en el modelo propuesto es el nivel de riesgo C-1, en el cual se identifica antes de la máxima separación de clientes buenos y malos (K-S = 56.68%) con un aumento drástico en la PI_C del 12.92%. Ambos resultados son mayores y mejores indicadores que el modelo vigente.

De igual manera se puede confirmar con esta investigación que el modelo propuesto satisface dos condiciones: la primera es que tiene una fuerte capacidad predictora (99.38%) y la segunda es que la estimación de los parámetros cuenta con alta relación de variabilidad de variables dependientes respecto a la variable independiente en un 97%. Sin embargo, esto abre las puertas a futuras investigaciones para la mejora continua de la predicción y con ello la mejora en rentabilidad para la banca de menudeo en México.

La validación del grado de rentabilidad para la banca de consumo en México mediante el indicador conocido por rentabilidad sobre activos (ROA) mostró resultados sumamente importantes. En primera instancia, el ahorro obtenido con el modelo propuesto de probabilidad de incumplimiento (PI_C) para la muestra analizada fue aproximadamente por MXN \$11 MM (4.8% respecto a la metodología vigente de la CNBV). Si esta tendencia es constante, a nivel nacional la banca hubiese tenido un ahorro aproximado de MXN \$1,395 MM. En segunda instancia, se pudo validar

también que el ahorro aumentaría el registro contable de Activos a MXN \$55,268 MM (versus MXN \$55,257 MM con modelo CNBV) y el aumento del Resultado Neto a MXN \$1,215 MM (versus MXN \$1,204 MM con modelo CNBV) al cierre de junio 2014; ayudados ambos casos por la disminución de pérdidas esperadas. Con esto, el ROA pasaría del 2.18% obtenido con el modelo PI de la CNBV, a un posible ROA por 2.20% con el modelo propuesto PIC.

La creación de reservas y el castigo de créditos son operaciones significativas que impactan sensiblemente en el nivel de capitalización; incluso las provisiones crediticias son el segundo gasto más importante después de sueldos y salarios, por lo que deben reportarse de inmediato desde el área de Finanzas o Riesgos en el momento que se originen. De esta manera, el Director (CEO) y los órganos facultados (Comité de Riesgos y Consejo de Administración) de la institución financiera tienen conocimiento del efecto financiero y con ello toman decisiones importantes respecto a los límites de crédito y al apetito de riesgo que la institución crediticia se permitirá tolerar con respecto de sus niveles de rentabilidad.

Ante la escasa investigación relacionada a modelos de *behavioral scoring* para medir el comportamiento del riesgo crediticio de la banca de consumo en México, se tiene que la propia banca y la CNBV tienen todavía un amplio margen de discusión en esta línea, sobre todo para actualizar la metodología vigente en la CUB en materia de provisiones al menudeo.

Referencias

- 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. y 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. Disponible en: <http://www.cnbv.gob.mx/Paginas/NORMATIVIDAD.aspx>
- Costa, T., Boj, E. y Fortiana, J. (2012). Bondad de ajuste y elección del punto de corte en regresión logística basada en distancias. Aplicación al problema de Credit Scoring. *Anales del Instituto de Actuarios Españoles*, 18, 19–40.
- Durand, D. (1941). *Risk Elements in Consumer Instalment Financing, Studies in Consumer Instalment Financing*. New York, 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. y Jacka, S. D. (1998). Consumer credit and statistics. En D. J. Hand y S. D. Jacka (Eds.), *Statistics in Finance* (pp. 69–81). London: Arnold.
- Hand, D. J. y 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. y Wang, C.-J. (2007). Credit scoring with a data mining approach based on support vector machines. *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. y 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*. Ohio: Thomson South Western, Mason.
- 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. y 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. y 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. y 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. y Edelman, D. B. (1992). *Credit Scoring and Credit Control*. Oxford: Oxford University Press.
- Thomas, L. C., Edelman, D. B. y 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. y 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 Creditworthiness Evaluation. En R. R. Ribeiro, H.-J. Zimmermann, R. R. Yager, y J. Kacprzyk (Eds.), *Soft Computing in Financial Engineering* (pp. 388–401). Heidelberg: Physica-Verlag.