

Determinantes de la pérdida esperada en la cartera de clientes de una institución prestadora de servicios de salud

Carlos Julio Correa Barragán¹

Resumen

La estimación de la pérdida esperada en las cuentas por cobrar de una empresa y la modelación de las variables que la determinan constituyen una importante herramienta de medición de riesgo, con efecto directo en el desempeño del patrimonio. En este trabajo se realiza la estimación mediante un modelo econométrico logit para datos agrupados, en una institución prestadora de servicios de salud ubicada en Colombia. Se presenta una selección de cuatro modelos basados en la metodología logit para datos agrupados, encontrando que la pérdida esperada en las cuentas por cobrar de la entidad está condicionada por diez variables, que incluyen indicadores financieros y características propias de cada cliente, al igual que indicadores operativos de la cartera analizada. Los resultados obtenidos muestran un adecuado ajuste y significancia global, al igual que índices de capacidad predictiva superiores al 89%. De manera complementaria, a partir del modelo se determinan focos de gestión en los procesos operativos, comerciales y contractuales que permitan optimizar este riesgo para la entidad.

Palabras clave: cartera en servicios de salud, modelo *logit* datos agrupados, pérdida esperada, deterioro de instrumentos financieros

Abstract

The estimate of the expected loss on the receivable accounts of a company and the modeling of the variables that determine it, are an important tool for measuring risk with direct effect on the performance of owner's equity. In this paper makes the estimate, using an econometric model logit for grouped data, in an institution providing health services; located in Colombia. This paper present a choice of four models based on logit methodology for grouped data, stating that the expected loss on accounts receivable for the company are conditioned by ten variables that include financial indicators and own characteristics of each client as well as indicators operatives of the portfolio that is analyzed. The results show a proper fit and an overall significance as too predictive capabilities rates superior to 89%. In a complementary way, to optimize this risk for the entity are determined foci of work on the operational, commercial and contractual processes from the model.

¹ Ingeniero Industrial de la Universidad del Valle, Especialista en Finanzas de la Universidad EAFIT, candidato a Magíster en Administración Financiera de la Universidad EAFIT. Correo electrónico: cjcorrea79@hotmail.com

Asesor: Alfredo Trespalacios Carrasquilla. Ingeniero Electricista de la Universidad Nacional de Colombia. MSc Finanzas de la Universidad EAFIT. Profesional de Mercados EPM, profesor de la Maestría en Administración Financiera de la Universidad EAFIT, de la Especialización en Estadística Aplicada en la Escuela de Ingeniería de Antioquia, y de Ingeniería Financiera en el Instituto Tecnológico Metropolitano. Correo electrónico: alfredo.trespalacios@gmail.com

Keywords: portfolio in health services, logit model grouped data, expected loss, deterioration of financial instruments

Introducción

En el mundo empresarial, como resultado de las transacciones de compraventa de productos y servicios se generan cuentas por cobrar, que representan para su poseedor el derecho a recibir flujos de caja futuros. Dependiendo de diversos factores, las cuentas por cobrar pueden sufrir un deterioro, representado, o por la probabilidad de que la deuda sea atendida, o por la diferencia entre el valor inicial de la obligación y el valor esperado como recaudo real. Con la adopción en Colombia de las Normas Internacionales de Información Financiera NIIF (IASB, 2005), entre ellas la NIC 39 y la NIIF 9, se hace de obligatorio cumplimiento la estimación del deterioro de los activos financieros, como es el caso de las cuentas por cobrar.

Los modelos para cuantificar el deterioro de la cartera han sido abordados desde los años sesenta. Trabajos como el de Altman (1968) plantean la metodología Z-score, para predecir la quiebra de las empresas por medio de un análisis discriminante y de cinco indicadores financieros. Posteriormente, Altman, Haldeman & Narayanan (1977) modifican el modelo original incluyendo otros indicadores financieros, tales como el retorno sobre activos, la estabilidad en utilidades (error estándar del ROA), el servicio de deuda y los indicadores de capitalización, entre otros. Wiginton (1980) realiza un análisis comparativo entre los modelos logísticos y el análisis discriminante para predecir el comportamiento de pagos de créditos, encontrando mejores resultados con los modelos logit estimados por máxima verosimilitud en cuanto a la precisión de los resultados. Kolesar & Showers (1985) realizan una comparación de metodologías de análisis de variables binarias, entre ellas el modelo discriminatorio lineal enfocado en la predicción del comportamiento de créditos, destacando en este su robustez, efectividad y flexibilidad.

Para los últimos diez años, Mures, García y Vallejo (2005) realizan una comparación entre análisis discriminante y modelo logístico en el estudio de la morosidad en entidades financieras, encontrando que ambas son adecuadas, sin llegar a conclusiones contundentes acerca de cuál metodología es mejor. Támara, Aristizábal y Velásquez (2010), Fernández y Pérez (2005), López

(2013), Salazar (2013) y Vélez (2009) presentan la estimación de provisiones por pérdida esperada en carteras crediticias bancarias, para los cuatro primeros casos, y en una empresa funeraria, en el quinto caso, a través de modelos logit para la variable binaria de pago, encontrando en la aplicación del modelo resultados significativos y similares, al compararlos con modelos como probit, log-log y z-score. Ochoa, Agudelo y Galeano (2010) construyen un modelo de *scoring* para otorgamiento de crédito en una entidad financiera, mediante análisis discriminatorio, catalogado como consistente y con un alto valor predictivo. Bolton (2010); Rayo, Lara y Camino (2010); Moreno (2013); y Rivera (2011) presentan modelos de *credit scoring* en entidades financieras y cooperativas mediante modelos logit, encontrando adecuados niveles de predicción y un ajuste aceptable en cuanto a su validación. Arrobo y Montoya (2014) evalúan el riesgo crediticio de cartera hipotecaria a través de cadenas de Markov, encontrando que estas permiten generar alertas tempranas sobre el deterioro financiero de una cartera de créditos. Sandoval (2015) propone un modelo de gestión de cartera para empresas de medicina prepagada, a través del modelo de Markowitz; sin embargo, no propone un modelo de estimación de pérdida esperada.

Estos autores estudian la cartera de deudores como variables dicotómicas individuales, cuyo resultado es el pago (valor 1) o el no pago (valor 0), para la obligación de un cliente, y por consiguiente estiman la probabilidad de que la variable dicotómica tome uno u otro valor, condición que no se ajusta totalmente al caso objeto de este trabajo, por cuanto en las instituciones prestadoras de servicios de salud cada cliente tiene un gran número de obligaciones (facturas) con la entidad, y su comportamiento de pago puede ser diferente para cada una de ellas, presentándose incluso casos en los cuales la obligación es parcialmente atendida. Bajo estas condiciones, se hace necesario proponer un modelo que evalúe datos de manera agrupada, estimando la proporción esperada de recuperación del conjunto de obligaciones que posee cada cliente con la entidad.

El problema se aborda mediante la aplicación del modelo logit, encontrando como variables más representativas tres indicadores financieros de liquidez y rentabilidad, tres características propias del cliente y cuatro indicadores operativos propios de la cartera analizada. Igualmente, se encuentra que las variables que describen el subsector al cual pertenece el cliente no son significativas, mientras que los resultados para las variables de antigüedad de la cartera no son

contendientes; sin embargo, no se descarta el uso de estas variables para investigaciones posteriores, al considerarse representativas por experiencia previa del autor.

La investigación permite realizar recomendaciones en la gestión de las cuentas por cobrar de la entidad objeto de estudio, en cuanto a puntos de control y estructura en sus procesos operativos, tanto de facturación como de la gestión de cobro, al igual que aspectos comerciales y contractuales con sus clientes, incluyendo la verificación de indicadores financieros como parte de la formalización de convenios.

Como caso de estudio se tomó una clínica privada ubicada en Colombia, en el departamento del Valle del Cauca. Se aborda, inicialmente, el modelo econométrico logit y su tratamiento para datos agrupados y, a continuación, la aplicación al caso de estudio, el análisis de resultados, la interpretación y las conclusiones.

Default, pérdida esperada y deterioro

Los estudios de Arrobo y Montoya (2014) definen la probabilidad de incumplimiento (*default*) descomponiéndola en dos elementos: el *default real*, que tiene lugar cuando la altura de la mora crece de manera indefinida, y para el cual se hacen efectivas las garantías, llegando hasta las últimas instancias judiciales; sin embargo, esta definición no es práctica para efectos de cálculos de provisiones, puesto que los tiempos de resolución varían de un caso a otro, dependiendo de varios factores. El segundo elemento es el *default técnico*, basado en una altura de mora por encima de la cual se considera que la obligación no será atendida en las condiciones convenidas. Este último enfoque también es utilizado en el trabajo de Támara, Aristizábal y Velásquez (2010) quienes definen la pérdida esperada como el monto de capital que podría perder una institución como resultado de la exposición crediticia en un horizonte de tiempo dado. Finalmente, Moreno (2013) define el *default* en términos de la mencionada altura de la mora, estimada en 90 días, según lo establecido en el Capítulo II de la Circular Externa 100 de la Superintendencia Financiera de Colombia (2011) y en el Acuerdo Capital de Basilea II (2004).

Estos conceptos permiten definir el deterioro como la pérdida esperada del valor de un instrumento financiero en un horizonte de tiempo definido, medido en términos del monto por el cual el instrumento será realizado en comparación con el valor inicial de la transacción.

En cuanto al deterioro de los activos financieros, la NIC 39 (IASB, 2005) hace referencia a las características de los modelos de pérdida histórica, a través de diferentes variables que describan el comportamiento, y no solamente a través de la morosidad, mientras la NIIF 9 (IFRS Foundation, 2012) define como objetivo de los requerimientos del deterioro de valor:

Reconocer las pérdidas crediticias esperadas durante el tiempo de vida del activo de todos los instrumentos financieros para los cuales ha habido incrementos significativos en el riesgo crediticio desde el reconocimiento inicial —evaluado sobre una base colectiva o individual— considerando toda la información razonable y sustentable, incluyendo la que se refiera al futuro.

En la tabla 1 se presentan las variables explicativas estudiadas por otros autores, para estimar el *default*.

Tabla 1. Variables estudiadas por autores para estimar el *default*

CU: Cuantitativa, CL: Cualitativa, D: Discreta, C: Continua, B: Binaria

Variables	CU	CL	D	C	B	Autor
Activos	X			X		Aristizábal, Támara & Velásquez (2010), Rivera (2011), Vélez (2009)
Administración – Estructura		X	X			Fernández & Pérez (2005)
Administración – Experiencia		X	X			Fernández & Pérez (2005)
Antigüedad como cliente	X		X			Agudelo, Galeano & Ochoa (2010), Camino, Lara & Rayo (2010)
Antigüedad de la empresa	X		X			Fernández & Pérez (2005)
Antigüedad laboral	X		X			Agudelo, Galeano & Ochoa (2010); Moreno (2013); Wiginton (1980)
Apalancamiento (total pasivo/total patrimonio)	X			X		Camino, Lara & Rayo (2010), Vélez (2009)
Calificación crediticia		X			X	Agudelo, Galeano & Ochoa (2010), Camino, Lara & Rayo (2010), Moreno (2013), Rivera (2011)
Capacidad de pago	X		X			Agudelo, Galeano & Ochoa (2010)
Capital	X			X		Rivera (2011)
Capital de trabajo/activo corriente	X			X		Altman (1968)
Cargas	X		X			López (2013), Moreno (2013), Rivera (2011)
Cartera vencida	X			X		Rivera (2011)
Costos	X			X		Vélez (2009)
Crecimiento del activo	X			X		Fernández & Pérez (2005)
Crecimiento en utilidades	X			X		Fernández & Pérez (2005)

Tabla 1. Variables estudiadas por autores para estimar el *default* (continuación)

Variab les	CU	CL	D	C	B	Autor
Crecimiento en ventas	X			X		Fernández & Pérez (2005)
Créditos concedidos con anterioridad	X		X			Camino, Lara & Rayo (2010), Moreno (2013)
Créditos negados con anterioridad	X		X			Camino, Lara & Rayo (2010)
Cuentas por cobrar	X			X		Vélez (2009)
Cuentas por pagar	X			X		Vélez (2009)
Cuota mensual	X			X		Vélez (2009)
Cuotas por pagar	X		X			Rivera (2011)
Cupo en millones de pesos	X			X		Fernández & Pérez (2005)
Dependencia o endeudamiento (total pasivo/(total pasivo + total patrimonio))	X			X		Camino, Lara & Rayo (2010)
Duración retrasos	X		X			García, Mures & Vallejo (2005)
Edad	X		X			Agudelo, Galeano & Ochoa (2010), Aristizábal, Támara & Velásquez (2010), Camino, Lara & Rayo (2010), Moreno (2013), Rivera (2011)
Egresos	X			X		Rivera (2011)
Tipo de vinculación laboral		X			X	Vélez (2009)
Endeudamiento (pasivo total/activo total)	X			X		Aristizábal, Támara & Velásquez (2010), Moreno (2013), Vélez (2009)
Endeudamiento financiero /Ventas	X			X		Fernández & Pérez (2005)
Endeudamiento total/ventas,	X			X		Fernández & Pérez (2005)
Estado civil		X	X		X	Agudelo, Galeano & Ochoa (2010), Camino, Lara & Rayo (2010), García, Mures & Vallejo (2005), Moreno (2013), Rivera (2011)
Estrato social		X	X			Agudelo, Galeano & Ochoa (2010), Moreno (2013)
Forma de pago		X			X	Agudelo, Galeano & Ochoa (2010).
Garantía		X			X	Agudelo, Galeano & Ochoa (2010), Camino, Lara & Rayo (2010), García, Mures & Vallejo (2005), Moreno (2013), Rivera (2011)
Gastos generales	X			X		Vélez (2009)
Género		X			X	Agudelo, Galeano & Ochoa (2010), Camino, Lara & Rayo (2010), Moreno (2013), Rivera (2011), Salazar (2013)
Histórico de morosidad	X		X			Arrobo & Montoya (2014), Rivera (2011), Camino, Lara & Rayo (2010)
Impagos anteriores	X		X			García, Mures & Vallejo (2005), López (2013)
Ingreso total	X		X			Agudelo, Galeano & Ochoa (2010), Aristizábal, Támara & Velásquez (2010), López (2013), Moreno (2013), Rivera (2011), Vélez (2009)
Inventarios	X			X		Vélez (2009)

Tabla 1. Variables estudiadas por autores para estimar el *default* (continuación)

Variab les	CU	CL	D	C	B	Autor
IPC	X			X		Camino, Lara & Rayo (2010)
Liquidez (capacidad de pago/activo total)	X			X		Camino, Lara & Rayo (2010), Vélez (2009)
Lugar de solicitud		X	X			Agudelo, Galeano & Ochoa (2010), Camino, Lara & Rayo (2010), Moreno (2013), Rivera (2011)
Margen bruto	X			X		Fernández & Pérez (2005), Vélez (2009)
Margen neto	X			X		Aristizábal, Támara & Velásquez (2010), Vélez (2009)
Margen operativo (utilidad operativa/ingresos)	X			X		Aristizábal, Támara & Velásquez (2010)
Monto	X		X			Agudelo, Galeano & Ochoa (2010), Camino, Lara & Rayo (2010), Moreno (2013), Salazar (2013)
Nivel de regulación		X	X			Fernández & Pérez (2005)
Nivel de riesgo ambiental		X	X			Fernández & Pérez (2005)
Nivel educativo		X	X			Agudelo, Galeano & Ochoa (2010), Moreno (2013), Rivera (2011)
Nueva residencia		X			X	García, Mures & Vallejo (2005), López (2013)
Número de meses transcurridos desde el último crédito	X		X			Moreno (2013)
Obligaciones financieras	X			X		Vélez (2009)
Ocupación		X	X			Agudelo, Galeano & Ochoa (2010), Aristizábal, Támara & Velásquez (2010), López (2013), Moreno (2013), Rivera (2011), Wiginton (1980)
Otros egresos	X			X		Vélez (2009)
Otros ingresos	X			X		Vélez (2009)
Otros ingresos/utilidad neta	X			X		Fernández & Pérez (2005)
Pasivo	X			X		Rivera (2011), Vélez (2009)
Pasivo/ingreso	X			X		Aristizábal, Támara & Velásquez (2010)
Patrimonio	X			X		Rivera (2011), Vélez (2009)
Patrimonio al valor de mercado/pasivos	X			X		Altman (1968).
Personas a cargo	X		X			Agudelo, Galeano & Ochoa (2010).
PIB	X			X		Camino, Lara & Rayo (2010)
Plan exequial		X	X			Vélez (2009)
Plazo	X		X			Agudelo, Galeano & Ochoa (2010), Camino, Lara & Rayo (2010), Moreno (2013), Rivera (2011), Salazar (2013)
Presencia de otras deudas		X			X	Salazar (2013)

Tabla 1. Variables estudiadas por autores para estimar el *default* (continuación)

Variab les	CU	CL	D	C	B	Autor
Productos		X	X			Fernández & Pérez (2005)
Promedio de mora últimos dos años	X			X		Moreno (2013)
Pronóstico del analista sobre la situación del crédito a su vencimiento		X			X	Camino, Lara & Rayo (2010)
Proveedores	X			X		Vélez (2009)
Prueba ácida	X			X		Vélez (2009)
Reestructuraciones anteriores		X			X	Agudelo, Galeano & Ochoa (2010)
Referencias de socios/administradores		X	X			Fernández & Pérez (2005)
Rentabilidad sobre activos (ROA)	X			X		Altman (1968), Camino, Lara & Rayo (2010), Fernández & Pérez (2005)
ROE (utilidad neta/patrimonio)	X			X		Camino, Lara & Rayo (2010), Fernández & Pérez (2005)
Rotación de activo total	X			X		Vélez (2009)
Rotación de cartera	X			X		Fernández & Pérez (2005), Vélez (2009)
Rotación de inventarios	X			X		Vélez (2009)
Rotación de proveedores	X			X		Fernández & Pérez (2005)
Rotación liquidez (capacidad de pago/ingresos por ventas x 360)	X			X		Camino, Lara & Rayo (2010)
Saldo capital	X			X		Fernández & Pérez (2005), Rivera (2011)
Sector económico		X	X			Camino, Lara & Rayo (2010), Moreno (2013), Vélez (2009)
Tamaño de la empresa		X	X			Vélez (2009)
Tasa de cambio	X			X		Camino, Lara & Rayo (2010)
Tasa de empleo	X			X		Camino, Lara & Rayo (2010)
Tasa de interés	X			X		Salazar (2013)
Tasa de variación tasas de interés	X			X		Camino, Lara & Rayo (2010)
Tasa de variación de precios servicios públicos	X			X		Camino, Lara & Rayo (2010)

Tabla 1. Variables estudiadas por autores para estimar el *default* (continuación)

Tipo de contrato laboral		X	X		Agudelo, Galeano & Ochoa (2010), Camino, Lara & Rayo (2010)
Tipo de crédito		X		X	Camino, Lara & Rayo (2010), López (2013), Salazar (2013)
Tipo de vivienda		X	X		Agudelo, Galeano & Ochoa (2010), Moreno (2013), Rivera (2011)
Total empleados	X			X	Vélez (2009)
Utilidad	X			X	Vélez (2009)
Utilidad bruta/costos	X			X	Camino, Lara & Rayo (2010)
Utilidades antes de intereses e impuestos/activos totales	X			X	Altman (1968)
Utilidades retenidas/activos totales	X			X	Altman (1968)
Ventas/activos totales	X			X	Altman (1968), Camino, Lara & Rayo (2010)

Fuente: elaboración propia.

Planteamiento del modelo

Selección del modelo

Para el caso de este estudio, como modelo econométrico se utiliza el logit con datos agrupados, puesto que, a diferencia de las investigaciones encontradas, en las cuales se plantean modelos donde la variable dependiente toma el valor cero o uno (variable binaria), se plantea la variable dependiente como la proporción total de la cartera que es recaudada, con base en un volumen de facturas por cada cliente, siendo esta una variable acotada entre cero y uno.

El modelo logit para datos agrupados

Los estudios presentados por Gujarati y Porter (2010), Medina (2003) y Wiginton (1980) muestran en los modelos logit una mejor capacidad predictiva y mayor ajuste que en metodologías como el modelo lineal general, el análisis discriminante y los árboles de decisión;

asimismo, mayor facilidad de aplicación e interpretación que con los modelos probit, logit mixtos y tobit, entre otros.

El modelo logit parte de la expresión

$$P_i = \frac{1}{1 + e^{-(\beta_1 + \beta_2 X_i)}} = \frac{1}{1 + e^{-Z_i}} \quad (1)$$

Donde $Z_i = (\beta_1 + \beta_2 X_i)$, Z_i se encuentra en un rango de $-\infty$ a $+\infty$, P_i se encuentra entre 0 y 1, y no está relacionado linealmente con Z_i ; sin embargo, es posible linealizar la ecuación (1) llegando a

$$L_i = \ln\left(\frac{P_i}{1-P_i}\right) = (\beta_1 + \beta_2 X_i + u_i) \quad (2)$$

donde el logaritmo de la razón de las probabilidades L_i se denomina logit y es lineal tanto en X como en los parámetros. En este modelo, aunque L_i es lineal en X , las probabilidades no lo son; es decir, no aumentan linealmente con X . Se pueden tener tantas variables independientes como sea necesario, para lograr un modelo estadísticamente significativo. La expresión u_i corresponde al término de error estocástico.

En el *modelo logit para datos agrupados*, se cuenta con observaciones repetidas para cada valor de X_i , se tiene un valor N_i correspondiente al número total de observaciones asociadas a un mismo valor de la variable independiente y una cantidad n_i , que corresponde al número de observaciones que cumplen con la condición definida para que la variable dicotómica tome el valor de 1, donde $n_i \leq N_i$.

Según Gujarati y Porter (2010) y Medina (2003), calculando la frecuencia relativa:

$$\hat{P}_i = \frac{n_i}{N_i} \quad (3)$$

esta se puede utilizar en la ecuación (2) como una estimación de la P_i correspondiente a cada X_i . Teniendo en cuenta que la probabilidad de un evento es el límite de la frecuencia relativa, a medida que el tamaño de la muestra se hace infinitamente grande, si el número de observaciones N_i en cada nivel de X es razonablemente grande, se tiene una estimación relativamente buena del verdadero logit, utilizando el logit estimado.

De manera complementaria, Medina (2003) indica que la generalización del modelo a k variables explicativas implica la existencia de observaciones repetidas de \hat{P}_i para cada combinación de las k variables explicativas, pudiéndose calcular las proporciones o probabilidades muestrales de la misma forma que en el caso de una sola variable exógena.

Selección de las variables

La variable dependiente que se va a modelar es P_i , correspondiente a la proporción de recaudo de la cartera para el cliente i . Se encuentra acotada entre 0 y 1.

Partiendo de las propuestas de otros autores recopiladas en la tabla 1 y atendiendo las condiciones del caso de estudio, se definieron las 24 variables explicativas evaluadas que se describen a continuación:

1. *Convenio* (C_i): existencia de convenio con la entidad cliente. Se trata de una variable dicotómica, cuyo valor es 1 si existe convenio, o 0 si no existe convenio. Para el caso de las atenciones de aseguradoras SOAT (seguro obligatorio para accidentes de tránsito), por reglamentación no es necesario establecer convenios entre las aseguradoras y las IPS, así que a estos clientes se les asigna el valor 1 aunque no exista convenio. Se espera un coeficiente positivo por cuanto la existencia de un convenio facilita el proceso de autorización, aceptación y pago de las facturas.
2. *Tipo Estado* (TE_i): variable dicotómica que asigna 1 si el cliente es el Estado, o 0 si es una empresa diferente al Estado. Se espera un signo negativo por dificultades de liquidez del sistema de salud estatal.
3. *Tipo IPS* (IPS_i): variable dicotómica, que asigna 1 si el cliente es una institución prestadora de servicios de salud- IPS, o 0 si no lo es. Se espera un signo positivo, pues estas atenciones corresponden a alianzas estratégicas entre IPS.
4. *Régimen subsidiado* (RS_i): variable dicotómica, que asigna 1 si el cliente es una EPS del régimen subsidiado, o 0 si no lo es. Se espera un signo negativo por dificultades de liquidez de estas entidades.

5. *Régimen excepción* (RE_i): variable dicotómica, que asigna 1 si el cliente es una aseguradora del régimen de excepción, o 0 si no lo es. Se espera un signo positivo, por cuanto estas entidades cuentan con presupuestos de destinación específica para atender este tipo de contratos.
6. *Régimen contributivo* (RC_i): variable dicotómica, que asigna 1 si el cliente es una EPS del régimen contributivo, o 0 si no lo es. Se espera un signo positivo, por cuanto son entidades privadas cuya eficiencia en la administración debe permitir la atención de sus obligaciones.
7. *Seguros* (SE_i): variable dicotómica, que asigna 1 si el cliente es una compañía de seguros, o 0 si no lo es. Se espera un signo positivo, por cuanto estas entidades cuentan en el sector con una reputación de eficiencia en su gestión financiera.
8. *Medicina prepagada* (MP_i): variable dicotómica, que asigna 1 si el cliente es una compañía de medicina prepagada, o 0 si no lo es. Se espera un signo positivo, por cuanto son empresas con un modelo de negocio que se conoce como rentable y de alta liquidez.
9. *Domicilio* (D_i): variable dicotómica, que toma el valor 1 para clientes cuya sede para trámites relacionados con el cobro de la cartera se ubica en Palmira o Cali, y 0 para otras ciudades. Se espera un coeficiente positivo, considerando que una ubicación geográfica cercana del deudor facilita la gestión de recaudo.
10. *Vinculada económica* (V_i): variable dicotómica, cuyo valor es 1 si el cliente es una empresa del mismo grupo económico de la IPS, y 0 si no lo es. Se espera un coeficiente positivo, puesto que la pertenencia al mismo grupo económico del cliente favorece los procesos entre las partes.
11. *Naturaleza* (N_i): variable dicotómica, que toma el valor 1 para entidades de naturaleza pública, y 0 para entidades de naturaleza privada. Se espera un coeficiente negativo, por cuanto las entidades públicas son catalogadas en el sector como de difícil recaudo.
12. *Glosas* (G_i): proporción de la cartera objeto de análisis que se encuentra en estado glosa o devolución y que tiene antigüedad mayor a 180 días. Su valor puede oscilar entre 0 y 1. Se espera un coeficiente negativo, por cuanto una mayor proporción de cartera glosada o devuelta implica un aumento del riesgo de no aceptación del cobro.

13. *Facturación sin radicar* (R_i): proporción de la cartera objeto de análisis, que se encuentra sin radicar ante la entidad y tiene antigüedad mayor a 180 días. Su valor puede oscilar entre 0 y 1. Se espera un coeficiente negativo, por cuanto una mayor proporción de cartera sin radicar ante el cliente implica un aumento del riesgo de no aceptación de la cuenta, por vencimiento de términos o soportes incompletos para el cobro.
14. *Cartera radicada* (RA_i): proporción de la cartera objeto de análisis, que se encuentra efectivamente radicada ante la entidad sin glosas ni devoluciones. Se espera un coeficiente positivo, por cuanto es un requisito previo indispensable para ser programado el pago por parte del cliente.
15. *Índice de liquidez corriente en el año inmediatamente anterior* (L_i): variable continua calculada como activo corriente/pasivo corriente del cliente i , al cierre del año inmediatamente anterior. Se espera un coeficiente positivo, por cuanto una mayor liquidez mejora las posibilidades de pago por parte del cliente.
16. *Período medio de cobranza en el año inmediatamente anterior* (CO_i): variable continua calculada como *cuentas por cobrar/ingresos operacionales * 360 del cliente i* para el año inmediatamente anterior. Se espera un coeficiente negativo, por cuanto un mayor período promedio de cobranza implica dificultades de liquidez en el cliente.
17. *Índice de endeudamiento en el año inmediatamente anterior* (E_i): variable continua calculada como *total pasivo/total activo del cliente i* para el año inmediatamente anterior. Se espera un coeficiente negativo, por cuanto un mayor endeudamiento implica mayores compromisos que atender con sus activos y, por consiguiente, menores posibilidades de pago.
18. *Margen neto en el año inmediatamente anterior* (M_i): variable continua calculada como *utilidad neta/ingreso del cliente i* para el año inmediatamente anterior. Se espera un coeficiente positivo, por cuanto una mayor rentabilidad es indicio de bienestar financiero del cliente que mejore su comportamiento en el pago de sus compromisos.
19. *Margen de solvencia patrimonial en el año inmediatamente anterior* (S_i): variable continua calculada como *patrimonio total/activo total del cliente i* para el año inmediatamente anterior. Se espera un coeficiente positivo, por cuanto una mayor solvencia es indicio de

bienestar financiero del cliente que mejore su comportamiento en el pago de sus compromisos.

20. *Porcentaje de cartera que supera 90 días de vencimiento (W_i):* variable continua entre 0 y 1, que cuantifica la proporción de cartera que cuenta con más de 90 días de vencimiento (en total 180 días de antigüedad, por cuanto el plazo de vencimiento es de 90 días). Mide el nivel inicial de mora. Se espera un coeficiente negativo, por cuanto la morosidad es un indicio de posibles dificultades para obtener el pago.
21. *Porcentaje de cartera que supera 180 días de vencimiento (X_i):* variable continua entre 0 y 1, que cuantifica la proporción de cartera que cuenta con más de 180 días de vencimiento (en total 270 días de antigüedad, por cuanto el plazo de vencimiento es de 90 días). Mide el nivel intermedio de mora. Se espera un coeficiente negativo.
22. *Porcentaje de cartera que supera 270 días de vencimiento (Y_i):* variable continua entre 0 y 1, que cuantifica la proporción de cartera que cuenta con más de 270 días de vencimiento (en total 360 días de antigüedad, por cuanto el plazo de vencimiento es de 90 días). Mide el nivel alto de mora. Se espera un coeficiente negativo.
23. *Porcentaje de cartera que supera 360 días de vencimiento (Z_i):* variable continua entre 0 y 1, que cuantifica la proporción de cartera que cuenta con más de 360 días de vencimiento (en total 450 días de antigüedad, por cuanto el plazo de vencimiento es de 90 días). Mide el nivel crítico de mora. Se espera un coeficiente negativo.
24. *Porcentaje de participación sobre la cartera total (PAR_i):* variable continua entre 0 y 1, que cuantifica la participación de la cartera de cada cliente sobre la cartera total. Se espera un coeficiente negativo, por cuanto un alto volumen de cartera implica una mayor exigencia operativa.

Especificación del modelo

El modelo logit propuesto se construye a partir de (1), incluyendo las 24 variables explicativas ya presentadas en la sección *Selección de las variables*:

$$P_i = \frac{1}{1 + e^{-(\beta_1 + \beta_2 C_i + \beta_3 TE_i + \beta_4 IPS_i + \dots + \beta_{23} Y_i + \beta_{24} Z_i + \beta_{25} PAR_i)}} \quad (4)$$

Tratamiento de la data

En este estudio se analiza el comportamiento del recaudo de 54.028 facturas, correspondientes a 49 clientes personas jurídicas, las cuales contaban con saldos pendientes en la cartera al corte del 30 de junio de 2013. Se calcula la proporción de recaudo de cartera por cada cliente dos años después de esta fecha, asumiendo que los valores no recaudados a esa altura corresponden a glosas, o aceptadas, o que no serán pagados por otras circunstancias, siendo esta proporción la variable endógena por estimar. Los 49 clientes incluidos en la modelación concentran el 99,9 % de la cartera de la entidad.

Adicionalmente, se consideró una variante del modelo en la cual la variable endógena es la proporción de recaudo de la cartera con antigüedad mayor a 180 días de emitida. En este caso, la base de datos consta de 15.394 facturas, las cuales se agruparon en 40 clientes personas jurídicas, que representaban el 99,7% de la cartera de la entidad con una antigüedad mayor a 180 días. Las fechas de corte y los horizontes de tiempo son iguales al modelo expuesto inicialmente.

Es importante aclarar que, para los modelos considerados, los valores de Pi se calculan sobre el valor inicial de la factura, con independencia del saldo de la cartera en el momento del análisis. En el anexo 1 se pueden observar, a manera de ejemplo para cinco clientes, los cálculos realizados para la variable endógena, cuando Pi es el porcentaje recaudado.

A su vez, se plantea una variante adicional, estimando para ambos casos la variable Pi como el porcentaje de cartera no recaudado (el *default*).

Para la estimación de los modelos se utilizó el software E-Views 9.

En los anexos 2 y 3 se presenta el resumen de estadística descriptiva de las variables analizadas, incluyendo la variable dependiente.

Resultados obtenidos

Después de la estimación y calibración utilizando el modelo lineal generalizado, en el anexo 4 se presentan los reportes de resultados, y en la tabla 2, el resumen para los cuatro modelos seleccionados. Las variables explicativas que no aparecen en la tabla 2 fueron descartadas

durante el proceso de ajuste, por resultar con signo contrario al esperado, según intuición del analista, por no ser estadísticamente significativas o no adicionar capacidad predictiva al modelo.

Tabla 2. Resumen de resultados para modelos

Variables		Porcentaje recaudado (pago)					Porcentaje no recaudado (no pago)				
		Signo esper.	Cartera total		Cartera > 180 días		Signo esper.	Cartera total		Cartera > 180 días	
			Modelo 1		Modelo 2			Modelo 3		Modelo 4	
			coefic.	prob. z est.	coefic.	prob. z est.		coefic.	prob. z est.	coefic.	prob. z est.
Constante			-5,188	0,070	-3,097	0,358		-4,247	0,004	3,005	0,218
Convenio	C	+	1,657	0,112	6,763	0,257	-			-7,056	0,102
Período cobranza	CO	-	-1,878	0,005			+	1,560	0,025		
Domicilio	D	+	3,609	0,012	9,581	0,124	-	-1,375	0,200	-8,989	0,006
Endeudamiento	E	-					+	2,479	0,035		
Margen neto	M	+	19,471	0,012			-	-2,372	0,118		
Participación	PAR	-	-7,023	0,052	-207,409	0,006	+	1,642	0,660	212,225	0,002
Cartera radicada	RA	+	7,258	0,028	5,129	0,149	-			-5,064	0,063
Tipo Estado	TE	-	-14,393	0,008	-0,863	0,889	+	4,591	0,007		
Vinculada económica	V	+			104,304	0,010	-			-107,259	0,002
Vencido 90	W	-	-13,523	0,021			+	1,252	0,309		
Vencido 180	X	-	19,730	0,016			+				
Prob(LR statistic)			0,000		0,000			0,000		0,000	
Criterio Akaike			-0,392		-0,278			0,050		-0,331	
Criterio Schwarz			-0,005		0,018			0,358		-0,078	
Deviance			1,259		1,229			2,143		1,230	
p value deviance			1,000		1,000			1,000		1,000	
Pearson SSR (Sum squared resid)			1,259		1,229			2,143		1,230	
Pearson statistic			0,032		0,037			0,052		0,036	
p Pearson statistic			1,000		1,000			1,000		1,000	
Especificidad			0,975		1,000			1,000		1,000	
Sensibilidad			0,667		0,800			0,444		0,800	
Capacidad predictiva			0,918		0,975			0,898		0,975	
pc óptimo			0,300		0,500			0,300		0,500	
Errores inferiores a 5%			0,898		0,902			0,837		0,902	

Fuente: elaboración propia.

Análisis de resultados

Los criterios de análisis incluyen la significancia global de los modelos, el signo y significancia de las variables explicativas y la bondad de ajuste de los modelos a los datos reales. En la tabla 3, se presentan los conceptos revisados y los autores que los proponen.

Tabla 3. Conceptos de bondad de ajuste evaluados

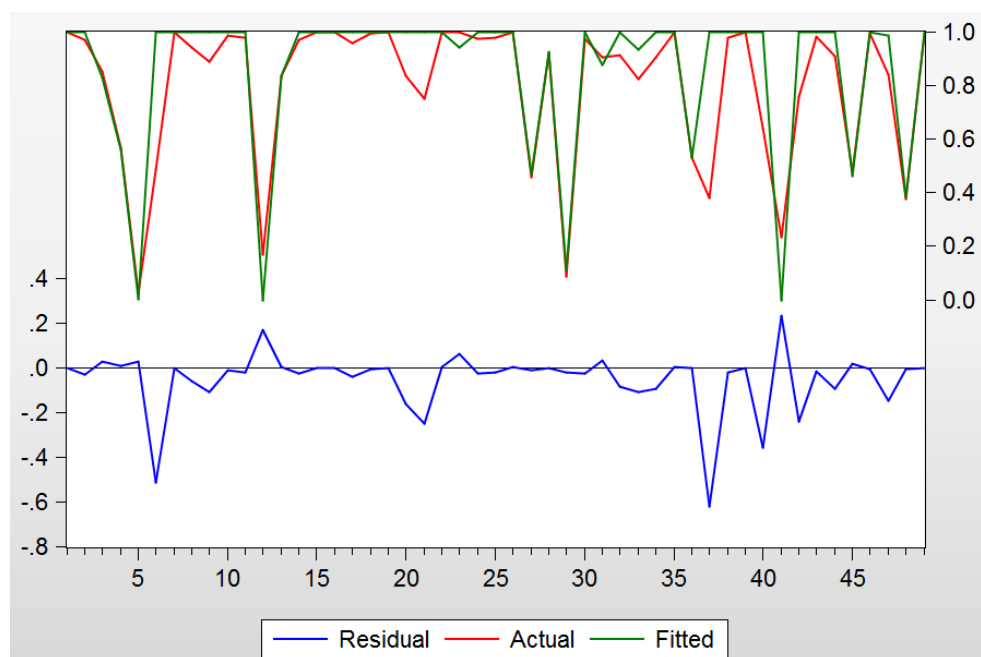
Análisis	Interpretación	Autor
Prob. LR statistic: Probabilidad estadístico razón de verosimilitud	Prueba la hipótesis nula respecto de que todos los coeficientes de las variables regresoras son simultáneamente iguales a cero.	Gujarati (2010) Bolton (2010), Camino, Lara & Rayo (2010)
Signo coeficientes	Se evalúan los signos esperados de los coeficientes de regresión y su importancia práctica y(o) estadística.	Gujarati (2010), Salazar, A. (2010)
Prob. z-Statistic: Probabilidad estadístico z	Se evalúa la importancia estadística de los coeficientes, según el nivel de significancia definido.	Gujarati (2010), Camino, Lara & Rayo (2010), Moreno (2014)
	Se pueden considerar valores p menores a 0,25, puesto que el uso de niveles tradicionales como 0,05 puede no permitir la identificación de variables reconocidas como importantes.	Moreno (2014), Hosmer, Lesmeshow & Sturdivant (2000)
Criterio información de Akaike (CIA) y criterio de información de Schwarz (CIS)	Al comparar dos o más modelos, se prefiere el de menor valor de CIA y CIS.	Gujarati (2010), Moreno (2014)
Devianza	La devianza sigue una distribución Chi-cuadrado con n-k-1 grados de libertad. Se evalúa la hipótesis nula: Ho: el modelo es de buen ajuste a los datos. Valores p por debajo del nivel de significancia implican rechazar Ho.	Turner, H. (2008), Bolton (2010)
Pearson SSR y Pearson statistic	El estadístico de Pearson sigue una distribución Chi-cuadrado con n-k grados de libertad. Se evalúa la hipótesis nula: Ho: el error es cero. Valores p por debajo del nivel de significancia implican rechazar Ho.	Medina (2003), Camino, Lara & Rayo (2010).
Sensibilidad y especificidad	Se buscan los mayores niveles de sensibilidad y especificidad al igual que el umbral óptimo.	Camino, Lara & Rayo (2010) , García, Mures & Vallejo (2005), Medina (2003), Moreno (2014),
Errores inferiores al 5%	Se busca la mayor tasa de errores inferiores al umbral definido (5%).	Elaboración propia.

Fuente: elaboración propia.

A continuación, se analizan los resultados a nivel de variables explicativas para cada modelo. Asimismo, se presenta el gráfico de residuales y ajuste de estimación para cada caso.

Modelo 1: se estima el porcentaje recaudado sobre la cartera total. Las variables que cumplen con el signo esperado y tienen significancia estadística para un nivel de confianza del 95% son CO, D, M, PAR, RA, TE, W. La variable C cumple con el criterio de signo esperado, pero es significativa en un nivel de confianza del 88%. La variable X, aunque es significativa para un nivel de confianza por encima del 95%, presenta un coeficiente con signo contrario al esperado, situación que indica que la edad de la cartera con vencimiento por encima de 180 días no necesariamente es un indicio de que esta no será pagada. Esta última variable no fue retirada del modelo, pues se generaba desajuste general del mismo, en signos y significancias, de las demás variables (gráfico 1).

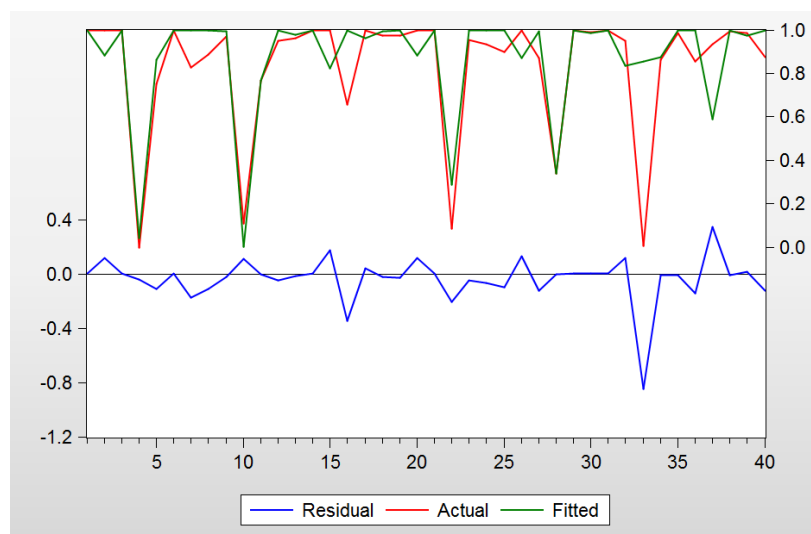
Gráfico 1. Ajuste de estimación y residuales – Modelo 1



Fuente: elaboración propia.

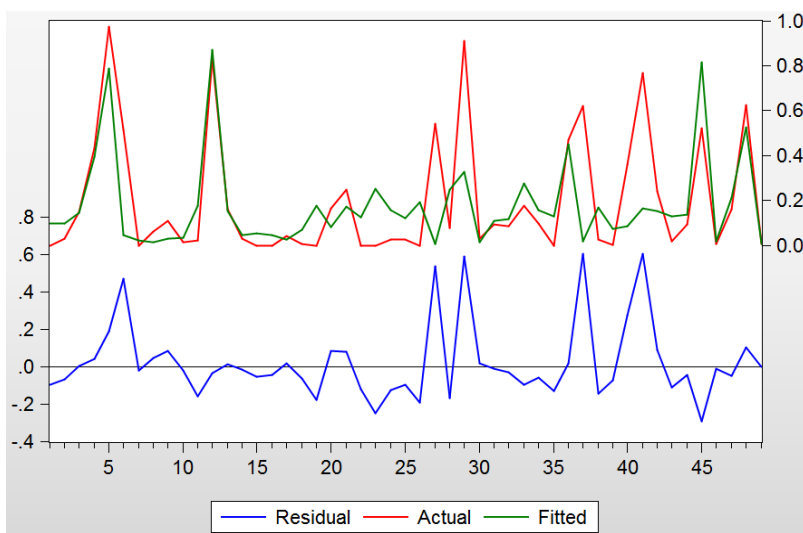
Modelo 2: se estima el porcentaje de recaudo para la cartera con antigüedad mayor a 180 días. Se encuentra que las variables C, D, PAR, RA, TE y V cumplen con el criterio de signo esperado; sin embargo, C y TE no son significativas si se toma un nivel de confianza del 75%, mientras que D y RA son significativas con un nivel de confianza del 85% y las variables PAR y V son significativas con un nivel de confianza del 99% (gráfico 2). Al explorar retirando las variables no significativas se genera desajuste del modelo y pérdida de capacidad predictiva, además de considerarse importantes según la intuición del analista.

Gráfico 2. Ajuste de estimación y residuales – Modelo 2



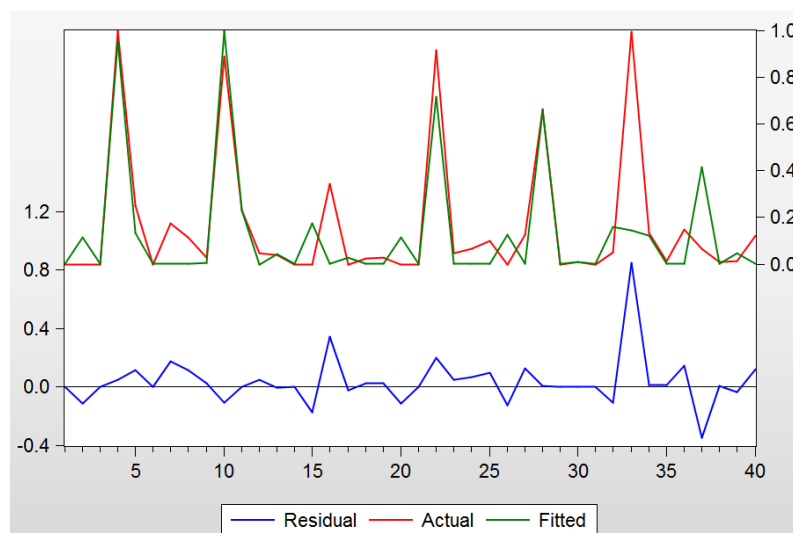
Fuente: elaboración propia.

Modelo 3: se estima el porcentaje de no recaudo para la cartera total. Se encuentra que las variables CO, D, E, M, PAR, TE, W cumplen con el criterio de signo esperado; sin embargo, PAR y W no son significativas si se toma un nivel de confianza del 75%, mientras que D y M son significativas para un nivel de confianza del 80% y las variables CO, E y TE son significativas con un nivel de confianza del 95% (gráfico 3). En cuanto a las variables no significativas, estas se conservaron debido a la intuición del analista, y por pérdida de bondad de ajuste global y capacidad predictiva cuando se intenta retirarlas.

Gráfico 3. Ajuste de estimación y residuales – Modelo 3

Fuente: elaboración propia.

Modelo 4: se estima el porcentaje de no recaudo para la cartera con antigüedad mayor a 180 días. Las variables C, D, PAR, RA y V cumplen con el criterio de signo esperado. En cuanto a la significancia estadística, las variables D, PAR, y V son significativas para un nivel de confianza del 95%, mientras que C y RA lo son para un nivel de confianza del 89% (gráfico 4).

Gráfico 4. Ajuste de estimación y residuales – Modelo 4

Fuente: elaboración propia.

En cuanto a la significancia global de los modelos, evaluada mediante la probabilidad del estadístico de la razón de verosimilitud - Prob (LR statistic), se encuentran resultados satisfactorios para niveles de confianza del 99%. Los criterios de Akaike y Schwarz muestran un mejor desempeño en los modelos 1 y 4, indicando en estos un mayor ajuste en relación con el número de variables explicativas. El análisis de *deviance* muestra resultados satisfactorios, por lo que se acepta la hipótesis nula de bondad del ajuste para niveles de confianza del 99%. En cuanto a la suma de los cuadrados de los residuales de Pearson, se observan resultados similares para los modelos 1, 2 y 4, y mejores que para el modelo 3, mientras que el estadístico de Pearson y su probabilidad muestran bondad de ajuste mediante la aceptación de la hipótesis nula con un nivel de confianza del 99% para todos los casos.

Finalmente, se considera importante hacer referencia al análisis de especificidad, sensibilidad y capacidad predictiva, criterios expuestos por Moreno (2013). En la tabla 4 se presenta como ejemplo lo evaluado para el modelo 1.

Tabla 4. Capacidad predictiva para el modelo 1

Punto corte		0,5		Estimación modelo		Total	
		No Default	Default	No Default	Default		
Observado	No Default	39	1	40	0,975	Especificidad	
	Default	3	6	9	0,667	Sensibilidad	
Total		42	7	49	0,918	Capacidad predictiva	
Promedio especificidad y sensibilidad		0,821		0,300		pc óptimo	

Fuente: elaboración propia, con base en Moreno (2013).

Para los modelos presentados, la especificidad (proporción de los clientes catalogados como no *default* cuya predicción por el modelo es acertada) se ubica por encima del 97%, mientras que la sensibilidad (proporción de los clientes catalogados como *default* cuya predicción por el modelo fue acertada) está por encima del 60%, con excepción del modelo 3. El total de aciertos representado por la capacidad predictiva (total predicciones acertadas) es superior al 89%, incluso con modelos que llegan al 97%. Cabe aclarar que la clasificación de cada observación/estimación como no *default* o *default* se hizo con un punto de corte de 0,5. Se calcula, adicionalmente, el punto de corte óptimo para cada caso; es decir, aquel que optimiza el promedio de los índices de especificidad y sensibilidad, encontrando 0,3 para los modelos basados en la cartera total y 0,5 para los modelos basados en la cartera mayor a 180 días.

Finalmente, la tasa de predicciones con error inferior al 5% se ubicó por encima del 83%, incluso llegando a 90%.

Interpretaciones

En la interpretación de los coeficientes obtenidos para las variables explicativas se tiene en cuenta que en la ecuación (2) estas son lineales para L_i pero no para P_i . A partir de la ecuación (2), Gujarati y Porter (2010) exponen, mediante cálculo diferencial, que la variación de P_i con respecto al cambio en las variables explicativas depende tanto del coeficiente de cada variable como del valor P_i a partir del cual se mide el cambio:

$$\frac{\partial P_i}{\partial X_i} = \beta_i P_i (1 - P_i) \quad (5)$$

Para la interpretación de resultados obtenidos mediante esta expresión, se requiere conocer, además, el valor inicial de las variables explicativas que generan el valor P_i a partir del cual se mide el cambio.

Como ejemplo de la aplicación de este concepto, se toman el modelo 3 y el cliente 13, cuya estimación de P_i (proporción de no recaudo) es 0,151 y los resultados se resumen en la tabla 5.

Tabla 5. Variación de P_i con respecto al cambio en variables explicativas en el modelo 3

Variables		Modelo 3					
		$P_i = 0,151$					
		Signo esperado	β_i	$\frac{\partial P_i}{\partial X_i}$	Valor inicial variable explicativa	Unidades variación variable explicativa	Variación P_i
Período cobranza	CO	+	1,560	0,200	0,274	1,000	0,200
Domicilio	D	-	-1,375	-0,176	1,000	1,000	-0,176
Endeudamiento	EI	+	2,479	0,318	0,862	0,100	0,032
Margen neto	M	-	-2,372	-0,304	-0,008	0,100	-0,030
Participación	PAR	+	1,642	0,210	0,590	0,100	0,021
Tipo Estado	TE	+	4,591	0,588	0,000	1,000	0,588
Vencido 90	W	+	1,252	0,160	0,273	0,100	0,016

Fuente: elaboración propia.

Los impactos generados por variaciones en las variables explicativas en el modelo 3 para el cliente seleccionado se presentan a continuación:

Período cobranza: el aumento de una unidad (un día) en el indicador de período de cobranza del cliente aumenta en un 20% el porcentaje de no recaudo esperado. El aumento de los días de cuentas por cobrar del cliente implica una menor liquidez y, por consiguiente, mayor dificultad para atender sus compromisos, entre ellos la cartera adeudada con la institución.

Domicilio: para el cliente tomado como ejemplo, la ubicación geográfica en Cali o Palmira disminuye en un 17,6% el porcentaje de no recaudo esperado. Dado que el cliente ya cuenta con dicha ubicación geográfica, un cambio de la misma aumenta la proporción esperada de no recaudo en la cifra mencionada. Esto es consistente, conociendo que la ubicación geográfica es importante, por cuanto el cobro de cartera requiere gestión directa en las instalaciones del cliente. Dado que generalmente existirán clientes ubicados en otras ciudades, se puede considerar la ubicación de oficinas de gestión de cobranza o gestores externos de cartera ubicados geográficamente cerca de estos clientes.

Endeudamiento: un incremento de 0,1 en el indicador de endeudamiento del cliente implica un aumento del 3,2 % en el porcentaje esperado de no recaudo.

Margen neto: el incremento de 10% en el margen neto del cliente disminuye en un 3% el porcentaje esperado de no recaudo.

Participación sobre el total: si la participación de la cartera del cliente sobre la cartera total aumenta en un 10%, se aumenta en un 2,1% el porcentaje esperado de no recaudo. Volúmenes grandes de facturas en cartera implican un mayor esfuerzo operativo, el cual, al no poder ser atendido, da origen a negociaciones generales en condiciones desfavorables para la institución. Como alternativa, se pueden considerar estructuras operativas exclusivas para la gestión de las carteras más grandes de la entidad.

Tipo Estado: si el cliente es el Estado, el porcentaje de no recaudo esperado aumenta en un 58,8%. Al respecto, se debe minimizar la atención de pacientes cuyo pagador es el Estado, y en caso de presentarse alguna obligatoriedad, debe ser objeto de especial seguimiento.

Vencido 90: un incremento del 10% en el porcentaje de cartera con más de 90 días de vencimiento aumenta en un 1,6% el porcentaje esperado de no recaudo. La interpretación de este resultado implica que la antigüedad de la cartera no es una condición concluyente para considerar que la cartera no será recaudada y que, por consiguiente, deba ser castigada en su totalidad.

A continuación se comentan otras variables analizadas que no fueron incluidas en el modelo 3.

La variable *Convenio* resultó ser estadísticamente significativa, cumpliendo además con el signo esperado en los modelos 1-2 y 4. En los tres casos, presenta un efecto positivo en el porcentaje de recaudo esperado. Se encuentra que la formalización de un contrato con el cliente permite definir procesos y tarifas de manera clara, con lo cual se garantice el pago de los servicios prestados por la institución.

Las variables financieras *Liquidez* y *Solvencia* no resultaron significativas ni presentaron los signos esperados, según la intuición del analista. La situación financiera del cliente medida por estos indicadores no ofrece información con la que se pueda determinar el porcentaje esperado de recaudo de la cartera.

La variable *Cartera radicada* resultó significativa y acorde con el signo esperado en los modelos 1-2 y 4, con un efecto positivo en la proporción esperada de recaudo. Como requisito indispensable para el pago de la cartera, esta debe cumplir con la condición que se mide con esta variable, con lo que se corrobora la necesidad de contar en la institución con procesos eficientes que garanticen la calidad de la facturación.

Las variables *Glosas* y *Sin radicar* no cumplieron con los criterios de significancia y signo esperado; sin embargo, variables relacionadas como *Cartera radicada* y *Participación sobre total* se incluyeron en los modelos seleccionados, lo cual indica que el estado operativo de la cartera tiene incidencia en el recaudo de la institución.

Las variables *IPS*, *Medicina prepagada*, *Naturaleza*, *Régimen contributivo*, *Régimen excepción*, *Régimen subsidiado* y *Seguros* no resultaron significativas ni cumplieron con los signos esperados. Esto indica que, para la base de datos utilizada en la modelación, el tipo de entidad no aporta información que permita estimar la proporción de recaudo de cartera. Este resultado

difiere del criterio del analista y de su conocimiento del sector, por lo que las variables no se deben descartar en futuros estudios sobre el tema.

En cuanto a la variable *Vinculada económica*, esta resultó ser significativa, y cumplió con el criterio de signo esperado solamente para las modelaciones de cartera mayor a 180 días, indicando así que para la cartera de un cliente que pertenece al mismo grupo empresarial de la institución se espera un mayor porcentaje de recaudo. Este resultado cumple de manera parcial con el criterio del analista, basado en el conocimiento de la institución y del sector, puesto que se espera que esta variable tenga influencia en todas las modelaciones que se planteen al respecto.

En cuanto a las variables *Vencido 180*, *Vencido 270* y *Vencido 360*, su significancia estadística y su signo del coeficiente en los modelos evaluados no permiten concluir de manera contundente, en cuanto al efecto de la edad de la cartera sobre el porcentaje de recaudo esperado, puesto que la intuición del analista indica que poseer cartera vencida antigua es indicio de deterioro de la misma. Se espera que los resultados obtenidos en futuros estudios que incluyan estas variables permitan aclarar el mencionado efecto.

Conclusiones

En este trabajo se presenta una metodología para estimar la pérdida esperada en las cuentas por cobrar de una institución prestadora de servicios de salud (IPS) y para determinar las variables que influyen en ella, utilizando el modelo econométrico logit con datos agrupados. De esta manera, se atiende un requerimiento originado por la adopción de las Normas Internacionales de Información Financiera NIIF, en referencia al deterioro de los activos financieros.

Los estudios consultados en esta investigación presentan el modelo logit como apropiado para la estimación de variables binarias, por su facilidad de aplicación e interpretación; sin embargo, las aplicaciones están enfocadas en el caso de datos individuales. En el caso de estudio, al aplicar la metodología logit con datos agrupados se logra una alta significancia estadística global, buen ajuste a los datos, alta capacidad predictiva y signos de los coeficientes de las variables explicativas acordes con la experiencia previa del autor.

Los autores consultados utilizan una amplia gama de variables, las cuales se utilizaron como punto de partida para esta investigación, teniendo en cuenta las características del caso de estudio y las bases de datos disponibles, para lo cual se evaluaron en total 24 variables. Con respecto a las variables para clasificar por tipo de cliente, estas resultan no tener influencia en el modelo; sin embargo, no se descarta su uso para futuros estudios. En cuanto a las variables financieras y operativas, aquellas que no son influyentes en el modelo pueden ser evaluadas en estudios posteriores, teniendo en cuenta que otras variables relacionadas sí resultaron tener influencia en el modelo.

Las variables que resultan tener influencia en la estimación de la pérdida esperada incluyen algunas que miden la liquidez y rentabilidad del cliente, como son *Período de cobranza*, *Endeudamiento* y *Margen neto*; otras que miden características propias del cliente, como *Domicilio*, *Tipo Estado* y *Vinculada económica*; finalmente, variables que miden condiciones operativas, como *Convenio*, *Participación sobre el total*, *Cartera radicada*, *Vencido 90* y *Vencido 180*. En referencia a la antigüedad de la cartera medida por las dos últimas variables, no se logra corroborar de manera contundente la intuición del analista acerca de su relación directa con el incremento de la pérdida esperada, situación que debe ser evaluada en estudios posteriores.

En cuanto a la gestión de cuentas por cobrar, esta investigación permite generar recomendaciones a la institución objeto de estudio, tales como: la importancia de contar con procesos efectivos de facturación con los cuales se garantice la mayor proporción de cartera radicada y aceptada por los clientes (libre de glosas y devoluciones), la adopción de estructuras operativas segmentadas y suficientes para la gestión de cartera de los clientes más grandes, la ubicación de oficinas de gestión de cobro en las ciudades domicilio de sus clientes, o la contratación de asesores externos ubicados en dichas ciudades, control exhaustivo a los servicios prestados a pacientes cubiertos por el Estado, el fortalecimiento y ampliación del portafolio ofrecido a los clientes pertenecientes al mismo grupo económico y la formalización de convenios contractuales con todos los clientes, evaluando además sus indicadores financieros como parte del proceso de negociación y contratación.

Tanto el tipo de cartera analizada como la aplicación del modelo econométrico utilizado no son temas comúnmente encontrados en la literatura disponible, por lo cual este hecho se constituye

en uno de los aportes más importantes de la presente investigación, y deja la puerta abierta a futuros estudios relacionados.

Referencias bibliográficas

- Altman, E. I. (1968). Financial ratios, discriminant analysis and the prediction of corporate bankruptcy. *The journal of finance*, 23(4), 589-609.
- Altman, E., Haldeman, R. & Narayanan, P. (1977). ZETATM analysis A new model to identify bankruptcy risk of corporations. *Journal of Banking & Finance*, 1(1), 29-54.
- Arrobo, E., y Montoya, L. (2014). *Modelo para el análisis de riesgo crediticio de la cartera de vivienda basado en matrices de transición de calificación para el sector de bancos privados nacionales* (tesis inédita de maestría). Universidad de las Fuerzas Armadas ESPE, Salgonquí, Ecuador.
- Bolton, C. (2010). *Logistic regression and its application in credit scoring*. University of Pretoria, Sudáfrica.
- Consejo de Normas Internacionales de Contabilidad IASB (2005). Norma Internacional de Contabilidad 39: Instrumentos financieros, Reconocimiento y valoración. Recuperado de <http://nicniif.org>
- Fernández, H. y Pérez, F. (2005). El modelo logístico: una herramienta estadística para evaluar el riesgo de crédito. *Revista Ingenierías Universidad de Medellín*, 4(6), 55-75.
- Gujarati, D., y Porter, D. (2010). *Econometría*. México D. F., México: McGraw Hill.
- Hosmer, D., Lesmeshow, S., y Sturdivant, R. (2000). *Applied Logistic Regression*. New York, United States of America: John Wiley & Sons, Inc.

IFRS Foundation (2012). Norma Internacional de Información Financiera 9. Instrumentos financieros.

Recuperado de <http://www.ifrs.org>

Kolesar, P., & Showers, J. L. (1985). A robust credit screening model using categorical data.

Management Science, 31(2), 123-133.

López, C. (2013). Modelo predictivo de riesgo de morosidad para créditos bancarios usando datos simulados. *Epistemus*, 14(7), 22-30. Recuperado de

<http://www.epistemus.uson.mx/revistas/articulos/14-MODELO%20PREDICTIVO.pdf>

Medina, E. (2003). Modelos de elección discreta. Recuperado de

http://www.uam.es/personal_pdi/economicas/eva/pdf/logit.pdf

Moreno, S. (2013). *El modelo logit mixto para la construcción de un scoring de crédito* (tesis inédita de maestría). Universidad Nacional de Colombia, Medellín, Colombia.

Mures, M., García, A., y Vallejo M. (2005). Aplicación del Análisis Discriminante y Regresión Logística en el estudio de la morosidad en las entidades financieras. Comparación de resultados.

Pecunia, 01, 175-199.

Ochoa, J., Agudelo, L., y Galeano, W. (2010). Construcción de un modelo de scoring para el otorgamiento de crédito en una entidad financiera. *Perfil de Coyuntura Económica*, 16, 191-222.

Rayo, S., Lara, J., y Camino, D., (2010). Un Modelo de Credit Scoring para instituciones de microfinanzas en el marco de Basilea II. *Journal of Economics, Finance and Administrative Science*, 15(28), 89-124.

Rivera, J. (2011). *Modelo de evaluación de crédito —scoring— para la cartera de consumo de la Cooperativa de Ahorro y Crédito “Riobamba”* (tesis inédita de maestría). Universidad Andina Simón Bolívar, Quito, Ecuador.

- Salazar, A. (2010). *Modelos de respuesta discreta en R y aplicación con datos reales* (tesis inédita de maestría). Universidad de Granada, Granada, España.
- Salazar, F. (2013). Cuantificación del riesgo de incumplimiento en créditos de libre inversión: un ejercicio econométrico para una entidad bancaria del municipio de Popayán, Colombia. *Estudios Gerenciales*, 29(129), 416-427.
- Sandoval, L. (2015). *Modelo de gestión de la cartera para las empresas de medicina prepagada de Quito* (tesis inédita de maestría). Escuela Politécnica Nacional, Quito, Ecuador.
- Támara, A., Aristizábal, R., y Velásquez, H. (2010). Estimación de las provisiones esperadas en una institución financiera utilizando modelos logit y probit. *Revista Ciencias Estratégicas*, 18(24), 259-270.
- Turner, H. (2008). *Introduction to Generalized Linear Models*. University of Warwick, Coventry, Inglaterra. Recuperado de http://statmath.wu.ac.at/courses/heather_turner/
- Vélez, C. (2009). Modelo de riesgo crediticio para la empresa funeraria. *Revista Ciencias Estratégicas*, 17(21), 33-47.
- Wiginton, J. C. (1980). A note on the comparison of Logit and discriminant models of consumer credit behavior. *Journal of Financial and Quantitative Analysis*, 15(03), 757-770.

Anexos

Anexo 1. Cálculo de la variable endógena para cinco clientes

La variable endógena se plantea en términos de la proporción de la cartera que es recaudada para cada cliente al cabo de dos años, tomando como base el valor inicial facturado. Se plantea un modelo para el total de la cartera y un modelo para la cartera mayor a 180 días. Asimismo, se plantean dos modelos adicionales, estimando como variable endógena la proporción de la cartera no recaudada, tanto para la cartera total como para la cartera mayor a 180 días. A continuación se muestran los cálculos para cinco clientes, en los casos de proporción de cartera recaudada para el caso cartera total y el caso cartera mayor a 180 días.

	Modelo cartera total			Modelo cartera mayor 180 días		
	Facturado	Recaudo	P_i	Facturado	Recaudo	P_i
Cliente 1	16.409.709	16.409.709	1,000	15.956.402	15.956.402	1,000
Cliente 2	6.908.893	6.699.635	0,970	32.700	32.700	1,000
Cliente 3	27.847.294	23.745.614	0,853	20.778.663	20.778.663	1,000
Cliente 4	291.890	165.305	0,566	71.888	-	0,000
Cliente 5	158.204.758	4.158.614	0,026	5.539.795	4.158.614	0,751

Fuente: elaboración propia.

Anexo 2. Estadística descriptiva de las variables evaluadas (variables cuantitativas)

Variable		Media	Mediana	Desviación estándar	Curtosis	Mínimo	Máximo	Coefficiente de variación
Porcentaje de recaudo	P	0,801	0,923	0,273	1,150	0,026	1,000	0,341
Glosas	G	0,066	0,003	0,165	19,843	0,000	0,972	2,501
Facturación sin radicar	R	0,012	0,000	0,038	11,119	0,000	0,171	3,253
Índice de liquidez	L	7,904	1,393	16,199	13,043	0,014	87,527	2,049
Período cobranza	CO	0,292	0,208	0,391	23,268	0,008	2,540	1,337
Endeudamiento	E	0,752	0,750	0,317	2,346	0,026	1,756	0,421
Margen neto	M	0,017	0,006	0,153	18,619	-0,249	0,851	8,873
Margen de solvencia	S	0,248	0,250	0,315	2,231	-0,731	0,974	1,269
Vencido 90	W	0,281	0,219	0,286	0,344	0,000	1,000	1,018
Vencio 180	X	0,191	0,094	0,250	2,611	0,000	1,000	1,309
Vencido 270	Y	0,146	0,063	0,226	6,439	0,000	1,000	1,554
Vencido 360	Z	0,111	0,009	0,218	9,446	0,000	1,000	1,976
Participación sobre el total	PAR	0,020	0,002	0,085	44,434	0,000	0,590	4,166
Radicación sin glosas	RA	0,922	0,997	0,174	15,355	0,028	1,000	0,188

Fuente: elaboración propia.

Anexo 3. Estadística descriptiva de las variables evaluadas (variables cualitativas dicotómicas)

Estas variables hacen referencia al tipo de entidad, a la existencia de relación contractual formal y a la cercanía geográfica del cliente.

Variable		Sí		No	
Convenio	C	26	53,06%	23	46,94%
Tipo Estado	TE	2	4,08%	47	95,92%
Tipo IPS	IPS	5	10,20%	44	89,80%
Regimen subsidiado	RS	6	12,24%	43	87,76%
Régimen excepción	RE	4	8,16%	45	91,84%
Régimen contributivo	RC	11	22,45%	38	77,55%
Seguros	SE	16	32,65%	33	67,35%
Medicina prepagada	MP	5	10,20%	44	89,80%
Domicilio	D	23	46,94%	26	53,06%
Vinculada económica	V	2	4,08%	47	95,92%
Naturaleza (pública)	N	10	20,41%	39	79,59%

Fuente: elaboración propia.

Anexo 4. Reportes de resultados en E-Views para los modelos seleccionados

Modelo 1

Dependent Variable: PI
 Method: Generalized Linear Model (Newton-Raphson / Marquardt steps)
 Date: 02/14/16 Time: 01:03
 Sample: 1 49
 Included observations: 49
 Family: Normal
 Link: Logit
 Dispersion computed using Pearson Chi-Square
 Convergence achieved after 9 iterations
 Coefficient covariance computed using observed Hessian

Variable	Coefficient	Std. Error	z-Statistic	Prob.
C	-5.187823	2.862106	-1.812589	0.0699
CI	1.656943	1.043382	1.588051	0.1123
COI	-1.877709	0.672356	-2.792731	0.0052
DI	3.609393	1.437046	2.511676	0.0120
MI	19.47140	7.715480	2.523680	0.0116
PAR	-7.023280	3.619241	-1.940539	0.0523
TE	-14.39265	5.386860	-2.671806	0.0075
WI	-13.52345	5.845537	-2.313466	0.0207
XI	19.72954	8.180527	2.411769	0.0159
RA	7.257796	3.297478	2.201014	0.0277

Mean dependent var	0.800746	S.D. dependent var	0.272928
Sum squared resid	1.258706	Log likelihood	19.59220
Akaike info criterion	-0.391518	Schwarz criterion	-0.005433
Hannan-Quinn criter.	-0.245038	Deviance	1.258706
Deviance statistic	0.032275	Restr. deviance	3.575493
LR statistic	71.78375	Prob(LR statistic)	0.000000
Pearson SSR	1.258706	Pearson statistic	0.032275
Dispersion	0.032275		

Modelo 2

Dependent Variable: PI
 Method: Generalized Linear Model (Newton-Raphson / Marquardt steps)
 Date: 02/07/16 Time: 21:05
 Sample: 1 40
 Included observations: 40
 Family: Normal
 Link: Logit
 Dispersion computed using Pearson Chi-Square
 Convergence achieved after 12 iterations
 Coefficient covariance computed using observed Hessian

Variable	Coefficient	Std. Error	z-Statistic	Prob.
C	-3.096898	3.371288	-0.918610	0.3583
CI	6.762717	5.966205	1.133504	0.2570
DI	9.581170	6.227464	1.538535	0.1239
PAR	-207.4094	75.15576	-2.759728	0.0058
TE	-0.862577	6.205295	-0.139007	0.8894
VI	104.3043	40.57443	2.570691	0.0101
RA	5.129083	3.557636	1.441711	0.1494

Mean dependent var	0.833959	S.D. dependent var	0.292196
Sum squared resid	1.228747	Log likelihood	12.55272
Akaike info criterion	-0.277636	Schwarz criterion	0.017918
Hannan-Quinn criter.	-0.170773	Deviance	1.228747
Deviance statistic	0.037235	Restr. deviance	3.329751
LR statistic	56.42592	Prob(LR statistic)	0.000000
Pearson SSR	1.228747	Pearson statistic	0.037235
Dispersion	0.037235		

Modelo 3

Dependent Variable: PI
 Method: Generalized Linear Model (Newton-Raphson / Marquardt steps)
 Date: 02/09/16 Time: 20:04
 Sample: 1 49
 Included observations: 49
 Family: Normal
 Link: Logit
 Dispersion computed using Pearson Chi-Square
 Convergence achieved after 8 iterations
 Coefficient covariance computed using observed Hessian

Variable	Coefficient	Std. Error	z-Statistic	Prob.
C	-4.247066	1.472928	-2.883417	0.0039
COI	1.560491	0.695574	2.243457	0.0249
DI	-1.375469	1.072701	-1.282248	0.1998
EI	2.479139	1.177057	2.106217	0.0352
MI	-2.372290	1.519033	-1.561711	0.1184
PAR	1.641801	3.728444	0.440345	0.6597
TE	4.591014	1.708923	2.686495	0.0072
WI	1.251809	1.230099	1.017648	0.3088

Mean dependent var	0.199254	S.D. dependent var	0.272928
Sum squared resid	2.142617	Log likelihood	6.784845
Akaike info criterion	0.049598	Schwarz criterion	0.358467
Hannan-Quinn criter.	0.166782	Deviance	2.142617
Deviance statistic	0.052259	Restr. deviance	3.575493
LR statistic	27.41877	Prob(LR statistic)	0.000280
Pearson SSR	2.142617	Pearson statistic	0.052259
Dispersion	0.052259		

Modelo 4

Dependent Variable: PI
 Method: Generalized Linear Model (Newton-Raphson / Marquardt steps)
 Date: 02/08/16 Time: 22:31
 Sample: 1 40
 Included observations: 40
 Family: Normal
 Link: Logit
 Dispersion computed using Pearson Chi-Square
 Convergence achieved after 12 iterations
 Coefficient covariance computed using observed Hessian

Variable	Coefficient	Std. Error	z-Statistic	Prob.
C	3.004537	2.439850	1.231443	0.2182
CI	-7.056078	4.318482	-1.633926	0.1023
DI	-8.989210	3.287080	-2.734710	0.0062
PAR	212.2254	67.98826	3.121502	0.0018
RA	-5.063809	2.724571	-1.858571	0.0631
VI	-107.2587	34.89153	-3.074061	0.0021

Mean dependent var	0.166041	S.D. dependent var	0.292196
Sum squared resid	1.229972	Log likelihood	12.62984
Akaike info criterion	-0.331492	Schwarz criterion	-0.078160
Hannan-Quinn criter.	-0.239895	Deviance	1.229972
Deviance statistic	0.036176	Restr. deviance	3.329751
LR statistic	58.04398	Prob(LR statistic)	0.000000
Pearson SSR	1.229972	Pearson statistic	0.036176
Dispersion	0.036176		

Fuente: elaboración propia.