



**CREACIÓN DE UN MODELO DE RIESGO CREDITICIO A TRAVÉS DE
UN MODELO LOGIT PARA ESTIMAR LA PROBABILIDAD DE
DEFAULT DE LOS CLIENTES**

ALEXANDER OVALLE PAÍS

25 de diciembre de 2021

Profesor Guía

Pablo Méndez Montenegro. Ph.D.

Instituto de Estadística, Universidad de Valparaíso

Profesor Co-Guía

Rolando Rubilar. Ph.D.

Departamento de Estadística, Universidad de Valparaíso

Proyecto de titulación para optar al:

grado académico de: *Licenciado en Estadística*

título profesional de: *Ingeniero en Estadística*

minor en: *Finanzas*

Resumen

El presente trabajo busca construir un modelo que permita estimar la probabilidad de default de la empresa Tec Tec S.A. así de esta manera poder calcular las provisiones necesarias que deberá tener de respaldo la compañía, las cuales serán de necesidad si el mercado entra en crisis, se elaboró un modelo para cada segmento del negocio y cada uno de éstos se determinó según la categorización de sus clientes.

El cálculo de los modelos utiliza una base de datos de la institución financiera que lleva como nombre Tec Tec S.A., donde la cartera corresponde a los créditos comerciales de la totalidad de clientes lo cuales son personas naturales. La creación de este modelo permite tener un parámetro objetivo y cuantitativo para cada segmento, ayudando a la clasificación de riesgo.

Los resultados obtenidos muestran diferencias evidentes con un segmento en particular, este es el de clientes que repactan su deudas, por lo tanto en este segmento es donde la empresa debe arriesgar una mayor cantidad de recursos financieros.

Lo conseguido en esta tesis fue desarrollar una metodología cuantitativa que complementa la labor de la evaluación por segmento y determinar cuantos son los fondos que la compañía está arriesgando por segmento. Dado el desarrollo de esta herramienta, es posible realizar un análisis cuantitativo, con el fin de tener una mejor gestión de riesgo, como así también aportar en el cálculo de provisiones por riesgo de crédito.

El aporte del modelo aplicado es que disminuye la cantidad de variables por lo tanto ya no se tiene colinealidad entre las variables que componen el nuevo modelo a aplicar, por otro lado la correlación entre variables en la ecuación resulta ser muy baja.

Por último al aplicar el modelo de provisiones se muestra que la empresa debe reservar para provisionar un 16,63% de su capital.

Algunas palabras

Agradecimientos a mi familia por el apoyo cuando se necesitó.

A mi hija April la que es mi motivación para lograr siempre otro objetivo.

A mi profesor guía Pablo Méndez por su apoyo cuando se necesitó.

A mi profesor co-guía Rolando Rubilar por su excelente e inmediata disposición a resolver todas las dudas sobre mi tesis.

Al Profesor Carlos Felipe por su paciencia y disponibilidad para resolver cualquier duda durante la carrera.

A todos los profesores que alguna vez aportaron con conocimientos para mi formación.

A mi amigo Pancho por aportar con sus conocimientos en mis primeras asignaturas.

Índice General

Resumen	2
Algunas palabras	3
Índice General	4
Introducción	8
Capítulo 1	11
1. Preliminares	11
Capítulo 2	13
2. Estado del Arte	13
Capítulo 3	21
3. Objetivos	21
Capítulo 4	22
4. Marco Teórico	22
4.1. Riesgo Crediticio	22
4.2. Criterios y herramientas de evaluación del riesgo crediticio	26
4.3. Concepto de Quiebra	31
4.4. Modelos de Riesgo Crediticio	33
4.5. Modelo Logit	37
4.6. Modelo de regresión logística	40
4.7. Modelo Forward Looking	46
4.8. Variables WOE	47

4.9.	Comprensión intuitiva de IV	48
4.10.	WOE	49
4.11.	Modelo probit. McFadden (1973)	51
Capítulo 5		53
5.	Metodología	53
5.1.	Estructura metodológica	53
5.2.	Levantamiento de información	54
5.3.	Determinar información estadística a utilizar	54
5.4.	Determinar las variables que influyen en el modelo a proponer	55
5.5.	Determinar ecuación lineal	57
5.6.	Comprobación del modelo propuesto.	57
Capítulo 6		58
6.	Desarrollo del Modelo	58
6.1.	Supuestos Generales	58
6.2.	Definición de Categorías	59
6.3.	Antecedentes Generales	60
Capítulo 7		66
7.	Análisis de datos y resultado de modelos	66
7.1.	Segmentación de Cartera	67
7.2.	Desarrollo de los modelos estadísticos	68
7.3.	Definición de desempeño y horizonte de observación para los Modelos de PD	69
7.4.	Construcción de bases de muestreo.	69

7.5.	Construcción y análisis de variables	78
7.6.	Matrices de correlación	80
7.7.	Regresiones Logísticas	82
7.8.	Capacidad de discriminación de los modelos de PD	84
7.9.	Relación PD vs Logit	85
7.10.	Modelamiento macroeconómico	86
7.11.	PD Lifetime	90
7.12.	PD Point in time	92
7.13.	Cálculo de LGD:	93
7.14.	Definiciones conceptuales	96
7.15.	Aplicación del resultado de modelo de provisiones	102
Anexo		103
Definición de las variables		103
Variables existentes		106
Variables de Fuentes Externas		124
Nuevas Variables		125
Detalle calculo K-S		129
Categorización en WOE de variables no seleccionadas.		133
Matrices de correlaciones		136
Tabla de correlaciones modelo Consumo		136
Tabla de correlaciones modelo Renegociado		136
Tabla de correlaciones modelo Transaccional		137

Tabla de correlaciones modelo Normal	137
Gráficos de las probabilidades de default por segmento	138
Referencias	139

Introducción

El riesgo de crédito tiene relación con el incumplimiento de un convenio contraído. Al respecto, Garcia (2015), señala que el mismo se considera como “la posibilidad de sufrir pérdidas por el incumplimiento que un deudor o contraparte hace de sus obligaciones contractuales” (p.22). Los mercados le ponen un precio a este riesgo, en lo cual, Chatterjee (2016), refiere que el mismo se incluye en el precio de compra del mercado para un pago contratado. Este incumplimiento de los clientes efectivamente generan la aparición del riesgo crediticio, lo cual es una realidad a la que se enfrentan las entidades financieras, y cuyo costo es incluido en los presupuestos y proyecciones que las mismas realizan. Sin embargo, cuando este incumplimiento es mayor a los valores estimados genera dificultades en la liquidez de la entidad financiera, así como posibles pérdidas con las que no se contaba.

El motivo para crear modelos que permitan predecir el riesgo crediticio radica en la necesidad de calcular el capital con que debe contar la entidad financiera para sustentar los riesgos asociados al otorgamiento de créditos. Por esta razón, las entidades financieras requieren evaluar el riesgo de los clientes que soliciten créditos, para que su selección dependa de la atención que los mismos presten a sus pagos.

Para que los créditos sean suministrados, la evaluación debe cumplir con los términos convenidos y aprobados por el ente crediticio para ello. Esto permitiría que la entidad tengan el respaldo financiero para atender las obligaciones adquiridas con los que suministran los fondos, que permiten realizar las negociaciones respectivas y la entrega de créditos. Dentro de ellos se pueden mencionar los titulares de cuentas de ahorro o corriente, depositantes de certificados de depósito a término, otros establecimientos de crédito, y el propio accionista de la compañía.

En este orden de ideas, es de gran relevancia que las entidades financieras cuenten con metodologías que permitan evaluar oportuna, concreta y eficiente, las solicitudes para el

otorgamiento de créditos de diferentes tipos. En este sentido, se debe procurar prestar la mejor atención posible a la mayor cantidad de clientes, teniendo la seguridad de que todos los clientes atendidos con créditos generen al menos una rentabilidad mínima, así tengan elevados riesgos crediticios.

Existen diferentes modelos que permiten valorar la probabilidad de incumplimiento a partir de variables presentadas por las entidades financieras. En los mismos se pueden incluir las particularidades de los clientes solicitantes de créditos, para valorar el riesgo crediticio de cada uno de ellos. En este sentido, se pueden considerar para la evaluación del riesgo crediticio algunas variables sociodemográficas y de comportamiento disponibles en la entidad crediticia o en alguna central de riesgo. Esta última, es una entidad que se encarga de recolectar y consolidar información de clientes con el fin de clasificarlos según la probabilidad de morosidad de sus créditos. Esto permite que las posibles pérdidas que se puedan generar al hacer el otorgamiento de un crédito, se encuentre dentro de valores tolerables y previamente estimados por la predicción del riesgo crediticio.

La utilización de modelos estadísticos para identificar perfiles según probabilidad de incumplimiento, es una práctica que satisface plenamente la necesidad planteada. La evaluación de la probabilidad de ocurrencia es lo que puede colaborar a establecer los valores permitidos o tolerables para asumir el riesgo crediticio, permitiendo desarrollar estrategias para mitigar las consecuencias del riesgo asumido.

La orientación hacia la gestión interna de los riesgos crediticios, considera la constante evaluación de diferentes modelos, para ubicar los más acertivos y representativos y cuya información de la cual se alimentan, sean de fácil acceso para las entidades financieras. La regresión logística es la herramienta más utilizada en las actividades de modelado para predecir eventos de incumplimiento en riesgo de crédito, dado que es el mejor ajuste que tiene la distribución frente a la de los eventos de incumplimiento observados

En función de lo planteado anteriormente, se formula como problemática a resolver, la creación de un modelo de riesgo crediticio a través de un modelo *Logit*, para estimar la probabilidad de default de los clientes, es decir, calcular el puntaje que refleja el perfil de riesgo crediticio, donde generalmente, se asignan valores bajos a los clientes con mayor nivel de morosidad y/o valores de endeudamiento por encima de lo criterios asumidos como normales (González, 2015).

La presente investigación esta estructurada del siguiente modo: en el capítulo 1 se presentan los preliminares, en el cual se muestra el planteamiento de la situación a estudiar con el presente trabajo.

En el capítulo 2 se presenta el estado del arte, en el cual se contextualiza y se proporciona una revisión de investigaciones anteriores y bases para continuar líneas de otros autores, permitiendo optimizar el trabajo en la creación de un modelo nuevo de riesgo crediticio.

En el capítulo 3 se presentan los objetivos generales y los objetivos específicos que sustentan la presente investigación.

El capítulo 4 constituye el marco teórico, en el cual se explican las definiciones básicas para comprender el desarrollo del modelo, donde se desarrolla la concepción y definición de riesgo crediticio en conjunto con las diferentes definiciones de quiebra y la diferencia entre insolvencia e incumplimiento.

En el capítulo 5 se presenta la metodología necesaria para la identificación de las variables que se requieren para evaluar el modelo, lo cual debe demostrar el cumplimiento o no cumplimiento del pago del crédito que el cliente puede asumir.

En el capítulo 6, referido al desarrollo del modelo, que es el objetivo final de la presente investigación, se muestra el proceso de desarrollo del mismo, precisando los supuestos del caso acorde a la entidad financiera cuyo nombre ficticio es TEC TEC S.A.

Capítulo 1

1. Preliminares

Al considerar las finanzas como una disciplina científica que se dedica al estudio de la forma en que se asignan los recursos escasos a lo largo del tiempo bajo condiciones de incertidumbre (Paula Nicole Roldán, 2017), es que, se han suscitado elementos de control por parte de agencias internacionales hacia países en desarrollo, intentificando su estudio e investigación. En este sentido, el riesgo es una de las variables de mayor interés para el mercado financiero, ya que determina una amenaza a la generación de valor por parte de las organizaciones, especialmente por sus implicancias en aspectos como el mercado, la liquidez, las operaciones y otros (Saldaña, Palomo y Blanco, 2017).

Uno de los riesgos más típicos que enfrentan las empresas financieras, es el riesgo de crédito. Si bien, este es propio de la actividad económica, puede provocar que se incurra en pérdidas potenciales que pueden desencadenar en crisis financieras de diversos tipos. Como ejemplo de ello, se puede mencionar el estallido de la crisis subprime o burbuja inmobiliaria en el año 2007, cuyo primer síntoma real fue el aumento en las tasas de interés en el año 2005 y el aumento del desempleo en el año 2007; estos elementos desataron la caída definitiva de la burbuja inmobiliaria (Quitral, 2012).

La gestión adecuada del riesgo crediticio permite obtener rentabilidad acorde al nivel de riesgo aceptable por cada organización, constituyendo, junto a la solvencia, los pilares fundamentales de la gestión en las entidades crediticias (Thomas, Crook y Edelman, 2017). Estudiar la probabilidad de ocurrencia de los riesgos y las consecuencias de los mismos, permite establecer niveles de referencia que se pueden clasificar como aceptables o no

aceptables, (Vargas y Mostajo, 2014). A fin de desarrollar estrategias y procurar tomar medidas para mitigar las repercusiones negativas es que se ha dado espacio a la creación de mecanismos de control que atenúen el riesgo existente de manera tanto sistemática, como procedimental, como por ejemplo la central de riesgos. Según Rossi (2013), refiere que el alza de los precios del petróleo afectó la recuperación económica entre finales del año 2011-2012, a nivel del mundo desarrollado. Esta falta de recuperación económica, llevó al mercado de acciones a un comportamiento errático aumentando la volatilidad e inestabilidad del mercado financiero a nivel generalizado junto con el comportamiento de las tasas de interés, potenciado fuertemente por el fenómeno de la globalización.

La orientación hacia la gestión interna de los riesgos crediticios y de mercado implica constantes iteraciones para conseguir un modelo con resultados representativos con información disponible para las entidades financieras, como también, debe establecer claramente los principios y criterios bajo los cuales opera, en conjunto con las políticas que rigen las evaluaciones. Es necesario potenciar la migración hacia el uso de sistemas internos de riesgo crediticio considerando los parámetros a los que se encuentra sujeta cada entidad, determinando sistemas de rating internos (Fernández y Pérez, 2005).

La propuesta a presentar responde a la necesidad de toma de decisiones que enfrentan las entidades financieras en torno a la selección de sus clientes, estableciendo el riesgo asociado a su captación con productos de crédito. Dado que es frecuente que estas elecciones se realicen de manera automatizada, el modelo requiere facilidad de uso y entendimiento para el usuario, aunque con un profundo desarrollo lógico detrás que permite determinar la decisión de otorgar o no un crédito evaluando debidamente la probabilidad de que el cliente devuelva íntegramente el monto pactado y sus intereses, ya que de otra manera, se estaría realizando una asignación ineficiente de recursos al dar acceso a clientes sin capacidad de pago.

Capítulo 2

2. Estado del Arte

A fin de contextualizar, el estado del arte proporciona una revisión de investigaciones anteriores y bases para continuar líneas de otros autores, permitiendo optimizar el trabajo en la creación de un modelo nuevo.

Los modelos utilizados generalmente para la gestión del riesgo crediticio son bajo el amparo de modelos de puntuación de crédito, más reconocidos en inglés como *credit scoring*. Estos, tratan de modelos multivariantes que utilizan como entrada los principales indicadores financieros de una empresa o potencial cliente, atribuyendo a cada uno de ellos un peso que refleja su importancia relativa en la previsión de la morosidad (Gurný y Gurný, 2013).

Entre estas nuevas soluciones al problema de las entidades financieras, se han generado propuestas cuantitativas, en las que se da énfasis a los modelos econométricos de regresión logística o *Logit* (Hachuel, Boggio y Harvey, 2012). El modelo econométrico de regresión logística se enfoca en la predicción de un resultado binario, como es si un cliente es sujeto, o no, de crédito. Este modelo genera incidencia en una variable dependiente del tipo binaria, considerando un código 0 para un sujeto que no entrará en mora y un código 1 para un sujeto que entrará en mora (Fernández y Pérez, 2005).

Los avances en la problemática se han extendido a múltiples variantes de los modelos tradicionales. Majnoni *et. al.* (2004) utilizan información de Argentina, Brasil y México para demostrar una capacidad predictiva superior al sustituir información negativa como moras y no pagos por unas de connotación positiva y descriptiva de las deudas preexistentes. A su vez, Miller y Rojas (2005), realizan un *credit scoring* para pequeñas y medianas empresas

(PYMES) de México y Colombia, que posteriormente Pailhé (2006) ampliaría con el establecimiento de puntos de cortes para las solicitudes aceptadas por las entidades en Argentina, determinando que los solicitantes de créditos con un puntaje menor al punto de corte suponen un rechazo automático, mientras que quienes se encuentren sobre el mismo, precisan complementar sus antecedentes.

Aunque las técnicas en las que se basan estos modelos fueron ideadas desde épocas cercanas al 1940, por autores como Fischer (1936) y Durand (1941), el auge de la promoción de estos modelos no sucede hasta pasado 1965, de la mano de los estudios de Beaver (1967) y Altman (1968) quienes se dedicaron a comparar muestras emparejadas de empresas en quiebra (*default*) y no quiebra, demostrando que las quiebras pueden predecirse a partir de la información contenida en los indicadores (*ratios*) financieros.

La gran mayoría de los modelos de calificación crediticia ya propuestos se derivaron de una muestra de instituciones no financieras, debido al hecho de que los impagos de las entidades financieras son relativamente escasos y no todos los datos están disponibles públicamente. Sin embargo, también era necesario identificar los factores clave de las empresas del rubro, para lo cual se utilizaron los estados financieros (Peresetsky y Karminsky, 2008; Gurný y Gurný, 2009).

Según Gurný y Gurný, (2013), los análisis de regresión son técnicas multivariantes, las cuales permiten estimar la probabilidad de que se produzca un evento, o bien, de que no se produzca dicho evento. Según el resultado lo predicen a partir de un conjunto de variables independientes en base al siguiente modelo:

$$Y_i = x_i\beta + \varepsilon_i, \quad (2.1)$$

donde Y_i , es la probabilidad de que " Y_i " sea igual a 0 si es un sujeto no apto a un crédito, mientras que si el default no ocurre se obtiene el valor de 1, con probabilidad $1-Y_i$, por lo que el modelo estima la probabilidad de *default*, quiebra o impago. Cabe destacar que x_i corresponde a un vector de indicadores financieros individuales y que tanto ε_i como β' son un grupo de parámetros a ser estimados.

Entre las formas de estimar la probabilidad P_i , se utilizan los modelos logísticos o *logit* (Peresetsky y Karminsky, 2008). Este modelo utiliza la denominada transformación logística, bajo la formula (2.2).

$$P_i = \frac{\exp(\alpha + \beta' x_i)}{1 + \exp(\alpha + \beta' x_i)} = \frac{1}{1 + \exp(-\alpha - \beta' x_i)}, \quad (2.2)$$

Si bien en esta estimación se consideran los mismos parámetros que el modelo lineal, hacen unos cambios interesantes al aplicar una distribución exponencial a la función.

Gurný y Gurný (2013), trabajan con el modelo *Logit* y una muestra de 298 bancos comerciales estadounidenses, dividiéndose en dos: el primer grupo de organizaciones, se encuentra en orden con sus compromisos financieros, mientras que el segundo grupo se encuentra en default, considerando este estado como instituciones financieras que han entrado en proceso de liquidación o bien, que se han sometido a procesos de reestructuración financiera profundas, como su adquisición por el gobierno u otros accionistas mayoritarios. Ambas muestras se eligen de manera aleatoria y conforme a la información pública disponible, para posteriormente, basados en el trabajo de Peresetsky y Karminsky (2008), identificar los *ratios* financieros a partir de los estados financieros de cada organización.

Usando como base la investigación empírica con el modelo de Z-metrics de un año realizado por Altman (2010), los autores determinan que es pertinente la revisión de los estados financieros de dos años previo a la situación de mora o default. Por este motivo se estudian los *ratios* financieros de cada banco concreto para el año $t-2$, considerando como t al año de

default o bien, como el año de estimación para las instituciones no morosas, suponiendo la estimación de modelos predictivos para un periodo de 1 a 2 años. Adicionalmente, se toman variables de mercado, como los datos durante la crisis financiera 2007-2010, los que se vinculan con tamaño de la institución, rentabilidad, eficiencia, calidad de los activos y adecuación del capital (Gurný y Gurný, 2013).

Como resultado de un análisis comparativo, los autores (Peresetsky y Karminsky) determinan que el modelo logit es el de mejor resultado para la muestra de instituciones financieras, siendo el más adecuado en la predicción de morosidad al evaluarse junto al modelo probabilístico o PROBIT y LDA. Entre las razones, se tiene que el resultado del test de razón de verosimilitud implica el rechazo de la hipótesis nula, con valores menores a 0,05, rechazando la hipótesis nula, es decir, al menos uno de los parámetros del modelo es diferente de cero, por lo que el modelo es estadísticamente significativo al nivel del 5%, como también, un coeficiente de determinación (R^2) de 0,9630 con el uso de 3 variables: Logaritmo de los activos totales, Rendimiento de los fondos propios medios (%) y razón de Préstamos problemáticos / Préstamos brutos (%).

En el caso de Fernández y Pérez (2005), el modelo *Logit* es utilizado para el análisis de una cartera comercial, utilizando un procedimiento definido desde la determinación de las unidades objeto de análisis, es decir, las empresas, y sus principales variables, categorizándolas en cuantitativas y cualitativas. Las del primer tipo, corresponden a indicadores financieros como el año de información, crecimiento de las ventas, rentabilidad bruta, utilidad operativas, razón utilidad neta y ventas, crecimiento en utilidades, ciclo operativo, razón utilidad neta y patrimonio, crecimiento del patrimonio, endeudamiento, razón endeudamiento y ventas, apalancamiento financiero, razón pasivo corriente y pasivo total, crecimiento del activo, rentabilidad del activo, prueba ácida, razón corriente, rotación de cartera, rotación de proveedores, rotación de inventarios, razón corrección monetaria y utilidad neta, calificación interna de default, cupo en millones de pesos, año de calificación,

saldo capital promedio año y máximo días de vencimiento en base al balance general, entre otras, cuyo interés se encuentra en el nivel de correlación entre estas variables por posibles problemas de multicolinealidad, lo que es evitado mediante el análisis de componentes principales que permiten pasar de un espacio linealmente dependiente a uno independiente. En tanto, las variables cualitativas corresponden a la antigüedad de la empresa, antigüedad como cliente, socios, referencias de los socios, administradores, experiencia, estructura de administración, auditor de los estados financieros, disponibilidad de los estados financieros, competencia, productos, estabilidad de la cartera de clientes, proveedores, barreras de entrada y salida, instalaciones, activos, recursos humanos, nivel de regulación, nivel de riesgo ambiental, fuentes de financiamiento y alternativas de financiamiento.

En este caso, se utiliza como base la ecuación 2.3:

$$Y_i = \frac{1}{1 + \exp(-z)} + u_i. \quad (2.3)$$

Considerando a Y_i como variable dependiente con valores entre 0 y 1. Por su parte, z aplica como scoring logístico en forma matricial como $z = \beta x$, mientras que u_i responde a una variable aleatoria distribuida normalmente como $N = (0, \sigma^2)$. A través de una serie de transformaciones, se utiliza como modelo final el siguiente.

$$\text{logit}(P(y = 1|x)) = \beta_0 + \sum_{i=1}^k \beta_i x_i. \quad (2.4)$$

Utilizando dichos parámetros en el modelo de la ecuación 2.4 y a través de tablas de contingencia, se realizan pruebas de asociación para la estimación de probabilidad de *default*,

destacando que la regresión logística es la herramienta estadística más adecuada para estimar la probabilidad de *default* del cliente.

La propuesta Calixto y Casaverde (2011) es estimar la quiebra a través de una función, por lo que proponen un modelo Logit con las siguientes variables: Experiencia en la misma organización, experiencia en el sector financieros, sexo, edad, estado civil, número de hijos, nivel de estudios, tipo de residencia, actividad económica, endeudamiento en la misma institución y deuda, último saldo acreedor.

El modelo de Calixto y Casaverde (2011) concluye que variables como el crédito a la misma empresa, el plazo del préstamo, el saldo de la deuda del sistema financiero, el domicilio, la edad y el estado civil son determinantes para la evaluación del riesgo crediticio. Es por esto que el pago del préstamo en conjunto con las estimaciones de la función logística binaria y crediticia muestran que el quiebre aumenta cuando las solicitudes de préstamo tienen un vencimiento prolongado. Dichas fluctuaciones en el número de préstamos concedidos indican un aumento en la cantidad de préstamos, el riesgo crediticio y la antigüedad del negocio hasta el momento, por lo que es menos probable que sea predeterminado.

Por otro lado, Sotomayor (2012), propone un modelo para encontrar la pérdida esperada en los microcréditos a partir de la creación de nuevos criterios de elegibilidad utilizados por las empresas financieras. Dicho modelo utiliza el método de regresión logística ya que es más flexible que el modelo lineal. De esta manera, encuentra la probabilidad de que un cliente cancele un préstamo en función de una variable en particular y utiliza un modelo logístico con una variable dependiente para determinar si el cliente es un buen pagador o un mal pagador.

Las variables a considerar son: Estado civil, edad, índice de refinanciamiento, número medio de días diferidos, número máximo de días diferidos, flujo de caja bancario, gastos, dependencia de otros ingresos, asignación total propuesta, dependencia comercial, número

de microcréditos, deuda líquida y estado, monto máximo de la cuota, número de préstamos cancelados. Con esto se busca obtener la eficiencia de las variables utilizadas para determinar la posibilidad de que no pague.

Otra propuesta es la de Moreno (2013), en donde busca comparar el modelo Logit mixto con los modelos Logit y Probit para mostrar que este modelo es más adecuado para estimar la probabilidad de incumplimiento de pagos. Para esto se levanta un modelo Logit mixto que incluye variables de género, personas activas, dependientes, tipo de vivienda, antecedentes matrimoniales, antecedentes educativos, situación laboral, clase, garantía, mes de préstamo más cercano, sector y región.

Dicho modelo se basa en la recopilación de datos aleatorios y no correlacionados. En el modelo, la tasa promedio de incumplimiento, la duración, el saldo del préstamo, la evaluación de riesgos, el nivel de educación, la situación laboral, las personas activas, las garantías, los meses desde el último préstamo, la antigüedad y el sector son todos importantes. Luego concluyó que todos los modelos eran muy discriminatorios y que el Logit mixto era superior a otros modelos solo en términos de sensibilidad, pero predijo el número máximo de falsos positivos, no obstante, el modelo Logit más restringido en términos de las variables, pero sus resultados son más certeros.

Otra propuesta es la de Correa (2016), el cual propone un modelo logístico utilizando datos agregados enfocados en organizaciones con baja probabilidad de quiebra, en donde este utiliza una variable dicotómica como variable dependiente, por lo cual el período de cobro, la vivienda, la deuda, el margen de ganancia, las acciones, los contratos y el estado como variables independientes de tipo, dicho lo anterior la cartera tiene 90 días de antigüedad y más de 180 días de vencimiento, Por lo tanto, concluye que el uso de un modelo logístico para datos agrupados es efectivo en este caso debido a su alta significancia estadística general.

Fernandez, & Pérez (2005) a diferencia del resto de autores mencionados, presenta un fuerte empeño por añadir los progresos más actuales sobre el cálculo del riesgo para determinar el mínimo capital requerido por las empresas del sector financiero. Esta propuesta se basa en un modelo logit con las variables cuantitativas: año, información, monto de la deuda, calificación interna, tiempo de vencimiento de deuda, rentabilidad bruta, incremento de ventas, ganancia neta entre ventas y ganancia operativa, aumento de utilidades, y crecimiento del patrimonio.

A raíz de este análisis, se concluye que el riesgo y la posibilidad para obtener resultados no esperados, se convierte en indispensable considerar el uso de matemáticas en este tipo de modelos.

Karacula (2009), propone un ejemplo de utilización del modelo LOGIC entorno a un hipotecario, cuyo propósito es encontrar el perfil predeterminado esperado para hipotecas utilizando métodos de estimación *logit* y *probit*. El modelo *Logit* utilizó en las siguientes variables: Quiebra, Relación pago / ingresos, fecha de vencimiento, monto de la hipoteca, fecha de inicio del crédito, valor del préstamo, estado civil, sexo, años, educación, límite de crédito, monto de pago, tasa de interés, nivel de ingresos y el monto acumulado.

A través de estas variables trata de establecer parámetros que inciden en las tasas de morosidad hipotecaria, como la información personal sobre los clientes del banco y las características del préstamo o hipoteca.

Consistente con lo anterior, se procede al desarrollo del marco teórico para sustentar el modelo a desarrollar.

Capítulo 3

3. Objetivos

El objetivo general corresponde a:

Crear un modelo de riesgo crediticio a través de un modelo *Logit*, para estimar la probabilidad de *default* de los clientes en una empresa de servicios financieros, encargada de entregar créditos y tarjetas de créditos a sus clientes y cuyo nombre ficticio es TEC TEC S.A.

Este objetivo, es desarrollado mediante los siguientes objetivos específicos.

1. Identificar aplicaciones anteriores de modelos *Logit* en ambientes similares, a fin de otorgar una base teórica y metodológica.
2. Definir las variables independientes que permitan la estimación de la probabilidad de *default* en contexto de riesgo crediticio.
3. Crear un modelo de regresión logística que permita explicar la probabilidad de *default* de los clientes.

Capítulo 4

4. Marco Teórico

En el presente capítulo, se explican las definiciones básicas para entender el desarrollo del modelo. Se desarrolla la concepción y definición de riesgo crediticio, en conjunto con las diferentes definiciones de quiebra y la diferencia entre insolvencia e incumplimiento, a fin de elaborar un modelo riesgo crediticio en base a una metodología propuesta.

4.1. Riesgo Crediticio

El riesgo se define como la oportunidad o probabilidad del surgimiento de algún evento desfavorable, generalmente ligado a cualquier hecho económico en el que las contingencias puedan ocasionar pérdidas (Sánchez, 2007).

Se puede representar la definición de pérdida dentro del concepto de riesgo a cualquier situación en que la contraparte, entendiendo esta como a la que se le otorga un crédito, no cumpla con algunas de las condiciones previamente establecidas a través de un contrato. Vela & Caro (2015) definen el término de crédito como un proceso en el cual una persona o empresa, recibe un financiamiento de otra persona natural o jurídica, con el propósito de recibir dividendos dados por el capital prestados, es decir, en un periodo posterior se devuelva el financiamiento recibido conjuntamente con la tasa de interés correspondiente, acordada previamente. Es por ello, que los créditos, representan riesgos que están asociados al impago de los mismos.

A modo de ejemplo, esto ocurre cuando una entidad financiera otorga un crédito a alguna persona natural o jurídica, siendo la entidad la que asume el riesgo de que, ante la eventual

posibilidad de que el cliente no cumpla con las condiciones de retorno del valor pactadas al momento de suscribir el acuerdo, pueda incurrir en morosidad y finalmente, generar pérdidas financieras para el banco (Salas, 1998).

Otro ejemplo para respaldar el concepto, ocurre en el caso que un inversionista decida adquirir un bono corporativo: Existe un riesgo en el que el emisor de dicho bono no pueda pagar lo comprometido, por lo cual, el riesgo crediticio es asumido por los particulares, empresas e instituciones financieras, entrando la posibilidad de que el deudor no se haga responsable de los pagos previamente acordados (Gárate, 2016).

En este mismo orden de ideas, Vela & Caro (2015) mencionan los diferentes tipos de carteras crediticias, dentro de las cuales se tienen los créditos directos e indirectos; los primeros consideran la emisión de sumas de dinero en cualquier moneda, entre personas naturales o jurídicas, los cuales tienen la obligación de devolver el dinero más la tasa de interés, a la persona o ente prestamista; los segundos se refieren a préstamos otorgados por personas jurídicas mediante avales, cartas de crédito, fianzas, etc., estos préstamos pueden usarse en negocios y contratos, cuyo incumplimiento de la parte deudora, genera que el emisor se debe hacer responsable del compromiso adquirido por el deudor.

De igual forma, cuando los créditos se clasifican de acuerdo a su tipo, se presenta la siguiente variedad: los créditos corporativos, los cuales se dan a las empresas que registran un cierto nivel de ventas en el año. Asimismo, se tienen los créditos a pequeñas y medianas empresas, las cuales son clasificadas de acuerdo a la capacidad de endeudamiento de las mismas. Igualmente, se tienen los créditos a grandes empresas, clasificadas de esta manera, por el capital de ingresos en un año, así como, consideran ciertas características en el endeudamiento.

Otro tipo de crédito, es el hipotecario para la adquisición de viviendas, el cual es otorgado a personas naturales para la compra de una vivienda, construcción, remodelación, ampliación,

etc., es decir, que considere involucrada de una u otra manera, una vivienda. Igualmente, existen otros tipos de créditos que se dan a personas naturales con la finalidad de usarse en pagos de bienes, servicios o gastos, que no tienen relación con alguna actividad empresarial.

Continuando con la clasificación de los créditos, Vela & Caro (2015) establecen los llamados “por deudores minoristas o no minoristas”, los cuales consideran tanto personas naturales como jurídicas, y su clasificación depende de su uso y a quien es otorgado, es decir, tiene relación si el mismo es suministrado a microempresas o grandes empresas y su uso es empresarial o particular. Asimismo, se tiene otra clasificación denominada por producto crediticio, relacionado con el tipo de producto para el cual es dado el crédito, por ejemplo, el crédito prendario, el crédito personal directo, crédito convenio descuento por planillas, o crédito por microconsumo, crédito vehicular, por capital de trabajo, créditos para Micropyme, de viviendas, agropecuario, ganadero, agrícola, etc.

Otra clasificación relevante del crédito considera una clasificación del tipo de deudor, el cual puede ser clasificado como normal, con problemas potenciales, deficiente, dudoso, o categoría de pérdida ésta última clasificación es una de las que considera el riesgo crediticio dentro de su asignación. Asimismo, existe una clasificación denominada contable, la cual considera si la cartera es vigente, vencida, judicial, refinanciada o reestructurada.

De acuerdo a las diferentes clasificaciones de los créditos, es importante considerar el riesgo que se asume en crédito, a lo cual se le denomina el riesgo crediticio. En este sentido, Vela & Caro (2015) lo definen “como aquella posibilidad de pérdidas por la incapacidad o falta de voluntad de los deudores, contrapartes, o terceros obligados, para cumplir sus obligaciones contractuales registradas dentro o fuera del balance” (p.33). Es por ello, que es relevante la gestión eficiente del riesgo crediticio, puesto que cualquier resultado negativo en función de la recuperación del mismo con sus correspondientes intereses, son considerados como ingresos en suspensos que no se pueden considerar como ingresos financieros. En este

sentido, se pueden tener créditos en situación de vencidos, considerados así desde la primera cuota de vencimiento; también se tienen los créditos en situación de cobranza judicial, créditos en situación de refinanciamiento, créditos en situación de reestructurado, créditos de clasificación dudosa, créditos clasificados de pérdida y los denominados créditos castigados.

Según Salas (1998), la morosidad es la falta de puntualidad o retraso, en especial en el pago de una cantidad establecida o en la devolución de una cosa previamente acordada. Se considera que la morosidad se encuentra fuertemente ligada al ciclo económico, sobre todo en épocas de recesión: En situaciones de incertidumbre económica, en donde las empresas ven reducidas sus ventas, y, a través de reformas, deben reestructurarse o incluso suspender pagos o quebrar, debido a que ya no son solventes para seguir cancelando la deuda acordada. Si bien existe una relación negativa entre la morosidad y el ciclo económico, se puede inferir que el riesgo crediticio que soportan las entidades bancarias nace debido a la morosidad.

Es importante resaltar que la situación de incumplimiento en el pago o morosidad de un crédito, puede suceder únicamente por dos casos: en primer lugar se tiene la falta de capacidad de pago del deudor ya que no posee los recursos para cumplir con los compromisos adquiridos; en segundo lugar, se tiene la falta de voluntad de pago por parte del deudor, es decir, tiene el dinero para pagar pero no tiene la disposición de hacerlo. Estas son variables fundamentales que se deben evaluar al momento de asumir el riesgo de dar crédito, tanto a personas naturales como a personas jurídicas.

Otro efecto que provoca la morosidad en el riesgo crediticio es la perspectiva que se obtiene del cliente que debe cancelar, por lo cual entran una serie de variables del perfil estadístico del cliente que se puede resumir en:

1. Cliente responsable: Paga las deudas en el momento concertado en la venta a crédito

2. Cliente no responsable: No paga las deudas en el momento concertado en la venta a crédito.

Estos factores son evaluados internamente en la entidad bancaria, dentro de los clientes no responsables, o bien, declarados como “clientes morosos”. El análisis se comienza desde la premisa de que los “morosos” pueden regularizar su situación dentro de un periodo determinado. Si bien, dicho comportamiento queda registrado en la información de la entidad bancaria, no obstante, la información puede influir en la entidad si otorgarle un crédito o no, considerándolo como una variable crítica.

Según Ríos (2008), toda teoría financiera y lo que respecta su estudio, se debe plantear a partir de una hipótesis y una prueba, con el fin de determinar si a mayor riesgo hay un mayor rendimiento del capital, en referencia al crédito financiero. De todas maneras, entre más alta sea la posibilidad de que el rendimiento real se encuentre por debajo del esperado, este implica un mayor riesgo asociado al hecho de ser un propietario activo.

Si se considera el crédito ideal, este se define como la acción financiera de entregar seguridad total o riesgo nulo (Gurný y Gurný, 2013). No obstante, esto es difícil de conseguir ya que no existe un crédito sin riesgo, por lo cual, la entidad bancaria no solo busca recuperar el capital que otorgan en forma de créditos, sino que también se espera tener un rendimiento económico por la función de intermediarios y la gestión de aceptar el riesgo de no pago.

4.2. Criterios y herramientas de evaluación del riesgo crediticio

Dentro de los criterios utilizados en la evaluación de riesgos crediticios, Vela & Caro (2015) señalan que existen dos criterios básicos, que al ser aplicados de manera eficiente y de forma conjunta, ayuda en la toma de decisión con relación al otorgamiento del crédito. Uno de ellos es la capacidad de pago del cliente, el cual se considera un criterio objetivo. La capacidad de

pago del cliente toma en consideración varias herramientas que permiten definir la capacidad de pago del solicitante del crédito, dentro de las cuales se tienen el diseño y análisis del flujo de caja de la persona que desea adquirir un crédito.

Otro criterio se refiere a los antecedentes crediticios del solicitante del crédito, el cual se considera un criterio subjetivo. En este sentido, se realiza una investigación con relación al cumplimiento de los compromisos contractuales con sus acreedores, por parte del solicitante, ya sean crediticios, comerciales, tributarios, etc. Una de las formas para determinar los antecedentes del solicitante, es a través de las centrales de riesgos. Las mismas son entidades que registran información, tanto de las personas naturales como jurídicas, con relación al cumplimiento de las obligaciones contraídas con entidades financieras, comercios, cooperativas, etc. Las centrales de riesgo, ayudan a generar confianza en el proceso de otorgamiento de los créditos, tanto a personas como a empresas.

No se debe descartar, la posibilidad de la visita a las entidades comerciales con la cual el solicitante ha tenido relación. Lo anterior forma parte de una fuente de información de tipo cualitativo, de la que se puede obtener una idea de los antecedentes de la persona que ayudará a tomar mejores decisiones con relación al riesgo crediticio que se observa. Garcia (2015), presenta una serie de elementos que se deben considerar al momento de llevar a cabo el otorgamiento de un crédito. Según el mismo autor, las entidades deben tener políticas relacionadas con la evaluación de la documentación que se requiere y que sustente la decisión de otorgar un crédito, así como, en el caso de renovaciones y refinanciación de los mismos. De igual forma, debe existir una clasificación de deudores de acuerdo a la normativa, además de existir un registro de auditoría en el cual quede constancia del proceso seguido en el otorgamiento del crédito.

De igual forma, Vela & Caro (2015) mencionan que existen otros criterios que permiten evaluar el riesgo crediticio, complementarios al criterio de capacidad de pago y antecedentes

crediticios, los cuales consideran al tipo de deudor con el que se espera trabajar, es decir, si son deudores no minoristas o deudores minoristas. En el caso de deudores no minoristas, se pueden mencionar los siguientes criterios de evaluación:

Evaluación de la situación económica y financiera; Evaluación del cambio de patrimonio neto; Evaluación de proyectos futuros y aplicación de los indicadores del presupuesto de capital; Evaluación del entorno económico, sectorial y regional; Evaluación de la capacidad de hacer frente a sus obligaciones ante variaciones cambiarias o de su entorno comercial, político o regulatorio; Evaluación de las garantías; Evaluación de los posibles efectos de los riesgos financieros relacionados a los descortes en moneda, plazos y tasas de interés de los estados financieros de la empresa deudora y que pueden repercutir en su capacidad de pago, incluyendo a las operaciones con instrumentos financieros derivados (Vela & Caro, 2015, p. 44)

En el mismo orden de idea, los criterios adicionales que se pueden aplicar a los deudores minoristas son los siguientes: cálculo del monto de las diferentes obligaciones adquiridas y el importe de las cuotas a pagar a la empresa que hace el préstamo. Es importante señalar que la aplicación de las herramientas de gestión se deben complementar con otro tipo de información, tales como, el entorno social y económico de la persona que desea el crédito, el flujo de ingreso y egreso de la familiar, con el fin de evaluar la potencialidad del cliente en relación a la voluntad de cumplimiento de las obligaciones adquiridas.

Con relación a las herramientas de evaluación del régimen crediticio, Vela & Caro (2015) describen su concepto como “el medio por el cual los criterios de evaluación se van a llevar a cabo” (p.45). El uso de las herramientas y una buena aplicación de las mismas, es un factor que contribuye de manera positiva en una evaluación del riesgo crediticio, permitiendo que las decisiones tomadas al respecto, minimicen el riesgo asumido por el ente crediticio al dar un préstamo.

En este sentido, se presentan dos tipos de herramientas a utilizar en la aplicación de los criterios mencionados anteriormente, los cuales se clasifican en cualitativas y cuantitativas.

Dentro de las herramientas cualitativas se pueden mencionar las siguientes: revisión de record crediticio del solicitante del préstamo en la entidad en la que se solicita el crédito; revisión de la central de riesgo de la entidad financiera; revisión de las referencias comerciales y personales de la persona que solicita el crédito; revisión de la central de riesgo fuera de la entidad financiera; revisión de modelos analíticos que permitan medir el riesgo crediticio, tales como, modelos basados en reglas de decisión, modelos de alerta temprana, entre otros; validación de la documentación presentada por el solicitante del crédito; evaluación de las clasificaciones que han dado otras empresas o entidades financieras al solicitante del crédito. Una herramienta que no se debe desechar, es el uso de páginas web existentes que posean información valiosa que aporte a la evaluación del solicitante.

Continuando con el tipo de herramientas que se utilizan en la aplicación de los criterios de evaluación de riesgo crediticio, se tienen las herramientas cuantitativas. Dentro de ellas se pueden mencionar las siguientes: el flujo de caja, el cual permite conocer las fuentes de ingreso a través del conocimiento de sus fuentes de ingreso y el uso de los mismos; con ello se busca evaluar la capacidad de pago del solicitante del crédito; generalmente el flujo de caja se proyecta a un tiempo futuro determinado. De igual forma se tiene como herramientas el análisis de los estados financieros, con el cual se busca conocer la situación económica y financiera del solicitante del crédito. Con ello, se busca conocer el estado de resultados y el balance general del solicitante del crédito. Igualmente se tiene como herramienta cuantitativa, la evaluación de proyectos; esta permite evaluar la viabilidad de proyectos de capital, utilizando ciertos indicadores como es el valor actual neto (VAN), tasa interna de retorno (TIR), periodo de recuperación de capital, la tasa de retorno promedio contable y el Índice de beneficio - costo .

Profundizando un poco más en las herramientas cuantitativas de evaluación de riesgo crediticio, se pueden mencionar ciertas características del flujo de caja, el cual esté constituido por la entradas de efectivo (considera los ingresos al contado más los ingresos a

crédito más los ingresos de otros negocios más otros ingresos extraordinarios). De igual forma se tienen las salidas de efectivo, las cuales corresponden al pago por costo de reposición, gastos administrativos, gastos de ventas, gastos financieros, impuestos, amortización del préstamo y otros gastos. Asimismo, también se tiene el saldo inicial de caja, el cual corresponde al monto con que cuenta la empresa inicialmente en ahorro. De manera resumida se expresa el flujo de caja como sigue:

Flujo de caja final = Entrada de efectivo – salidas de efectivo + saldo inicial de caja

Dentro de las ventajas que tiene el flujo de caja se pueden mencionar las siguientes: permite determinar la capacidad de pago del solicitante de un crédito. Con ello se puede verificar que la cuota de pago se pueda cubrir, lo cual se calcula mediante la diferencia entre los ingresos y los egresos de efectivo. Otra de las ventajas que es que el flujo de caja muestra organizadamente los ingresos y los egresos de efectivo en el periodo en consideración. Asimismo, con el flujo de caja se pueden predecir los requerimientos de financiamiento y la forma de cumplir con el compromiso adquirido, se puede conocer la viabilidad económica, operativa y financiera del proyecto para el cual se está solicitando el crédito.

Dentro de las desventajas que tiene el flujo de caja es que se requiere de la sinceridad de los números referidos a los ingresos y egresos de efectivo, por lo que, si esto no se cumple, los resultados de omitirlos, puede generar distorsiones en el proceso de toma de decisiones. Por otro lado, para el cálculo del flujo de caja se requiere de la utilización de condicionantes, los cuales pueden generar distorsiones en el proceso de toma de decisiones, si no se hace de manera adecuada.

Con relación a la herramienta referida a los estados financieros, los mismos constituyen una herramienta cuantitativa que permite gestionar el riesgo crediticio. Se conocen como estados financieros los siguientes elementos: el balance general, el estado de resultados, el estado de cambio en el patrimonio neto y el estado de flujo de efectivo. El balance general está

constituido por los activos (son los resultados de transacciones que generan beneficios económicos a futuro), los pasivos (consideran las deudas que se espera generen salida de recursos) y el patrimonio neto (se refiere a los recursos que posee una empresa, posterior al descuento de las deudas que puede tener); se describe como la diferencia entre el activo total menos el pasivo total.)

Según Vela & Caro (2015), el análisis de los estados financieros permite el uso de la información, con la finalidad de sustentar la toma de decisiones. Algunos métodos utilizados para hacer el análisis de los estados financieros son los métodos de análisis horizontal, método de análisis vertical y método de *ratios* financieros.

4.3. Concepto de Quiebra

Según Cinca (1993), no existe una definición propiamente tal de quiebra en los modelos de predicción, ya que se determina como una variable dependiente denominada “quiebra”, aunque esta definición varía dependiendo el modelo y del concepto que se pretende definir.

Por otra parte, Pastena y Ruland (1986) plantean 3 formas de describir la quiebra:

- Insolvencia contable/económica.
- Incumplimiento.
- Quiebra legal: “Procedimiento de liquidación ordenada de los activos de una empresa en situación de insolvencia”.

Dentro de las formas planteadas por Pastena y Ruland, hay dos conceptos que se tienden a confundir: El incumplimiento de pagos y la insolvencia contable y/o económica. Según (García, 2017), el incumplimiento se define cuando una entidad no cumple con el compromiso financiero establecida previamente, por lo cual incurre en una falta si es que no realiza los pago en el plazo determinado, o bien, incurre en no pago del monto acordado, por lo cual la entidad bancaria asume el riesgo.

Por otra parte, el concepto de insolvencia se debe distinguir en contable y económico: En la primera, se define cuando el patrimonio arroja valores negativos, por ende, el valor contable de los activos es menor al valor contable de los pasivos y en el caso de la insolvencia económico, es cuando los activos son menores al valor económico de los pasivos.

El incumplimiento e insolvencia normalmente se tratan como sinónimos al realizar un análisis de quiebra, incluso son considerados indicadores, no obstante, tal como plantea García (2017), el incumplimiento de alguna obligación, es considerado como un indicio de insolvencia, por lo cual puede conllevar a la quiebra, reflejando la poca liquidez de la empresa, incluso se puede ir una empresa a quiebra sin incumplir alguna de sus obligaciones por falta de liquidez.

Con el fin de poder establecer parámetros para medir y cuantificar el riesgo crediticio se empezó a incluir los ratios financieros, así como lo señala Sunder (1979), una de las principales razones de la tasa de inclusión es controlar la influencia del tamaño de cada empresa en las variables explicativas.

En términos generales, y a fin de consensuar la definición para las etapas sucesivas del desarrollo del texto, la quiebra se refiere a una situación de desequilibrio entre los valores realizables y las prestaciones exigibles de una firma que conllevan a que la empresa no pueda hacerse cargo de sus obligaciones, conocido como una situación de insolvencia o, alternativamente, cuando se genera un desequilibrio en la identidad contable básica, siendo sustancialmente superior las obligaciones que los recursos disponibles:

$$\text{Activos} = \text{Pasivos} + \text{Patrimonio}. \quad (4.1)$$

Cabe hacer el alcance de que el concepto de quiebra es un nuevo concepto en la legislación chilena, específicamente aplicado en la Ley de Reorganización y Liquidación de Empresas y Personas, que está establecida en la ley N° 20.720, que establece el régimen general de los procedimientos destinados a reorganizar y/o liquidar los pasivos y activos de deudores de

empresa, y a repactar los pasivos y/o liquidar los activos de deudores persona natural de fecha 09 de octubre del año 2014. Ésta abarca todas las situaciones críticas de las empresas, y no especifica exactamente la definición, debido a que se considera como un proceso multidimensional donde participan diversos actores, por lo que no es posible afirmar una definición en conformidad con lo legal dentro del territorio nacional.

4.4. Modelos de Riesgo Crediticio

Un ejemplo de la variabilidad del concepto de quiebra, son las distintas perspectivas que abordan los autores para concebir modelos para su identificación o predicción. En general, existen modelos de tipo univariado que estudian el riesgo crediticio, los que se pueden dividir en dos categorías: Los modelos estadísticos (Altman, Ohlson, Shumway, Duffie y Wang) y aquellos modelos teóricos.

Los modelos estadísticos intentan identificar variables que puedan explicar la quiebra, es decir, variables que se comporten de diferentes formas en el proceso que implica la quiebra. La identificación de estas variables se realiza mediante diferentes técnicas estadísticas, con modelos empíricos que no parten de la teoría que se quiere contrastar.

A diferencia de los modelos estadísticos, los modelos teóricos se basan en una teoría formal y desde ahí comienzan el análisis. Estos modelos aplican la teoría de opciones de Black & Scholes (1973) y la teoría Merton (1974), sobre el análisis de riesgo corporativo, en la cual las acciones de una firma son vistas como una opción sobre los activos de la misma.

García (2017) menciona que, en general, la probabilidad de quiebra estimada por modelos teóricos es mayor que el modelo estadístico. Esto sucede porque la probabilidad estimada por el modelo teórico indica la probabilidad de quiebra asumiendo que las demás condiciones de la empresa pueden ser invariables. En otras palabras, si una empresa no realiza ninguna acción para mejorar su situación (las demás condiciones son constantes), existe la posibilidad de quebrar.

Altman (1968) fue uno de los autores que comenzó los modelos de riesgo crediticio. Creó su modelo de puntaje Z, utilizando variables entre capital de trabajo en activos totales, ganancias retenidas en activos totales y ganancias antes de intereses e impuestos para encontrar la probabilidad de quiebre de la empresa. Inclusive, el más utilizado y confiable es el modelo de Altman denominado Z-Score, en donde, a través de una ecuación matemática se puede predecir el estado de quiebra. Este considera que las empresas quebradas son aquellas que realizan una declaración legal, no obstante, esto depende del resultado de la siguiente ecuación:

$$\text{Altman } Z - \text{score} = 1,2 \times T1 + 1,4 \times T2 + 3,3 \times T3 + 0,6 \times T4 + 1,0 \times T5, \quad (4.2)$$

donde:

T1: (Capital Circulante/Activos Totales)

T2: (Beneficios no distribuidos/Activos Totales)

T3: (EBITDA/Activos Totales)

T4: (Capitalización Bursátil/Deuda Total)

T5: (Ventas Netas/Activos Totales)

Las principales conclusiones del modelo reportan que la relación entre ventas y activos totales ayuda a identificar si la empresa puede ir a la quiebra, así como también, la tasa de rentabilidad de los activos.

De igual forma, Salamanca & Benitez (2018), identifican dos tipos de modelos que estudian el riesgo crediticio, los cuales se concentran en los estructurales y los de forma reducida. Los modelos estructurales pueden determinar la probabilidad de incumplimiento en el pago del crédito, así como, las posibles situaciones que pudieran hacer que una empresa o persona con un crédito, incumpla con sus compromisos de pagos. Los mismos, consideran como

variables la información completa del mercado y del valor de los activos y las firmas. Asimismo, los modelos de forma reducida, “Se llaman modelos de intensidad, en la que se define de manera exógena un proceso estocástico para el evento de incumplimiento” (Jarrow y Protter, 2004, p.1-10.). Dentro de estos últimos, se encuentran los métodos de estimación *Logit* y *probit*.

Merton (1974) propuso un modelo econométrico para predecir la probabilidad de incumplimiento de la empresa a través del precio de mercado de la empresa. El modelo considera los activos y pasivos, así como la volatilidad de los activos y sus rendimientos esperados.

Merton utiliza el porcentaje de activos y pasivos para mostrar el grado de apalancamiento de la empresa y calcula la probabilidad de que la empresa no pueda pagar intereses con base en esta información. Es precisamente por este modelo que Merton (1974) concluyó que cuando los activos de la empresa son menores que sus pasivos, la empresa incurrirá en incumplimiento y cuanto mayor sea el apalancamiento, mayor será la probabilidad de incumplimiento.

Ohlson (1980), fue el primero en incluir un método de estimación de máxima verosimilitud para encontrar la probabilidad de incumplimiento. Para ello, utiliza el modelo *Logit* condicional, que incluye activos totales entre el índice de precios, pasivos totales / activos totales, capital de trabajo en activos totales, activos circulantes en pasivos circulantes, si la variable ficticia con un pasivo total de 1 excede los activos totales. , de lo contrario Si es 0, la utilidad neta en los activos totales, la utilidad operativa en los activos totales, si la utilidad neta en los últimos dos años es negativa, entonces es 1, de lo contrario es 0, y el cambio en el beneficio neto.

Además, Ohlson (1980) señaló que, al utilizar una muestra de población más grande, se pueden hacer menos suposiciones sobre la distribución de las variables del modelo y

posteriormente, llegó a la conclusión de que el modelo logit utiliza una transformación lineal de los ratios de la empresa para tener un fuerte poder predictivo.

Lennox (1999), tiene como objetivo explicar y analizar la confiabilidad y la capacidad de presentación de informes correspondientes de auditorías, utilizando modelos *Logit* y *Probit* para comparar las variables de índice financiero de flujo de efectivo, deuda y rentabilidad a fin de identificar la probabilidad de que una empresa esté al borde de la quiebra. Los resultados del informe de auditoría, reflejan con mayor precisión la probabilidad de quiebra de la empresa y saca como conclusión que el modelo de predicción de la probabilidad predeterminada es más fiable que el informe de quiebra. Otro factor, es que la empresa ha incrementado su probabilidad de incumplimiento al presentar altos niveles de endeudamiento, baja rentabilidad y problemas de flujo de caja.

Lando (2004) describe los diferentes modelos utilizados para medir el riesgo crediticio y cómo aplicar específicamente cada modelo de manera similar, considerando un método estadístico para analizar pérdidas o defectos. El autor describe la fase de estimación del modelo *logit*, que se considera una variable dependiente, y su aplicación al riesgo crediticio, sea o no el crédito por default. Por lo tanto, muestra que la probabilidad de incumplimiento es una función de una característica particular (covariable) y concluye que el modelo logit puede usarse de manera efectiva para encontrar la probabilidad de incumplimiento de un préstamo.

Gupton (1997), proporcionó una descripción general de la metodología y las matemáticas detrás de la estimación estadística del riesgo crediticio y una documentación detallada del análisis que produce los conjuntos de datos que proporciona. Para ello se maneja el modelo de series de Márkov, que utiliza las calificaciones de cada deudor, y lo muestra en la matriz. Esta es una herramienta que permite determinar la probabilidad de llegar al estado en el que existe la deuda. Desde un cierto *período* t hasta un cierto período t_1 .

Ching, Huang, Ng, & Sui (2013) son autores representativos de las cadenas de Markov, las cuales son secuencias de variables aleatorias que corresponden al estado de un cierto sistema, el cual depende solo del estado en el período anterior. Si bien, Ching busca proponer un modelo de Markov con un algoritmo eficiente para estimar los parámetros del modelo y mostrar una aplicación de la modelización de tipos de interés, las calificaciones crediticias y los datos predeterminados contando con el permiso de diversas instituciones importantes. Dicho modelo, muestra los resultados de las estimaciones de una matriz de transición con datos del incumplimiento de bonos trimestrales en cuatro sectores (consumo, energía, tecnología y transporte) en donde se asume que la transición del estado de riesgo oculto de cada sector depende del número actual observado de bonos que incumplen en cada sector, dando como resultados que 6 o más incumplimientos de 40 observaciones, con aumento de riesgo en el sector de consumo.

4.5. Modelo Logit

La regresión logística es una técnica estadística multivariante que permite estimar la relación entre una variable de tipo dependiente no métrica, en especial dicotómica, respecto un conjunto de variables independientes, sean estas métricas o no.

El modelo *Logit* corresponde a un modelo de ajuste no lineal. En general, se define como una característica señalada X_i , un modelo cualitativo en el que se supone que un individuo es parte de uno de los dos grupos particulares representados por Y_i utilizando variables explicativas. Esto se puede interpretar como que es una variable dicotómica dependiente de la calidad de cliente, determinando un valor igual a 1 si está en el grupo exitoso e igual a 0 si está en el grupo negativo, es decir, es una variable dependiente binaria.

El modelo logit se considera un modelo de tipo logístico, en el cual la variable dependiente se puede representar de la siguiente manera: $Y = 1$ con probabilidad p ó $Y = 0$ con probabilidad

1-p. Este modelo utiliza el método de máxima verosimilitud siendo una función de colas más anchas, en el cual hay mayor probabilidad de éxito en los extremos.

El modelo presenta la siguiente forma:

$$Y = F(\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \dots + \beta_k X_k). \quad (4.3)$$

Recordando que en este modelo, la variable dependiente es una variable dicotómica, la probabilidad de que la misma tome valores entre 1 y 0 es la siguiente:

$$P_r \left(Y = \frac{1}{X_{2, \dots, x_k}} \right) = E \left(\frac{Y_i}{X = x} \right) = ezi_1 + ezi, \quad (4.4)$$

$$P_r \left(Y = \frac{0}{X_{2, \dots, x_k}} \right) = 1 - ezi_1 + ezi = 11 + ezi. \quad (4.5)$$

Para evaluar el peso de las variables, se utiliza el estadístico de Wald, el cual a su vez, considera el hecho de que las variables sean categóricas o no sean categóricas. Esto conduce a las siguientes ecuaciones:

Si la variable es categórica: $Wald(\beta_i) = \left(\frac{\beta^{\wedge}_i}{\sigma(\beta_i)} \right)^2$,

considerando una distribución ji cuadrado, con número de grados de libertad igual al número de parámetros por estimar.

Si la variable no es categórica: $Wald_i = \beta_{i2} \hat{\sigma} \beta_{12}$, con una distribución ji al cuadrado con un grado de libertad.

Posteriormente, se debe aplicar una prueba de hipótesis. Para ello se considera la hipótesis nula como la variable X_i no es relevante para establecer las variaciones en el modelo, y la hipótesis alterna es lo opuesto a la hipótesis nula.

Es decir:

Ho: $\beta_i = 0$; la variable X_i no es relevante para explicar variaciones en el modelo.

H1: $\beta_i \neq 0$; la variable X_i es relevante para explicar las variaciones en el modelo.

Es en este punto que surge el concepto de regresión logística, el cual simplifica la discriminación a los solicitantes de crédito en buenos y malos prospectos, dicho esto, la regresión logística es un método de clasificación que se utiliza comúnmente en el *credit scoring*. Este modelo no requiere de los supuestos de la regresión lineal, como el supuesto de normalidad de los errores de observación. Es en este aspecto que la regresión logística se utiliza tanto en datos que son gaussianos como datos que no lo son, por tanto, este modelo es útil cuando la variable de respuesta no está distribuida normalmente y tanto las variables dependiente como las independientes tienen valores discretos, categóricos ordinales o no ordinales.

La regresión lineal no es aplicable a este tipo de variables, ya que la variable respuesta solo presenta dos valores. Esta capacidad predictiva del modelo logístico se valora mediante la comparación entre el grupo de pertenencia observado y el pronosticado por el modelo. Dicho lo anterior, el modelo debe ser capaz de clasificar a los individuos en cada uno de los dos grupos: buenos o malos, basado en las variables que definen las características de los individuos en función del pago de sus obligaciones financieras.

Esta clasificación está a cargo de una distribución de probabilidad que separa a la población en dos grupos, la separación está basada en un punto de corte preestablecido en el rango de cero a uno. Esta distribución de probabilidad sirve para estimar el valor de y que dependiendo de su valor asignara al individuo a un grupo; por ejemplo, se espera que para buenos clientes se obtengan valores muy cercanos a uno y para malos clientes valores cercanos a cero.

El modelo *Logit*, puede ser de variable con respuesta dicotómica o múltiple. De igual forma, este modelo tiene la media cero y su varianza es diferente. Asimismo, es un modelo más práctico, desde el punto de vista matemático.

4.6. Modelo de regresión logística

Si bien en las variables explicativas, no se establece ninguna restricción, logrando ser cualitativas o cuantitativas y a su vez tanto continuas como discretas y categóricas, con dos o más valores. La variable respuesta y tendrá los valores cero y uno, podemos definir $y = 1$ si se trata de un buen cliente e $y = 0$ si se trata de un mal cliente. Con la regresión logística se modela la probabilidad de que, y sea igual a uno,

dado los valores observados de las variables predictoras contenidas en el vector \mathbf{x} , $P\left(y = \frac{1}{x}\right)$.

Para discriminar a un individuo se estima su probabilidad con $P^{\wedge}\left(y = \frac{1}{x}\right)$ y si por ejemplo

$P_b = \left(y = \frac{1}{x}\right) > 0,5$ se clasifica al cliente como bueno y si $P_b = \left(y = \frac{1}{x}\right) < 0,5$ se clasifica como malo.

Para justificar el modelo de regresión logística, consideremos una muestra de n datos donde \mathbf{x}_i es el vector de variables explicativas binarias de la forma que tiene asociada una variable de respuesta binaria dependiente y_i que toma el valor de $y = 0$, si el cliente i es malo, e $y = 1$ si el cliente i es bueno. Sea $P(y = 1|x_i) = p_i$ la probabilidad de que $y = 1$, dado el vector \mathbf{x}_i de datos observados. Se define una relación entre p_i y un modelo lineal mediante una función monótona y creciente g , llamada función enlace (*link*).

$$\mathbf{x}_i^T = [X_{i1}, X_{i2} \dots \dots, X_{ip}], \quad i = 1, 2, \dots, n, \quad (4.6)$$

$$g(P_i) = \beta_0 + \beta_1^T x_i, \quad (4.7)$$

tal que:

$$\beta_1^T = [\beta_1, \beta_2, \dots, \beta_p]. \quad (4.8)$$

El vector de parámetros de coeficientes de las variables explicativas del modelo y β_0 la ordenada al origen. La función *link* que se aplica se conoce como la transformación logito y es el logaritmo del cociente de probabilidades de p_i y $(1 - p_i)$

$$g(p_i) = \log \left(\frac{p_i}{1-p_i} \right) = \beta_0 + \beta_1^T x_i. \quad (4.9)$$

El modelo en términos de $g(p_i)$ puede escribirse como $g(p_i) = \beta_0 + \beta_1^T x + \varepsilon$, con ε variable aleatoria tal que $E(\varepsilon) = 0$ y $V(\varepsilon) = \sigma^2$. La función de distribución logística dada por la transformación inversa de g se escribe como

$$p_i = \frac{e^{\beta_0 + \beta_1^T x_i}}{1 + e^{\beta_0 + \beta_1^T x_i}}, \quad (4.10)$$

que satisface $0 \leq p_i \leq 1$. Así

$$1 - p_i = \frac{1}{1 + e^{\beta_0 + \beta_1^T x_i}}. \quad (4.11)$$

Los coeficientes del modelo logístico sirven para calcular un parámetro de cuantificación de riesgo conocido como *odds ratio* (OR), también denominado razón de probabilidades. El *odds* asociado a un evento, es el cociente entre la probabilidad de que ocurra con la probabilidad de que no ocurra.

$$odds = \frac{p_i}{1-p_i} . \quad (4.12)$$

Un caso particular es cuando todas las variables explicativas pueden ser representadas de tal modo que todas sean binarias independientes. Cada variable se le asocia una probabilidad según sea la población a la que pertenecen. Sean $P(\mathbf{x}_i|y = 1)$ y $P(\mathbf{x}_i|y = 0)$ las probabilidades del vector \mathbf{x}_i dado que el individuo pertenece a la población uno y dos respectivamente. Bajo el supuesto que las coordenadas de $\mathbf{x}_i = (x_{i1}, x_{i2}, \dots, x_{ip})$ son binarias e independientes, se tiene que

$$P(\mathbf{x}_i|y = 1) = P[(x_{i1}, x_{i2}, \dots, x_{ip})|y = 1] = \prod_{j=1}^p P(x_{ij}|y = 1) = \prod_{j=1}^p p_{1j}^{x_{ij}} (1 - p_{1j})^{1-x_{ij}} \quad (4.13)$$

y

$$P(\mathbf{x}_i|y = 0) = \prod_{j=1}^p p_{2j}^{x_{ij}} (1 - p_{2j})^{1-x_{ij}}, \quad (4.14)$$

$j=1$.

Suponiendo que las probabilidades a priori son las mismas, es decir

$$P(y = 0) = P(y = 1)$$

y considerando que nuestro modelo es construido sobre la misma cantidad de datos para las dos poblaciones. Se tiene que la probabilidad condicionada está dada por

$$P(y = 1|x_i) = \frac{P(y = 1|x_i) P(y = 1)}{P(x_i)} = \frac{P(y = 1)}{P(x_i)} \prod_{j=1}^p p_{1j}^{x_{ij}} (1 - p_{1j})^{1-x_{ij}} \quad (4.15)$$

y para $1 - P(y = 1|x_i) = P(y = 0|x_i)$

$$P(y = 0|x_i) = \frac{P(x_i | y = 0) P(y = 0)}{P(x_i)} = \frac{P(y = 0)}{P(x_i)} \prod_{j=1}^p p_{2j}^{x_{ij}} (1 - p_{2j})^{1-x_{ij}}, \quad (4.16)$$

calculamos la distribución logística

$$\frac{P(y = 1|x_i)}{1 - P(y = 1|x_i)} = \frac{\frac{P(y = 1)}{P(x_i)} \prod_{j=1}^p p_{1j}^{x_{ij}} (1 - p_{1j})^{1-x_{ij}}}{\frac{P(y = 0)}{P(x_i)} \prod_{j=1}^p p_{2j}^{x_{ij}} (1 - p_{2j})^{1-x_{ij}}} = \prod_{j=1}^p \left(\frac{p_{1j}}{p_{2j}} \right)^{x_{ij}} \left(\frac{1 - p_{1j}}{1 - p_{2j}} \right)^{1-x_{ij}}, \quad (4.17)$$

calculamos la transformación función logito,

$$g_i(x_i) = \log \frac{P(y = 1|x_i)}{1 - P(y = 1|x_i)}, \quad (4.18)$$

$$g_i(x_i) = \sum_{j=1}^p \log \left(\frac{p_{1j}(1 - p_{1j})}{p_{2j}(1 - p_{2j})} \right)^{x_{ij}} + \sum_{j=1}^p \log \left(\frac{1 - p_{1j}}{1 - p_{2j}} \right). \quad (4.19)$$

Se observa que $g_i(x_i)$ es una función lineal que coincide con la ecuación mencionada anteriormente, donde

$$\beta_0 = \sum_{j=1}^p \log \left(\frac{1 - p_{1j}}{1 - p_{2j}} \right) \quad (4.20)$$

y

$$\beta_A^T = \left[\log \frac{p_{11}(1-p_{11})}{p_{21}(1-p_{21})}, \dots, \left[\log \frac{p_{1p}(1-p_{1p})}{p_{2p}(1-p_{2p})} \right] \right]. \quad (4.21)$$

Para el modelo $g(P_i) = \beta_0 + \beta_1^T x_i$. El parámetro β_0 nos da la ordenada al origen, y $\beta_1 = (\beta_1, \dots, \beta_p)^T$ es el vector de pendientes. Así

$$\log\left(\frac{p_i}{1-p_i}\right) = \beta_0 + \beta_{1x_{i1}} + \dots + \beta_{px_p}, \quad (4.22)$$

y β_j con $j = 0, 1, \dots, p$, nos sirve para analizar la cantidad de cambio del *ratio* de probabilidades cuando se incrementa una variable de predicción x_j en una unidad.

$OR = \exp(\beta_j)$.

El modelo de regresión $g(P_i) = \beta_0 + \beta_1^T x_i$, es una transformación lineal en los parámetros del modelo por lo que podemos utilizar algunas técnicas aplicadas en el análisis de regresión lineal, como por ejemplo la selección de variables mediante el proceso de *backward* y *forward*.

La expresión general del logit de tipo dicotómico se modeliza como una ecuación cuyo resultado se interpreta como la probabilidad de pertenecer a un grupo codificado con valor 1. Ésta se visualiza a continuación en la ecuación 4.23:

$$Prob(Y_i = 1) = \frac{1}{1 + e^{-(\alpha + \beta_k \chi_{ki})}} = \frac{e^{\alpha + \beta_k \chi_{ki}}}{1 + e^{\alpha + \beta_k \chi_{ki}}}. \quad (4.23)$$

Una variable dependiente binaria es una clase limitada de variables dependientes, también conocida como VDL, y se caracteriza por un rango bastante limitado de valores posibles. Dado que es una variable binaria, se limita a elegir solo los valores 0 y 1.

Es posible sostener que el análisis de regresión logística se sustenta en principios similares que el análisis de regresión lineal múltiple, siendo diferente en que el Logit considera que la variable dependiente es métrica, por lo que, si bien su uso práctico es semejante, su enfoque matemático difiere sustancialmente. El modelo *Logit* aplica modelos de regresión múltiple a las variables dependientes binarias, a diferencia del modelo de probabilidad lineal. La

probabilidad lineal (MPL), supera algunas de las deficiencias de una variable explicativa, como que las probabilidades ajustadas pueden ser mayores o menores que el valor 1, y el efecto parcial de esta variable explicativa es constante, por lo cual en *Logit*, solo las variables explicativas toman valores de 1 a 0.

Por otro lado, la mayor ventaja de la no linealidad del modelo *Logit* es que elimina la necesidad de un supuesto inicial, dado que las variables se consideran como si estuvieran distribuidas normalmente. Esto mejora la aplicación de variables cualitativas o categóricas. Su estimador es más eficiente que el estimador MPL porque las propiedades estadísticas del modelo Logit son más adecuadas.

Según Wooldridge (2009), otro punto importante con diferencia con los modelos de probabilidad lineal es que estos pueden estimarse mediante el método de mínimos cuadrados ordinarios o MCO o también en algunas ocasiones con mínimos cuadrados ponderados o MCP. Por otro lado, la no linealidad del modelo logístico hace que sea indispensable la utilización de la estimación de máxima verosimilitud, el cual figura con las siglas EMV, como método de estimación de los parámetros.

En donde la estimación de máxima verosimilitud de β hace que se maximice la log-verosimilitud. Además, dicha función de log-verosimilitud es globalmente cóncava, permitiendo así una mayor facilidad en la maximización numérica (McFadden, 1974). Debido a esta maximización, el hecho de que haya variables omitidas da pie a que la log-verosimilitud disminuya, o por lo menos no incremente. Esta situación conduce a tener que determinar si dicha disminución es lo suficientemente significativa como para llegar a la conclusión de que las variables omitidas son sustanciales. Esto es posible realizarlo mediante la prueba de razón de verosimilitud o prueba chi-cuadrado, la cual usa un estadístico de razón de verosimilitud igual a dos veces la diferencia de la log-verosimilitud del modelo restringido con la log-verosimilitud del modelo sin restricción.

Esta estadística es necesariamente positiva y, dado que, como toda ciencia, la estadística tiene sus limitaciones, inevitablemente se duplica para tener una distribución de chi-cuadrado con un grado de libertad. Esto significa que, si está utilizando un proceso *stepwise*, se puede usar la prueba de chi-cuadrado para comprobar si agregar uno o más predictores puede mejorar significativamente el modelo. El modelo *Logit* no necesita aplicar varias pruebas econométricas para evaluarlo, pero la prueba aplicada al modelo *Logit* es marginal y compara los resultados con el modelo *probit*.

Finalmente, los principales problemas que conlleva el modelo *Logit* es, en primer punto, la heterocedasticidad en las varianzas que permite mitigarse con la aplicación de Mínimos Cuadrados Generalizados. En segundo lugar, se suma el rango de variación de la estimación no se encuentra acotado entre 0 y 1, lo que genera un problema de sentido e interpretación al tener una resolución un valor que permitiría rankear a uno por sobre otro, pero sin definir el resultado por si mismo y la categoría de cliente que se es. En otras palabras, este segundo problema equivale a la dificultad de declarar que un cliente cuyo resultado es, por brindar un ejemplo aleatorio, igual a "0,68", entonces es un cliente del grupo regular o moroso, siendo necesario complementar con herramientas como rangos de aceptabilidad o límites críticos. Finalmente, el tercer y último factor es la no normalidad de la perturbación aleatoria, para este último, no se han identificado medidas paliativas.

4.7. Modelo Forward Looking

El modelo *Forward Looking* es una herramienta que permite la identificación de escenarios extremos bajo cierta probabilidad, aunque sea ínfima, por lo que se pueden someter diversas carteras de clientes a pruebas de estrés para evaluar su posible comportamiento en condiciones desfavorables.

A fin de comprender el funcionamiento del modelo, se precisa aportar ciertas observaciones respecto la representación generalizada de los riesgos preexistentes en el mercado,

utilizando, usualmente, un gráfico basado en la campana de Gauss donde es medida la curtosis, entendiendo ésta como “el coeficiente que indica la cantidad de datos cercanos a la media de la distribución de manera que, a mayor grado de curtosis, la curva tiene mayor pendiente” (Wooldridge, 2009).

A rasgos generales, se observa que, en escenarios concretos, la distribución normal de los resultados no es tal, sino más bien se encuentra distorsionada y por ende, se percibe que las colas de la curva son más aplanadas, equivalente a mencionar que es factible tener casos, con una probabilidad ínfima de ocurrencia en sus extremos. Esto abre paso a la crítica a los sistemas tradicionales de evaluación de riesgo crediticio, donde se asume independencia de los retornos del mercado, especialmente financiero, como también, se asume que se distribuyen idéntica y normalmente, desestimando, de manera significativa, la probabilidad de eventualidades extremas.

4.8. Variables WOE

Las variables WOE corresponden a la transformación de ponderación de las pruebas (Weight of Evidence), las cuales, conforman un proceso secuencial que, además, otorga un resultado acorde al modelo que tiene especial significancia para ser explicado, además de ser de fácil implementación. Por otra parte, las variables WOE permiten convertir relaciones de tipo no lineal entre la variable dependiente y las múltiples variables independientes en una relación lineal.

El nombre completo de IV es Valor de información o cantidad de información, que en ingles se escribe *information*. Cuando usamos regresión logística, árbol de decisiones y otros métodos de modelo para construir modelos de clasificación, a menudo se necesita filtrar variables independientes.

El proceso de selección de variables en el modelo, es un proceso relativamente complejo. Hay muchos factores que deben tenerse en cuenta, como la capacidad predictiva de las variables,

la correlación entre variables, la simplicidad de las variables (fácil de generar y usar) y la robustez de las variables (no es fácil de identificar), la interpretabilidad de las variables en los negocios, etc. Sin embargo, la medida más importante y directa es la capacidad predictiva de las variables.

El término "capacidad de predicción de variables" es muy general, subjetivo y no cuantitativo. Al seleccionar variables, nunca podemos decir: "Creo que esta variable tiene una fuerte capacidad de predicción, por lo que tiene que entrar en el modelo". Para esto se necesitan algunos indicadores cuantitativos específicos para medir la capacidad predictiva de cada variable independiente, y según el tamaño de estos indicadores cuantitativos, para determinar qué variables entran en el modelo. IV que es un indicador de este tipo, se puede utilizar para medir la capacidad predictiva de variables independientes. Los indicadores similares incluyen la ganancia de información, el coeficiente de Gini, etc.

4.9. Comprensión intuitiva de IV

Desde la lógica intuitiva, la cuestión de "usar IV para medir la capacidad predictiva de variables" puede entenderse aproximadamente de la siguiente manera: Suponemos que en un problema de clasificación, hay dos tipos de variables objetivo: Y_1 , Y_2 . Para predecir un individuo A , para determinar si A pertenece a Y_1 o Y_2 , necesitamos cierta información. Suponiendo que la cantidad total de información es IV , la información requerida está contenida en todas las variables independientes $C_1, C_2, C_3, \dots, C_n$, entonces, para una de las variables C_i , cuanta más información contenga, mayor será su contribución a juzgar si A pertenece a Y_1 o Y_2 , y mayor será el valor de información de C_i . Cuanto mayor sea el IV , más debe ingresar en la lista de variables en el modelo.

Anteriormente se describe el IV desde el nivel perceptivo y lógico, luego volviendo al nivel matemático, no obstante, para introducir el método de cálculo de IV , primero necesitamos conocer y comprender otro concepto, WOE , porque el cálculo de IV se basa en WOE .

4.10. WOE

El nombre completo de WOE es *weight of evidence*, es decir, peso de la evidencia. WOE es una forma de codificación de la variable independiente original. Para realizar la codificación WOE en una variable, primero debe agrupar la variable (también llamado discretización, agrupamiento, etc., todos tienen el mismo significado). Después de agrupar, para el i-ésimo grupo, la fórmula de cálculo de WOE es la siguiente:

$$WOE_i = \ln\left(\frac{py_i}{pn_i}\right) = \ln\left(\frac{\frac{\#y_i}{\#y_T}}{\frac{\#n_i}{\#n_T}}\right). \quad (4.24)$$

Entre ellos, py_i es la proporción de clientes que respondieron en este grupo (en el modelo de riesgo, corresponde al cliente predeterminado, en resumen, se refiere al individuo cuya variable predictora es "sí" o 1 en el modelo) en todas las muestras. , pn_i es la proporción de clientes que no responden en este grupo a todos los clientes que no responden en la muestra, $\#y_i$ es el número de clientes que responden en este grupo, $\#n_i$ es el número de clientes que no responden en este grupo, $\#y_T$ son todos los clientes que responden en la muestra, $\#n_T$ es el número de todos los clientes que no responden en la muestra.

A partir de esta fórmula, podemos entender que WOE en realidad representa la diferencia entre "la proporción de clientes que respondieron en el grupo actual de todos los clientes que respondieron" y "la proporción de clientes que no respondieron en el grupo actual".

Se puede obtener una simple transformación de esta fórmula:

$$WOE_i = \ln\left(\frac{py_i}{pn_i}\right) = \ln\left(\frac{\frac{\#y_i}{\#y_T}}{\frac{\#n_i}{\#n_T}}\right) = \ln \frac{\#y_i/\#n_i}{\#y_T/\#n_T}. \quad (4.25)$$

Después de la transformación, podemos ver que WOE también se puede entender de esta manera, lo que representa es el ratio actual de clientes que respondieron a clientes que no respondieron en este grupo, y la diferencia entre éste ratio en todas las muestras. Esta diferencia se expresa por la razón de estas dos razones y luego el logaritmo. Cuanto mayor sea el WOE, mayor será la diferencia y mayor será la probabilidad de que respondan las muestras de este grupo. Cuanto menor sea el WOE y menor la diferencia, es menos probable que respondan las muestras de este grupo.

Con la introducción anterior, podemos dar formalmente la fórmula para el cálculo de IV. Para una variable agrupada, el WOE del i -ésimo grupo se ha introducido anteriormente y se calcula de la siguiente manera:

$$WOE_i = \ln\left(\frac{py_i}{pn_i}\right) = \ln\left(\frac{\frac{\#y_i}{\#n_i}}{\frac{\#y_T}{\#n_T}}\right) = \ln\frac{\#y_i/\#n_i}{\#y_T/\#n_T}. \quad (4.26)$$

De manera similar, para el grupo i , habrá un valor IV correspondiente, la fórmula de cálculo es la siguiente:

$$IV_i = (py_i - pn_i) * WOE_i = \left((py_i - pn_i) * \ln\left(\frac{py_i}{pn_i}\right) \right) = \left(\frac{\#y_i}{\#y_T} - \frac{\#n_i}{\#n_T} \right) * \ln\frac{\#y_i/\#n_i}{\#y_T/\#n_T}. \quad (4.27)$$

Con el valor de IV de cada grupo de una variable, podemos calcular el valor de IV de toda la variable. El método es muy simple, que consiste en sumar el IV de cada grupo:

$$IV_i = \sum_{i=1}^n IV_i. \quad (4.28)$$

Entre ellos, n es el número de grupos de variables, cabe destacar, que la aplicación de ponderación de las pruebas es de especial uso en modelados de regresiones logísticas, puesto que se sustentan ambos en el cálculo de log-odds. En complemento, su aplicación normaliza las variables independientes, por lo que los parámetros son directamente comparables al aplicar una regresión logística, aunque, en contraparte, solo considera el riesgo relativo que cada entidad financiera define para cada variable, sin incorporar de manera directa aquella proporción de este riesgo respecto el mercado o similares, siendo necesario valorizar la información disponible que se pueda incorporar para la ponderación de cada variable, a fin de evaluar su contribución relativa en el modelo.

Otro punto a comentar sobre las variables WOE, tienen que ver con que a través de la regresión logística visto como la técnica más común en la asignación de puntajes y valoraciones de riesgo crediticio, es que permite resolver problemas de clasificación binaria, aunque también, precisa que, previamente el ajuste del modelo, se realice otra iteración de la selección de variables, a fin de comprobar si estas son explicativas tras la transformación con WOE.

4.11. Modelo probit. McFadden (1973)

Este modelo es usado para hacer estimaciones del riesgo crediticio. El mismo se basa en la teoría de utilidad o de la perspectiva de selección racional, con fundamento en el comportamiento. Surgió de una función de distribución acumulativa normal, cuyas variables son discretas, las cuales se derivan del número de alternativas incluidas en la variable endógena. La misma esta representada por la siguiente ecuación:

$$P\left(y = \frac{1}{x}\right) = (\beta_0 + \beta_{1x1} + \dots + \beta_{kxk}) = G(\beta_0 + \beta_x). \quad (4.29)$$

En este caso, G es una función dicotómica, la cual adopta valores entre 0 y 1 para todos los números reales de z . El modelo probit es de variable dependiente limitada. Es por ello que el

cálculo de las variables se hace por el método de máxima verosimilitud. En este caso, los valores de los parametros son estimados para que los mismos maximicen el logaritmo de la función de verosimilitud. La función esta dada, para la observación, como sigue:

$$\lambda_i(\beta) = y_i \log[G(x_i\beta)] + (1 - y_i) \log[1 - G(x_i\beta)]. \quad (4.30)$$

El logaritmo de la función de verosimilitud para una muestra de tamaño n está dada por:

$$\mathcal{L} = \sum_{i=1}^n \lambda_i n_i = 1(\beta). \quad (4.31)$$

Para probar cada coeficiente se lleva a cabo la prueba de hipótesis, en la que: $H_0: \beta_i = 0$

De igual forma que el modelo de logit, también se puede utilizar el estadístico Wald y el de la razón de verosimilitud, para probar la significancia de variables de manera conjunta. Asimismo, se puede emplear la distribución ji cuadrado. El modelo *probit*, puede ser de variable con respuesta dicotómica o múltiple. El mismo se utiliza en los modelos de riesgo de crédito, ya que es utilizado cuando la variable endógena es clasificada de manera natural. Este es un modelo que en su funcionamiento se parece al modelo *logit*, ya que ambos son semejantes debido a que tienen colas un poco más anchas.

Capítulo 5

5. Metodología

El presente estudio tiene como objetivo proponer un modelo de riesgo crediticio basado en un modelo *Logit* para estimar la probabilidad de *default* de los clientes, por lo cual se aplicará un modelo econométrico comenzando por la identificación de las variables que puedan afectar al resultado de cumplimiento o incumplimiento del crédito, para esto se sigue la siguiente estructura metodológica presentada a continuación.

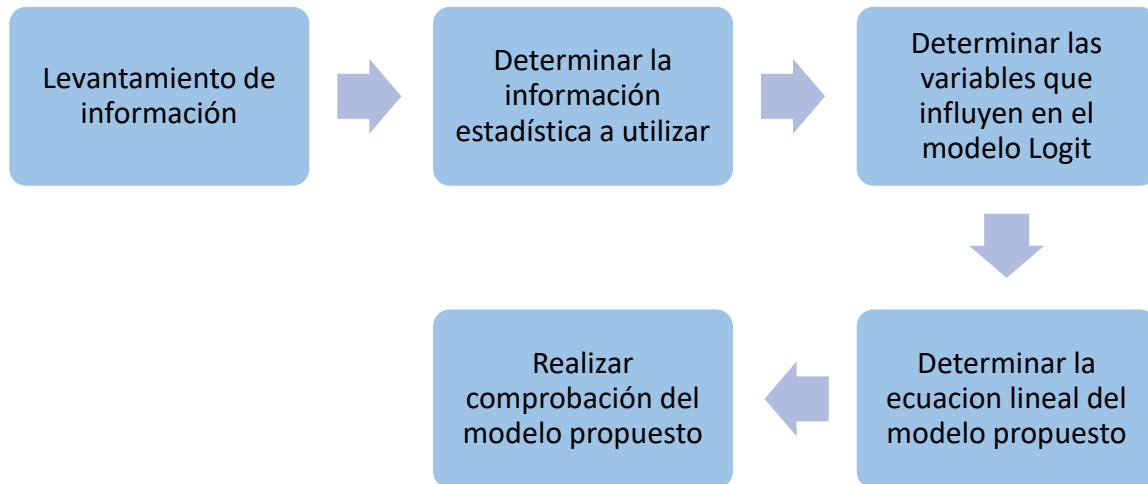
5.1. Estructura metodológica

En cuanto a la estructura metodológica, esta se ciñe a lo propuesto por Hernández, Fernández y Baptista, en el libro “Metodología de la investigación” (2014), por lo que se acota un estudio del tipo mixto, descriptivo y con objetivo de contextualización y correlación de las variables involucradas en el proceso de la investigación. Los métodos de recolección de datos son principalmente digitales, en base a la información y datos disponibles en los servidores de Tec Tec S.A.

Una vez recopilado los datos necesarios se elabora un filtro estadístico con el fin de establecer el modelo econométrico, determinando las variables que influyen en el modelo *Logit*.

A continuación, se realiza el paso a paso, lo describe la siguiente figura:

Figura 1. Estructura metodológica



5.2. Levantamiento de información

El levantamiento de información se realiza desde fuentes primarias y secundarias, generadas con el fin de proponer y/o determinar el modelo de riesgo crediticio a basado en un modelo *Logit*.

Para este fin, se ha determinado una entidad financiera, la cual se refiere a una empresa de servicios financieros, encargada de entregar tarjetas de créditos a sus clientes. Su nombre ficticio es TEC TEC S.A.

5.3. Determinar información estadística a utilizar

A través de la recopilación de información histórica de los clientes, se analizan las distintas variables que arrojan dicha base datos, considerando la población, el comportamiento, la variable sexo del cliente, el estrato socio económico, la tasa de interés cobrada por el crédito,

situación del estado de cuenta y del crédito. Dicha información será traspasada a distintos archivos .txt con el fin de elaborar una base de datos para ser analizada en los software estadísticos RStudio con *data table* para agilizar los tiempos de cálculos y SPSS modeler.

5.4. Determinar las variables que influyen en el modelo a proponer

Se utiliza la información recopilada anteriormente, por el cual se analizará las distintas variables, y se determinaran cuáles son fundamentales para el modelo a proponer. En principio se tienen 43 variables las cuales se categorizaron de manera general como sigue:

Categoría	Variable
Variables financieras	Producto asociado a cada cuenta
	Variable asociada al cupo asignado a la cuenta
	Estado financiero actual de la cuenta en el sistema de Intellect Card
	Clasificación de cartera: Nuevo, avance o refinanciamiento
	Saldo
	Relación entre deuda y cupo
	Proceso de cierre mensual y el calculo de la suma de los montos pagados en los últimos 3 meses
	Proceso de cierre mensual y el calculo de la suma de los montos pagados en los últimos 6 meses
	Total de pagos registrados efectivamente en un horizonte de 12 meses
	Suma de los montos facturados
	Suma de los montos facturados al cierre
	Montos facturados en los ultimos 12 meses
	Resumen de transacciones
	Relación cuota / pagos
	Máximo valor que ha adoptado el cupo del cliente
	Relacion cuota / mayor cupo

Categoría	Variable
Cliente	Tipo de cliente: Normal, Consumo, Renegociado y Transaccional Antigüedad Comportamiento del cliente dentro del último año Evaluación temporal de 6 meses Número de documentos Monto de documentos

Categoría	Variable
Mora	Días de mora Días de mora proyectados Tramo de mora nuevos Mora máxima en los últimos 3 meses Máximo periodo de mora que sostuvo el cliente Meses con pagos en un periodo de tres meses Meses con pagos en los últimos 12 meses Porcentaje de montos pagados sobre los totales facturados en los últimos 3 meses Porcentaje de monto asociado a pagos respecto el total facturado en los últimos 6 meses Porcentaje de monto asociado a pagos respecto el total facturado en los últimos 12 meses Necesidad de reestructuración a contar de la primera ocurrencia del evento Mora de origen de la renegociación Segmento de origen de la renegociación Número de cuotas de la renegociación Número de cuotas de la renegociación Porcentaje pagado del convenio Antigüedad de la renegociación Monto de documentos

Categoría	Variable
Crédito	Tipo de apertura Score de apertura Score de behavior

5.5. Determinar ecuación lineal

Se determina en base a los resultados del modelo, que las variables financieras de tasa de interés, rango de morosidad y rango de monto son significativas para el modelo logístico propuesto como herramienta útil para evaluar su incidencia en el default de los créditos de los clientes.

5.6. Comprobación del modelo propuesto.

Se determinan todas las variables que son significativas para el modelo logístico propuesto como herramienta útil para evaluar su incidencia en el default de los clientes, explicando la importancia de cada variable. La comprobación se realizará considerando los valores reales de la base de datos original y los valores con la base de datos de prueba. Es importante resaltar que la base de datos original se refiere a la base de datos de los clientes de la empresa TEC TEC S.A., y la base de datos de prueba se obtiene a través de un muestreo aleatorio tomado de la base de datos de la empresa.

Capítulo 6

6. Desarrollo del Modelo

Para comenzar el proceso de desarrollo del modelo, es necesario precisar los supuestos del caso acorde a la entidad financiera.

6.1. Supuestos Generales

Se determina que actualmente, dicha entidad sostiene problemas de cobranza con sus clientes, puesto que la valoración del riesgo crediticio no se encuentra estadísticamente diseñada para considerar escenarios extremos como los que han surgido con la pandemia a causa de Covid-19, por lo que se utilizan variables económicas bajo supuestos generales y sin considerar los casos extremos, lo que se ha de incorporar en el nuevo modelo al aplicar un enfoque *forward looking*. Este consiste en describir una organización con visión de futuro, y con ideas modernas.

Por otra parte, la entidad financiera tiene divisiones estructurales y formales para la revisión de la previsión del riesgo que supone cada cliente, llamada “Gerencia de Riesgo” y cuya principal función, es establecer el cálculo de provisiones en conformidad con las normativas. Actualmente, las normativas que sustentan la gestión de la división corresponden a IFRS 9: Instrumentos financieros y las que determine la Comisión de Mercado Financiero, ex Superintendencia de Bancos e Instituciones Financieras.

Cabe señalar que para desarrollar en conformidad el modelo, es preciso tener una revisión de la normativa, especialmente de la IFRS 9, ya que esta proporciona los estándares y parámetros para el reconocimiento, medición, deterioro y desreconocimiento acorde a la

contabilidad de cobertura. Los principales elementos a incorporar en el desarrollo del modelo de riesgo acorde a la norma, equivalen a que se utiliza un modelo de deterioro basado en las pérdidas de crédito esperadas, aplicando a los instrumentos de deuda con costo amortizado o bien, al denominado “FVTOCI”, por sus siglas en ingles “*Fair Value Trought Other Comprehensive Income*”, es decir, el valor razonable a través de otros ingresos comprensivos; adicionalmente, se utilizan cuentas por cobrar de arriendos, activos bajo contrato y otros compromisos de préstamo suscritos y/o, contratos asociados a garantías financieras.

Otro elemento a revisar de la norma IFRS 9, es que la provisión por pérdida, sean estas de crédito esperadas a corto plazo o bien, a largo plazo, deben sostener exposiciones dentro de los requerimientos de deterioro para ser reconocidas, siendo la contabilidad de cobertura aquello que refleje la administración del riesgo con sus respectivos instrumentos y elementos que califiquen como óptimos y aptos acorde lo normativo, determinando que ya no se precisa de una valoración retrospectiva sino más bien prospectiva, principal motivo por el cual se incorpora el enfoque Forward Looking, basados en el principio de relación económica.

Finalmente, se determina que la modificación requerida por la Gerencia de Riesgo tiene por objetivo la modificación de la estructura lógica para el cálculo de provisiones, utilizando un modelo de regresión logística.

6.2. Definición de Categorías

Acorde a las principales características de los clientes de la entidad financiera ficticia ya definida y lo normativo, se identifican cuatro segmentos de clientes:

1. Normal
2. Consumo
3. Renegociado
4. Transaccional

A su vez, cada una de estos segmentos tiene tres subsegmentos dependientes:

1. Nuevo
2. Avance
3. Refinanciamiento

Finalmente, cabe comentar que cada cálculo es realizado de manera independiente para cada segmento y subsegmento, por lo que es prioritario que la pertinencia de cada categorización se mantenga a lo largo del horizonte temporal que la condición exista. Como método de ejemplificar esta situación, se tiene el caso de un cliente del segmento normal que se encuentra solicitando un crédito de consumo, cuyo subsegmento correcto sería el de “avance”, se debe mantener en el subsegmento mientras su transacción esté vigente, para posteriormente, cuando la transacción finalice y se concrete, sea derivado al segmento consumo en subsegmento nuevo.

En concreto, los clientes tienen factibilidad en transitar entre segmentos y subsegmentos. No obstante, deben mantenerse en su categorización de origen hasta que se concreten los cambios.

6.3. Antecedentes Generales

El *ratio* de pérdida en caso de incumplimiento, en adelante, LGD, es entregada por segmento y subsegmento, además, la exposición en el momento de incumplimiento (EAD) es calculada por cada uno de los clientes. Ambas pueden ser diferentes según los parámetros establecidos por las exigencias normativas de la CMF y la IFRS 9, por lo que un mismo cliente podría, eventualmente, llegar a tener dos LGD y dos EAD.

Por otra parte, el enfoque *forward looking* se utiliza sobre las variables macroeconómicas, a fin de aportar valor en la identificación y definición de escenarios, determinando tres de ellos: Pesimista, Neutro y Optimista. De manera complementaria, cada escenario tiene un factor

multiplicativo asociado que, a su vez, se define de manera distinta para la probabilidad de incumplimiento (PD) y para LGD, según como los factores macroeconómicos de cada escenario impacten en los resultados de riesgo, lo que precisa de definir e incorporar unas cuantas nuevas variables al modelo, a fin de integrar los componentes de los respectivos escenarios.

Para cualquier compromiso financiero, entendiendo éstas como deudas, existe la opción de realizar un pago adelantado de un monto parcial o total, lo que es definido como la acción de “prepagar”. El prepago es un factor que debe ser incorporado al modelo de riesgo, ya que permite asignar probabilidades de que el cliente baje la exposición antes de los periodos formales establecidos y pactados, siendo este añadido en la PD con un nivel de detalle por segmento y subsegmento de cliente. Este se ciñe a la siguiente ecuación:

$$PD_f = PD_m \times f_{pp}, \quad (6.1)$$

donde:

PD_f = Probabilidad de incumplimiento ajustada por prepago, equivalente a PD.

PD_m = Probabilidad de incumplimiento orgánica acorde las variables macroeconómicas.

f_{pp} = Factor de prepago

A su vez, el factor de prepago se obtiene desde cada uno de los subsegmentos, manifestando como ejemplo el caso del factor de prepago de consumo, a continuación:

$$FP_{Consumo} = FP_{Nuevo} + FP_{Avance} + FP_{Refinanciado}. \quad (6.2)$$

Esto, se replica para los segmentos de Renegociado, Normal y Transaccional.

Adicionalmente y en línea con la existencia de escenarios de menor probabilidad, se incluye la condición de “cambio significativo en el riesgo” y de “bloqueo”, ambas diseñadas con el fin de aplicar un factor “*life time*” sobre la PD.

Acorde la ecuación 4.3, que da origen al calculo de la previsión ajustada por factor de prepago, se precisa definir los siguientes elementos:

$$PD_{Lt} = PD_f \times f_{Lt\ pp} , \quad (6.3)$$

siendo:

PD_{Lt} = Probabilidad de incumplimiento ajustada por Life time.

PD_f = Probabilidad de incumplimiento ajustada por factor de prepago.

$f_{Lt\ pp}$ = Factor por Life time para factor de prepago.

Este último componente se calcula a partir de:

$$f_{Lt\ pp} = P \times C_{Lt} , \quad (6.4)$$

cuyos elementos se definen como:

P = Plazo promedio ponderado por saldo

C_{Lt} = Cambio significativo en el riesgo.

Acorde al cálculo de EAD, este se obtiene a partir de la suma del riesgo que se asume como dispuesto de la operación y un porcentaje determinado al riesgo que no se está dispuesto, es decir, porcentaje sobre el saldo no dispuesto que se espera sea utilizado previo el evento de incumplimiento, por lo que la estimación del EAD se sustenta en el factor de conversión que se utiliza para definir la porción de riesgo asumida. En cualquier caso, se evalúa también la posibilidad de incorporar un porcentaje adicional asociado al límite máximo para

operaciones que sobrepasan la fecha de referencia proporcionada, siendo esta definida en las políticas de cada producto de la entidad.

La ecuación que define la EAD responde a la siguiente:

$$EAD_s = U + D \times F_s, \quad (6.5)$$

definiendo:

EAD_s = Exposición en el momento de incumplimiento otorgada por el modelo.

U = Riesgo dispuesto, equivalente al saldo utilizado.

D = Riesgo no dispuesto.

F_s = Factor de conversión de riesgo no dispuesto.

La IFRS 9 define de igual manera la EAD, visualizando el cambio en el subíndice en la ecuación 12:

$$EAD_9 = U + D \times F_9. \quad (6.6)$$

En aquellos clientes que cumplan con la condición de cambio significativo en el riesgo, se determina un factor que depende del periodo máximo de las transacciones del cliente y precisa de definir cuando la EAD será representada únicamente por el saldo utilizado en caso de bloqueo. Esto se manifiesta en la siguiente ecuación:

$$EAD_s = EAD_9 = U. \quad (6.7)$$

Esta condición se da ya que la cuenta se encuentra bloqueada, es decir, se estar presente la variable “C_b”, se cumple dicha ecuación.

Adicionalmente, para el cálculo de las provisiones, se utiliza la tasa ponderada de colocaciones mensual a fin de traer a valor presente la provisión, obteniéndose dos resultados, una acorde al modelo propuesto y otra, acorde a la IFRS 9.

La ecuación relacionada a lo descrito en el párrafo anterior, equivale a:

$$Provisión_s = PD_f \times EAD_s \times LGD , \quad (6.8)$$

cuyos componentes son:

$Provisión_s$ = Valor de la provisión bajo el modelo.

PD_f = Probabilidad de incumplimiento ajustada por factor de prepago.

EAD_s = Exposición al momento de incumplimiento calculada bajo el modelo.

$LG D$ = Ratio de pérdida en caso de incumplimiento.

Para el método de cálculo acorde a la IFRS 9, se tiene la siguiente fórmula:

$$Provisión_9 = \sum_{i=1}^n \frac{PD_f \times EAD_9 \times LGD}{(1 + t_p)^p} . \quad (6.9)$$

En caso de tener la condición de tener un cambio significativo del riesgo, es decir, estando presente C_{Lt} , se cumple que:

$$Provisión_9 = \sum_{p=1}^n \frac{PD_{Lt} \times EAD_9 \times LGD}{(1 + t_p)^p} . \quad (6.10)$$

Siendo la principal modificación, que la probabilidad de incumplimiento ahora es acorde al *Life time* en lugar del ajustado por valor prepago.

Capítulo 7

7. Análisis de datos y resultado de modelos

A modo de resumen se exponen los supuestos más relevantes para el desarrollo de los modelos de PD:

- Periodo de observación de 55 meses más 12 meses desde el mes 56.
- Se definen segmentos y subsegmentos de interés el Negocio.
- Desempeño calculado en 12 meses de observación (life time para el caso de los modelos IFRS 9 que cumplen la condición de cambio significativo en el riesgo).
- Definición de Desempeño normativa, esto es alguna vez 90 o más días de mora o Renegociaciones realizadas con 60 o más días de mora.
- Se trabaja con RUT único.
- Se reserva un 20% de la Población para validación de los modelos desarrollados con la Población de entrenamiento.
- Se desarrolla por Regresión Logística.
- Se vigila que no existan correlaciones significativas (superiores al 55%) para no sobre ajustar los modelos.

- Se mide calidad de los modelos por test estadísticos más usados: KS, Estabilidad Poblacional; Chi_2, otros.
- Se trabajó con 35 variables que han sido elegidas (criterio experto) y certificadas (Gerencia de operaciones y sistemas) como buenos indicadores del desempeño de pago de los clientes en procesos anteriores de modelación para el proceso de provisiones, estas no fueron modificadas ya que aún presentan buen poder predictivo en el comportamiento futuro de los clientes.

7.3. Segmentación de Cartera

Si bien la cartera de colocaciones de Tec Tec S.A. se origina con el otorgamiento de una línea de crédito para uso en compras y avances, una utilización adicional del producto se da cuando se oferta al cliente una línea temporal adicional a la principal la cual permite realizar giros pagaderos en cuotas (en adelante, Súper Avance).

Considerando que los clientes que giran esta línea de Súper Avance tienen características propias y los plazos asociados a estas colocaciones son más largos, esta es una segmentación de cartera que se evaluará de forma independiente y a la que llamamos Consumo.

En complemento a lo anterior, existe una diferencia significativa entre las tasas de malos de cada segmento, por lo tanto, se tomó la decisión de estratificar los modelos en función de grupos homogéneos diferenciados por la tasa de malos que presentan en 12 meses de observaciones.

Tabla (7.1) de segmentación de cartera de Tec Tec S.A.

MODELO	TASA DE MALOS
CONSUMO	12,10%
NORMAL	18,30%
TRANSACCIONAL	4,10%
RENEGOCIADOS	64,40%

Consumo: Clientes que una parte o todo el saldo vigente corresponde a un giro de Súper Avance.

Renegociado: Clientes que tienen vigente una renegociación de deuda de origen moroso.

Transaccionales: Clientes al día al momento de la clasificación, no registran transacciones de avance en los últimos 90 días, no tienen transacciones de refinanciamiento o renegociación vigentes o vencidas y registran compras en los últimos 90 días por un monto mayor a \$100.000 con saldo actual menor al monto de compras.

Normal: Son los clientes que quedan sin clasificar en alguna de las 3 nominaciones anteriores.

7.4. Desarrollo de los modelos estadísticos

Datos para el desarrollo de los modelos de PD

Para el desarrollo de estos modelos se ha trabajado sobre las carteras de Enero de 2012 hasta Julio de 2016, 55 meses de historia. Los resultados obtenidos son estadísticamente satisfactorios y se presentarán mas adelante.

7.5. Definición de desempeño y horizonte de observación para los Modelos de PD

Para clasificar a un cliente como incumplimiento se busca si durante los 12 meses posteriores a la fecha de observación presenta al menos un cierre de mes con morosidad mayor a 89 días o bien realizó una renegociación con 60 o más días de mora. En ambos casos, estos son los clientes “Deteriorados”.

7.6. Construcción de bases de muestreo.

Para la creación de las bases a RUT único el procedimiento fue el siguiente:

1. Se identifica la cantidad de clientes malos y buenos de tal forma que al final del proceso las tasas de malos y la proporción de casos con respecto al total de clientes del muestreo coincida con la base total.
2. Se toma el primer periodo (201201) y se selecciona una muestra aleatoria utilizando un muestreo aleatorio simple, de tal forma que tenga la proporción de buenos y malos obtenidos en el paso 1.
3. Se toma el siguiente periodo (201202) y procedemos de igual forma que en (2), teniendo en cuenta que debemos excluir los clientes que ya tomamos en el paso anterior.
4. Se procede respetivamente para todos los periodos. En el caso de muestrear el modelo, por ejemplo, consumo y tener ya la muestra del modelo normal, todos los pasos anteriores se deben realizar excluyendo los clientes que ya tomamos de la base del modelo normal.

Para el procedimiento se utilizó un muestreo aleatorio y se guardó cada base.

Se detallan a continuación tasas de malos y pesos para las bases completas y muestras de cada modelo.

Tabla (7.2) Modelo de segmento Consumo

base completa					base de muestra			
Periodo	Malos	Casos	PD	Peso	Malos	Casos	PD	Peso
201201	19.399	105.883	18,3%	1,8%	358	1.956	18,3%	1,8%
201202	18.683	104.871	17,8%	1,8%	345	1.937	17,8%	1,8%
201203	17.535	102.833	17,1%	1,7%	324	1.899	17,1%	1,7%
201204	16.281	100.567	16,2%	1,7%	301	1.857	16,2%	1,7%
201205	14.895	97.146	15,3%	1,7%	275	1.794	15,3%	1,7%
201206	13.822	94.692	14,6%	1,6%	255	1.749	14,6%	1,6%
201207	13.143	94.094	14,0%	1,6%	243	1.738	14,0%	1,6%
201208	12.856	94.400	13,6%	1,6%	237	1.743	13,6%	1,6%
201209	12.459	94.231	13,2%	1,6%	230	1.740	13,2%	1,6%
201210	12.095	93.806	12,9%	1,6%	223	1.732	12,9%	1,6%
201211	11.989	95.009	12,6%	1,6%	221	1.755	12,6%	1,6%
201212	11.775	95.677	12,3%	1,6%	217	1.767	12,3%	1,6%
201301	11.491	95.974	12,0%	1,6%	212	1.773	12,0%	1,6%
201302	11.432	96.093	11,9%	1,6%	211	1.775	11,9%	1,6%
201303	11.405	95.918	11,9%	1,6%	211	1.772	11,9%	1,6%
201304	11.183	94.800	11,8%	1,6%	207	1.751	11,8%	1,6%
201305	11.203	94.218	11,9%	1,6%	207	1.740	11,9%	1,6%
201306	11.289	93.132	12,1%	1,6%	208	1.720	12,1%	1,6%
201307	11.290	95.296	11,8%	1,6%	209	1.760	11,9%	1,6%
201308	11.231	94.659	11,9%	1,6%	207	1.748	11,8%	1,6%
201309	11.499	94.352	12,2%	1,6%	212	1.743	12,2%	1,6%
201310	11.555	96.911	11,9%	1,6%	213	1.790	11,9%	1,6%
201311	11.651	97.423	12,0%	1,7%	215	1.799	12,0%	1,7%
201312	11.873	98.716	12,0%	1,7%	219	1.823	12,0%	1,7%
201401	12.047	98.603	12,2%	1,7%	222	1.821	12,2%	1,7%
201402	12.442	99.310	12,5%	1,7%	230	1.834	12,5%	1,7%
201403	12.561	99.813	12,6%	1,7%	232	1.843	12,6%	1,7%
201404	12.519	99.802	12,5%	1,7%	231	1.843	12,5%	1,7%
201405	12.423	99.000	12,5%	1,7%	229	1.828	12,5%	1,7%
201406	12.604	99.817	12,6%	1,7%	233	1.844	12,6%	1,7%
201407	12.856	102.426	12,6%	1,7%	237	1.892	12,5%	1,7%
201408	13.160	105.064	12,5%	1,8%	243	1.940	12,5%	1,8%
201409	13.327	106.555	12,5%	1,8%	246	1.968	12,5%	1,8%

201410	13.222	107.129	12,3%	1,8%	244	1.979	12,3%	1,8%
201411	13.033	108.377	12,0%	1,8%	241	2.002	12,0%	1,8%
201412	12.646	108.714	11,6%	1,8%	234	2.008	11,7%	1,8%
201501	12.339	109.430	11,3%	1,9%	228	2.021	11,3%	1,9%
201502	12.147	109.798	11,1%	1,9%	224	2.028	11,0%	1,9%
201503	11.926	110.292	10,8%	1,9%	220	2.037	10,8%	1,9%
201504	11.722	111.054	10,6%	1,9%	216	2.051	10,5%	1,9%
201505	11.460	111.583	10,3%	1,9%	212	2.061	10,3%	1,9%
201506	11.553	113.223	10,2%	1,9%	213	2.091	10,2%	2,0%
201507	11.649	115.008	10,1%	2,0%	215	2.124	10,1%	2,0%
201508	12.015	117.036	10,3%	2,0%	222	2.162	10,3%	2,0%
201509	12.208	118.548	10,3%	2,0%	225	2.189	10,3%	2,0%
201510	12.288	119.480	10,3%	2,1%	227	2.207	10,3%	2,1%
201511	12.665	122.915	10,3%	2,1%	234	2.270	10,3%	2,1%
201512	13.188	126.495	10,4%	2,2%	244	2.336	10,4%	2,2%
201601	13.551	128.996	10,5%	2,2%	250	2.382	10,5%	2,2%
201602	13.997	131.506	10,6%	2,3%	259	2.429	10,7%	2,3%
201603	14.554	133.426	10,9%	2,3%	269	2.464	10,9%	2,3%
201604	14.929	135.559	11,0%	2,3%	276	2.504	11,0%	2,3%
201605	15.514	137.035	11,3%	2,3%	287	2.531	11,3%	2,4%
201606	15.878	138.729	11,4%	2,4%	293	2.562	11,4%	2,4%
201607	16.181	139.294	11,6%	2,4%	299	2.573	11,6%	2,4%
TOTALES	714.638	5.884.718	12,2%	100,0%	13.195	108.685	12,2%	100,0%

Tabla (7.3) Modelo segmento Normal

Periodo	base completa				base de muestra			
	Malos	Casos	PD	Peso	Malos	Casos	PD	Peso
201201	84.809	526.810	16,1%	2,2%	1.569	9.745	16,1%	2,2%
201202	83.834	523.754	16,0%	2,2%	1.551	9.689	16,0%	2,2%
201203	81.579	522.868	15,6%	2,2%	1.509	9.672	15,6%	2,2%
201204	80.213	517.912	15,5%	2,1%	1.484	9.581	15,5%	2,1%
201205	78.072	383.649	20,3%	1,6%	1.444	7.097	20,3%	1,6%
201206	79.910	394.432	20,3%	1,6%	1.478	7.296	20,3%	1,6%
201207	78.966	395.534	20,0%	1,6%	1.461	7.317	20,0%	1,6%
201208	79.454	405.624	19,6%	1,7%	1.470	7.503	19,6%	1,7%
201209	78.144	402.376	19,4%	1,7%	1.445	7.443	19,4%	1,7%
201210	76.060	391.583	19,4%	1,6%	1.407	7.244	19,4%	1,6%
201211	75.068	388.693	19,3%	1,6%	1.389	7.190	19,3%	1,6%
201212	72.199	385.463	18,7%	1,6%	1.336	7.131	18,7%	1,6%
201301	69.391	377.461	18,4%	1,6%	1.284	6.982	18,4%	1,6%
201302	68.038	372.841	18,2%	1,5%	1.259	6.897	18,3%	1,5%
201303	67.259	371.622	18,1%	1,5%	1.244	6.874	18,1%	1,5%
201304	65.820	364.284	18,1%	1,5%	1.218	6.739	18,1%	1,5%
201305	63.253	359.250	17,6%	1,5%	1.170	6.646	17,6%	1,5%
201306	63.351	361.354	17,5%	1,5%	1.172	6.685	17,5%	1,5%
201307	62.411	359.153	17,4%	1,5%	1.155	6.644	17,4%	1,5%
201308	62.847	364.083	17,3%	1,5%	1.163	6.735	17,3%	1,5%
201309	63.209	366.263	17,3%	1,5%	1.169	6.775	17,3%	1,5%
201310	63.515	365.314	17,4%	1,5%	1.175	6.758	17,4%	1,5%
201311	64.478	369.361	17,5%	1,5%	1.193	6.833	17,5%	1,5%
201312	67.777	381.397	17,8%	1,6%	1.254	7.055	17,8%	1,6%
201401	70.494	387.744	18,2%	1,6%	1.304	7.173	18,2%	1,6%
201402	72.789	395.729	18,4%	1,6%	1.346	7.320	18,4%	1,6%
201403	75.368	408.446	18,5%	1,7%	1.394	7.556	18,4%	1,7%
201404	76.186	414.394	18,4%	1,7%	1.409	7.666	18,4%	1,7%
201405	76.252	419.779	18,2%	1,7%	1.410	7.765	18,2%	1,7%
201406	79.371	433.769	18,3%	1,8%	1.468	8.024	18,3%	1,8%
201407	80.122	437.870	18,3%	1,8%	1.482	8.100	18,3%	1,8%
201408	81.265	438.967	18,5%	1,8%	1.503	8.120	18,5%	1,8%
201409	82.227	444.169	18,5%	1,8%	1.521	8.216	18,5%	1,8%

201410	82.850	446.954	18,5%	1,8%	1.533	8.268	18,5%	1,8%
201411	84.727	458.322	18,5%	1,9%	1.567	8.478	18,5%	1,9%
201412	84.490	466.340	18,1%	1,9%	1.563	8.627	18,1%	1,9%
201501	84.539	464.241	18,2%	1,9%	1.564	8.588	18,2%	1,9%
201502	85.463	471.307	18,1%	1,9%	1.581	8.718	18,1%	1,9%
201503	86.388	481.392	17,9%	2,0%	1.598	8.905	17,9%	2,0%
201504	86.561	481.007	18,0%	2,0%	1.601	8.898	18,0%	2,0%
201505	85.799	481.719	17,8%	2,0%	1.587	8.911	17,8%	2,0%
201506	88.282	489.036	18,1%	2,0%	1.633	9.045	18,1%	2,0%
201507	89.816	493.140	18,2%	2,0%	1.661	9.122	18,2%	2,0%
201508	92.501	500.598	18,5%	2,1%	1.711	9.260	18,5%	2,1%
201509	93.798	506.873	18,5%	2,1%	1.735	9.376	18,5%	2,1%
201510	94.527	507.314	18,6%	2,1%	1.749	9.385	18,6%	2,1%
201511	95.674	516.424	18,5%	2,1%	1.770	9.553	18,5%	2,1%
201512	95.559	489.973	19,5%	2,0%	1.768	9.064	19,5%	2,0%
201601	95.331	492.968	19,3%	2,0%	1.763	9.119	19,3%	2,0%
201602	96.641	507.149	19,1%	2,1%	1.788	9.382	19,1%	2,1%
201603	98.061	518.955	18,9%	2,1%	1.814	9.600	18,9%	2,1%
201604	98.322	517.017	19,0%	2,1%	1.819	9.564	19,0%	2,1%
201605	101.180	515.535	19,6%	2,1%	1.872	9.537	19,6%	2,1%
201606	103.788	518.957	20,0%	2,1%	1.920	9.600	20,0%	2,1%
201607	104.906	517.009	20,3%	2,1%	1.941	9.564	20,3%	2,1%
TOTALES	4.452.934	24.274.178	18,4%	100,0%	82.374	449.035	18,4%	100,0%

Tabla (7.4) Modelo segmento Transaccional

base completa					base de muestra			
Periodo	Malos	Casos	PD	Peso	Malos	Casos	PD	Peso
201201	3.474	98.959	3,5%	1,7%	64	1.821	3,5%	1,7%
201202	3.284	99.026	3,3%	1,7%	60	1.822	3,3%	1,7%
201203	2.912	91.388	3,2%	1,6%	54	1.681	3,2%	1,6%
201204	3.268	96.576	3,4%	1,7%	60	1.777	3,4%	1,7%
201205	3.593	96.463	3,7%	1,7%	66	1.775	3,7%	1,7%
201206	3.433	96.272	3,6%	1,7%	63	1.771	3,6%	1,7%
201207	3.452	94.283	3,7%	1,6%	64	1.735	3,7%	1,6%
201208	3.179	93.741	3,4%	1,6%	58	1.725	3,4%	1,6%
201209	3.248	95.677	3,4%	1,7%	60	1.760	3,4%	1,7%
201210	3.404	96.583	3,5%	1,7%	63	1.777	3,5%	1,7%
201211	3.500	96.997	3,6%	1,7%	64	1.784	3,6%	1,7%
201212	4.144	103.276	4,0%	1,8%	76	1.900	4,0%	1,8%
201301	3.782	101.282	3,7%	1,8%	70	1.863	3,8%	1,8%
201302	3.814	102.256	3,7%	1,8%	70	1.881	3,7%	1,8%
201303	3.191	93.117	3,4%	1,6%	59	1.713	3,4%	1,6%
201304	3.449	96.940	3,6%	1,7%	63	1.783	3,5%	1,7%
201305	3.554	96.937	3,7%	1,7%	65	1.783	3,6%	1,7%
201306	3.335	95.941	3,5%	1,7%	61	1.765	3,5%	1,7%
201307	3.257	95.432	3,4%	1,7%	60	1.756	3,4%	1,7%
201308	3.121	94.429	3,3%	1,6%	57	1.737	3,3%	1,6%
201309	2.982	94.824	3,1%	1,6%	55	1.744	3,2%	1,6%
201310	3.001	94.669	3,2%	1,6%	55	1.742	3,2%	1,6%
201311	2.868	93.573	3,1%	1,6%	53	1.721	3,1%	1,6%
201312	3.221	97.845	3,3%	1,7%	59	1.800	3,3%	1,7%
201401	3.017	95.374	3,2%	1,7%	56	1.755	3,2%	1,7%
201402	2.907	95.536	3,0%	1,7%	53	1.758	3,0%	1,7%
201403	2.671	88.796	3,0%	1,5%	49	1.634	3,0%	1,5%
201404	2.859	92.212	3,1%	1,6%	53	1.696	3,1%	1,6%
201405	3.267	94.323	3,5%	1,6%	60	1.735	3,5%	1,6%
201406	3.622	97.296	3,7%	1,7%	67	1.790	3,7%	1,7%
201407	4.001	99.678	4,0%	1,7%	74	1.834	4,0%	1,7%
201408	4.269	100.447	4,3%	1,7%	79	1.848	4,3%	1,7%
201409	4.232	101.538	4,2%	1,8%	78	1.868	4,2%	1,8%

201410	4.458	104.236	4,3%	1,8%	82	1.918	4,3%	1,8%
201411	4.462	105.455	4,2%	1,8%	82	1.940	4,2%	1,8%
201412	5.114	112.837	4,5%	2,0%	94	2.076	4,5%	2,0%
201501	4.668	112.677	4,1%	2,0%	86	2.073	4,1%	2,0%
201502	4.658	111.303	4,2%	1,9%	86	2.048	4,2%	1,9%
201503	4.189	103.249	4,1%	1,8%	77	1.899	4,1%	1,8%
201504	4.639	109.829	4,2%	1,9%	85	2.021	4,2%	1,9%
201505	5.183	114.023	4,5%	2,0%	95	2.098	4,5%	2,0%
201506	5.207	117.914	4,4%	2,1%	96	2.169	4,4%	2,1%
201507	5.480	117.744	4,7%	2,0%	101	2.166	4,7%	2,0%
201508	5.586	116.802	4,8%	2,0%	103	2.149	4,8%	2,0%
201509	5.558	115.424	4,8%	2,0%	102	2.123	4,8%	2,0%
201510	5.760	117.815	4,9%	2,0%	106	2.167	4,9%	2,0%
201511	5.986	124.035	4,8%	2,2%	110	2.282	4,8%	2,2%
201512	6.982	133.561	5,2%	2,3%	128	2.457	5,2%	2,3%
201601	6.646	132.891	5,0%	2,3%	122	2.445	5,0%	2,3%
201602	6.326	121.557	5,2%	2,1%	116	2.236	5,2%	2,1%
201603	5.816	110.425	5,3%	1,9%	107	2.031	5,3%	1,9%
201604	6.550	118.079	5,5%	2,1%	120	2.172	5,5%	2,1%
201605	6.931	119.844	5,8%	2,1%	128	2.205	5,8%	2,1%
201606	7.087	121.593	5,8%	2,1%	130	2.237	5,8%	2,1%
201607	7.870	125.598	6,3%	2,2%	145	2.311	6,3%	2,2%
TOTALES	236.467	5.748.577	4,0%	100,0%	4.349	105.757	4,0%	100,0%

Tabla (7.5) Modelo segmento Renegociado

base completa					base de muestra			
Periodo	Malos	Casos	PD	Peso	Malos	Casos	PD	Peso
201201	1.965	3.497	56,2%	1,7%	40	71	56,3%	1,7%
201202	1.907	3.395	56,2%	1,6%	39	69	56,5%	1,6%
201203	1.891	3.333	56,7%	1,6%	39	68	57,4%	1,6%
201204	1.910	3.369	56,7%	1,6%	39	69	56,5%	1,6%
201205	1.881	3.222	58,4%	1,5%	39	66	59,1%	1,5%
201206	1.852	3.217	57,6%	1,5%	38	66	57,6%	1,5%
201207	1.771	2.980	59,4%	1,4%	36	61	59,0%	1,4%
201208	1.739	2.945	59,0%	1,4%	35	60	58,3%	1,4%
201209	1.649	2.789	59,1%	1,3%	34	57	59,6%	1,3%
201210	1.574	2.654	59,3%	1,3%	32	54	59,3%	1,3%
201211	1.521	2.548	59,7%	1,2%	31	52	59,6%	1,2%
201212	1.454	2.414	60,2%	1,2%	30	49	61,2%	1,1%
201301	1.429	2.413	59,2%	1,2%	29	49	59,2%	1,1%
201302	1.465	2.526	58,0%	1,2%	30	52	57,7%	1,2%
201303	1.554	2.652	58,6%	1,3%	32	54	59,3%	1,3%
201304	1.727	2.901	59,5%	1,4%	35	59	59,3%	1,4%
201305	1.698	2.849	59,6%	1,4%	35	58	60,3%	1,4%
201306	1.737	2.904	59,8%	1,4%	35	59	59,3%	1,4%
201307	1.787	2.923	61,1%	1,4%	37	60	61,7%	1,4%
201308	1.822	2.925	62,3%	1,4%	37	60	61,7%	1,4%
201309	1.850	2.910	63,6%	1,4%	38	59	64,4%	1,4%
201310	1.918	2.915	65,8%	1,4%	39	59	66,1%	1,4%
201311	1.991	2.997	66,4%	1,4%	41	61	67,2%	1,4%
201312	1.905	2.854	66,7%	1,4%	39	58	67,2%	1,4%
201401	1.861	2.823	65,9%	1,3%	38	58	65,5%	1,4%
201402	1.842	2.844	64,8%	1,4%	38	58	65,5%	1,4%
201403	1.844	2.927	63,0%	1,4%	38	60	63,3%	1,4%
201404	2.085	3.329	62,6%	1,6%	43	68	63,2%	1,6%
201405	2.054	3.253	63,1%	1,6%	42	66	63,6%	1,5%
201406	2.277	3.532	64,5%	1,7%	46	72	63,9%	1,7%
201407	2.298	3.494	65,8%	1,7%	47	71	66,2%	1,7%
201408	2.348	3.500	67,1%	1,7%	48	71	67,6%	1,7%
201409	2.376	3.470	68,5%	1,7%	49	71	69,0%	1,7%

201410	2.356	3.387	69,6%	1,6%	48	69	69,6%	1,6%
201411	2.308	3.365	68,6%	1,6%	47	69	68,1%	1,6%
201412	2.267	3.332	68,0%	1,6%	46	68	67,6%	1,6%
201501	2.247	3.389	66,3%	1,6%	46	69	66,7%	1,6%
201502	2.267	3.484	65,1%	1,7%	46	71	64,8%	1,7%
201503	2.383	3.732	63,9%	1,8%	49	76	64,5%	1,8%
201504	2.605	4.147	62,8%	2,0%	53	85	62,4%	2,0%
201505	2.674	4.238	63,1%	2,0%	54	86	62,8%	2,0%
201506	2.882	4.469	64,5%	2,1%	59	91	64,8%	2,1%
201507	2.959	4.458	66,4%	2,1%	60	91	65,9%	2,1%
201508	3.012	4.503	66,9%	2,2%	62	92	67,4%	2,2%
201509	3.017	4.564	66,1%	2,2%	61	93	65,6%	2,2%
201510	2.969	4.491	66,1%	2,1%	61	92	66,3%	2,2%
201511	3.397	5.191	65,4%	2,5%	69	106	65,1%	2,5%
201512	3.553	5.406	65,7%	2,6%	72	110	65,5%	2,6%
201601	3.858	5.913	65,2%	2,8%	79	121	65,3%	2,8%
201602	4.195	6.297	66,6%	3,0%	86	129	66,7%	3,0%
201603	4.562	6.739	67,7%	3,2%	93	138	67,4%	3,2%
201604	4.844	7.047	68,7%	3,4%	99	144	68,8%	3,4%
201605	4.966	7.074	70,2%	3,4%	101	144	70,1%	3,4%
201606	5.197	7.291	71,3%	3,5%	106	149	71,1%	3,5%
201607	5.242	7.301	71,8%	3,5%	107	149	71,8%	3,5%
TOTALES	134.742	209.122	64,4%	100,0%	2.752	4.267	63,6%	100,0%

7.7. Construcción y análisis de variables

El proceso de modelación de la PD (probabilidad de default) se realizó analizando un total de 35 variables, de las cuales 28 variables fueron categorizadas y a cada categoría se le calculó un WOE (*weight of evidence*) específico, las 7 variables no categorizadas se dejaron fuera dado que no tenían un aporte discriminatorio en ninguno de los 4 modelos. A continuación, se detallan tanto las variables seleccionadas para cada modelo como las que no lo fueron, mostrando además los "*information value*" (IV), indicador que permite saber cuánta información entregará una variable a la predicción del comportamiento del cliente, entre más alto IV tenga la variable, mayor es su aporte discriminatorio. Todo este proceso se realiza sobre las bases a RUT único (muestra de entrenamiento: 80% de la base).

El proceso de selección de variables se realiza observando las correlaciones entre ellas, así como también los *information value*.

Si dos variables correlacionan en más de un 0,55 o menos de un -0.55, se descarta aquella que posea el menor *information value*.

Tabla (7.6) lista de variables seleccionadas

Variable	Descripción	IV Consumo	IV Normal	IV Transaccional	IV Renegociado
AMAX3	Máximo días de mora en 3 meses	1,08			
PSOBREF3	Porcentaje de pago sobre lo facturado en 3 meses	0,76			
NPAGOS3	Número de pagos en 3 meses	0,64			0,60
MONINFCOM	Monto de deuda en infocom	0,59	0,49	0,30	
ANTCTA	Antigüedad de la cuenta	0,29	0,34	0,76	0,24
DIAS_MORA	Días de mora		1,03		0,84
NPAGOS6	Número de pagos en 6 meses		0,34		
PSOBREF6	Porcentaje de pago sobre lo facturado en 6 meses		0,48		
CTACUP	Cupo		0,36		
PAGOS12	Monto pagado en 12 meses			0,62	
AMAX12	Máximo días de mora en 12 meses			0,37	
PORCPAGCONV	Porcentaje pagado del convenio				0,04
SALDO	Saldo				0,04
RECUOMYCUPO	División entre cuota mayor cupo histórico				0,11
NUMIFOCOM	Número de informantes en infocom				0,16
PAGOS6	Monto pagado en 6 meses				0,22

Tabla (7.7) lista de variables no seleccionadas

Variable	Descripción	IV Consumo	IV Normal	IV Transaccional	IV Renegociado
UTILIZACION	Porcentaje utilizado de la tarjeta	0,33	0,84	1,27	
TRAMOMORANVO	Tramo de mora	1,07	1,03		0,84
AMAX6	Máximo días de mora en 6 meses	1,02	0,96	0,32	0,14
NPAGOS12	Número de pagos en 12 meses	0,43	0,27		0,30
PAGOS3	Monto pagado en 3 meses	0,32	0,13	0,32	0,047
FACTURADO6	Monto facturado en 6 meses		0,07	0,16	
FACTURADO12	Monto facturado en 12 meses		0,03	0,25	
PSOBREF12	Porcentaje de pago sobre facturado en 12 meses	0,42	0,47		0,13
TRAMO_MORA	Tramo de mora	1,07	1,02		
ANTRENE	Antigüedad de renegociación				0,11
FACTURADO3	Monto facturado en 3 meses	0,06	0,09	0,14	0,06
MAYORCUPO	Máximo cupo histórico	0,20	0,27	1,09	0,03

7.8. Matrices de correlación

Tabla (7.8) Matriz de correlación modelo Consumo

MODELO CONSUMO	AMAX3	PSOBREF3	NPAGOS3	MONINFCOM	ANTCTA
AMAX3	1	0,450	0,509	0,544	0,066
PSOBREF3	0,450	1	0,435	0,288	0,110
NPAGOS3	0,509	0,435	1	0,278	0,093
MONINFCOM	0,544	0,288	0,278	1	0,048
ANTCTA	0,066	0,110	0,093	0,048	1

Tabla (7.9) Matriz de correlación modelo Transaccional

MODELO TRANSACCIONAL	ANTCTA	PAGOS12	AMAX12	NUMIFOCOM
ANTCTA	1	0,383	0,133	0,105
PAGOS12	0,383	1	0,182	0,114
AMAX12	0,133	0,182	1	0,171
NUMIFOCOM	0,105	0,114	0,171	1

Tabla (7.10) Matriz de correlación modelo Renegociado

MODELO RENEGOCIADO	DIAS_MORA	NPAGOS3	NUMIFOCOM	NPAGOS6	RELCUOMYCUPO	ANTCTA	SALDO	PORCPAGCONV
DIAS_MORA	1	0,542	0,256	0,293	0,052	0,211	0,068	-0,047
NPAGOS3	0,542	1	0,235	0,480	0,058	0,182	0,003	0,171
NOMIFOCOM	0,256	0,235	1	0,094	-0,002	0,010	0,041	0,107
NPAGOS6	0,293	0,480	0,094	1	0,002	0,285	-0,243	0,066
RELCUOMYCUPO	0,052	0,058	-0,002	0,002	1	0,245	0,164	0,015
ANTCTA	0,211	0,182	0,010	0,285	0,245	1	-0,019	-0,052
SALDO	0,068	0,003	0,041	-0,243	0,164	-0,019	1	0,186
PORCPAGCONV	-0,047	0,171	0,107	0,066	0,015	-0,052	0,186	1

Tabla (7.11) Matriz de correlación modelo Normal

MODELO NORMAL	DIAS_MORA	NOMIFOCOM	ANTCTA	NPAGOS6	PSOBREF6	CTACUP
DIAS_MORA	1	0,419	0,085	0,22	0,311	0,174
NOMIFOCOM	0,419	1	0,022	0,096	0,2	0,214
ANTCTA	0,085	0,022	1	0,243	0,137	0,316
NPAGOS6	0,22	0,096	0,243	1	0,326	0,201
PSOBREF6	0,311	0,2	0,137	0,326	1	0,105
CTACUP	0,174	0,214	0,316	0,201	0,105	1

Teniendo ya las variables, se procede a desarrollar los modelos con la metodología estadística de Regresiones Logísticas

7.9. Regresiones Logísticas

Como buena práctica, se ha separado las poblaciones en un 80% para el entrenamiento y un 20% para la validación. Esta validación consiste en usar los modelos desarrollados a partir del entrenamiento, y ejecutar sobre la muestra de validación. Si los modelos están desarrollados acorde a las mejores prácticas, se espera resultados similares en las capacidades de Discriminación y que no se produzcan sobreajustes.

El proceso multivariado de selección de variables se realizó por el método Step Wise.

Se presentan a continuación las ecuaciones de las regresiones (betas) para cada segmento

Tabla (7.12) Ecuación de las regresiones betas modelo Consumo

MODELO CONSUMO	Estimate	Estimate	Std. Error	z value	Pr(> z)	
(Intercept)	-1,947	-158,47	<2e-16	***	<2e-16	***
ANTCTAdtt_WOE	-0,987	-43,25	<2e-16	***	<2e-16	***
MONINFCOMdtt_WOE	-0,369	-20,91	<2e-16	***	<2e-16	***
PSOBREF3dtt_WOE	-0,488	-33,11	<2e-16	***	<2e-16	***
NPAGOS3dtt_WOE	-0,269	-15,15	<2e-16	***	<2e-16	***
AMAX3dtt_WOE	-0,587	-39,83	<2e-16		<2e-16	***

Null deviance: 63990 on 86651 degrees of freedom
 Residual deviance: 49358 on 86646 degrees of freedom
 AIC: 49370

Tabla (7.13) Ecuación de las regresiones betas modelo Transaccional

MODELO TRANSACCIONAL	Estimate	Std. Error	z value	Pr(> z)	
(Intercept)	-3,168	0,02062	-153,62	<2e-16	***
ANTCTAdtt_WOE	-0,761	0,02438	-31,22	<2e-16	***
MONINFCOMdtt_WOE	-0,682	0,02615	-26,06	<2e-16	***
AMAX12dtt_WOE	-0,690	0,02566	-26,88	<2e-16	***
PAGOS12dtt_WOE	-0,538	0,03396	-15,84	<2e-16	***

Null deviance: 28829 on 84690 degrees of freedom
 Residual deviance: 24476 on 84686 degrees of freedom
 AIC: 24486

Tabla (7.14) Ecuación de las regresiones betas modelo Renegociado

MODELO RENEGOCIADO	Estimate	Std. Error	z value	Pr(> z)	
(Intercept)	0,631	0,04358	14,472	< 2e-16	***
ANTCTAdtt_WOE	-0,916	0,08746	-10,474	< 2e-16	***
NUMIFOCOMdtt_WOE	-0,543	0,1032	-5,261	1,43E-07	***
PAGOS6dtt_WOE	-0,315	0,11012	-2,857	0,00428	**
RELCUOMYCUPOdtt_WOE	-0,593	0,13598	-4,358	1,31E-05	***
NPAGOS3dtt_WOE	-0,505	0,0689	-7,336	2,20E-13	***
DIASMORAdtt_WOE	-0,849	0,06603	-12,851	< 2e-16	***
SALDOdtt_WOE	-0,986	0,19594	-5,031	4,87E-07	***
PORCPAGCONVdtt_WOE	-0,882	0,23919	-3,689	0,00023	***

Null deviance: 4428,3 on 3402 degrees of freedom

Residual deviance: 3506,7 on 3394 degrees of freedom

AIC: 3524.7

Tabla (7.15) Ecuación de las regresiones betas modelo Normal

MODELO NORMAL	Estimate	Std. Error	z value	Pr(> z)	
(Intercept)	-1,486	0,005294	-280,61	<2e-16	***
ANTCTAdtt_WOE	-0,874	0,009745	-89,72	<2e-16	***
MONINFCOMdtt_WOE	-0,517	0,0074	-69,89	<2e-16	***
DIAS_MORAdtt_WOE	-0,715	0,005473	-130,62	<2e-16	***
NPAGOS6dtt_WOE	-0,294	0,009639	-30,47	<2e-16	***
PSOBREF6dtt_WOE	-0,489	0,007756	-63,02	<2e-16	***
CTACUPdtt_WOE	-0,548	0,009586	-57,17	<2e-16	***

Null deviance: 342272 on 358835 degrees of freedom

Residual deviance: 254616 on 358829 degrees of freedom

AIC: 254630

Ya desarrollados los modelos se procede a revisar su validez estadística, usando el estadístico de Kolmogorov-Smirnov (KS), indicador que mide qué tan bien los modelos separan a los casos buenos de los malos. Entre más cercano al 100% mejor será el modelo.

7.10. Capacidad de discriminación de los modelos de PD

Para el proceso de modelación se separó la base en 2 partes:

Parte 1: 80% de la base utilizada para ajustar los modelos.

Parte 2: 20% de la base utilizada únicamente para la validación.

En la siguiente tabla se muestra los desempeños de los modelos calculados con la base de desarrollo (1) y con la base de validación (2).

Tabla (7.16) Modelo desarrollado en periodo enero 2012 a julio 2016 (1)

Modelo desarrollado en periodo enero 2012 a julio 2016 (1)				
	MODELO			
	TRANSACCIONAL	NORMAL	CONSUMO	RENEGOCIADO
KS	46,9%	48,3%	46,0%	42,9%
GIINI	62,1%	65,0%	61,4%	57,9%
EXACTITUD	95,9%	81,6%	87,9%	64,1%

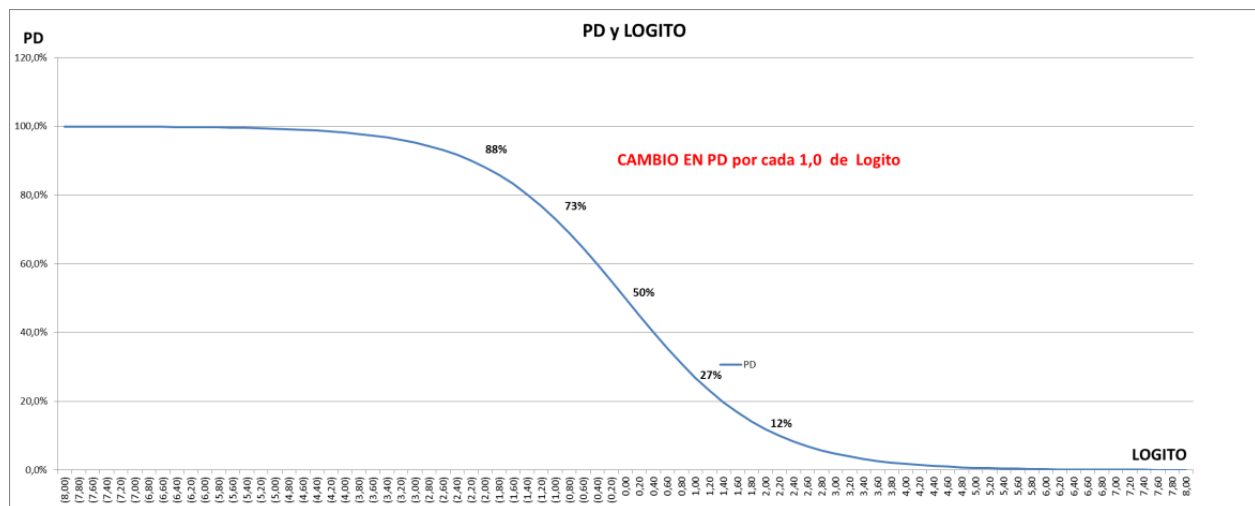
Tabla (7.17) Modelo desarrollado en periodo enero 2012 a julio 2016 (2)

Modelo desarrollado en periodo enero 2012 a julio 2016 (2)				
	MODELO			
	TRANSACCIONAL	NORMAL	CONSUMO	RENEGOCIADO
KS	47,7%	48,2%	44,5%	44,6%
GIINI	62,3%	64,9%	59,4%	58,3%
EXACTITUD	95,7%	81,7%	87,8%	64,2%

7.11. Relación PD vs Logit

A modo de entendimiento se presenta un gráfico que relaciona la PD de un modelo de regresión logística con su correspondiente Logit.

Gráfico (7.1) Probabilidad de default versus logito



7.12. Modelamiento macroeconómico

7.12.1. Análisis de variables macroeconómicas

Se utilizó la variable TASA DE PÉRDIDA MENSUAL (castigos del mes dividido en el outstanding del mes anterior) como la variable respuesta, para ser explicada por las variables macroeconómicas que se presentan a continuación y que representarán el estatus de riesgo mercado para el negocio.

- DESEMPLEO
- INDICE DE PRODUCCION INDUSTRIAL
- BCU 5
- BCP 5
- PRECIO DEL DÓLAR
- PRECIO DEL COBRE
- IMACEC
- IPC

7.12.2. Metodología

Selección de variables

Primero se calculó la correlación (1) entre la Tasa de Perdida Mensual y todas las variables desfasadas en 1, 3, 6, 9 y 12 meses.

Luego se dejaron las variables que tenían una correlación en valor absoluto mayor a 0,4, valor empírico que considera que las variables NO SON INDEPENDIENTES.

Estas fueron las siguientes:

- DESEMPLEO (desfasado en 6 meses)

- INDICE DE PRODUCCION INDUSTRIAL

Se analizó la existencia de causalidad entre la TASA DE PÉRDIDA y las 2 variables mencionadas y solamente el desempleo de 6 meses antes, cumplió el test.

Una vez se obtuvo la variable macroeconómica representativa para el análisis se determinó si enfrentábamos un escenario OPTIMISTA, PESIMISTA o NEUTRO, para esto se calcularon intervalos de confianza al 80%, 85%, 90% y 95% de la siguiente forma:

$$Ic: \bar{D} \pm z_{1-\frac{\alpha}{2}} \times \sqrt{VAR(D)}. \quad (7.1)$$

Al evaluar los escenarios para cada uno de los intervalos se eligió el intervalo de 85%, ya que esto nos permite tener más observaciones fuera del intervalo y poder medir la ponderación y el efecto que los escenarios tendrán sobre la PD.

Los límites resultantes al 85% de confianza fueron los siguientes:

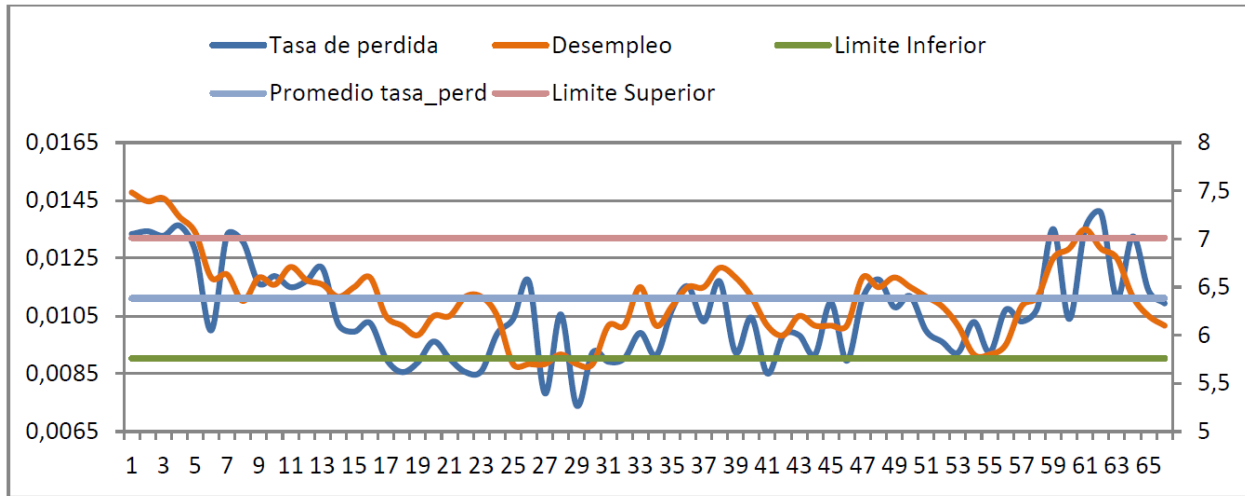
- Límite Inferior: 5,75
- Límite Superior: 7,01

Por lo tanto:

- ESCENARIO PESIMISTA: Tasa de desempleo mayor a 7,01
- ESCENARIO NEUTRO: Tasa de desempleo entre 5,75 y 7,01
- ESCENARIO OPTIMISTA: Tasa de desempleo menor a 5,75

Gráficamente queda de la siguiente forma:

Gráfico (7.2) de desempleo y tasa de perdida



La ponderación de los escenarios se realizó calculando la tasa de perdida promedio en los periodos en que cada escenario estaba presente:

Tabla (7.18) Tasa de perdida por escenario

ESCENARIOS	TASA PERDIDA PROMEDIO
PESIMISTA	1,33%
NEUTRO	1,06%
OPTIMISTA	0,93%

Luego se aplicó cada tasa sobre el saldo promedio de los últimos 5 años (\$361.962 MM).

Así la pérdida esperada para cada uno de los escenarios fue la siguiente:

Tabla (7.19) monto de perdidas por escenario

ESCENARIOS	MONTO DE PERDIDA PROMEDIO \$(MM)
PESIMISTA	\$4814,09
NEUTRO	\$3836,80
OPTIMISTA	\$3366,25

Luego se calculó la diferencia de los escenarios PESIMISTA (\$977,3MM) y OPTIMISTA (-\$470,55MM) con respecto al escenario NEUTRO y mediante un cociente se calcula el efecto sobre la provisión promedio de los últimos 5 años (\$28.651MM), resultando la variación porcentual de la provisión del negocio cuando se está en un escenario PESIMISTA u OPTIMISTA.

Tabla (7.20) Variación porcentual de la provisión según escenario

ESCENARIOS	R= DIFF/PROV_P	MULTIPLICADOR_PD (1+R)
PESIMISTA	0,0341	1,0341
OPTIMISTA	0,0164	0,9836

7.13. PD Lifetime

Para calcular la PD lifetime se utilizaron las bases a rut único utilizadas para modelar y se seleccionaron a todos los clientes que cumplieran la condición de “aumento significativo en el riesgo” (mayor a 30 y menor a 90 días de mora). Se hicieron 2 tablas cruzadas, agrupadas por Tramo de Mora (1), entre el segmento (grupos homogéneos) y los meses que tardo el cliente desde que fue observado hasta su default. El plazo máximo de lifetime son 67 meses, en los periodos posteriores los porcentajes se mantienen fijos.

Una vez teniendo el plazo máximo de cuotas pendientes, se debe separar el SALDO del cliente en saldos remanentes para cada mes desde el mes 13 hasta el final de la evaluación LT, las cuales se deben multiplicar por la PD de la “tabla de PD por periodo”. Esta tabla será de la siguiente forma:

Tabla (7.21) Tabla PD por periodo

TRAMO_MORA	SEGMENTO	A1	A2	A3	A4	A5	A6	A7	A8
3	CONSUMO	0,9642	0,9642	0,9642	0,9642	0,9642	0,9642	0,9642	0,9642
3	NORMAL	0,9260	0,9260	0,9260	0,9260	0,9260	0,9260	0,9260	0,9260
3	RENEGOCIADO	0,9667	0,9667	0,9667	0,9667	0,9667	0,9667	0,9667	0,9667
		A9	A10	A11	A12	A13	A14	A15	A16
3	CONSUMO	0,9642	0,9642	0,9642	0,9642	0,9652	0,9672	0,9672	0,9683
3	NORMAL	0,9260	0,9260	0,9260	0,9260	0,9274	0,9296	0,9306	0,9313
3	RENEGOCIADO	0,9667	0,9667	0,9667	0,9667	0,9667	0,9695	0,9695	0,9695

Por ejemplo:

Cliente 1:

Saldo = \$2.800.000; PPP = 14; Segmento = Normal; TRAMO_MORA = 3.

Primero: se separa el saldo en saldos remanentes.

Segundo: se multiplica la devolución de cada periodo por la PD que corresponda a ese periodo.

Tabla (7.22) Ejemplo de Remanente * PD

PERIODO	1	1...11	12	13	14	
REMANENTE	\$2.800.000	\$2.600.000...\$800.000	\$600.000	\$400.000	\$200.000	
PD MARGINAL	0,926	0...0	0	0,0014	0,0022	
REMANENTE*MG	\$2.592.800	\$	\$0	\$560	\$440	\$2.593.800

Luego la PD*EAD del Cliente 1 será la suma todos los remanentes ya multiplicados por la PD Marginal de cada periodo. En este caso:

$$PD*EAD = \$2.592.800 + \$560 + \$440 = \$2.593.800$$

Las camadas utilizadas son desde el periodo 201201 hasta 201607 y la observación va desde 201202 hasta 201707.

Definición variable "Tramo de Mora":

Tramo 0: Sin días de mora.

Tramo 1: 1 a 30 días de mora.

Tramo 2: 31 a 60 días de mora.

Tramo 3: 61 a 90 días de mora.

Tramo 4: 91 a 120 días de mora.

Tramo 5: 121 a 150 días de mora.

Tramo 6: 151 a 180 días de mora.

Posterior a eso, el cliente se castiga.

7.14. PD Point in time

El ajuste “Point In Time” de la PD se realiza calculando el efecto de incluir el último año con desempeño disponible a la fecha sobre la tasa de malos del periodo de modelación (201201-201607). Este efecto se calculó para cada segmento en cada tramo de mora y se calcula en base a esto un factor de ajuste que se aplica a la PD obtenida del modelo, los resultados fueron los siguientes:

Tabla (7.23) probabilidad de default in time

SEGMENTO	TRAMO MORA	TASA DE MALOS	TASA DE MALOS CON ÚLTIMO PERIODO	INCREMENTO PORCENTUAL
NORMAL	0	11,82%	12,27%	1,0381
NORMAL	1	43,33%	44,00%	1,0155
NORMAL	2	76,89%	77,62%	1,0095
NORMAL	3	92,70%	93,09%	1,0042
CONSUMO	0	8,31%	8,10%	0,9747
CONSUMO	1	39,59%	39,64%	1,0013
CONSUMO	2	79,76%	80,07%	1,0039
CONSUMO	3	95,99%	96,02%	1,0003
RENEGOCIADO	0	51,60%	54,36%	1,0535
RENEGOCIADO	1	72,94%	75,04%	1,0288
RENEGOCIADO	2	89,93%	91,11%	1,0131
RENEGOCIADO	3	97,76%	98,11%	1,0036
TRANSACCIONAL	0	4,03%	4,45%	1,1042

Monto de exposición al Incumplimiento (EAD)

Para calcular la EAD se utiliza el 100% del saldo adeudado a la fecha de cálculo de la provisión, más un porcentaje determinado según corresponda a SBIF (35%) o IFRS9 (28,05%).

Cálculo % de exposición por contingente IFRS 9 (CCF):

Para calcular el CCF se seleccionaron a todos los clientes que caen en default de la base utilizada para modelar, que tengan cupo disponible mayor a cero al momento de ser seleccionados y tengan hasta 5 días de mora. Para cada periodo, desde que fue observado el cliente hasta su default, se le calculó el porcentaje de utilización adicional, lo que corresponde al aumento de saldo entre un mes y otro sobre el disponible original y luego se utiliza el promedio de este valor como CCF. Esto es:

$$CCF = \text{PROMEDIO} \left(\frac{(SALDO_n - SALDO_{n-1})}{(CUP0_0 - SALDO_0)} \right). \quad (7.2)$$

Así finalmente el porcentaje de exposición por contingente a utilizar es 28,05% para todo segmento excepto RENEGOCIADO, dado que no poseen cupo disponible.

7.15. Cálculo de LGD:

La LGD fue calculada trayendo todos los pagos de los clientes caídos en default en el periodo de desarrollo del modelo a valor presente al momento en que cayó en default con la tasa ponderada por saldo que tenían de sus productos con la compañía antes de caer en default. Previo a traer el monto a valor presente, estos pagos fueron corregidos por un factor de costo de cobranza (0,7% costo de recupero promedio del periodo) y luego se le añadieron las ventas de cartera. Por último, el monto anterior se divide por el saldo total de los clientes al momento de caer en default y de esta forma obtenemos una tasa de recupero dado el default para cada segmento.

Matemáticamente:

$$LGD_k = \frac{\sum_{i=1}^n RECUPEROS_NETOS_i}{\sum_{i=1}^n SALDO_i} ; k=NORM, CONS, RENE, TRXL, \quad (7.3)$$

dónde,

n : Corresponde al total de clientes en default de la base de desarrollo.

$RECUPEROS_NETOS_i$: Corresponde a los pagos del cliente i traídos a valor presente al momento que cayó en default descontados del costo de recupero.

$SALDO_i$: Corresponde al saldo del cliente i al momento de caer el default.

Ejemplo: Un cliente del segmento normal cayó en default en el mes 1 teniendo un saldo de \$100.000 y su tasa de interés ponderada era del 2%. En el mes 4 registró un pago por \$10.000, y en el mes 5 registró otro pago por \$10.000.

El recupero de este cliente sería:

$$RECUPERO = \frac{10.000}{(1+0.02)^3} + \frac{10.000}{(1+0.02)^4} = 9.423,2 + 9.238,4 = 18.662$$

Luego, la LGD_{NORM} será la suma de todos los pagos en valor presente de los clientes del segmento normal dividido en la suma de todos los saldos del segmento normal.

Así finalmente la LGD por segmento será:

NORMAL Y TRANSACCIONAL = 70,26%.

CONSUMO = 80,1%.

RENE = 63,26%.

Para el cálculo de la LGD se utilizaron las bases a Rut único utilizadas para modelar y se descartaron a todos los clientes que no tuvieran una tasa de interés ponderada por saldo (ya sea por tener tasa cero o por solo comprar al contado).

Tabla (7.24) de resumen cálculo LGD

SEGMENTO	NÚMERO DE CLIENTES	SALDO AL DEFAULT	RECUPERO (1)	COSTO COBRANZA (2)	VENTAS CARTERA	RECUPERO FINAL
CONSUMO	78974	\$ 124.286.301.054	\$ 24.364.149.647	\$ 170.549.048	\$ 545.172.757	\$ 24.738.773.357
NORMAL	321479	\$ 153.444.948.733	\$ 45.226.698.700	\$ 316.586.891	\$ 720.103.516	\$ 45.630.215.325
RENEGOCIADO	5817	\$ 3.462.759.655	\$ 1.266.695.354	\$ 8.866.867	\$ 14.405.967	\$ 1.272.234.453

- Antes de ser sumados, el recuperado de cada cliente fue traído a valor presente con la tasa ponderada por saldo de cada cliente.
- El costo de cobranza fue aplicado multiplicando un factor de 1-0,7% al recuperado de cada cliente ANTES de ser traídos a valor presente.

Tabla (7.25) Factores de recuperado por segmento

SEGMENTO	SALDO DEFAULT	RECUPERO FINAL	FACTOR DE RECUPERO	LGD (1-FR)
CONSUMO	\$ 124.286.301.054	\$ 24.738.773.357	0,199	80,10%
NORMAL	\$ 153.444.948.733	\$ 45.630.215.325	0,297	70,26%
RENEGOCIADO	\$ 3.462.759.655	\$ 1.272.234.453	0,367	63,26%

La tasa de interés ponderada por saldo de cada cliente fue calculada multiplicando el saldo de cada producto que el cliente tenga vigente el mes previo a caer en *default*, por la tasa de interés que tenía el mismo producto, y luego dividiendo el resultado por la suma de los saldos.

Véase la siguiente tabla resumen de las tasas de interés por segmento:

Tabla (7.26) Resumen de las tasas de interés por segmento

	TASA DE INTERES PROMEDIO	NÚMERO DE CLIENTES
CONSUMO	3,00%	78.974
NORMAL	2,52%	321.479
RENEGOCIADO	2,82%	5.817

Las camadas utilizadas son desde el periodo 201201 hasta 201607 y la observación va desde 201202 hasta 201707.

7.16. Definiciones conceptuales

Definición Buckets:

Bucket 3: En este grupo se encontrarán todos los clientes en default y su PD será de 100%.

Bucket 2: En este grupo estarán todos los clientes que cumplan con la condición de cambio significativo en el riesgo y no han caído en default, estos clientes serán provisionados por todo el tiempo que duren sus créditos y su exposición corresponderá al 100% de su deuda.

Bucket 1: En este grupo estarán todos los clientes que no cumplan con las condiciones de pertenencia al Bucket 3 ni al Bucket 2. Su exposición también corresponderá al 100% de su deuda, además de un porcentaje de su saldo disponible según corresponda.

7.14.1 Criterios para cambiar de Bucket:

Traspaso del Bucket 1 al Bucket 2: Cumplir con la condición de cambio significativo en el riesgo.

Traspaso del Bucket 1 o Bucket 2 al Bucket 3: Cumplir con la condición de Default.

Traspaso del Bucket 3 al Bucket 2: No es posible llegar al Bucket 2 desde el Bucket 3, ya que un cliente al ser “curado” entra directamente al Bucket 1.

Traspaso del Bucket 2 al Bucket 1: Dejar de cumplir con la condición de cambio significativo en el riesgo.

Traspaso del Bucket 3 al Bucket 1: Cumplir con las condiciones de cura.

La definición de cura que utilizaremos para recuperar a los clientes deteriorados será la normativa de SBIF, esto es:

1. Los días de mora del deudor deben ser menor o igual a 30.
2. Estando en deterioro o con 60 o más días de mora, no haber normalizado la deuda con una renegociación.
3. Haber pagado el 100% de su facturación los últimos 4 meses dentro de los plazos establecidos.
4. No tener mora externa (se debe consultar las bases de morosos de casas comerciales cargada en el Sistema)

7.14.2. Análisis de factores de prepago

Durante el trabajo de desarrollo, se determinó que hay un evento en el cual se puede disminuir el monto a Provisionar. Esto ocurre cuando se produce un prepago total. A continuación, se presenta el análisis para evaluar su posible utilización.

Para establecer los factores de prepago del modelo de provisiones se siguieron los siguientes:

- Se obtuvo una base de precierres de operaciones, la cual cuenta con 1.852.371 registros, divididos en camadas desde 201201 a 201707. Esta base no es una base de

prepagos propiamente tal, pues son solicitudes que sólo algunas se llevaron a cabo, por ende se debió establecer qué créditos resultaron efectivamente siendo prepagos totales. Para esto, se estableció el siguiente criterio: Seleccionamos de la base de precierres aquellos créditos que hayan vencido dentro de los tres meses siguientes y que tengan cuotas pendientes mayores a 3, es decir, son créditos que terminaron anticipadamente, no de forma natural. Además, se impuso la condición que sean créditos sin mora y que el cliente haya pagado más de lo que tenía facturado.

- Luego de tener seleccionados los créditos que fueron prepagados completamente, se construyó una curva de prepagos, la cual consiste en presentar un evolutivo de la acumulación mensual de los factores de prepagos por segmento. Cada factor mensual se calcula del cociente entre el monto prepago promedio al mes correspondiente y el promedio de los saldos de la cartera (desde 201201 a 201607).

Ejemplo. Para el segmento consumo al mes 1 tenemos un factor de 0,069%, al mes 2 tenemos 0,062% y al mes 3, 0,074%. La curva de prepagos hasta el mes 3 es entonces los porcentajes acumulados, es decir:

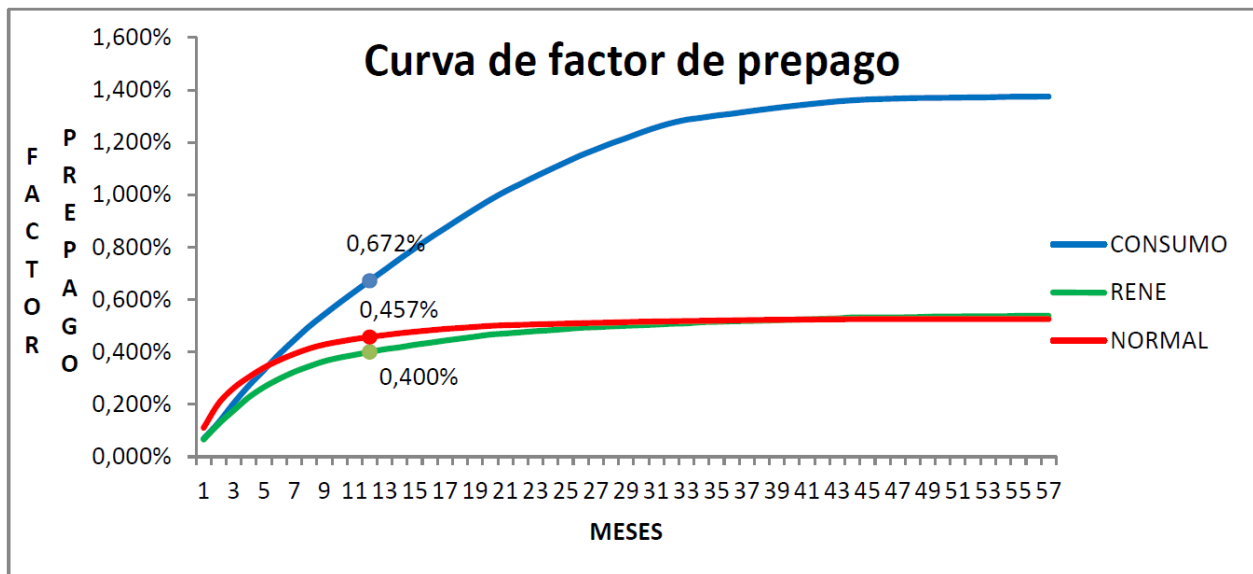
Mes 1: 0,069%.

Mes 2: 0,131% (0,069%+0,062%).

Mes 3: 0,205% (0,131%+0,074%).

Finalmente se aplicó el prepago acumulado al mes 12 para cada Segmento, dado que la PD se evalúa a 12 meses, se considera adecuado aplicar el factor de descuento de riesgo por prepago también en este tiempo, finalmente los resultados por segmento se muestran en el gráfico a continuación.

Gráfico (7.3) Curva de factor de prepago



7.14.3. Mediciones de desempeño futuro de los modelos

Criterios para Seguimiento de los Modelos

En el seguimiento de los modelos de provisiones es importante establecer controles periódicos que garanticen su correcto funcionamiento y eficacia de los modelos. Se revisa la Estabilidad de la Población y la capacidad para discriminar. Para lo anterior se debe considerar la revisión periódica (trimestral) de los siguientes informes:

Informe de Estabilidad Poblacional

Permite revisar si la población sobre la cual se utiliza el modelo presenta cambios significativos con respecto a la población de desarrollo, para esto se usa el test de Estabilidad Poblacional, que compara el Benchmark (desarrollo) y el análisis de las PD del mes a medir, lo que permite determinar si existe y a qué se debe alguna posible variación.

Objetivo: Generar alertas de cambios poblacionales, que pudiesen afectar la distribución de las Provisiones con respecto a su Benchmark (población del desarrollo).

Aplicación: Para cada modelo (PD) se debe generar un conteo del total de casos asignados, comparando la distribución del mes contra el Benchmark.

Periodicidad Trimestral

Valores de aceptación:

Estabilidad Poblacional

< 0,1 Sin cambio

>= 0,1 y <0,25 Algún nivel de cambio

>=0,25 Cambio significativo

Chi cuadrado >= 4,57

Informe de Discriminación:

Busca validar que los niveles de discriminación de cada modelo para diferenciar los clientes con Default y sin Default, se mantengan dentro de los parámetros de aceptación. Si se observa valores menores a estos rangos se requiere presentar al comité de Riesgos estudio sobre el origen de la desviación y propuesta de plan de acción para su aprobación.

Objetivo: Generar alertas de cambios en la capacidad de los modelos para diferenciar a los clientes Buenos (sin default) de los Malos, que pudiesen afectar el cálculo de las Provisiones contra el nivel de Discriminación al momento del desarrollo.

Aplicación:

Para cada modelo se revisa cual es la Probabilidad de Incumplimiento (PD) real y se calcula mediante el test KS, que la distancia entre Default y No default se mantenga dentro de los niveles aceptables.

Periodicidad trimestral**Valores de aceptación**

KS \geq 40 para Normal, Consumo, y Transaccional

KS \geq 30 para Renegociados

7.17. Aplicación del resultado de modelo de provisiones

El valor de la empresa en estudio es de \$545.701.387.046 y el modelo de provisiones indica que se debe provisionar por \$90.767.549.840 lo que corresponde a un 16.63% del patrimonio total de la empresa, esto significa que la empresa tiene holgura para provisionar y seguir realizando otras inversiones.

Anexo

Definición de las variables

La probabilidad de incumplimiento se calcula en base al modelo de regresión logística que se determina evaluando el desempeño de los clientes actuales de la entidad con un horizonte temporal de 12 meses, siendo utilizadas 43 variables existentes que proceden a ser categorizadas. Además, cada categoría es recodificada con un único valor WOE. De manera adicional, se utilizan variables complementarias acorde las necesidades del modelo y la normativa. A continuación se presenta un resumen de la nomenclatura utilizada en cada una de las variables, tanto las existentes (43), como las variables nuevas (10). Posteriormente, se dará una explicación detallada de cada una de ellas.

Tabla 1: Variables existentes y nuevas

Tipos de variables	Nomenclatura utilizada
Variables existentes	V1. CTAPRD: producto asociado a cada cuenta
	V2. CTACUP: cupo L1 asignado a la cuenta
	V3. CTAEFI: estado financiero actual de la cuenta
	V4. SEGMENTO: Normal, Consumo, Renegociado y Transaccional
	V5. SUBSEGMENTO: clasificación de cartera del módulo de cálculo de provisiones
	V6. DIAS_MORA: días de mora calculados bajo el modelo vigente
	V7. SALDO: cupo de L1 asignado a la cuenta
	V8. CTAENF: estado no financiero de la cuenta
	V9. ANTIGÜEDAD DE LA CUENTA
	V10. DIAS DE MORA PROYECTADO: número de días de mora con que la cuenta cerraría el mes si es que no se realizaran pagos en el periodo.
	V11. TRAMO DE MORA NUEVO: variable numérica que se trabaja por categorías
	V12. UTILIZACIÓN: relación entre deuda y cupo
	V13. AMAX3: mora máxima en los últimos 3 meses

V14. AMAX6: mora máxima en los últimos 6 meses

Tabla 1: Variables existentes y nuevas (Continuación)

Tipos de variables	Nomenclatura utilizada
Variables existentes	V15. AMAX12: comportamiento del cliente dentro del último año
	V16. NPAGOS3: meses con pagos en un periodo de tres meses
	V17. NPAGOS6: meses con pagos en un periodo de seis meses
	V18. NPAGOS12: los meses con pagos en los últimos 12 meses
	V19. PAGOS3: proceso de cierre mensual y el calculo de la suma de los montos pagados en los últimos 3 meses
	V20. PAGOS6: proceso de cierre mensual y el calculo de la suma de los montos pagados en los últimos 6 meses
	V21. PAGOS12: proceso de cierre mensual y el calculo de la suma de los montos pagados en los últimos 12 meses
	V22. FACTURADO3: suma de los montos facturados
	V23. FACTURADO6: suma de los montos facturados (en los días de facturación) en los últimos 6 meses
	V24. FACTURADO12: suma de los montos facturados (en los días de facturación) en los últimos 12 meses
	V25. PSOBREF3: porcentaje de montos pagados sobre los totales facturados en los últimos 3 meses
	V26. PSOBREF6: porcentaje de monto asociado a pagos respecto el total facturado en los últimos 6 meses
	V27. PSOBREF12: porcentaje de monto asociado a pagos respecto el total facturado en los últimos 12 meses
	V28. CONTADOR DE RENEGOCIACIONES HISTORICO: Resumen de transacciones
V29. CONTADOR DE RENEGOCIACIONES POS NORMALIZACION: veces en que se ha tenido que renegociar luego de una renegociación	
V30. MORA DE ORIGEN DE LA RENEGOCIACION: clientes que se encuentran en segmento de renegociado	
V31. SEGMENTO DE ORIGEN DE LA RENEGOCIACION: para los clientes pertenecientes al segmento renegociado	
V32. RELACIÓN CUOTA / PAGOS: relación del valor cuota pactado en la última operación de renegociación que sostuvo el cliente	
V33. MAYOR CUPO: máximo valor que ha adoptado el cupo del cliente, a nivel histórico	
V34. RELACION CUOTA / MAYOR CUPO: calculada solamente para clientes del segmento renegociado	

Tabla 1: Variables existentes y nuevas (Continuación)

Tipos de variables	Nomenclatura utilizada
Variables existentes	V35. NUMERO DE CUOTAS DE LA RENEGOCIACION: número de cuotas pactadas en la última operación de renegociación que tenga registrada el cliente
	V36. PORCENTAJE PAGADO DEL CONVENIO: porcentaje pagado de la última operación de renegociación registrada para el cliente de segmento renegociado
	V37. ANTIGÜEDAD DE LA RENEGOCIACION: la antigüedad de la última transacción de renegociación del Cliente
	V38. TIPO DE APERTURA
	V39. SCORE DE APERTURA
	V40. SCORE DE BEHAVIOR
	V41. NUMERO DE DOCUMENTOS
	V42. MONTO DE DOCUMENTOS
	V43. NUMERO DE INFORMANTES
Variables nuevas	V44. SALDO MAS INTERESES POR COBRAR: saldo completo del cliente, incluyendo intereses comprometidos, pero no capitalizados
	V45. CUPO TOTAL: cupo de la L1 del cliente más el cupo L20 convertido a pesos
	V46. CUPO DISPONIBLE EN TARJETA L1: Cupo total de la L1 del cliente menos el monto utilizado de la L1
	V47. CUPO DISPONIBLE EN TARJETA L20: Cupo total de la L20 del cliente menos el utilizado de la L20 traspasado a pesos
	V48. TASA EFECTIVA DE COLOCACION: tasa efectiva promedio ponderado por saldo de las colocaciones del cliente
	V49. ESCENARIO: estado del escenario actual macroeconómico, siendo estos definidos como optimista, neutro o pesimista
	V50. DESEMPLEO: obtenida del Banco Central
	V51. ASEER: aumento significativo del riesgo, corresponde a una marca equivalente al cliente
	V52. PPP: máximo valor de cuotas pendientes dentro de todos los productos asociados al cliente
V53. DIAS_MORA_CONV: días de mora con que se adquirió la renegociación o refinanciamiento	

VARIABLES EXISTENTES

Las variables existentes corresponden a 43, las cuales se proceden a identificar con el código "V#", donde # equivale al número correlativo del 1 al 43.

1. V1: CTAPRD

Corresponde al código de producto asociado a cada cuenta. Estas responden a 7 posibles tarjetas equivalentes a productos de crédito.

Actualmente, la entidad posee 4 tarjetas propias, además de 3 que poseen convenio con la entidad internacional Mastercard, la cuál es marca registrada. A su vez, se identifican cuatro segmentos de cliente: Preferente, Tradicional, Solo Cuotas y Crédito Universal. Esto es meramente una coincidencia respecto los segmentos definidos en el modelo, ya que esta clasificación es más bien comercial en lugar de la utilizada por el modelo, la cual exige una clasificación acorde al riesgo del cliente propiamente tal, siendo el producto una de las múltiples variables que afectan en su determinación.

A continuación, en la tabla 1, es posible visualizar los productos que componen esta variable.

Tabla 2: Elementos de variable CTAPRD

Código	Descripción
015	Tarjeta Presto Preferente
003	Tarjeta Presto Tradicional
044	Tarjeta Presto Solo Cuotas
066	Tarjeta Presto Crédito Universal
040	Tarjeta Líder Mastercard ® Preferente Internacional
041	Tarjeta Líder Mastercard ® Tradicional Internacional
042	Tarjeta Líder Mastercard ® Uso Nacional

2. V2: CTACUP

Variable asociada al cupo L1 asignado a la cuenta, estos corresponden a un valor relativo a montos en pesos chilenos diferenciados por cada producto, considerando parámetros por clasificación comercial, aunque estos pueden variar según comportamiento, viéndose aumentado o disminuido.

3. V3: CTAEFI

La tercera variable corresponde al estado financiero actual de la cuenta en el sistema de Intellect Card, siendo un campo con el mismo nombre de la interfaz Z012, también generada por Intellect Card.

4. V4: SEGMENTO

Esta variable corresponde a las categorías definidas previamente: Normal, Consumo, Renegociado y Transaccional. Estos son asignados por el proceso de clasificación de cartera del módulo de cálculo de provisiones.

5. V5: SUBSEGMENTO

Con una descripción similar a la anterior, la variable subsegmento es determinada por el proceso de clasificación de cartera del módulo de cálculo de provisiones y adquiere los valores de: Nuevo, Avance y Refinanciamiento.

6. V6: DIAS_MORA

Corresponde a los días de mora calculados bajo el modelo vigente actualmente en el módulo de provisiones. Los valores que componen esta variable son numéricos.

7. V7: SALDO

Equivale al campo CTACUP de la interfaz Z012 generada por Intellect Card, es decir, el cupo de L1 asignado a la cuenta.

8. V8: CTAENF

Corresponde al estado no financiero de la cuenta en el sistema de Intellect Card, siendo el equivalente al campo del mismo nombre de la interfaz Z012 generada por IC.

9. V9: ANTIGÜEDAD DE LA CUENTA

La novena variable se calcula a partir del campo CTAFIT de la interfaz Z012. Esta adquiere valores numéricos en cantidad de meses y se rige por la siguiente ecuación:

$$\textit{Antigüedad} = (A2 - A1) * 12 + (M2 - M1), \quad (8.1)$$

cuyos campos corresponden a:

$A1 =$ Año del campo CTAFIT

$A2 =$ Año de la fecha de proceso

$M1 =$ Mes del campo CTAFIT

$M2 =$ Mes de la fecha de proceso

10.V10: DIAS DE MORA PROYECTADO

Responde al número de días de mora con que la cuenta cerraría el mes si es que no se realizaran pagos en el periodo.

Los valores que puede adquirir esta variable corresponden a la siguiente formulación:

Si Dias de mora = 0;

No se proyecta → *Días de mora proyectado* = 0

Si Dias de mora > 0;

$$\textit{Se proyecta} \rightarrow \textit{Dias de mora proyectado} = \textit{Dias de mora} + D2 - D1, \quad (8.2)$$

donde:

D2 = Días que tiene el mes de proceso (28, 29, 30 o 31),

D1 = Día de proceso.

El dato de días de mora que requiere esta variable se obtiene a partir del módulo de cálculo de provisiones, equivalente a:

$$\textit{Días de mora} = \textit{Fecha de proceso} - (\textit{Fecha de inicio de mora} + 1). \quad (8.4)$$

11. V11: TRAMO DE MORA NUEVO

El tramo de mora nuevo es una variable numérica que se trabaja por categorías. Estas se visualizan en la tabla siguiente:

Tabla 3: Tramos asociados a variable V11.

Días de Mora	Tramo
0	0
1 a 31	1
32 a 61	2
62 a 91	3
92 a 121	4
122 a 151	5
152 a 181	6
Sobre 181	7

Al igual que en la variable V10, los días de mora se calculan bajo la ecuación 19.

12.V12: UTILIZACIÓN

Corresponde a la relación entre deuda y cupo, expresada como un porcentaje, es decir, el porcentaje de utilización de cupo que tiene el cliente.

Este se rige por el siguiente conjunto de ecuaciones:

$Si \ CTACUP \leq 0; \ Utilización = -999$

$CTAOUP \geq 0; \ Utilización = 0$

$CTACUP > 0 \ y \ CTAOUP > 0; \ Utilización = \min(1000; 100 * \frac{CTAOUP}{CTACUP})$. (8.5)

El sistema definido en la ecuación 8.6, manifiesta la fórmula de cálculo para la utilización:

$$Utilización = 100 * \frac{CTAOUP}{CTACUP}, \quad (8.6)$$

donde:

CTACUP= Cupo L1 asignado a la cuenta

CTAOUP= Deuda total del cliente

Ambos campos se obtienen a partir de la interfaz Z012, generada por IC.

Tal como evidencia la ecuación 20, si este resultado es mayor que 1000, se debe utilizar el límite máximo, es decir, el menor valor entre 1000 y la utilización calculada.

13.V13:AMAX3

La treceava variable corresponde a la mora máxima en los últimos 3 meses incluyendo el mes de curse, el que es calculado en el proceso del último día del mes. Se determina el valor máximo que obtuvo el campo "DIAS_MORA", que como fue señalado, se obtiene desde la interfaz Z012. Para ser considerado el mes, se reconoce como cierre al proceso realizado el último día del mes.

Para el resto de los días, es preciso considerar el máximo valor que resulte de la comparación del campo DIAS DE MORA calculado para el día de proceso y el "máximo valor del campo DIAS DE MORA calculado en los cierres de los últimos 3 meses".

14.V14: AMAX6

Similar a su variable predecesora, corresponde al máximo periodo de mora que sostuvo el cliente durante los últimos 6 meses, incluyendo el mes actual.

Se realiza el mismo alcance que en V13, destacando que, en el caso de los otros días, se debe considerar el máximo valor resultante al comparar el campo DIAS DE MORA calculado para

el día de proceso y el “máximo valor del campo DIAS DE MORA calculado en los cierres de los últimos 6 meses”

15.V15: AMAX12

Dado que los periodos de mora pueden deberse a cuestiones circunstanciales, tales como desempleo o desembolsos eventuales de cierto periodo, es preciso integrar una variable que incorpore el comportamiento del cliente dentro del último año, para evidenciar si este presenta mora recurrentemente o es más bien una situación puntual. Para estos efectos, se incorpora la variable AMAX12, considerando los días máximos de mora resultantes entre los días de mora calculados de proceso y el máximo valor del campo DIAS DE MORA que se revisa y calcula en el último día del mes de un horizonte temporal de 12 meses.

16.V16: NPAGOS3

La variable a continuación, corresponde al caso opuesto de los días de mora, visualizando los meses con pagos en un periodo de tres meses, considerando los más recientes respecto a la fecha de proceso.

En el cierre mensual, que como ya hemos definido, es el proceso que se realiza el último día del mes, se calcula la cantidad de meses calendario donde se han realizado pagos efectivos superiores a 0 en los últimos tres meses, incluyendo el periodo de curso. En el caso del resto de los días, se mantiene el valor calculado en el último cierre de mes, por lo que solamente se actualiza mensualmente.

Se considerarán pagos efectivos las transacciones registradas en el sistema Intellect Card con los siguientes códigos:

Tabla 4: Consideraciones de pagos para variable V16.

TIPO	DESCRIPCION	CODTRX	DESCRIPCION
5	Pagos	36001	PAGO PAC BANCO ESTADO
5	Pagos	62001	PAGO
5	Pagos	69001	CONVENIO PAGO AUTOMATICO
5	Pagos	63001	PAGO CON CHEQUE/VV

A modo de ejemplo, se entabla la siguiente situación:

Un cliente presenta los siguientes pagos realizados a su cuenta asociada:

Tabla 5: Caso de ejemplo de tabla de pagos.

Nro	Fecha de Pago	Monto
1	06-12-2020	30.000
2	31-12-2020	30.000
3	15-01-2021	70.000
4	01-02-2021	100.000
5	02-02-2021	50.000
6	04-04-2021	55.000
7	28-05-2021	30.000
8	06-06-2021	12.000

9	14-06-2021	50.000
10	02-07-2021	130.000
11	31-07-2021	45.000

En total, se identifican 11 partidas, las cuales se clasifican acorde fechas posibles de proceso, a modo de ejemplo, se obtendrían los siguientes resultados:

Tabla 6: Caso de ejemplo tabla de pagos por fecha de proceso

Fecha	15-03-	31-03-	20-04-	31-05-	28-06-	30-06-
Proceso	2021	2021	2021	2021	2021	2021
Nro	3	2	2	2	2	3
Meses						
Pagos						

Finalmente, el cliente presentaría, a junio de 2021, un valor de 3 para la variable NPAGOS3.

17.V17:NPAGOS6

La variable 17, corresponde a una similar al caso previo, aunque con una evaluación temporal de 6 meses. Se reitera que este utiliza el proceso del último día del mes, calculando el número de meses calendario donde exista presencia de pagos efectivos mayores de 0 en los últimos 6 meses, incluyendo el actual.

Al igual que en casos anteriores, se mantiene el valor calculado en el último cierre para días intermedios del periodo.

18.V18: NPAGOS12

Corresponde a los meses con pagos en los últimos 12 meses, con igual condiciones que las variables 16 y 17.

19.V19: PAGOS3

La variable PAGOS3, considera el proceso de cierre mensual y el calculo de la suma de los montos pagados en los últimos 3 meses incluyendo el mes actual. Esta variable viene a resolver situaciones donde se realizan prepagos sin aceleramiento de cuotas, ya que un cliente podría abonar el total adeudado en su tarjeta de crédito, pero al quedar saldo en cuotas y sin solicitar la aceleración respectiva, tendría pagos únicamente en un periodo, sin evidenciar que este fue realizado por un monto superior a la deuda del mes corriente.

Se considerarán pagos efectivos las transacciones registradas en el sistema Intellect Card con los siguientes códigos:

Tabla 7: Códigos aceptados para variable V19.

TIPO	DESCRIPCION	CODTRX	DESCRIPCION
5	Pagos	36001	PAGO PAC BANCO ESTADO
5	Pagos	62001	PAGO
5	Pagos	69001	CONVENIO PAGO AUTOMATICO
5	Pagos	63001	PAGO CON CHEQUE/VV

Al igual que en las variables de pago previas, se consideran únicamente 4 códigos válidos para determinar que el pago fue efectivo en el periodo.

20.V20: PAGOS6

Es el símil de la variable 19, aunque con un horizonte temporal de 6 meses, considerando, de igual forma, el proceso respectivo del último día del mes, por lo que se mantiene el valor de cierre en días intermedios.

21.V21:PAGOS12

Similares a las dos variables precursoras, equivalen a la variable del monto total de pagos registrados efectivamente en un horizonte de 12 meses, incluyendo el mes en curso.

Al igual que en casos anteriores, se mantendrá el valor calculado en el último cierre de mes anterior a los intermedios.

22.V22.: FACTURADO3.

En el proceso de cierre del mes, se calcula la suma de los montos facturados. No obstante, en este caso, existen distintos días de facturación para cada cliente, otorgando un abanico de 3 días posibles para efectuar la facturación: 5, 16 o 24 de cada mes. Este día, lo selecciona el cliente al suscribirse al producto de crédito, siendo pactado por contrato y modificable si es que el cliente lo solicitase.

Los montos facturados se obtienen desde la interfaz Z012 generada por el sistema Intellect Card, en cada periodo de cierre.

En esta variable, puede ocurrir la situación de tener desfases entre los pagos y facturados, especialmente en los primeros 3 meses de operación con la tarjeta de crédito, por lo que se requiere de esta variable en horizontes mayores para minimizar el efecto de la primera facturación fuera de plazo, ya que, un cliente podría eventualmente, tener su facturación el día 24 pero realizar el pago el 01 del mes siguiente.

23.V23:FACTURADO6

En el proceso del último día del mes (*cierre*) se calculará la suma de los montos facturados (en los días de facturación) en los últimos 6 meses incluyendo el mes actual. Los montos facturados se obtendrán de la interfaz Z012 generada por el sistema Intellect Card.

Para el resto de los días se mantendrá el valor calculado en el último cierre de mes.

24.V24: FACTURADO12

Al igual que las V22 y V23, esta variable manifiesta los montos facturados en los últimos 12 meses desde la fecha de proceso.

Se mantiene la condición de respetar el valor calculado en el proceso de cierre del mes anterior.

25.V25: PSOBREF3

La variable 25 corresponde al porcentaje de montos pagados sobre los totales facturados en los últimos 3 meses.

Este proceso, se lleva a cabo en conformidad con el sistema condicional a continuación:

$$\begin{aligned} & \text{Si } FACTURADO3 \leq 0; PSOBREF3 = -999 \\ & FACTURADO3 > 0; PSOBREF3 = \min \left(1000; 100 * \frac{PAGO3}{FACTURADO3} \right). \end{aligned} \quad (8.7)$$

El sistema definido en la ecuación 8.8, manifiesta la fórmula de cálculo:

$$PSOBREF3 = 100 * \frac{V19PAGOS3}{V22FACTURADO3}. \quad (8.8)$$

Lo anterior, manifiesta que, si la variable V22 es menor o igual a 0, entonces el valor de PSOBREF3, se define como -999, si no, se utiliza la segunda condicionante.

De utilizar la segunda condicionante, PSOBREF3 si tuviese un resultado inferior a 0, se debe considerar el valor mínimo de 0 y máximo de 1000, por lo que es preciso enmarcarse en dichos límites si PSOBREF3 calculado como porcentaje, fuese fuera de rango.

26.V26: PSOBREF6

El porcentaje de monto asociado a pagos respecto el total facturado en los últimos 6 meses se guía por igual método que la ecuación 4.17, aunque ahora considerando un horizonte temporal mayor.

27.V27: PSOBREF12

Se repite procedimiento de V25 y V26, con un periodo de 12 meses. Al igual que en casos anteriores, se consideran los datos respectivos al cierre del mes inmediatamente anterior.

28.V28: CONTADOR DE RENEGOCIACIONES HISTORICO

La renegociación es una operación que tiene lugar cuando un cliente sostiene la necesidad de reordenar su situación financiera, especialmente por sobre endeudamiento y/o incapacidad de pago, por ejemplo, por cambio de trabajo o situación de cesantía.

Para efectos del modelo, se define a la variable V28 como el equivalente al campo NUMREN que se calcula diariamente en el módulo de provisiones, usando el proceso de “Resumen de transacciones”. Actualmente, este campo se encuentra en la tabla llamada “*PROVISIONES.TBL_RESUMEN_TRX_RENE_REFI*”, este valor es el mismo que diariamente se graba en el archivo el archivo RESUTRXaaaammdd.txt que se deja en la carpeta compartida W:\Provisiones\Salidas.

29.V29: CONTADOR DE RENEGOCIACIONES POS NORMALIZACION

En línea con la renegociación, se aplica el contador de renegociaciones pos normalización, esto tiene relación con las veces en que se ha tenido que renegociar luego de una

renegociación, es decir, las veces que se ha repetido la necesidad de reestructuración a contar de la primera ocurrencia del evento.

En el módulo de provisiones, se obtiene a partir del campo CONREN, que es calculado en el mismo proceso de “Resumen de transacciones”, disponible en la tabla *“PROVISIONES.TBL_RESUMEN_TRX_RENE_REFI”*.

Cabe señalar que el valor que adquiere esta variable, es el equivalente al grabado diariamente en el archivo RESUTRXaaaammdd.txt

V30: MORA DE ORIGEN DE LA RENEGOCIACION

La variable 30 se calcula únicamente para los clientes que se encuentran en segmento de renegociado, ya que para los clientes de otro segmento su valor es automáticamente -999.

Esta variable, corresponde a los días de mora que se tienen en el proceso de cierre de mes inmediatamente anterior al curso de la última transacción asociada a una operación de renegociación del cliente, por lo que, en caso de no tener operaciones de este tipo, su valor es, por defecto, -999.

Para la obtención de esta variable, se precisa modificar el proceso actual de clasificación de transacciones, a fin de que, en el minuto de registrar una nueva operación de renegociación, esta se grabe en el proceso y se le otorgue como atributo, el tramo de mora que tenía el cliente al cierre del mes anterior, manteniendo los tramos disponibles en tabla 2.

Finalmente, cabe destacar que se requiere de un proceso que por una única vez pueble la información en la tabla *“PROVISIONES.TBL_HIST_TRX_RENE_REFI”* que actualmente es administrada en el módulo de provisiones vigente.

30.V31: SEGMENTO DE ORIGEN DE LA RENEGOCIACION

Al igual que en el caso de la variable 30, esta es sujeta de calcularse solo para los clientes pertenecientes al segmento renegociado, ya que, de otro modo, su valor por defecto es -999.

Este ítem corresponde al segmento de riesgo asignado al cliente al cierre de mes previo al curso de la última transacción de renegociación. Adicionalmente y en similitud con lo descrito en la variable previa, se requiere modificar el proceso actual de clasificación de transacciones, a fin de que esta sea grabada con la segmentación vinculada al cliente en el mes inmediatamente anterior a la operación de renegociación.

En consecuencia de lo anterior, se requiere de un proceso que complete por una vez estos datos en la tabla "PROVISIONES.TBL_HIST_TRX_RENE_REFI" que actualmente administra el módulo de provisiones.

31.V32: RELACIÓN CUOTA / PAGOS

Definido como la relación del valor cuota pactado en la última operación de renegociación que sostuvo el cliente respecto los montos pagados y número de periodos mensuales con pagos efectivos durante el año previo al curso de la operación.

La relación cuota/pagos es una variable relacionada a los clientes del segmento renegociado, puesto que, para su métrica, los clientes de otro segmento obtienen el valor por defecto de -999.

Como en las variables previas, se requiere la modificación del proceso actual de clasificación de transacciones a fin de que, al momento de registrar una nueva operación, esta sea grabada y relacionada con un atributo que represente el cálculo a continuación:

$$Relación \frac{cuota}{pagos} = \min(1000; \max(0; \frac{Valor \ cuota \ de \ renegociación}{\frac{PAGOS_AÑOANTERIOR}{NPAGOS_AÑOANTERIOR}} * 100)). \quad (8.9)$$

Este valor debe ser redondeado y sin decimales.

Los componentes de la fórmula corresponden a:

$$PAGOS_{AÑOANTERIOR} = \sum_0^{12} \text{Montos pagados previa negociación.}$$

$$NPAGOS_{AÑOANTERIOR} = \sum_0^{12} \text{Número de meses que registra pagos previa negociación.}$$

Finalmente, ambos datos sostienen la fórmula de cálculo de la variable, es decir:

$$\text{Relación } \frac{\text{cuota}}{\text{pagos}} = \frac{\text{Valor cuota de renegociación}}{\frac{PAGOS_{AÑOANTERIOR}}{NPAGOS_{AÑOANTERIOR}}}. \quad (8.10)$$

Al igual que en los casos previos, este valor también debe ser redondeado, sin decimales y ajustarse a los límites de entre 0 y 1000, por lo que, si el valor excede los márgenes, se debe utilizar el límite más cercano. Por otra parte, también se precisa de un proceso que incorpore esta información en la tabla PROVISIONES.TBL_HIST_TRX_RENE_REFI.

32.V33: MAYOR CUPO

La variable 33, corresponde al máximo valor que ha adoptado el cupo del cliente, a nivel histórico e informado debidamente en la interfaz Z012, la que, como se ha comentado, es generada por el sistema Intellect Card.

Todos los días se debe evaluar si el campo CTACUP de la interfaz Z012 es mayor que el mayor cupo registrado, en caso que se encuentre un cupo mayor se debe mantener el nuevo cupo como valor de la variable. Si el nuevo cupo es menor que el mayor registrado, se debe mantener el mayor cupo registrado.

Esta variable también requiere de realizar un proceso especial, a fin de determinar el mayor cupo histórico que se haya registrado para cada cliente en la tabla de clasificación histórica de cartera que se mantiene en el sistema actual, a fin de ir actualizando constantemente dicho valor.

V34: RELACION CUOTA / MAYOR CUPO

La V34 es debidamente calculada solamente para clientes del segmento renegociado, ya que en caso de que el cliente sea de otro segmento, se adopta automáticamente el valor de -999.

La variable se define como la relación del valor cuota de la última renegociación del cliente respecto el mayor cupo que se haya registrado históricamente en el sistema.

La ecuación asociada corresponde a:

$$\text{Relación} \frac{\text{cuota}}{\text{mayor cupo}} = \min(1000; \max(0; \frac{\text{Valor cuota de renegociación}}{\text{Mayor cupo}} * 100)). \quad (8.11)$$

El valor debe ser redondeado, sin decimales y ajustarse a los límites de entre 0 y 1000.

33.V35: NUMERO DE CUOTAS DE LA RENEGOCIACION

Se inscribe como el número de cuotas pactadas en la última operación de renegociación que tenga registrada el cliente. Este calculo aplica solo para clientes del segmento renegociado, ya que, los clientes de otro segmento adquieren inmediatamente el valor de -999.

Este valor se obtiene desde el campo NUMCUO de la tabla "PROVISIONES.TBL_HIST_TRX_RENE_REFI" del módulo de provisiones actualmente utilizado.

34.V36: PORCENTAJE PAGADO DEL CONVENIO

La variable a continuación se define como el porcentaje pagado de la última operación de renegociación registrada para el cliente de segmento renegociado, calculado diariamente. En caso de ser cliente de otro segmento, no se considera válido su cálculo y toma el valor de -999 por defecto.

$$\text{Porcentaje pagado} = \min(1000; \max(0; \frac{\text{PAGO_FPMENOS10} + \text{PAGO_POST}}{\text{MONTRX} + \text{PAGO_FPMENOS10}})). \quad (8.12)$$

El resultado debe ser redondeado, sin decimales y acotarse a los límites de entre 0 y 1000.

Se definen los componentes como:

PAGO_FPMENOS10: Pagos efectivos registrados desde 10 días antes de la fecha de posteo de la última transacción de renegociación de la cuenta hasta la fecha de posteo inclusive.

PAGO_POST: Pagos efectivos registrados desde la fecha de posteo de la última transacción de renegociación + 1 día hasta la fecha de proceso inclusive.

MONTRX: Monto de la última transacción de renegociación registrada para el Cliente en la tabla histórica de transacciones.

35.V37: ANTIGÜEDAD DE LA RENEGOCIACION

Corresponde a la antigüedad de la última transacción de renegociación del Cliente registrada en la tabla histórica de transacciones expresada en meses.

$$\textit{Antigüedad} = (A2 - A1) * 12 + (M2 - M1), \quad (8.13)$$

cuyos campos corresponden a:

A1 = Año del posteo

A2= Año de la fecha de proceso

M1 =Mes del posteo

M2 = Mes de la fecha de proceso

Este cálculo es válido solo para Clientes del segmento renegociado. Para los Clientes de los otros segmentos el valor de la variable será -999.

Las variables a continuación, no se encuentran disponibles en las bases del sistema de Intellect Card ni en las referentes al sistema actual de provisiones, siendo obtenidas desde otros sistemas de servicios financieros, por lo que no se calcularán, aunque se describen para efectos del modelo:

36.V38: TIPO DE APERTURA

Alfanumérico Largo 4 y se puebla en blanco

37.V39: SCORE DE APERTURA

Numérico de largo 3 y se puebla en cero

38.V40: SCORE DE BEHAVIOR

Numérico de largo 3 y se puebla en cero

Variables de Fuentes Externas

Se denominan como variables de fuentes externas a aquellas que se obtienen a partir de la Cámara de Comercio de Santiago a través de la interfaz INFOCOM, la que es utilizada en el módulo de provisiones y no precisan de cambios en su procesamiento, ni a nivel de carga ni de frecuencia.

39.V41. NUMERO DE DOCUMENTOS

La información disponible en el sistema financiero incorpora la posibilidad de revisar antecedentes comerciales fuera de la entidad, por lo que, se revisa el número de documentos informados en INFOCOM como morosos y/o protestos, siendo excluidos de este análisis aquellos documentos informados por la entidad financiera que se ha determinado para el caso de estudio.

40.V42: MONTO DE DOCUMENTOS

Como complemento de la variable 41, se identifica el monto de los documentos informados como protestos o morosos en la base de INFOCOM, excluyendo los informados por la entidad financiera que se ha definido y presentados en pesos chilenos.

41.V43: NUMERO DE INFORMANTES

Los documentos en condición de mora o protesto están sujetos a entidades que informan dichos documentos, por lo que se precisa de revisar el número de informantes, distintos a la entidad financiera, que se han manifestado con cargo al cliente. A modo de ejemplo, se definen como los acreedores que publican las morosidades, llamése Banco A, Banco B, Retail A, Retail B, y otros.

Nuevas Variables

Como se comentó en un inicio, el modelo puede precisar de variables que complementen el análisis de riesgo crediticio asociado a cada cliente. Se definen las siguientes a incorporar:

42.V44: SALDO MAS INTERESES POR COBRAR

Corresponde al saldo completo del cliente, incluyendo intereses comprometidos, pero no capitalizados.

43.V45: CUPO TOTAL

Se define como el cupo de la L1 del cliente más el cupo L20 convertido a pesos con el tipo de cambio ingresado por operaciones.

44.V46: CUPO DISPONIBLE EN TARJETA L1

Cupo total de la L1 del cliente menos el monto utilizado de la L1.

45.V47: CUPO DISPONIBLE EN TARJETA L20

Cupo total de la L20 del cliente menos el utilizado de la L20 traspasado a pesos con el tipo de cambio ingresado por operaciones.

46.V48: TASA EFECTIVA DE COLOCACION

Corresponde a la tasa efectiva promedio ponderado por saldo de las colocaciones del cliente que se encuentren vigentes.

Esta se define a través de la ecuación 4.14:

$$TE = \sum_0^n (Tasa\ de\ transacción_n * Saldo\ de\ transacción_n). \quad (4.14)$$

47.V49: ESCENARIO

Como se comentaba previamente, es preciso incorporar un factor que defina el estado del escenario actual macroeconómico, siendo estos definidos como optimista, neutro o pesimista. Esta variable puede tomar tres valores que serán el incremento o decremento de la probabilidad de default (PD) según corresponda el escenario, mientras que, las condiciones para definir el escenario serán entregadas por el área de riesgo de TEC TEC S.A. riesgo.

Tabla 8: Tabla de escenarios V49

Índice de desempleo	Escenario	Valor de factor
0<=ID<6	Optimista	0.689
6<=ID<7	Neutro	1
ID >7	Pesimista	1,03

50. V50: DESEMPLEO

Variable obtenida del Banco Central, ingresada por Operaciones, utilizada con desfase de 6 meses.

51. V51: ASEER

Definida como el aumento significativo del riesgo, corresponde a una marca equivalente al cliente que sostiene una mora superior a 30 días, valor máximo de cuotas pendientes superior a 12 y una probabilidad de incumplimiento menor a 1.

Esta condición es modificable de forma paramétrica.

52. V52: PPP

Corresponde al máximo valor de cuotas pendientes dentro de todos los productos asociados al cliente, siendo de carácter numérico.

53. V53: DIAS_MORA_CONV

Se determina la variable como los días de mora con que se adquirió la renegociación o refinanciamiento.

Para el nuevo modelo se requiere calcular diariamente las 43 variables ya definidas por la Gerencia de Riesgo y además incorporar las 10 siguientes:

Tabla 9: Resumen de nuevas variables.

Variable	Nombre Riesgo	Nombre Variable	Tipo de dato	Tipo Validación
44	Saldo más intereses por pagar	CTAOUPI	Numérico	Por Valores
45	CUPO TOTAL (L1 + L20) en pesos	CTACUPT	Numérico	Por Valores
46	Cupo Disponible en tarjeta L1	CTACUPDI1	Numérico	Por Valores
47	Cupo Disponible en tarjeta L20	CTACUPDI20	Numérico	Por Valores
48	Tasa Efectiva de colocación	TE	Numérico	Por Valores
49	Escenario	ESCENARIO	Numérico	Por Valores
50	Desempleo	M1	Numérico	Por Valores
51	ASEER	M2	Numérico	Por Valores
52	PPP	PPP	Numérico	Por Valores
53	Días de mora convenio mientras convenio este vigente	DIAS_MORA_CONV	Numérico	Por Valores

Detalle calculo K-S

Tabla cruzada Score_PDO*MarcaMalo									
Recuento	MarcaMalo		Total	MarcaMalo					
	0.0	1.0		% Malos	% Buenos	Malos ACC (a)	Buenos ACC (b)	abs((a)-(b))	
0	1	41	42	0,00%	0,20%	0,00%	0,20%	0,20%	
8	11	142	153	0,00%	0,90%	0,00%	1,10%	1,10%	
13	0	1	1	0,00%	0,00%	0,00%	1,10%	1,10%	
15	24	556	580	0,00%	3,40%	0,00%	4,50%	4,40%	
17	4	75	79	0,00%	0,50%	0,10%	4,90%	4,90%	
20	1	6	7	0,00%	0,00%	0,10%	5,00%	4,90%	
23	4	79	83	0,00%	0,50%	0,10%	5,50%	5,40%	
28	3	37	40	0,00%	0,20%	0,10%	5,70%	5,60%	
30	3	9	12	0,00%	0,10%	0,10%	5,70%	5,70%	
32	2	95	97	0,00%	0,60%	0,10%	6,30%	6,30%	
36	0	7	7	0,00%	0,00%	0,10%	6,40%	6,30%	
...	
565	314	66	380	0,40%	0,40%	23,80%	76,60%	52,90%	
566	75	27	102	0,10%	0,20%	23,90%	76,80%	52,90%	
567	54	14	68	0,10%	0,10%	23,90%	76,90%	53,00% <- MAX	
568	145	30	175	0,20%	0,20%	24,10%	77,10%	52,90%	
569	190	42	232	0,30%	0,30%	24,40%	77,30%	52,90%	
570	122	16	138	0,20%	0,10%	24,60%	77,40%	52,90%	
...	
958	271	2	273	0,40%	0,00%	99,00%	100,00%	1,00%	
961	2	0	2	0,00%	0,00%	99,00%	100,00%	1,00%	
968	5	0	5	0,00%	0,00%	99,00%	100,00%	1,00%	
969	25	1	26	0,00%	0,00%	99,00%	100,00%	1,00%	
973	267	2	269	0,40%	0,00%	99,40%	100,00%	0,60%	
977	19	0	19	0,00%	0,00%	99,40%	100,00%	0,60%	
985	130	1	131	0,20%	0,00%	99,60%	100,00%	0,40%	
1000	308	2	310	0,40%	0,00%	100,00%	100,00%	0,00%	
Total	73735	16465	90200				Max diferencia	53,00%	
	0,8174612	0,1825388					KS	53,00%	

Gráfico (8.1) Curvas de tasa de malos (PD PIT) segmento Normal

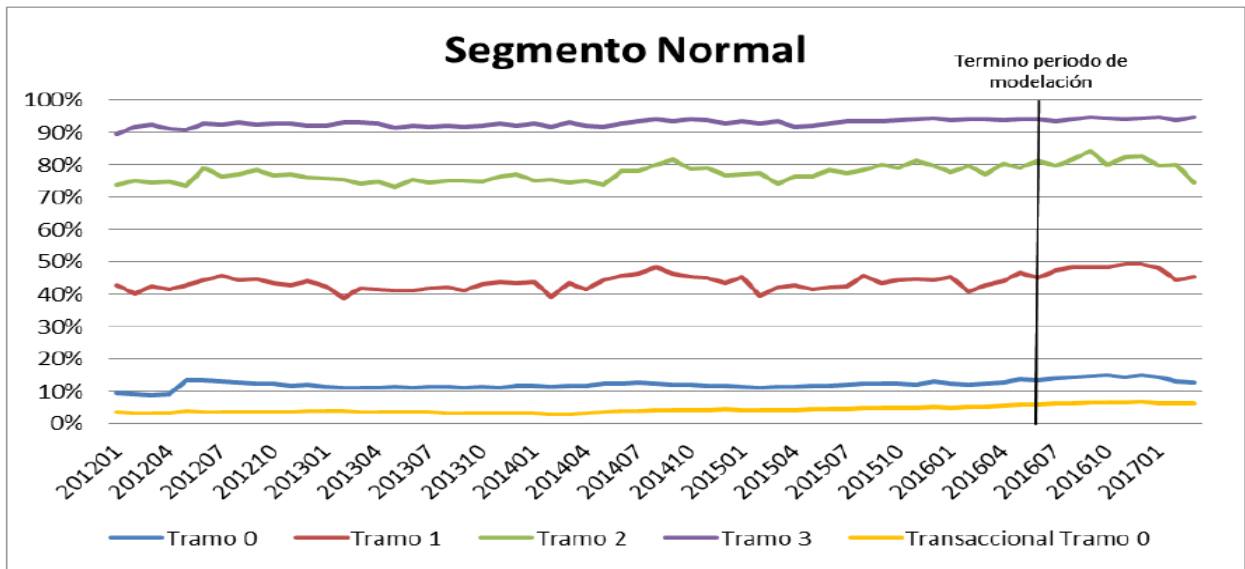


Gráfico (8.2) Curvas de tasa de malos (PD PIT) segmento Consumo

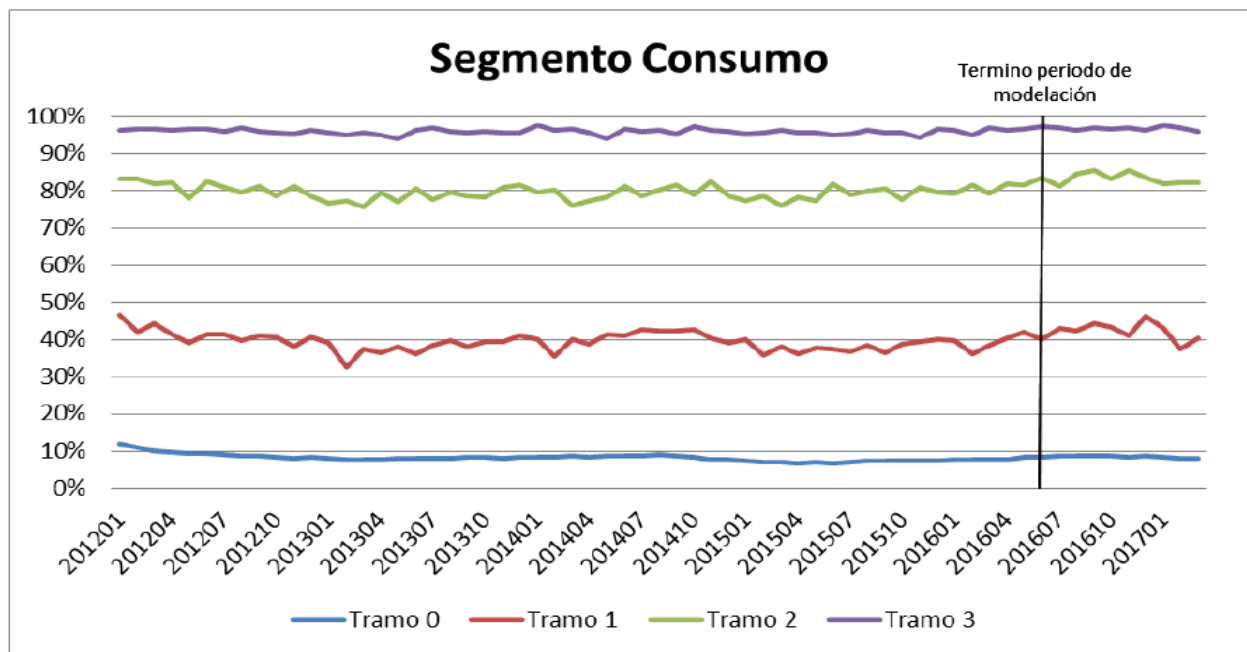
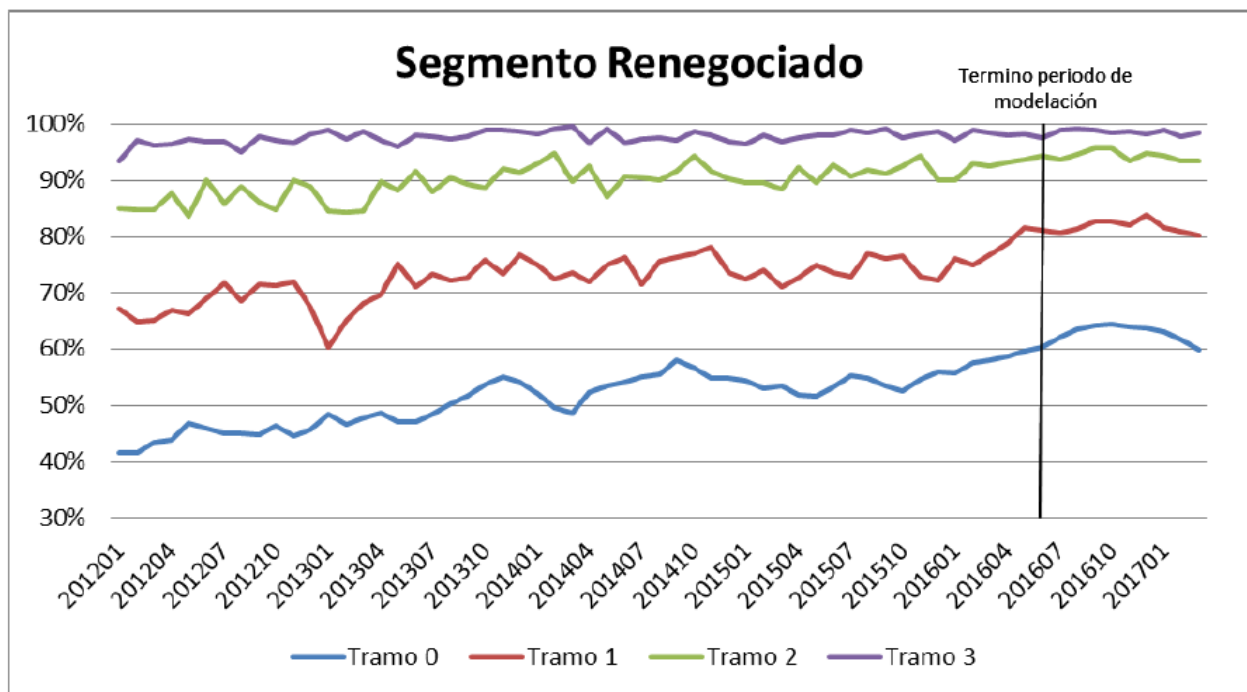


Gráfico (8.3) Curvas de tasa de malos (PD PIT) segmento renegociado



Categorización en WOE de variables no seleccionadas. Los WOE se calculan mediante la fórmula

$WOE = \left[\ln \left(\frac{dist.buenos}{dist.malos} \right) \right] * 100$ y la función de éstos es linealizar las variables según su efecto en el cálculo de la PD.

MODELO CONSUMO	
TRAMOMORANVO	WOE
0	0,42
1	-1,65
[2,3]	-3,90
AMAX6	WOE
0	0,55
[9,22]	-0,80
[23,42]	-1,42
[43,149]	-3,17
PAGOS3	WOE
[0,59000]	-1,50
[59036,103123]	-0,52
[103128,142940]	-0,22
[142991,236431]	0,01
[236482,503335]	0,27
[503337,719834]	0,46
[72000,7750306]	0,75
FACTURADO6	WOE
[-1129509,1455722]	0,07
[1455968,23225551]	-0,28
PSOBREF12	WOE
[0,21]	-1,45
[22,53]	-1,05
[55,77]	-0,59
[78,2700]	0,40
TRAMO_MORA	WOE
0	0,4
1	-1,6
[2,3]	-3,8
FACTURADO3	WOE
[-2042595,467376]	0,16
[467379,791919]	-0,03
[792042,1185279]	-0,33
[1185739,12062215]	-0,60
MAYORCUPO	WOE
[3000,138000]	-0,86
[139000,219200]	-0,58
[220000,449000]	-0,13
[540000,1000011]	0,17
[1000012,1499000]	0,50
[1500000,5500000]	0,79
DIASMORA	WOE
0	0,42
[9,29]	-1,61
[30,89]	-3,79
NPAGOS6	WOE
[0,3]	-1,62
4	-0,74
5	-0,19
6	0,70
CIACUP	WOE
[0,120000]	-0,91
[273000,120003]	-0,45
[273010,520000]	-0,02
[520021,963200]	0,27
[964000,1499000]	0,48
[1500000,5500000]	0,87
PAGOS12	WOE
[0,298270]	-0,63
[298272,445350]	-0,39
[445455,1145177]	-0,06
[1145361,2644049]	0,22
[2644150,42019190]	0,60
AMAX12	WOE
0	0,59
[9,12]	-0,41
[13,23]	-0,73
[24,45]	-1,22
[46,184]	-2,87

MODELO NORMAL	
TRAMOMORANVO	WOE
0	0,49
1	-1,22
[2,3]	-3,10
AMAX6	WOE
0	0,72
[9,20]	-0,40
[21,28]	-0,80
[29,55]	-1,68
[56,184]	-2,22
NPAGOS12	WOE
[0,10]	-0,21
11	0,94
12	1,63
PAGOS3	WOE
[0,47069]	-0,37
[47070,168862]	0,09
[168872,270000]	0,39
[270003,15836454]	0,78
FACTURADO6	WOE
[-8865711,23551]	0,57
[23552,260018]	0,15
[260023,24955051]	-0,24
FACTURADO12	WOE
[-12876596,31581]	0,39
[31583,604634]	0,04
[604661,53439902]	-0,15
PSOBREF12	WOE
[0,35]	-0,77
[36,56]	-0,42
[57,80]	0,10
[81,5223900]	0,81
TRAMO_MORA	WOE
0	0,49
1	-1,19
[2,3]	-3,03
FACTURADO3	WOE
[-4406556,15388]	0,79
[15394,390649]	0,00
[390651,20699372]	-0,38
MAYORCUPO	WOE
[0,99990]	-0,76
[100000,280000]	-0,37
[280001,347511]	-0,01
[348000,675000]	0,29
[675001,1140000]	0,64
[1140011,5500000]	1,09
AMAX3	WOE
0	0,64
[9,12]	-0,60
[13,23]	-0,92
[24,52]	-1,49
[53,184]	-2,59
PSOBREF3	WOE
[0,44]	-0,68
[45,72]	0,09
[73,100]	0,71
[101,25104900]	1,01
NPAGOS3	WOE
[0,1]	-0,60
2	-0,04
3	0,81
PAGOS12	WOE
[0,436065]	-0,16
[436071,1028223]	0,23
[1028249,81360689]	0,70

MODELO TRANSACCIONAL	
AMAX6	WOE
0	0,23
[9,20]	-1,08
[21,183]	-1,58
NPAGOS12	WOE
[0,8]	-0,67
9	-0,14
10	0,15
11	0,58
12	1,43
PAGOS3	WOE
[0,286159]	-0,32
[286194,481363]	-0,01
[481446,774986]	0,61
[775000,18320000]	1,64
FACTURADO6	WOE
[-601280,928112]	-0,16
[928189,1445022]	0,14
[1445856,18430728]	1,24
FACTURADO12	WOE
[-1377107,1216667]	-0,24
[1216686,1662918]	-0,01
[1663770,2682055]	0,39
[2687770,34075888]	1,51
PSOBREF12	WOE
[0,68]	-1,00
[69,86]	-0,56
[87,92]	-0,15
[93,1181475]	0,61
FACTURADO3	WOE
[-464050,367640]	-0,16
[367657,739120]	0,04
[740328,10925274]	1,17
MAYORCUPO	WOE
[0,200000]	-0,97
[200001,313000]	-0,65
[314000,432000]	-0,15
[433000,652000]	0,29
[653000,1139000]	0,78
[1140000,2053012]	1,72
[2060000,10000000]	2,55
AMAX3	WOE
0	0,09
[9,30]	-1,52
PSOBREF3	WOE
[0,49]	-0,96
[50,66]	-0,55
[67,78]	-0,24
[79,1181475]	0,22
NPAGOS3	WOE
[0,2]	0,06
3	0,04
NPAGOS6	WOE
[0,4]	-0,48
5	-0,16
6	0,53
PSOBREF6	WOE
[0,68]	-0,93
[69,79]	-0,49
[80,87]	-0,25
[88,99]	0,20
[100,1181475]	0,48

MODELO RENEGOCIACIONES	
TRAMOMORANVO	WOE
0	0,50
1	-0,42
2	-2,05
3	-3,23
AMAX6	WOE
[0,31]	0,24
[40,70]	0,07
[71,182]	-1,17
NPAGOS12	WOE
[0,5]	-0,61
6	-0,34
7	-0,02
[8,9]	0,44
[10,12]	1,04
PAGOS3	WOE
0	-1,96
[1500,28352]	-0,31
[28559,67726]	0,27
[68000,1300000]	0,59
FACTURADO6	WOE
[-565173,205195]	0,40
[205254,13792127]	-0,04
PSOBREF12	WOE
[0,26]	-0,26
[27,33]	-0,05
[34,62]	0,38
[63,200]	0,87
ANTRENE	WOE
[0,2]	0,22
3	-0,24
[4,46]	-0,65
FACTURADO3	WOE
[-701837,67990]	0,68
[68283,135366]	0,16
[135392,8258366]	-0,14
AMAX3	WOE
0	0,97
[9,59]	-0,11
[60,182]	-1,43
PSOBREF3	WOE
[0,9]	-0,41
[10,50]	0,02
[51,1000]	1,04
MONINFCOM	WOE
Sin Infocom	0,48
[0,112133]	-0,20
[112144,7146498]	-0,57
NPAGOS6	WOE
[0,1]	-1,32
2	-0,74
[3,4]	0,23
[5,6]	1,24
PSOBREF6	WOE
[0,16]	-0,25
[17,43]	0,23
[44,113]	0,74
PAGOS12	WOE
[0,162125]	-0,45
[162460,4840197]	0,19
AMAX12	WOE
[0,57]	0,14
[58,72]	-0,35
[73,182]	-0,89
MAYORCUPO	WOE
[0,404000]	-0,11
[405000,3204000]	0,26

MODELO CONSUMO	
PAGOS6	WOE
[0,156450]	-0,86
[156505,308258]	-0,31
[308270,985000]	0,08
[985059,1395792]	0,38
[1395907,21444300]	0,68
NUMIFOCOM	WOE
Sin Infocom	0,36
Con Infocom	-1,71
NPAGOS12	WOE
[0,8]	-0,96
9	-0,57
10	-0,24
11	0,12
12	0,83
PSOBREF6	WOE
[0,37]	-1,51
[38,66]	-0,76
[67,85]	0,01
[86,4200]	0,55
UTILIZACION	WOE
<91%	1,71
[91%-131%]	1,12
[131%-244%]	0,77
[244%-346%]	0,43
[346%-416%]	0,09
[416%-529%]	-0,25
>529%	-1,19

MODELO NORMAL	
AMAX12	WOE
0	0,73
[9,13]	-0,24
[14,41]	-0,52
[42,184]	-1,59
SALDO	WOE
[1,16883]	1,38
[16884,44610]	0,42
[44612,147931]	-0,04
[147932,10896388]	-0,30
NUMIFOCOM	WOE
Sin Infocom	0,398
Con Infocom	-1,26
PAGOS6	WOE
[0,178886]	-0,21
[178900,526036]	0,17
[526040,30748951]	0,76
UTILIZACION	WOE
<7%	1,71
[7%-23%]	1,12
[23%-42%]	0,77
[42%-63%]	0,43
[63%-81%]	0,09
[81%-95%]	-0,25
[95%-103%]	-0,64
>103%	-1,19

MODELO TRANSACCIONAL	
CTACUP	WOE
[30000,150000]	-1,19
[150013,304000]	-0,68
[305000,419010]	-0,14
[420000,636000]	0,36
[636012,1065000]	0,83
[1068000,2000001]	1,91
[2000011,10000000]	2,56
SALDO	WOE
[1,53320]	0,81
[53325,86559]	0,16
[86583,5417393]	-0,12
NUMIFOCOM	WOE
Sin Infocom	0,22
Con Infocom	-1,58
PAGOS6	WOE
[0,529298]	-0,38
[529382,679072]	-0,05
[679455,917970]	0,19
[918001,1503246]	0,63
[1503247,32730000]	1,88
UTILIZACION	WOE
<15%	1,80
[16%-27%]	1,21
[28%-41%]	0,61
[42%-58%]	-0,01
[59%-74%]	-0,33
[75%-88%]	-0,70
[89%-98%]	-0,99
99%+	-1,55

Matrices de correlaciones

Tabla de correlaciones modelo Consumo

Modelo Consumo	AMAX3	TRAMOMORAN	TRAMO_MO	DIASMORA	AMAX6	PSOBREF3	NPAGOS3	PSOBREF6	MONINFCOM	NUMIFOCOM	NPAGOS6	NPAGOS12	PSOBREF12	AMAX12	UTILIZACION	ANTCTA	PAGOS3	CTACUP	PAGOS6	MAYORCUPO	PAGOS12	FACTURADO3
AMAX3	1,000	0,856	0,853	0,851	0,901	0,450	0,509	0,397	0,544	0,538	0,419	0,304	0,311	0,788	0,121	0,066	0,278	0,155	0,162	0,125	0,115	0,116
TRAMOMORANVO	0,856	1,000	0,996	0,993	0,777	0,469	0,480	0,397	0,482	0,477	0,378	0,266	0,305	0,679	0,130	0,068	0,309	0,130	0,164	0,107	0,110	0,117
TRAMO_MORA	0,853	0,996	1,000	0,997	0,775	0,470	0,480	0,396	0,479	0,474	0,378	0,266	0,304	0,677	0,130	0,069	0,308	0,129	0,164	0,106	0,110	0,116
DIASMORA	0,851	0,993	0,997	1,000	0,773	0,469	0,480	0,396	0,477	0,472	0,378	0,266	0,304	0,677	0,130	0,069	0,308	0,130	0,164	0,107	0,110	0,116
AMAX6	0,901	0,777	0,775	0,773	1,000	0,412	0,460	0,385	0,514	0,509	0,430	0,323	0,313	0,868	0,113	0,061	0,263	0,160	0,161	0,136	0,118	0,100
PSOBREF3	0,450	0,469	0,470	0,469	0,412	1,000	0,435	0,798	0,288	0,286	0,346	0,253	0,682	0,377	0,115	0,110	0,349	0,112	0,209	0,093	0,145	0,325
NPAGOS3	0,509	0,480	0,480	0,480	0,460	0,435	1,000	0,319	0,278	0,275	0,788	0,580	0,249	0,416	0,100	0,093	0,443	0,160	0,265	0,140	0,200	0,011
PSOBREF6	0,397	0,397	0,396	0,396	0,385	0,798	0,319	1,000	0,259	0,256	0,314	0,231	0,816	0,350	0,113	0,099	0,248	0,103	0,196	0,081	0,142	0,355
MONINFCOM	0,544	0,482	0,479	0,477	0,514	0,288	0,278	0,259	1,000	0,992	0,234	0,172	0,207	0,460	0,093	0,048	0,168	0,144	0,100	0,112	0,070	0,074
NUMIFOCOM	0,538	0,477	0,474	0,472	0,509	0,286	0,275	0,256	0,992	1,000	0,231	0,170	0,204	0,456	0,093	0,046	0,165	0,142	0,098	0,111	0,069	0,075
NPAGOS6	0,419	0,378	0,378	0,378	0,430	0,346	0,788	0,314	0,234	0,231	1,000	0,770	0,243	0,389	0,106	0,101	0,405	0,193	0,337	0,171	0,276	-0,033
NPAGOS12	0,304	0,266	0,266	0,266	0,323	0,253	0,580	0,231	0,172	0,170	0,770	1,000	0,200	0,342	0,109	0,114	0,364	0,241	0,345	0,217	0,367	-0,093
PSOBREF12	0,311	0,305	0,304	0,304	0,313	0,682	0,249	0,816	0,207	0,204	0,243	0,200	1,000	0,308	0,089	0,086	0,207	0,092	0,161	0,071	0,145	0,325
AMAX12	0,788	0,679	0,677	0,677	0,868	0,377	0,416	0,350	0,460	0,456	0,389	0,342	0,308	1,000	0,111	0,055	0,273	0,189	0,177	0,154	0,135	0,071
UTILIZACION	0,121	0,130	0,130	0,130	0,113	0,115	0,100	0,113	0,093	0,093	0,106	0,109	0,089	0,111	1,000	0,256	0,020	0,348	0,018	0,328	0,044	0,038
ANTCTA	0,066	0,068	0,069	0,069	0,061	0,110	0,093	0,099	0,048	0,046	0,101	0,114	0,086	0,055	0,256	1,000	0,206	0,447	0,232	0,476	0,252	-0,134
PAGOS3	0,278	0,309	0,308	0,308	0,263	0,349	0,443	0,248	0,168	0,165	0,405	0,364	0,207	0,273	0,020	0,206	1,000	0,535	0,826	0,520	0,746	-0,413
CTACUP	0,155	0,130	0,129	0,130	0,169	0,112	0,160	0,103	0,144	0,142	0,193	0,241	0,092	0,189	0,348	0,447	0,535	1,000	0,596	0,933	0,625	-0,445
PAGOS6	0,162	0,164	0,164	0,164	0,161	0,209	0,265	0,196	0,100	0,098	0,337	0,345	0,161	0,177	0,018	0,232	0,826	0,596	1,000	0,585	0,859	-0,535
MAYORCUPO	0,125	0,107	0,106	0,107	0,136	0,093	0,140	0,081	0,112	0,111	0,171	0,217	0,071	0,154	0,328	0,476	0,520	0,933	0,585	1,000	0,616	-0,449
PAGOS12	0,115	0,110	0,110	0,110	0,118	0,145	0,200	0,142	0,070	0,069	0,276	0,367	0,145	0,135	0,044	0,252	0,746	0,625	0,859	0,616	1,000	-0,563
FACTURADO3	0,116	0,117	0,116	0,116	0,100	0,325	0,011	0,355	0,074	0,075	-0,033	-0,093	0,325	0,071	0,038	-0,134	-0,413	-0,445	-0,535	-0,449	-0,563	1,000

Tabla de correlaciones modelo Renegociado

Modelo Renegociado	SALDO	ANTCTA	TRAMOMORANVO	AMAX3	AMAX6	AMAX12	NPAGOS3	NPAGOS6	NPAGOS12	PAGOS3	PAGOS6	PAGOS12	FACTURADO3	FACTURADO6	PSOBREF3	PSOBREF6	PSOBREF12	ANTRENE	MAYORCUPO	RELCUOMYCUPO	PORCPAGCONV	MONINFCOM	NUMIFOCOM	DIASMORA
SALDO	1,000	-0,019	0,067	0,041	0,019	0,022	0,003	-0,056	-0,088	-0,069	-0,243	-0,238	0,375	0,445	0,032	0,025	0,020	0,020	-0,141	0,164	0,186	0,037	0,041	0,068
ANTCTA	-0,019	1,000	0,032	0,074	-0,024	-0,049	0,182	0,227	0,370	0,207	0,285	0,324	-0,024	-0,024	0,142	0,106	0,204	-0,109	0,379	0,245	-0,052	-0,007	0,010	0,011
TRAMOMORANVO	0,067	0,032	1,000	0,615	0,602	0,529	0,567	0,431	0,275	0,574	0,318	0,119	0,079	0,038	0,245	0,213	0,217	0,694	0,021	0,053	-0,048	0,272	0,272	0,948
AMAX3	0,041	0,074	0,615	1,000	0,627	0,558	0,463	0,299	0,235	0,379	0,189	0,113	0,294	0,055	0,477	0,169	0,180	0,290	0,038	0,020	0,089	0,314	0,326	0,595
AMAX6	0,019	-0,024	0,602	0,627	1,000	0,834	0,333	0,284	0,183	0,324	0,155	0,076	0,128	0,072	0,162	0,189	0,173	0,377	-0,005	-0,002	-0,036	0,167	0,172	0,551
AMAX12	0,022	-0,049	0,529	0,558	0,834	1,000	0,289	0,241	0,187	0,275	0,116	0,041	0,096	0,057	0,129	0,149	0,173	0,384	-0,029	-0,037	-0,057	0,128	0,129	0,513
NPAGOS3	0,003	0,182	0,567	0,463	0,333	0,289	1,000	0,731	0,516	0,834	0,480	0,259	0,080	-0,032	0,489	0,324	0,252	0,283	0,120	0,058	0,171	0,232	0,235	0,542
NPAGOS6	-0,056	0,227	0,431	0,299	0,235	0,379	0,731	1,000	0,644	0,603	0,617	0,352	-0,042	-0,090	0,314	0,397	0,297	0,173	0,164	0,049	0,087	0,138	0,134	0,413
NPAGOS12	-0,088	0,370	0,275	0,235	0,183	0,187	0,516	0,644	1,000	0,433	0,481	0,352	-0,052	-0,110	0,258	0,256	0,402	0,091	0,217	0,035	0,020	0,084	0,086	0,266
PAGOS3	-0,069	0,207	0,574	0,379	0,324	0,275	0,834	0,603	0,433	1,000	0,606	0,346	-0,069	-0,127	0,399	0,290	0,233	0,305	0,215	0,027	0,158	0,204	0,208	0,549
PAGOS6	-0,243	0,285	0,318	0,189	0,155	0,116	0,480	0,617	0,481	0,606	1,000	0,611	-0,307	-0,332	0,235	0,352	0,285	0,117	0,358	0,002	0,066	0,093	0,094	0,293
PAGOS12	-0,238	0,324	0,119	0,113	0,076	0,041	0,259	0,352	0,526	0,346	0,611	1,000	-0,233	-0,314	0,159	0,186	0,330	-0,004	0,329	0,024	0,019	0,005	0,015	0,108
FACTURADO3	0,375	-0,024	0,079	0,294	0,128	0,096	0,080	-0,042	-0,052	-0,069	-0,307	-0,233	1,000	0,511	0,384	0,069	0,055	-0,059	-0,191	0,151	0,218	0,145	0,152	0,084
FACTURADO6	0,445	-0,024	0,038	0,055	0,072	0,057	-0,032	-0,090	-0,110	-0,127	-0,332	-0,314	0,511	1,000	0,601	0,102	0,043	-0,012	-0,161	0,149	0,092	0,035	0,033	0,038
PSOBREF3	0,032	0,142	0,245	0,477	0,162	0,129	0,489	0,314	0,258	0,399	0,235	0,159	0,384	0,061	1,000	0,380	0,282	-0,113	0,101	0,063	0,369	0,221	0,236	0,250
PSOBREF6	0,025	0,106	0,213	0,169	0,189	0,149	0,324	0,397	0,256	0,290	0,352	0,186	0,069	0,102	0,390	1,000	0,496	0,006	0,051	0,036	0,323	0,064	0,075	0,206
PSOBREF12	0,020	0,204	0,217	0,180	0,173	0,173	0,252	0,297	0,402	0,233	0,285	0,330	0,055	0,043	0,282	0,496	1,000	0,112	0,100	0,050	0,119	0,093	0,104	0,218
ANTRENE	0,020	-0,109	0,694	0,290	0,377	0,384	0,283	0,173	0,091	0,305	0,117	-0,004	-0,059	-0,012	-0,113	0,006	0,112	1,000	-0,076	-0,005	-0,271	0,142	0,131	0,690
MAYORCUPO	-0,141	0,379	0,021	0,038	-0,005	-0,029	0,120	0,164	0,217	0,215	0,358	0,329	-0,191	-0,161	0,101	0,051	0,100	-0,076	1,000	0,316	-0,083	-0,017	-0,004	0,008
RELCUOMYCUPO	0,164	0,245	0,053	0,020	-0,002	-0,037	0,058	0,049	0,035	0,027	0,002	0,024	0,151	0,149	0,363	0,036	0,050	-0,005	0,316	1,000	0,015	-0,015	-0,002	0,050
PORCPAGCONV	0,186	-0,052	-0,048	0,089	-0,036	-0,057	0,171	0,087	0,020	0,158	0,066	0,019	0,218	0,092	0,369	0,323	0,119	-0,271	-0,083	0,015	1,000	0,092	0,107	-0,047
MONINFCOM	0,037	-0,007	0,272	0,314	0,167	0,128	0,232	0,138	0,084	0,204	0,093	0,005	0,145	0,035	0,221	0,064	0,093	0,142	-0,017	-0,015	0,092	1,000	0,935	0,259
NUMIFOCOM	0,041	0,010	0,272	0,326	0,172	0,129	0,235	0,134	0,086	0,208	0,094	0,015	0,152	0,033	0,236	0,075	0,104	0,131	-0,004	-0,002	0,107	0,935	1,000	0,256
DIASMORA	0,068	0,011	0,948	0,595	0,551	0,513	0,542	0,413	0,266	0,549	0,293	0,108	0,084	0,038	0,250	0,206	0,218	0,690	0,008	0,050	-0,047	0,259	0,256	1,000

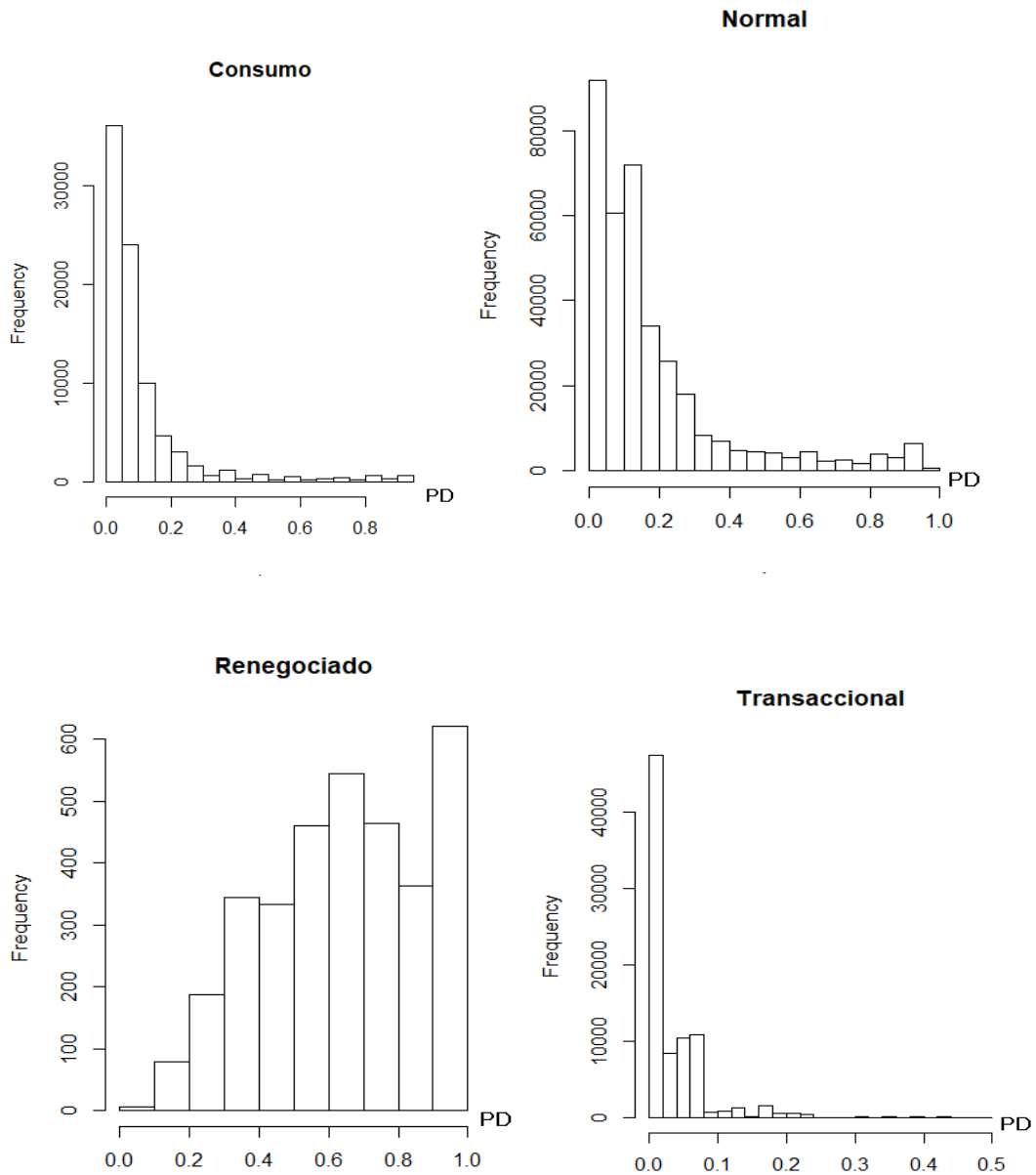
Tabla de correlaciones modelo Transaccional

Modelo transaccional	UTILIZACION	ANTICTA	PAGOS12	AMAX12	MONIFOCOM	SALDO	AMAX3	AMAX6	NPAGOS3	NPAGOS6	NPAGOS12	PAGOS3	PAGOS6	FACTURADOS3	FACTURADOS6	FACTURADOS12	PSOBREF3	PSOBREF6	PSOBREF12	MAYORCUPO	NUMFOCOM	CTACUP
UTILIZACION	1.000	0.484	0.321	0.182	0.209	0.381	0.095	0.156	0.030	0.071	0.248	0.225	0.258	0.079	0.104	0.160	0.204	0.242	0.305	0.621	0.208	0.649
ANTICTA	0.484	1.000	0.383	0.133	0.105	0.109	0.088	0.133	-0.002	0.044	0.339	0.277	0.299	0.212	0.226	0.294	0.113	0.150	0.229	0.588	0.105	0.575
PAGOS12	0.321	0.383	1.000	0.182	0.114	0.009	0.091	0.150	0.188	0.296	0.482	0.821	0.898	0.643	0.668	0.747	0.205	0.232	0.280	0.627	0.113	0.646
AMAX12	0.182	0.133	0.182	1.000	0.171	0.035	0.481	0.769	0.105	0.192	0.191	0.154	0.170	0.101	0.096	0.097	0.094	0.181	0.249	0.209	0.169	0.219
MONIFOCOM	0.209	0.105	0.114	0.171	1.000	0.045	0.180	0.172	0.042	0.057	0.066	0.100	0.105	0.061	0.067	0.071	0.068	0.098	0.114	0.189	0.995	0.199
SALDO	0.381	0.109	0.009	0.035	0.045	1.000	0.016	0.030	-0.011	-0.017	0.030	-0.003	-0.005	-0.232	-0.212	-0.228	0.152	0.146	0.166	0.088	0.045	0.084
AMAX3	0.095	0.088	0.091	0.481	0.180	0.016	1.000	0.626	0.177	0.150	0.111	0.080	0.087	0.043	0.047	0.052	0.089	0.119	0.122	0.114	0.178	0.118
AMAX6	0.156	0.133	0.150	0.769	0.172	0.030	0.626	1.000	0.112	0.222	0.189	0.123	0.142	0.081	0.074	0.081	0.087	0.195	0.215	0.180	0.170	0.186
NPAGOS3	0.030	-0.002	0.188	0.105	0.042	-0.011	0.177	0.112	1.000	0.737	0.465	0.238	0.192	0.102	0.101	0.198	0.287	0.287	0.150	0.089	0.041	0.100
NPAGOS6	0.071	0.044	0.296	0.192	0.057	-0.017	0.150	0.222	0.737	1.000	0.658	0.384	0.291	0.150	0.154	0.176	0.255	0.276	0.216	0.166	0.056	0.182
NPAGOS12	0.248	0.209	0.482	0.191	0.066	0.030	0.111	0.169	0.465	0.658	1.000	0.377	0.397	0.231	0.242	0.311	0.233	0.264	0.296	0.371	0.066	0.389
PAGOS3	0.225	0.277	0.821	0.154	0.100	-0.003	0.080	0.123	0.238	0.284	0.377	1.000	0.867	0.750	0.745	0.727	0.234	0.200	0.234	0.552	0.009	0.566
PAGOS6	0.258	0.299	0.898	0.170	0.105	-0.005	0.087	0.142	0.192	0.292	0.397	0.867	1.000	0.679	0.706	0.727	0.195	0.217	0.238	0.565	0.104	0.582
FACTURADOS3	0.079	0.212	0.643	0.101	0.061	-0.232	0.043	0.081	0.102	0.150	0.231	0.750	0.679	1.000	0.911	0.855	0.020	0.027	0.057	0.447	0.060	0.458
FACTURADOS6	0.104	0.226	0.658	0.096	0.067	-0.212	0.047	0.074	0.101	0.154	0.242	0.745	0.706	0.911	1.000	0.904	0.053	0.038	0.061	0.456	0.066	0.465
FACTURADOS12	0.160	0.294	0.747	0.097	0.071	-0.228	0.052	0.081	0.108	0.176	0.111	0.727	0.727	0.855	0.904	1.000	0.081	0.073	0.084	0.510	0.071	0.521
PSOBREF3	0.204	0.113	0.205	0.094	0.068	0.152	0.089	0.087	0.287	0.255	0.333	0.234	0.195	0.202	0.053	0.081	1.000	0.667	0.600	0.151	0.067	0.162
PSOBREF6	0.242	0.150	0.232	0.181	0.098	0.146	0.119	0.195	0.187	0.276	0.264	0.210	0.217	0.027	0.038	0.073	0.667	1.000	0.787	0.193	0.097	0.274
PSOBREF12	0.305	0.229	0.280	0.249	0.114	0.166	0.122	0.215	0.150	0.216	0.296	0.234	0.238	0.057	0.061	0.084	0.600	0.787	1.000	0.259	0.112	0.308
MAYORCUPO	0.621	0.588	0.627	0.209	0.189	0.088	0.114	0.180	0.089	0.166	0.371	0.552	0.565	0.447	0.456	0.510	0.151	0.193	0.259	1.000	0.188	0.930
NUMFOCOM	0.208	0.105	0.113	0.169	0.995	0.045	0.178	0.170	0.041	0.056	0.066	0.099	0.104	0.060	0.066	0.071	0.067	0.097	0.112	0.188	1.000	0.199
CTACUP	0.649	0.575	0.646	0.219	0.199	0.084	0.118	0.186	0.100	0.182	0.389	0.566	0.582	0.458	0.465	0.521	0.162	0.208	0.274	0.930	0.199	1.000

Tabla de correlaciones modelo Normal

Modelo Normal	MONIFOCOM	NUMFOCOM	DIASMORA	SALDO	ANTICTA	TRAMOMORANVO	UTILIZACION	AMAX3	AMAX6	AMAX12	NPAGOS3	NPAGOS6	NPAGOS12	PAGOS3	PAGOS6	PAGOS12	FACTURADOS3	FACTURADOS6	FACTURADOS12	PSOBREF3	PSOBREF6	PSOBREF12	MAYORCUPO	CTACUP	DIAS_MORA	TRAMO_MORA
MONIFOCOM	1.000	0.991	0.419	0.110	0.022	0.423	0.270	0.488	0.466	0.441	0.121	0.096	0.082	0.061	0.046	0.032	0.139	0.168	0.157	0.194	0.200	0.208	0.101	0.214	0.439	0.420
NUMFOCOM	0.991	1.000	0.415	0.110	0.023	0.420	0.268	0.484	0.463	0.438	0.120	0.096	0.081	0.060	0.045	0.032	0.138	0.167	0.156	0.192	0.198	0.206	0.101	0.211	0.415	0.417
DIASMORA	0.419	0.415	1.000	0.119	0.085	0.979	0.269	0.806	0.706	0.609	0.297	0.220	0.155	0.187	0.111	0.078	0.158	0.160	0.130	0.335	0.311	0.284	0.119	0.174	1.000	0.976
SALDO	0.110	0.110	0.119	1.000	0.077	0.117	0.639	0.098	0.086	0.077	-0.259	-0.219	-0.168	-0.381	-0.321	-0.287	0.528	0.523	0.447	0.022	0.041	0.084	-0.098	-0.065	0.119	0.117
ANTICTA	0.022	0.023	0.085	0.077	1.000	0.092	0.181	0.068	0.063	0.004	0.154	0.243	0.286	0.181	0.248	0.320	-0.020	-0.102	-0.180	0.108	0.137	0.143	0.509	0.316	0.085	0.092
TRAMOMORANVO	0.423	0.420	0.979	0.117	0.092	1.000	0.266	0.797	0.704	0.615	0.297	0.221	0.157	0.187	0.112	0.080	0.155	0.154	0.124	0.330	0.306	0.279	0.123	0.175	0.979	0.995
UTILIZACION	0.270	0.268	0.269	0.639	0.181	0.266	1.000	0.727	0.727	0.285	0.307	-0.080	-0.056	-0.019	-0.159	-0.125	-0.101	0.367	0.370	0.316	0.151	0.182	0.247	0.218	0.392	0.269
AMAX3	0.488	0.484	0.806	0.098	0.068	0.797	0.271	1.000	0.864	0.743	0.317	0.257	0.190	0.142	0.100	0.074	0.198	0.208	0.178	0.323	0.322	0.304	0.131	0.226	0.806	0.797
AMAX6	0.466	0.463	0.706	0.086	0.063	0.704	0.285	0.864	1.000	0.845	0.265	0.267	0.210	0.127	0.101	0.075	0.183	0.231	0.211	0.277	0.328	0.324	0.141	0.271	0.706	0.706
AMAX12	0.441	0.438	0.609	0.077	0.004	0.615	0.307	0.743	0.845	1.000	0.193	0.188	0.212	0.110	0.091	0.068	0.164	0.211	0.251	0.217	0.253	0.329	0.191	0.325	0.609	0.613
NPAGOS3	0.121	0.120	0.297	-0.259	0.154	0.297	-0.080	0.317	0.265	0.193	1.000	0.743	0.503	0.555	0.378	0.301	-0.388	-0.399	-0.398	0.457	0.377	0.299	0.175	0.150	0.297	0.297
NPAGOS6	0.096	0.096	0.220	-0.219	0.243	0.221	-0.056	0.257	0.267	0.188	0.743	1.000	0.703	0.446	0.408	0.379	-0.293	-0.351	-0.359	0.314	0.236	0.250	0.225	0.201	0.220	0.222
NPAGOS12	0.082	0.081	0.155	-0.168	0.286	0.157	-0.019	0.190	0.210	0.212	0.503	0.703	1.000	0.359	0.372	0.404	-0.208	-0.260	-0.279	0.230	0.238	0.235	0.265	0.242	0.155	0.157
PAGOS3	0.061	0.060	0.187	-0.381	0.181	0.187	-0.159	0.142	0.127	0.110	0.555	0.446	0.359	1.000	0.779	0.678	-0.508	-0.594	-0.544	0.377	0.300	0.248	0.416	0.366	0.187	0.186
PAGOS6	0.046	0.045	0.111	-0.321	0.248	0.112	-0.125	0.100	0.101	0.091	0.378	0.428	0.372	0.779	1.000	0.835	-0.431	-0.557	-0.534	0.213	0.230	0.196	0.503	0.439	0.111	0.112
PAGOS12	0.032	0.032	0.078	-0.287	0.320	0.080	-0.101	0.074	0.075	0.068	0.301	0.379	0.404	0.678	0.835	1.000	-0.390	-0.511	-0.523	0.149	0.165	0.175	0.552	0.466	0.078	0.080
FACTURADOS3	0.139	0.138	0.158	0.538	-0.000	0.155	0.367	0.198	0.183	0.164	-0.388	-0.351	-0.308	-0.508	-0.431	-0.390	1.000	0.727	0.691	-0.033	0.007	0.045	-0.187	-0.148	0.158	0.155
FACTURADOS6	0.168	0.167	0.160	0.523	-0.102	0.154	0.370	0.208	0.231	0.211	-0.398	-0.359	-0.320	-0.544	-0.511	-0.511	0.727	1.000	0.821	-0.029	0.001	0.037	-0.261	-0.192	0.160	0.154
FACTURADOS12	0.157	0.156	0.130	0.447	-0.180	0.124	0.316	0.178	0.211	0.211	-0.398	-0.359	-0.279	-0.544	-0.511	-0.523	0.691	0.821	1.000	-0.066	-0.064	-0.022	-0.294	-0.190	0.130	0.124
PSOBREF3	0.194	0.192	0.335	0.022	0.108	0.330	0.151	0.322	0.277	0.217	0.457	0.314	0.230	0.377	0.213	0.149	-0.033	-0.029	-0.066	1.000	0.786					

Gráficos de las probabilidades de default por segmento



Referencias

Altman, E. (1960). Altman Z-score.

Altman, E. I. (1968), "Financial Ratios, Discriminant Analysis and the Prediction of Corporate Bankruptcy." *Journal of Finance*, September, pp. 589-609.

Altman, E. I. (2010), *The Z-Metrics Methodology For Estimating Company Credit Ratings And Default Risk Probabilities*. RiskMetrics Group.

Beaver, W. (1967), "Financial Ratios as Predictors of Failures. Empirical Research in Accounting: Selected Studies – 1966." In *Supplement to Journal of Accounting Research* 4, pp. 71-111.

Black, F., & Scholes, M. (1973). The Pricing of Options and Corporate Liabilities. *The Journal of Political Economy*, Vol. 81, 637-654.

Calixto , M., & Casaverde , L. (2011). Variables determinantes de la probabilidad de incumplimiento de un microcrédito en una entidad microfinanciera del Perú, una aproximación bajo el modelo de regresión logística binaria.

Chatterjee S. (2016). Modelos del riesgo de crédito. *Boletín Cemla*. Julio-Septiembre. Autorizado por el Handbook, núm. 34, del Centre for Central Banking Studies, Banco de Inglaterra.

Cinca, C. (1993). Predicción de la quiebra bancaria mediante el empleo de redes neuronales artificiales.

- Correa, A. (2016). Prediction of clinical deterioration in hospitalized adult patients with hematologic malignancies using a neural network model. .
- Durand, B. (1941), "Risk Elements in Consumer Installments Financing." Working Paper, NBER.
- Fernández, H., & Pérez, F. (2005). El modelo logístico: una herramienta estadística para evaluar el riesgo de crédito. Revista Ingenierías Universidad de Medellín, 4(6),55-75. [fecha de Consulta 12 de junio de 2021]. ISSN: 1692-3324. Disponible en: <https://www.redalyc.org/articulo.oa?id=75040605>
- Fisher, R. (1936), "The Use of Multiple Measurements in Taxonomic Problems." Annals of Eugenics 7, pp. 179-188.
- Gárate, J. (2016). Relación entre Gestión del Riesgo Crediticio y Morosidad en clientes del segmento empresa del BBVA Continental.
- García, C. (2017). RIESGO FINANCIERO: UNA APROXIMACIÓN CUALITATIVA AL INTERIOR DE LAS PYMES .
- Garcia, G. (2015). Gestión de riesgo crediticio en entidades financieras PICP. Universidad Abierta Interamericana. Rosario. Argentina.
- González, B. (2015). Desarrollo de un modelo estadístico que permita estimar la probabilidad de incumplimiento de los clientes que teniendo experiencia crediticia, solicitan cupo de tarjeta de crédito. Fundación Universitaria Los Libertadores. Bogotá. Colombia.
- Gupton, N. (1997). A Model for Preparing Automotive Technicians: Training for a Qualified Workforce.
- Gurný, P., & Gurný, M. (2013). Comparison of credit scoring models on probability of default estimation for us banks.

- Gurný, P., Gurný, M. (2009), "Estimation of PD of Financial Institutions within Linear Discriminant Analysis." *Mathematical Methods in Economics*
- Hachuel, L., Boggio, G., & Harvey, G. (2012). Estudio comparativo de Métodos de Estimación en un Modelo Logit Mixto.
- Hernández-Sampieri, Fernández-Collado, & Pilar-Baptista. (2014). *Metodología de la Investigación*. Santiago: Mc Graw Hill.
- Jarrow, R. A., & Protter, P. (2004). Structural versus reduced-form models: a new information based perspective. *Journal of Investment Management*, 2(2), 1-10. En línea. Disponible en <http://forum.johnson.cornell.edu/faculty/jarrow/100%20Structural%20vs%20Reduced%20Review%20JIM%202004.pdf>
- Karacula, E. (2009). An application of logit and probit models over the default probabilities of retail banking mortgage credits .
- Lando, S. (2004). *Graphs on surfaces and their applications* .
- Lennox, C. (1999). Are large auditors more accurate than small auditors?
- Majnoni, G., Miller, M., Mylenko, N., & Powell, A. (2004). Improving credit information, bank regulation and supervision: On the role and design of public credit registries. WORKING PAPER, World Bank Policy Research.
- McFadden, D. (1974). The measurement of urban travel demand.
- Merton R. (1974). On the Pricing of Corporate Debt-The risk Structure of Interest Rates, *The Journal of Finance*, Vol. 29, Issue 2, Papers and Proceedings of the thirty-second Annual Meeting of the American Finance Association, New York.

- Miller, M. & Rojas, D. (2005). Improving access to credit for SMEs: An empirical analysis of the Feasibility of pooled data small business credit scoring models in Colombia y Mexico. WORKING PAPER, World Bank, Washington,
- Moreno, D. (2013). Estudio de un kit de diseño mixto con Softcore Microblaze e IP's, para el control ESP, ASR y ABS.
- Ohlson, J. (1980). Financial ratios and the probabilistic prediction of bankruptcy.
- Pailhé, C. (2006). Sistemas de información para la administración del riesgo de crédito. Relevamiento en el sistema financiero argentino. WORKING PAPER, Central Bank, October.
- Pastena, V., & Ruland, W. (1986). The merger/bankruptcy alternative.
- Peretsky, A., Karminsky, A. (2008), "Models for Moody's Bank Rating." BOFIT Discussion Papers 17.
- Quitral, M. (2012). La crisis subprime y los objetivos del milenio en América Latina. Política y Cultura, primavera, núm. 37, pp. 11-33
- Ríos, L. (2008). Evolución de la Teoría Financiera en el Siglo XX.
- Salamanca, A. & Benitez, J. (2018). Estimación De La Probabilidad De Incumplimiento De Créditos Para Una Empresa Del Sector Siderúrgico En Colombia. Universidad Escuela de Administracion, Finanzas e Instituto Tecnológico (EAFIT).
- Salas, J. (1998). Determinantes de la morosidad .
- Saldaña, J., Palomo, M., & Blanco, M. (2017). Los Modelos CAPM y APT para la valuación de empresas de Telecomunicaciones con parámetros operativos. Innovaciones de negocios, 4(8).

- Sánchez, I. (2007). Análisis del riesgo crediticio bancario en la economía cubana.
- Sunder, S. (1979). Methodological issues in the use of financial ratios.
- Thomas, L., Crook, J., & Edelman, D. (2017). Credit scoring and its applications. Society for industrial and Applied Mathematics.
- Vargas, A., & Mostajo, S. (2014). Medición del riesgo crediticio mediante la aplicación de métodos basados en calificaciones internas. *Investigación & Desarrollo*, 2(14), 5-25.
- Vela, S. & Caro, A. (2015). Herramientas financieras en la evaluación del riesgo de crédito. Fondo editorial Universidad Inca Garcilaso de la Vega.
- Wooldridge, M. (2009). An introduction to multiagent systems.
- Rossi, G. (2013). La volatilidad en mercados financieros y de commodities un repaso de sus causas y la evidencia reciente.