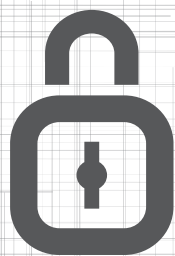




NIVELES ÓPTIMOS DEL SEGURO DE DEPOSITOS



COSEDE

Corporación del Seguro de Depósitos,
Fondo de Liquidez y
Fondo de Seguros Privados

Determinación de los niveles óptimos del
Seguro de Depósitos del Ecuador a partir de
su función de distribución de pérdidas
esperadas y de la estimación de las
probabilidades de riesgo alto de las entidades
financieras

Autor: Andrea Bayancela Espinel

Coordinación y edición: Eugenio Paladines y David Villamar

Corporación del Seguro de Depósitos,
Fondo de Liquidez y Fondo de Seguros Privados

Diciembre 2016

Determination of the target levels of the Deposit Insurance in Ecuador
from its expected loss distribution function and the estimation of the
probabilities of high risk rating of the financial entities

Resumen

El presente estudio determina los niveles óptimos del Seguro de Depósitos del Sector Financiero Privado en el Ecuador, con datos desde enero de 2007 hasta junio de 2016, a través de modelos de riesgo crediticio para la estimación tanto de las probabilidades de que las entidades financieras pasen a una calificación de riesgo alto de liquidación forzosa (incumplimiento), como de la función de distribución de pérdidas esperadas, medidas de riesgo y provisiones del Seguro de Depósitos.

Los modelos de riesgo crediticio utilizan indicadores macroeconómicos y propios de las entidades financieras para estimar las probabilidades de las variaciones en las calificaciones de riesgo de las entidades financieras. Del análisis de los indicadores macroeconómicos se desprende que incrementos en la tasa de interés activa están relacionados con una mayor probabilidad de que una entidad financiera pase a una calificación de riesgo alto. Mientras que de los indicadores propios de las entidades, se observa que aumentos en *rentabilidad*, *liquidez corriente*, y *cobertura de cartera* se asocian con una menor probabilidad de que una entidad financiera pase a una calificación de riesgo alto.

A partir de las probabilidades obtenidas de que una entidad financiera pase a una calificación de riesgo alto, se estimó la función de distribución de pérdidas esperadas del Seguro de Depósitos. De esta última, se determinó que los niveles óptimos del mismo equivalen a una pérdida esperada de 4 % de los depósitos asegurados y 12 % de los depósitos cubiertos; mientras que, si se consideran los valores extremos, dichos niveles de pérdida van desde 8 % de los depósitos asegurados y 28 % de los depósitos cubiertos con 90 % de confianza; hasta 17 % de los depósitos asegurados y 71 % de los depósitos cubiertos al 99 % de confianza.

Con el objetivo de mantener un enfoque más conservador, a los valores de pérdida esperados se debería añadir el valor de provisiones correspondiente a las entidades financieras con menor nivel de riesgo. Dicho valor representa 1.03 % de los depósitos asegurados y 3.56 % de los depósitos cubiertos.

Palabras clave: seguro de depósitos, versión asintótica modelo Vasicek, probabilidad de riesgo de incumplimiento, función de distribución de pérdidas, Value at Risk, Tail Value at Risk.

Código JEL: C15, C53, G21, G28, G33, F34.

Abstract

This study determines the target levels for the Deposit Insurance of the Private Financial Sector in Ecuador, with data from January 2007 to June 2016, through credit risk models for estimating the likelihood that financial institutions move to a high-risk rating of compulsory liquidation (default), and its distribution function of expected losses, risk measures and provisions of the Deposit Insurance.

Credit risk models use macroeconomic and financial entities' indicators to estimate probabilities of variations in the entities' risk ratings. Analysis of macroeconomic indicators shows how increases in the interest rates are related to a greater probability that a financial entity moves to a high-risk rating, whereas the entities' indicators show how increases in profitability, current liquidity and portfolio coverage ratio are associated with a lower probability that a financial entity moves to a high-risk rating.

Based on the calculated probabilities that a financial entity moves to a high-risk rating, the distribution function of expected losses of the Deposit Insurance was estimated. Using this function, it was determined that the optimum levels of the Deposit Insurance are equal to an expected loss of 4% of preferred deposits and 12% of covered deposits; whereas if the extreme values of the loss distribution are considered, these optimum levels range from 28% of covered deposits and 8% of preferred deposits with 90% of confidence; to 71% of covered deposits and 17% of preferred deposits with 99% confidence.

To take a more conservative approach, to the expected loss values should be added the value of the provisions regarding the financial entities with lower risk ratings. That value represents 1.03% of preferred deposits and 3.56% of covered deposits.

Keywords: deposit insurance, asymptotic version of Vasicek model, default probability, loss distribution function, Value at Risk, Tail Value at Risk.

Código JEL: C15, C53, G21, G28, G33, F34.

Índice general

Resumen	2
Índice general	4
1 Marco teórico	13
1.1 Modelos de riesgo crediticio	13
1.2 Modelo asintótico unifactorial de Vasicek	16
1.3 Modelos de variable dependiente binaria <i>logit</i>	18
1.4 Bondad de ajuste y especificación del modelo	22
1.5 Metodología <i>stepwise</i>	24
1.6 Estimación de la pérdida esperada	25
1.7 Generación de la función de distribución empírica	26
1.8 Medidas de riesgo: valor en riesgo y valor en riesgo condicional	28
1.8.1 Valor en riesgo (VaR)	28
1.8.2 Valor en riesgo condicional (TVaR)	28
1.9 Metodologías para la definición del porcentaje de recuperación	29
2 Datos y resultados	31
2.1 Definición de variables	32
2.1.1 Variables independientes	32
2.1.1.1 Variables financieras	32
2.1.1.2 Variables macroeconómicas	33
2.1.2 Variable dependiente	34
2.2 Análisis exploratorio de los datos	35
2.3 Estimación de las probabilidades de riesgo alto	39
2.3.1 Matrices de correlación de las entidades financieras	46
2.3.1.1 Matriz de correlaciones: Sector Financiero Privado	46
2.3.1.2 Matriz de correlaciones: bancos	47
2.3.1.3 Matriz de correlaciones: mutualistas	47
2.3.1.4 Matriz de correlaciones: sociedades financieras	48

2.4	Definición de los parámetros para la estimación de la función de pérdidas	48
2.5	Construcción del proceso de simulación	49
2.6	Funciones probabilísticas de las pérdidas	50
2.6.1	Función de densidad de las pérdidas	50
2.6.2	Función de distribución empírica de las pérdidas	51
2.7	Tasa de transición a calificación de riesgo alto (incumplimiento)	52
2.8	Determinación del valor en riesgo y valor en riesgo condicional	55
2.9	Determinación del porcentaje de recuperación de cartera	56
2.10	Provisión de pérdidas para entidades con menor nivel de riesgo	56
2.11	Resumen de resultados	58
3	Conclusiones y recomendaciones	61
3.1	Conclusiones	61
3.2	Recomendaciones	65
	Referencias	67
Apéndice 1	Marco teórico	69
1.1	Teorema: probabilidades de incumplimiento	69
1.2	Portafolio homogéneo grande (<i>LHP</i>)	70
1.2.1	Lema: tasa de incumplimiento	71
1.2.1.1	Prueba del lema	71
1.2.1.2	Teorema sobre la tasa de incumplimiento	72
1.2.1.3	Prueba del teorema	72
1.3	Modelos espacio-estado: filtro de Kalman	72
Apéndice 2	Estimación de los modelos <i>logit</i> para las entidades financieras	79
2.1	Entidad financiera X_2	79
2.2	Entidad financiera X_3	84
2.3	Entidad financiera X_4	89
2.4	Entidad financiera X_5	94
2.5	Entidad financiera X_6	99
2.6	Entidad financiera X_7	103
2.7	Entidad financiera X_8	107
2.8	Provisiones bancos	111
2.9	Provisiones mutualistas	112

2.10 Provisiones sociedades financieras	113
---	-----

Índice de figuras

1.1 Función de densidad de las pérdidas	20
1.2 Funciones de densidad de las distribuciones normal y logística	23
1.3 Funciones de distribución acumulada normal y logística	23
2.1 Morosidad, improductividad y cobertura de cartera	42
2.2 Liquidez, margen operativo y tasa de interés	42
2.3 Rentabilidad del capital, variación del precio del petróleo, y patrimonio técnico constituido / activos y contingentes ponderados por riesgo	42
2.4 Índice de Actividad Económica Coyuntural, rentabilidad y apalancamiento del activo	42
2.5 Probabilidades de riesgo alto	47
2.6 <i>Effect plot</i> : improductividad de activos vs. nivel de riesgo	48
2.7 <i>Effect plot</i> : cobertura de cartera vs. nivel de riesgo	49
2.8 <i>Effect plot</i> : rentabilidad vs. nivel de riesgo	49
2.9 <i>Effect plot</i> : PTAC vs. nivel de riesgo	50
2.10 Función de densidad empírica de las pérdidas	54
2.11 Función de distribución empírica de las pérdidas	55
2.12 Función de densidad empírica de las tasas de riesgo alto	57
2.13 Función de distribución empírica de las tasas de riesgo alto	58
2.1 X_2 : probabilidades de riesgo alto	86
2.2 X_2 : apalancamiento del activo vs. nivel de riesgo	87
2.3 X_2 : rentabilidad del capital vs. nivel de riesgo	87
2.4 X_2 : variación del Índice de Actividad Económica Coyuntural vs. nivel de riesgo	88
2.5 X_3 : probabilidades de riesgo alto	91
2.6 X_3 : improductividad activos vs. nivel de riesgo	92
2.7 X_3 : liquidez corriente vs. nivel de riesgo	92
2.8 X_3 : variación del Índice de Actividad Económica Coyuntural vs. nivel de riesgo	92

2.9	X_3 : rentabilidad vs. nivel de riesgo	92
2.10	X_3 : tasa de interés vs. nivel de riesgo	93
2.11	X_4 : probabilidades de riesgo alto	96
2.12	X_4 : cobertura de la cartera vs. nivel de riesgo	97
2.13	X_4 : improductividad de activos vs. nivel de riesgo	97
2.14	X_4 : margen operativo vs. nivel de riesgo	97
2.15	X_4 : rentabilidad vs. nivel de riesgo	97
2.16	X_4 : liquidez corriente vs. nivel de riesgo	98
2.17	X_5 : apalancamiento del activo vs. nivel de riesgo	101
2.18	X_5 : tasa de interés vs. nivel de riesgo	101
2.19	X_5 : liquidez corriente vs. nivel de riesgo	102
2.20	X_5 : cobertura de la cartera vs. nivel de riesgo	102
2.21	X_5 : margen operativo vs. nivel de riesgo	102
2.22	X_6 : probabilidades de riesgo alto	105
2.23	X_6 : variación del Índice de Actividad Económica Coyuntural vs. nivel de riesgo	106
2.24	X_6 : rentabilidad vs. nivel de riesgo	106
2.25	X_6 : tasa de interés vs. nivel de riesgo	107
2.26	X_7 : probabilidades de riesgo alto	109
2.27	X_7 : apalancamiento del activo vs. nivel de riesgo	110
2.28	X_7 : variación del Índice de Actividad Económica Coyuntural vs. nivel de riesgo	110
2.29	X_7 : cobertura de cartera vs. nivel de riesgo	111
2.30	X_8 : cobertura de cartera vs. nivel de riesgo	114
2.31	X_8 : patrimonio técnico constituido / activos y contingentes pon- derados por riesgo vs. nivel de riesgo	114
2.32	X_8 : rentabilidad vs. nivel de riesgo	115
2.33	X_8 : variación del Índice de Actividad Económica Coyuntural vs. nivel de riesgo	115

Índice de cuadros

2.1	VARIABLES explicativas financieras	36
2.2	Niveles de riesgo	39
2.3	Resumen de estadísticas - parte 1	39

2.4	Resumen de estadísticas - parte 2	40
2.5	Matriz de correlaciones de las variables independientes	41
2.6	X_1 : modelo <i>logit</i> con variable dependiente multinomial	44
2.7	Estadísticos de las variables independientes	45
2.8	Validación cruzada: nivel de riesgo CAMEL vs. nivel de riesgo predicho por el modelo	46
2.9	Pseudo R^2	46
2.10	X_1 : resumen de las probabilidades de riesgo alto	48
2.11	Matriz de correlaciones: entidades con probabilidad de riesgo alto	50
2.12	Matriz de correlaciones: bancos	51
2.13	Matriz de correlaciones: mutualistas	51
2.14	Matriz de correlaciones: sociedades financieras	52
2.15	Resumen de las tasas de riesgo alto	56
2.16	Medidas de riesgo VaR y TVaR	59
2.17	Provisiones de pérdidas según el nivel de riesgo	61
2.18	Provisiones por tipo de entidad	62
2.1	X_2 : modelo <i>logit</i> con variable dependiente binaria	84
2.2	Validación cruzada: nivel de riesgo CAMEL vs. nivel de riesgo predicho por el modelo	85
2.3	Pseudo R^2	85
2.4	X_2 : resumen de las probabilidades de riesgo alto	86
2.5	X_3 : modelo <i>logit</i> con variable dependiente multinomial	89
2.6	Validación cruzada: nivel de riesgo CAMEL vs. nivel de riesgo predicho por el modelo	90
2.7	Pseudo R^2	90
2.8	X_3 : resumen de las probabilidades de riesgo alto	91
2.9	X_4 : modelo <i>logit</i> con variable dependiente multinomial	94
2.10	Validación cruzada: nivel de riesgo CAMEL vs. nivel de riesgo predicho por el modelo	95
2.11	Pseudo R^2	95
2.12	X_4 : resumen de las probabilidades de riesgo alto	96
2.13	X_5 : modelo <i>logit</i> con variable dependiente binaria	99
2.14	Validación cruzada: nivel de riesgo CAMEL vs. nivel de riesgo predicho por el modelo	100
2.15	Pseudo R^2	100
2.16	X_5 : resumen de las probabilidades de riesgo alto	101
2.17	X_6 : modelo <i>logit</i> con variable dependiente multinomial	103
2.18	Validación cruzada: nivel de riesgo CAMEL vs. nivel de riesgo predicho por el modelo	104
2.19	Pseudo R^2	104

2.20	X_6 : resumen de las probabilidades de riesgo alto	105
2.21	X_7 : modelo <i>logit</i> con variable dependiente multinomial	108
2.22	Validación cruzada: nivel de riesgo CAMEL vs. nivel de riesgo predicho por el modelo	108
2.23	Pseudo R^2	109
2.24	X_7 : resumen de las probabilidades de riesgo alto	110
2.25	X_8 : modelo <i>logit</i> con variable dependiente binaria	112
2.26	Validación cruzada: nivel de riesgo CAMEL vs. nivel de riesgo predicho por el modelo	113
2.27	Pseudo R^2	113
2.28	X_8 : resumen de las probabilidades de riesgo alto	114
2.29	Provisiones bancos	116
2.30	Provisiones mutualistas	117
2.31	Provisiones sociedades financieras	117

Glosario

AIC: criterio de información de Akaike (*Akaike Information Criterion*).

Algoritmo EM: algoritmo esperanza-maximización (*Expected Maximization Algorithm*).

BIC: criterio de información de Bayes (*Bayesian Information Criterion*).

BLUE: mejor estimador lineal insesgado (*Best Linear Unbiased Estimator*).

CAMEL rating: calificación de riesgo calculada con base en 4 aspectos: adecuación de capital, activos, gestión, rentabilidad y liquidez. (*Capital adequacy, Assets, Management Capability, Earnings, Liquidity*).

Depósitos asegurados: depósitos a la vista o a plazo fijo efectuados por personas naturales o jurídicas en las entidades financieras de los Sectores Financieros Privado y Popular y Solidario, bajo la forma de cuentas corrientes, de ahorros, depósitos a plazo fijo u otras modalidades legalmente aceptadas, de acuerdo con la ley vigente en el Ecuador para el pago del seguro.

Depósitos cubiertos: depósitos que el Seguro de Depósitos cubre para cada persona natural o jurídica, los cuales son diferenciados por cada uno de los sectores financieros asegurados; para el caso de las entidades financieras privadas, los depósitos cubiertos son iguales a dos veces la fracción básica exenta vigente del impuesto a la renta, pero en ningún caso inferior a US\$ 32.000,00 (treinta y dos mil dólares de los Estados Unidos de América).

IADI: Asociación Internacional de Aseguradores de Depósitos (*International Association of Deposit Insurers*).

LGD: pérdida dado el incumplimiento (*Loss Given Default*).

LHP: portafolio homogéneo grande (*Large Homogeneous Portfolio*).

PTAC: patrimonio técnico constituido respecto a activos y contingentes ponderados por riesgo.

ML: máxima verosimilitud (*Maximum Likelihood*) .

Metodología stepwise: metodología paso a paso.

OLS: mínimos cuadrados ordinarios (*Ordinary least squares*).

RR: tasa de recuperación (*Recovery Rate*).

RSS: suma de cuadrados de residuos (*Residual Sum of Squares*).

Seguro de Depósitos: sistema establecido para proteger (de forma limitada) los depósitos efectuados por personas naturales y jurídicas en las entidades de los Sectores Financieros Privado y Popular y Solidario autorizadas por los respectivos organismos de control. El Seguro de Depósitos se compone de los siguientes recursos, en lo que a cada sector financiero corresponde: las contribuciones que realizan las entidades de los Sectores Financieros Privado y Popular y Solidario, de conformidad con lo previsto en la ley vigente; el rendimiento de las inversiones y las utilidades líquidas de cada ejercicio anual; las donaciones que reciba; los provenientes de préstamos o líneas contingentes obtenidos para el financiamiento de sus actividades; los provenientes de préstamos entre los fideicomisos del Seguro de Depósitos; y, los pagos derivados de la liquidación forzosa de una entidad financiera.

TVaR: valor en riesgo condicional (*Tail Value at Risk*).

VaR: valor en riesgo (*Value at Risk*).

Introducción

Un tema ampliamente discutido en la actualidad es la gestión de los niveles de riesgo que enfrentan las entidades financieras, para lo cual las autoridades emplean mecanismos de supervisión y control, con el objetivo de procurar la estabilidad y sostenibilidad de los sistemas financieros. Para controlar los costos asociados con posibles problemas financieros, las autoridades implementan redes de seguridad financiera que buscan reducir la probabilidad de que ocurra la liquidación forzosa de una entidad financiera y minimizar su impacto en la economía.

Los principales mecanismos de seguridad financiera que estas redes suelen incluir son: supervisión financiera, fondo de liquidez, fortalecimiento patrimonial, resolución financiera, seguro de depósitos, entre otros.

En este sentido, el establecimiento de un Seguro de Depósitos explícito ha llegado a ser considerado como uno de los pilares fundamentales sobre el cual debe construirse un sistema financiero moderno. Es así que el número de jurisdicciones que ofrecen un Seguro de Depósitos explícito ha aumentado rápidamente, pasando de 12 en 1974 a 113 en el año 2014, según datos de la Asociación Internacional de Aseguradores de Depósitos (IADI). Hoy en día la mayoría de los países de la Organización para la Cooperación y el Desarrollo Económico (OCDE) y un creciente número de países emergentes cuentan con un sistema de Seguro de Depósitos explícito.

Dada la popularidad que ha tenido el Seguro de Depósitos explícito, se podría pensar que el diseño y la operación de un sistema eficiente es relativamente sencillo. Sin embargo, en la práctica los administradores de este mecanismo de seguridad financiera se enfrentan a una constante disyuntiva que hace su tarea bastante compleja ya que no solo deben procurar la estabilidad y sostenibilidad financiera, sino también evitar que las entidades financieras asuman mayores riesgos y otras prácticas que podrían resultar imprudentes.

Uno de los principales desafíos a los que se enfrentan los gestores de los seguros de depósitos consiste en establecer un adecuado equilibrio entre la prevención de una crisis y el control respecto al nivel de riesgo asumido por parte de las entidades financieras.

Con el fin de que las entidades financieras mantengan niveles de riesgo adecuados se han desarrollado diversos instrumentos como: requerimientos de calidad de activos, niveles mínimos de patrimonio, control de pasivos, entre otros.

Además, un enfoque alternativo más reciente consiste en utilizar modelos de riesgo de crédito en la administración de los Seguros de Depósitos, con el objeto de probar si estos modelos permiten cuantificar el riesgo de una cartera compuesta por las entidades financieras que asegura.

Este nuevo instrumento sugiere que “si el seguro de depósitos es visto como un servicio que utilizan las entidades financieras, la pregunta es ¿cuál es el precio que debe tener este servicio?”. (Maccario, Sironi, & Zazzara, 2003, p. 2)

Una posible respuesta consiste en que el precio debe reflejar el riesgo que la entidad financiera representa para el sistema de Seguro de Depósitos. Este enfoque de pérdida esperada de los precios se ajusta con las mejores prácticas que se han desarrollado en la industria financiera en los últimos años, esencialmente los principios de Basilea.

Adicionalmente conviene destacar que en el caso de Ecuador, la crisis financiera acaecida en 1999 develaría la necesidad de un esquema de supervisión estricto por parte de los organismos de control sobre las entidades financieras del país, con el objetivo de proteger los recursos de los depositantes a través de la regulación de las operaciones de colocación de créditos. En tal sentido a partir del año 2007 se diseña una nueva arquitectura financiera nacional, desarrollándose varios mecanismos de seguridad financiera, que junto a un oportuno y efectivo control estatal de carácter preventivo (basado en un enfoque de riesgos), permita definir el conjunto de procedimientos y medidas para resolver situaciones de liquidación de entidades financieras, beneficiando primordialmente a los depositantes.

Como parte de la nueva arquitectura financiera nacional, en el año 2008 se expide la Ley de Creación de la Red de Seguridad Financiera, la cual parte del esquema de dolarización adoptado a finales del siglo XX y se ampara en el artículo 308 de la Constitución de la República, el cual determina que las actividades financieras son un servicio de orden público que deberá tener como finalidad fundamental la preservación de los depósitos y atender los requerimientos de financiamiento para la consecución de los objetivos de desarrollo del país.

Dicha ley dictó la creación de la Corporación del Seguro de Depósitos; que más adelante, en el año 2014, con la expedición del Código Orgánico Monetario y Financiero, se transformó en la Corporación del Seguro de Depósitos, Fondo de Liquidez y Fondo de Seguros Privados (COSEDE), siendo ahora esta institución la encargada de la administración general de algunos de los principales mecanismos de seguridad financiera.

Considerando lo expuesto, la presente investigación tiene como objetivo determinar los niveles óptimos del patrimonio del Seguro de Depósitos del Sector Financiero Privado. Para ello se emplea modelos de riesgo crediticio que estiman las probabilidades de que las entidades financieras presenten un riesgo alto de liquidación forzosa (incumplimiento) y la función de distribución de pérdidas esperadas del Seguro de Depósitos.

El presente trabajo se divide en tres capítulos. El capítulo 1 describe el marco teórico mediante la explicación de los modelos econométricos y las técnicas estadísticas utilizadas. El capítulo 2 describe la base de datos, define las variables y presenta los resultados de estimación, tanto de los modelos como de la pérdida esperada y medidas de riesgo. Finalmente, el capítulo 3 presenta las conclusiones del estudio y recomendaciones para futuros trabajos de investigación.

Capítulo 1

Marco teórico

1.1 Modelos de riesgo crediticio

En esta sección se describe las diferentes metodologías de los modelos de riesgo de crédito y su aplicabilidad dentro de un sistema de Seguro de Depósitos.

El riesgo de crédito se puede definir como la probabilidad de que una entidad financiera incurra en pérdidas debido a la falta de pago de un deudor, sea de un préstamo u otra línea de financiamiento. Este evento se produce cuando el deudor es incapaz de cumplir con su obligación legal de acuerdo con el contrato de financiamiento. Por ejemplo, falta de pago de títulos, quiebra corporativa, falta de pago de la tarjeta de crédito o la ejecución de una hipoteca.

Los modelos de medición del riesgo de crédito se pueden agrupar en dos categorías principales, cuya diferencia radica en la disponibilidad de los datos: mediciones directas de crédito y precios observados en el mercado de crédito.

Las dos categorías de modelos de medición de riesgo de crédito pueden ser caracterizadas por los tipos de probabilidad de incumplimiento, siendo el primer tipo de probabilidad el observado y el otro implícito. Dicho de otro modo, la primera categoría de riesgo de crédito corresponde a las observaciones directas de los valores de incumplimiento (riesgo explícito). Mientras que la segunda categoría de riesgo de crédito se refiere a la probabilidad de incumplimiento tomada de datos financieros e información macroeconómica disponible (riesgo implícito). (Zhang, 2009)

Actualmente la literatura académica de riesgo de crédito se ha inclinado mayoritariamente al estudio de las probabilidades de incumplimiento, es decir, al riesgo implícito de las entidades financieras a través de la estimación de modelos estadísticos.

El principal resultado de un modelo de riesgo de crédito es la función de densidad de las pérdidas esperadas¹ del portafolio. A partir del análisis de dicha distribución de pérdidas, una entidad financiera puede estimar la pérdida esperada de su cartera de crédito. La misma es igual a la media (incondicional) de la distribución de pérdidas, que representa la cantidad que la entidad puede esperar perder dentro de un período específico de tiempo (generalmente un año).

Las entidades financieras aplican modelos de riesgo de crédito para evaluar el capital económico necesario para hacer frente a los riesgos asociados con sus carteras de crédito. En este sentido, las provisiones para pérdidas de crédito deben cubrir las pérdidas esperadas. Del mismo modo se puede aplicar esta lógica a las exposiciones de crédito de los sistemas de Seguro de Depósitos explícitos. En el caso de un portafolio de crédito bancario, la pérdida se produce cuando un prestatario no paga, mientras que en el caso de un Seguro de Depósitos la pérdida se da cuando una entidad asegurada cae en liquidación forzosa, generándose un pago por parte del Seguro de Depósitos.

Siguiendo este enfoque, un Seguro de Depósitos puede ser comparado con una entidad financiera en la medición de riesgo de incumplimiento de sus contrapartes. Mientras para las entidades financieras las contrapartes son los prestatarios, para las entidades de Seguro de Depósitos son las propias entidades financieras, con exposiciones equivalentes a los depósitos cubiertos. El Seguro de Depósitos puede entenderse como el administrador de una cartera de exposiciones de crédito contingente para las entidades que asegura. (Maccario et al., 2003, p. 4-6)

Por otra parte, las distribuciones empíricas de las pérdidas de las carteras de crédito son generalmente asimétricas con mayor densidad en los valores más bajos de pérdidas. Es decir, la probabilidad de incurrir en pérdidas

¹Una función de densidad es una función matemática de una variable aleatoria continua cuyo área bajo la curva representa la probabilidad de que el valor de esa variable caiga en un intervalo determinado.

extremadamente altas tiende a cero. Sin embargo, es mayor que la probabilidad que se obtendría de una distribución normal estándar en estos niveles.

Por tanto, al enfrentarse el Seguro de Depósitos a una situación similar a la distribución de pérdidas de una cartera de crédito, la probabilidad de ocurrencia de pérdidas con valores bajos es relativamente alta debido al incumplimiento de entidades financieras pequeñas, mientras que a medida que las pérdidas son mayores se reduce la probabilidad asintóticamente hacia cero, lo cual se daría en caso de que el incumplimiento sea por parte de las entidades financieras grandes.

En esta investigación, por incumplimiento se entiende la probabilidad de que una entidad financiera pase a una calificación de riesgo alto, esta definición difiere del concepto tradicional de incumplimiento que se entiende como la probabilidad de liquidación de una entidad.

La probabilidad de incumplimiento es mayor a aquella que se generaría de una distribución normal, lo que se ilustra en la figura 1.1, en la cual se observa la asimetría de la distribución hacia la derecha, siendo la cola evidentemente más densa que la de una distribución normal estándar². Con lo cual, en la medida en que la distribución sea asimétrica, determinar los cuantiles extremos de la distribución se torna un aspecto clave.

²Datos simulados de una distribución empírica.

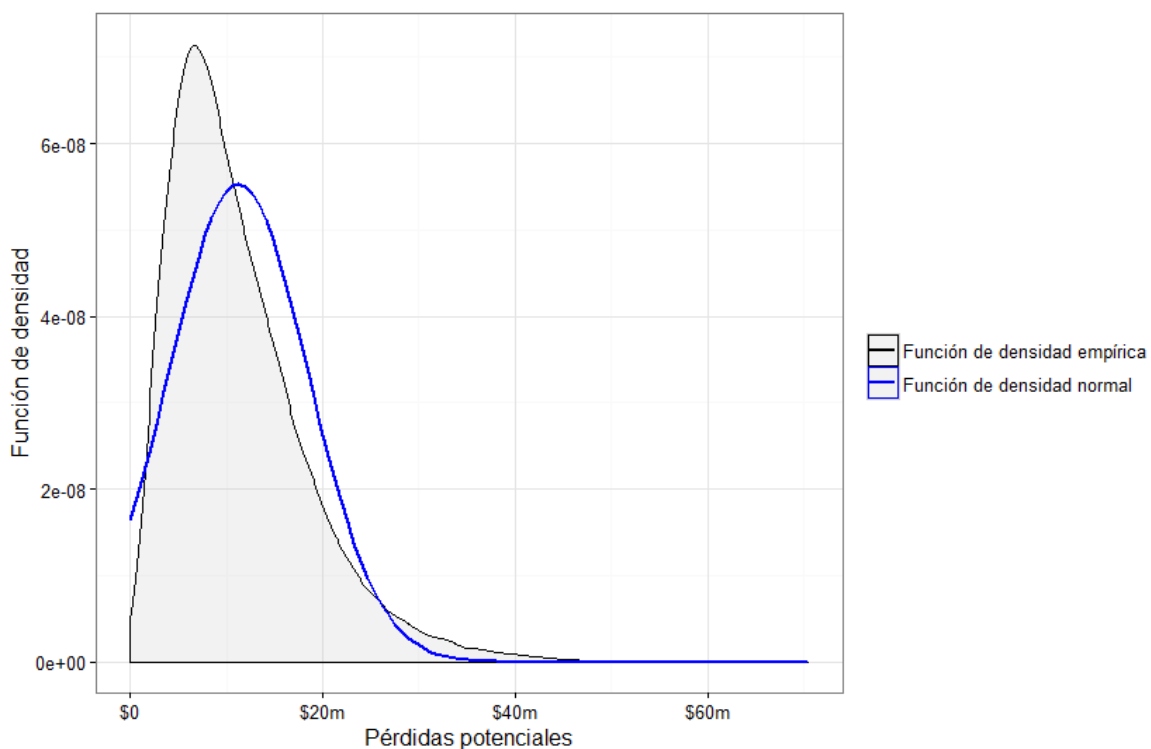


Figura 1.1: Función de densidad de las pérdidas

El objetivo del presente trabajo es aproximar la distribución de pérdidas del Seguro de Depósitos para el caso de Ecuador, a partir de la estimación de las probabilidades de que las entidades pasen a riesgo alto y de la simulación de las pérdidas esperadas del Seguro de Depósitos empleando la versión asintótica del modelo de Vasicek.

Esta distribución resulta útil para identificar el nivel apropiado de los recursos necesarios para hacer frente a las posibles pérdidas potenciales del Seguro de Depósitos.

1.2 Modelo asintótico unifactorial de Vasicek

El modelo asintótico unifactorial de Vasicek resulta atractivo dentro del campo de la estadística actuarial, en razón de que su objetivo es ofrecer una forma práctica para obtener la función de distribución tanto de la morosidad de la cartera como de las tasas de pérdida de una entidad

financiera.(Analytics, 2014)

Para obtener la tasa de morosidad de la cartera no solo es necesario conocer las probabilidades de incumplimiento $(p_{1,t}, \dots, p_{N,t})$, sino también la estructura de correlación de estas probabilidades, dado que la única variable aleatoria que afecta el estado de cada entidad financiera n en el tiempo t (incumplimiento o no incumplimiento) es $X_{n,t}$, la cual representa las pérdidas. De esta manera, la estructura de correlación entre las probabilidades de incumplimiento de las entidades se debe introducir a través de la variable aleatoria normal $X_{1,t}, \dots, X_{N,t}$. Se asume que el coeficiente de correlación para cada par de variables aleatorias $X_{n,t}$ y $X_{m,t}$ es $\rho_{n,m,t}$.(Vasicek, 1987)

Por otro lado, se puede pensar que existe un factor aleatorio o fuente de incertidumbre que afecta a todas las entidades financieras exactamente de la misma manera. Así, se pueden expresar las variables aleatorias $X_{n,t}, \dots, X_{m,t}$ de la siguiente forma:

$$X_{n,t} = \sqrt{\rho_t} * Y_t + \sqrt{1 - \rho_t} * \epsilon_{n,t} \quad (1.1)$$

Para cada $n = 1, \dots, N$, donde Y_t y $\epsilon_{1,t}, \dots, \epsilon_{N,t}$ son independientes e idénticamente distribuidas como una distribución normal estándar.

La fórmula 1.1 se puede interpretar tal que la variable aleatoria que representa las pérdidas ($X_{n,t}$) se expresa como la suma de dos factores de riesgo, por un lado el riesgo común (Y_t) y por otro el riesgo idiosincrático ($\epsilon_{n,t}$) que es particular para cada entidad financiera.

Por tanto:

- Y_t es el riesgo sistemático común que afecta a todas las entidades financieras (estado de la economía).
- $\epsilon_{n,t}$ es el factor idiosincrático, particular para cada entidad financiera.
- ρ_t definido como $Corr(X_{n,t}, X_{m,t})$ es el coeficiente de correlación entre cada par de entidades financieras, que para este modelo se asume igual en todo n y m .

1.3 Modelos de variable dependiente binaria *logit*

Los modelos econométricos *logit* y *probit* son apropiados para modelar una variable dependiente dicotómica o multinomial (más de dos categorías de respuesta) en relación a un conjunto de variables categóricas o numéricas (escalas de intervalo o ratios). La necesidad de utilizar estos modelos surge en virtud de que la aproximación clásica a través del modelo de regresión lineal presenta varios problemas en cuanto a la estimación.³

Cuando la variable dependiente es categórica, la estimación por *Mínimos Cuadrados Ordinarios* ya no produce el mejor estimador lineal insesgado (*BLUE*). Por consiguiente, se han desarrollado varios modelos de regresión para trabajar con este tipo de variables. (Leo, 2014)

Los modelos *logit* y *probit* pretenden resolver estos problemas mediante el ajuste de una función no lineal, mientras que las estimaciones con el modelo clásico de regresión se obtienen a partir de la siguiente función:

$$\hat{Y} = X\beta + \mu \quad (1.2)$$

Además estos modelos asumen que la relación entre las variables no es lineal y por tanto la función adopta otra forma.

$$\hat{Y} = f(X\beta + \mu) \quad (1.3)$$

Dicha función no es una línea recta como en el caso del modelo de regresión clásico, sino una curva en forma de *S* que: 1) respeta los límites de la variable dependiente; 2) admite diferentes tasas de variación en los extremos inferior y superior; y 3) elimina el problema de heterocedasticidad, asumiendo especificación correcta de las variables independientes.

La principal diferencia entre los modelos *logit* y *probit* consiste en la forma de la distribución de los errores (perturbaciones) que asumen ambos modelos. Por un lado, el modelo *logit* utiliza la función de distribución acumulada de la distribución logística, cuya función de densidad viene dada

³La forma funcional de la técnica conocida como *OLS* puede conducir a predicciones que no correspondan a probabilidades y por tanto se encuentren fuera del intervalo $[0, 1]$, asumiendo el mismo efecto marginal sobre la variable dependiente. Además, esta técnica suele revelar problemas de heterocedasticidad con este tipo de variables.

por $\lambda(\epsilon) = e^\epsilon / (1 + e^\epsilon)^2$, con media 0 y varianza $\frac{\pi^2}{3}$. Mientras que el modelo *probit* utiliza la función de distribución acumulada de la distribución normal estándar, $\Phi(\epsilon) = \frac{1}{\sqrt{2\pi}} e^{-\frac{\epsilon^2}{2}}$, con varianza 1. (Park, 2009)

Ambas funciones transforman los valores de tal manera que se encuentren en el intervalo $[0,1]$. Por lo tanto, cualquiera que sea el valor de $X\beta + \mu$, se puede transformar para obtener una probabilidad.

Como se puede apreciar en las siguientes figuras, un punto importante a destacar sobre la forma de estas funciones consiste en que la función de densidad (PDF) de la distribución normal estándar tiene un pico más alto y colas más delgadas que la distribución de probabilidad logística estándar. En la función de distribución logística estándar el pico es más bajo, pero las colas son más densas (*heavy tails*). Como resultado, la función de densidad acumulada (CDF) de la distribución normal estándar es más pronunciada para los valores centrales que la CDF de la distribución logística estándar y se acerca más rápidamente a cero en ambas colas. (Fox, 2010)

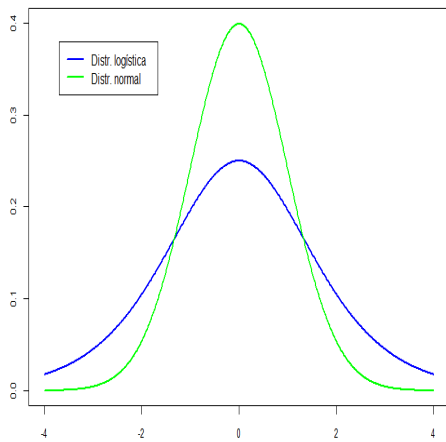


Figura 1.2: Funciones de densidad de las distribuciones normal y logística

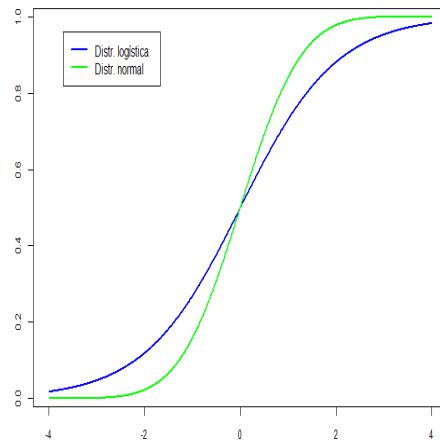


Figura 1.3: Funciones de distribución acumulada normal y logística

Los dos modelos producen diferentes estimaciones de parámetros. En los modelos de respuesta discreta, las estimaciones de un modelo *logit* son aproximadamente $\pi/\sqrt{3}$ veces más grandes que las del modelo *probit*. Sin embargo, los estimadores de las variables independientes tienen los mismos efectos estandarizados en ambos modelos.

La elección entre modelos *logit* o *probit* está relacionada con la estimación y la familiaridad con cada modelo, más que a aspectos teóricos o de interpretación. En general, los modelos *logit* alcanzan convergencia bastante bien. Actualmente, la potencia de cálculo ha mejorado y se han desarrollado nuevos algoritmos que hacen que la importancia de esta decisión sea cada vez menor.⁴

Estimación del modelo econométrico *logit*

Asumiendo que se tienen k ensayos de Bernoulli, y_1, \dots, y_k , y que la observación i -ésima puede ser tratada como la realización de una variable aleatoria Y_i . Se supone que Y_i tiene una distribución binomial. (Rodríguez, 2007)

$$Y_i \sim \text{Binomial}(n_i, \pi_i) \quad (1.4)$$

Con denominador binomial n_i , probabilidad de éxito π_i y datos individuales $n_i = 1$ para todo i . Además, se supone que el *logit* de la probabilidad subyacente π_i es una función lineal de los predictores.

$$\text{logit}(\pi_i) = x_i' \beta, \quad (1.5)$$

Donde x_i es un vector de covariables y β es el vector de coeficientes de la regresión.

El modelo definido en las ecuaciones 1.4 y 1.5 es un modelo lineal generalizado con respuesta de tipo *binomial* y función de enlace *logit*.

Los coeficientes de regresión pueden interpretarse de la misma manera que en los modelos lineales, considerando que la variable dependiente es un *logit* y no una media. Por lo tanto, β_j representa el cambio en el *logit* de la probabilidad asociado a una unidad de cambio en la j -ésima variable predictora, manteniendo todo lo demás constante. Así, también es posible expresar los resultados en la escala *logit*, lo cual resulta más sencillo de interpretar.

Si se obtiene la exponencial de la ecuación 1.5 se tendría que la probabilidad para la i -ésima unidad está dada por:

⁴Los modelos *probit* y *logit* son estimados por el método de Máxima Verosimilitud (*ML*), en lugar del método de *Mínimos Cuadrados Ordinarios* (*MCO*).

$$\frac{\pi_i}{1 - \pi_i} = \exp(x'_i\beta) \quad (1.6)$$

Esta expresión define un modelo multiplicativo de las probabilidades. Por ejemplo, si se tuviera que cambiar el predictor *i-ésimo* en una unidad manteniendo las demás variables constantes, se multiplicarían las probabilidades por $\exp(\beta_j)$.

Bajo el supuesto de que el predictor $x'_i\beta$ es lineal y se incrementa x_j en una unidad, se obtiene $x'_i\beta + \beta_j$. Si se calcula su exponencial, esta expresión sería $\exp(x'_i\beta) * \exp(\beta_j)$. Por lo tanto, el coeficiente exponencial $\exp(\beta_j)$ es lo que se conoce como *odds ratio* o razón de probabilidades. Transformar los resultados en un efecto multiplicativo de los *odds ratio*, es a menudo bastante útil ya que los resultados se expresan en una escala más sencilla.

Despejando la probabilidad π_i en el modelo *logit* de la ecuación 1.6 se obtiene lo siguiente:

$$\pi_i = \frac{\exp(x'_i\beta)}{1 + \exp(x'_i\beta)} \quad (1.7)$$

Mientras que el lado izquierdo de la igualdad está expresado en términos de probabilidad, el lado derecho es una función no lineal de los predictores. De esta forma, se expresa el efecto en la probabilidad dado un aumento de una unidad en la variable predictora. Es posible obtener una respuesta aproximada tomando la derivada con respecto a x_j , que evidentemente tiene sentido únicamente para variables predictoras continuas.

Para ello se utiliza la regla del cociente:

$$\frac{d\pi_i}{dx_{ij}} = \beta_j\pi_i(1 - \pi_i) \quad (1.8)$$

Por lo tanto, el efecto del predictor *j-ésimo* en la probabilidad π_i depende del coeficiente β_j .

1.4 Bondad de ajuste y especificación del modelo

En este apartado se estudia los principales criterios estadísticos empleados para medir tanto la bondad de ajuste como la especificación de los modelos vistos en la sección 1.1.

Para los modelos econométricos basados en variables dependientes categóricas existen dos enfoques para determinar si el modelo estimado se ajusta al conjunto de datos. El primer enfoque consiste en emplear medidas que permitan evaluar la capacidad predictiva del modelo. El segundo enfoque prueba si el modelo debe ser más complejo (por ejemplo, adición o diferenciación de variables explicativas), si la especificación es correcta o se necesita incluir interacciones para representar satisfactoriamente los datos.

La capacidad predictiva del modelo se evalúa a través de una tabla cruzada que contrasta los valores observados de la variable respuesta, con los valores estimados por el modelo para cada categoría.

Respecto a la bondad de ajuste, en modelos *OLS* es posible obtener una medida conocida como R^2 , la cual representa: el porcentaje de la varianza explicada por los regresores del modelo, el cuadrado del coeficiente de correlación entre los datos ajustados y predichos y la transformación del estadístico F . Sin embargo, al estimar modelos con variable respuesta de tipo categórica, por ejemplo los modelos *logit*, no existe una medida estadística equivalente al R^2 , ya que las estimaciones de los modelos de regresión logística se realizan por máxima verosimilitud a través de un proceso iterativo, es decir, no se calculan con el objetivo de minimizar la varianza de los errores. Sin embargo, hay una serie de medidas conocidas como pseudo- R^2 que se han desarrollado con el fin de evaluar el ajuste de estos modelos.

Entre ellas se encuentran:

- R^2 **McFadden**: se conoce como el índice de la razón de verosimilitud. Esta medida compara la verosimilitud del modelo incluyendo todos los predictores versus el modelo considerando únicamente el intercepto. Para un modelo de regresión logística, el logaritmo de la verosimilitud es siempre negativo (el logaritmo de una probabilidad será un valor igual o menor a cero). Si el modelo que incluye tanto el intercepto como las variables regresoras no predice la variable respuesta mejor que el modelo que considera únicamente el intercepto, el logaritmo

del modelo completo no será mucho más alto que el logaritmo del modelo sin variables independientes, con lo cual el ratio entre estos dos será de aproximadamente uno. De esta manera, según la forma de cálculo ($1 -$ razón de verosimilitud), el resultado será cercano a cero. Este indicador tiende a ser más pequeño que el R^2 , valores entre 0.2 y 0.4 se consideran satisfactorios.

$$R_{McFadden}^2 = 1 - \frac{L_\beta}{L_\alpha}$$

- **R^2 Verosimilitud máxima:** expresa el ajuste del modelo como una transformación de la razón de verosimilitud χ^2 , de forma análoga a la de R^2 en regresión por mínimos cuadrados ordinarios, que se expresa como una transformación del estadístico F .

$$R_{VM}^2 = 1 - \left(\frac{L_\alpha}{L_\beta} \right)^{2/N}$$

$$R_{VM}^2 = 1 - \exp(-G^2/N), \text{ donde:}$$

$$G^2 = -2\ln(L_\alpha/L_\beta)$$

- **R^2 Cragg-Uhler:** Cragg-Uhler propusieron un índice relativo que puede alcanzar el valor de uno. Si el modelo predice perfectamente y tiene probabilidad de uno, el indicador de Cragg y Uhler será igual a uno, si la razón de verosimilitudes es uno, el valor del R^2 será igual a cero.

$$R_{Cragg-Uhler}^2 = \frac{1 - (L_\alpha/L_\beta)^{2/N}}{1 - L_\alpha^{2/N}}$$

Se debe tener presente que no existe un criterio uniforme relacionado con el valor de R^2 y pseudo- R^2 a partir del cual se puede considerar un “buen ajuste”, eso depende principalmente del objetivo del investigador y del tipo de modelo, el cual puede ser explicativo o predictivo. En el presente estudio interesan tanto las relaciones entre las variables (explicativo) como la capacidad predictiva del modelo, con lo que valores pseudo- R^2 muy bajos no serían adecuados para estimar las probabilidades de riesgo alto. Por otro lado, al calcular cada una de estas medidas (pseudo- R^2) con base en criterios distintos, en muchas ocasiones resulta conveniente evaluar varios modelos para comparar los valores, de este modo, mientras más alto sea el valor del pseudo R^2 , mejor será la predicción del modelo respecto a la

variable respuesta.

En el presente estudio se empleó los dos enfoques, es decir, se realizó la validación cruzada del modelo para evaluar su capacidad de predicción y se calculó las medidas conocidas como pseudo- R^2 para determinar la bondad de ajuste del modelo.

1.5 Metodología *stepwise*

En este apartado se revisará el uso de la metodología *stepwise*, la misma que será empleada para la selección del modelo.

La metodología *stepwise* suele ser utilizada cuando existe un número relativamente grande de posibles covariables. Con n covariables existen 2^n modelos posibles. De esta forma, esta regresión se construye a partir de un proceso iterativo que implica la elección de las variables de predicción mediante un procedimiento automático basado en un criterio específico. Esta metodología tiene dos enfoques:

- **Selección hacia adelante**, implica iniciar sin variables en el modelo. Luego, se prueba la adición de cada variable y se va seleccionando la variable que más contribuya a mejorar el modelo. Este proceso se repite hasta que no se produzca alguna mejora significativa con base en el criterio específico.
- **Selección hacia atrás**, implica iniciar con todas las variables en el modelo. Luego, se prueba la eliminación de cada variable; la variable se elimina siempre que su omisión mejore el modelo, dicho proceso se repite hasta que no se produzca mejoría en el modelo por la eliminación de alguna variable ⁵.

Esta metodología suele utilizar el “Criterio de información Akaike” (AIC)⁶ ($AIC = -2 \ln L + 2k$), el cual es un poco más flexible respecto a la inclusión de una variable en el modelo, en comparación con el “Criterio de Información de Bayes” (BIC) ⁷ ($BIC = -2 \ln L + k \ln N$), este criterio es más

⁵En el presente documento se aplicó la regresión paso a paso utilizando ambos métodos.

⁶Sea L el valor máximo de la función de verosimilitud y k el número de parámetros estimados en el modelo.

⁷Sea L el valor máximo de la función de verosimilitud, k el número de parámetros libres que se estima y N el número de datos observados.

riguroso en la selección de una variable. En general la elección de uno u otro criterio depende de la finalidad para la cual se desarrolle el modelo.

Como se mencionó, esta metodología es ampliamente utilizada cuando se tiene un gran número de posibles variables explicativas. El procedimiento se utiliza principalmente en el análisis de regresión. Sin embargo, el enfoque básico es aplicable en muchas formas de selección del modelo.

Con esta sección concluye el marco teórico referente a la primera parte de la presente investigación, la cual corresponde a los modelos econométricos empleados para la estimación de las probabilidades de incumplimiento (riesgo alto), la segunda parte comprende la estimación de la pérdida esperada basada en la técnica de simulación.

1.6 Estimación de la pérdida esperada

En esta sección se describe la segunda parte del marco teórico del presente trabajo, que abarca el proceso de simulación sobre el cual se estimará tanto la pérdida esperada como las diversas medidas de riesgo para el Seguro de Depósitos en el plazo de un año.

En un modelo de incumplimiento (*default-mode model*) la pérdida esperada de la cartera (EL_p) es igual al producto de las exposiciones al riesgo individuales para cada una de las n entidades (EXP_i), las probabilidades de incumplimiento (p_i) de cada entidad, es decir, la proporción de deudores que podrían caer en incumplimiento dentro de un marco de tiempo dado (generalmente de un año) y las pérdidas dado el incumplimiento (LGD_i). Tal que: (Matjaž, 2014)

$$EL_p = \sum_{i=1}^n EXP_i * p_i * LGD_i \quad (1.9)$$

- EXP : la exposición del sistema de Seguro de Depósitos para cada entidad se define como el monto de sus depósitos cubiertos.
- LGD : $1 - RR$, donde RR corresponde a la tasa de recuperación, (*Recovery Rate*), la cual se puede calcular con la base de datos de recuperaciones. Otra alternativa para el cálculo de dicho porcentaje es utilizar el método establecido por el Comité de Supervisión de Basilea (2006) en el contexto de los requisitos de capital. La Comisión

propone un LGD^8 de 75 % de todos los préstamos, dicha aproximación tiene sentido dado que la mayor exposición del sistema de Seguro de Depósitos se da en el sector bancario.

- p : estas probabilidades pueden ser estimadas como se menciona en la sección 1.3, a través de modelos econométricos *logit*, en donde es necesario definir una variable respuesta categórica que represente el evento de incurrir en incumplimiento.

1.7 Generación de la función de distribución empírica

Una vez que se ha estimado las probabilidades de incumplimiento se debe construir el proceso iterativo a partir de la ecuación 1.10, correspondiente a la versión asintótica del modelo de Vasicek, la cual se expresa a continuación:

$$X_{n,t} = \sqrt{\rho_t} * Y_t + \sqrt{1 - \rho_t} * \epsilon_{n,t} \quad (1.10)$$

Donde el crédito i se asume en incumplimiento cuando la variable latente X_i toma un valor menor al límite de crédito especificado. Para efectos de la simulación se generan m iteraciones, con m valores aleatorios para la variable *Threshold*, que se define como un número aleatorio que representa el límite para determinar si la entidad cae o no en incumplimiento. (Hernández, 2013)

$$X_i < Threshold \quad (1.11)$$

El principal rol de una entidad financiera es brindar el servicio de intermediación. Cuando los activos de la entidad no permiten hacer frente a los pasivos exigibles de la misma, se le dificulta ofrecer el servicio referido y podría incurrir en incumplimiento. De esta manera, se generan suficientes eventos que simulen esta situación y así permitan aproximar la función de densidad y de distribución de las pérdidas.

⁸Según el Comité de Supervisión de Basilea, cuando no existan medidas para el cálculo del LGD se sugiere considerar 45 % en créditos preferentes y 75 % en créditos subordinados. En este caso se decidió tomar 75 %, con el objetivo de realizar el análisis bajo el escenario más conservador, es decir, aquel que asume el mayor porcentaje de pérdidas.

A nivel macroeconómico se conoce que hay un aspecto que afecta a todas las entidades financieras independientemente de su tamaño, aunque en diferente magnitud, que es el estado de la economía Y_t . Para reflejar este hecho, se define la variable Y_t como un número aleatorio proveniente de una distribución normal estándar (media 0 y desviación típica 1), tal que corresponda al estado actual de la economía. Un número alto implicaría una buena situación económica, un número bajo implicaría un mal estado económico.

Con el fin de reflejar la idiosincrasia de una entidad en particular, se define $\epsilon_{n,t}$ también como un número aleatorio específico para la entidad i , proveniente de una función de densidad normal estándar.

Para ponderar el efecto macroeconómico se multiplica Y_t por las correlaciones de la cartera (es decir, la fuerza de la relación entre la cartera de créditos) y para ponderar la “singularidad” de la entidad se multiplica por $1 - \rho$ (es decir, lo contrario de su relación con otras entidades financieras).

Para cada iteración se establece un nuevo Y_t estado económico, que todas las entidades compartirán para esa iteración. Un nuevo $\epsilon_{n,t}$ se genera para cada entidad dentro de cada iteración. Una vez que se ha calculado suficientes simulaciones los escenarios de pérdida que se obtienen finalmente convergen en la pérdida total esperada de la cartera, lo cual da lugar a la generación de la función de distribución empírica y las probabilidades asociadas para toda la gama de pérdidas. (Charpentier, 2015)

1.8 Medidas de riesgo: valor en riesgo y valor en riesgo condicional

1.8.1 Valor en riesgo (VaR)

En su forma más general, el VaR mide la pérdida potencial en el valor de un activo de riesgo o de un portafolio durante un período definido para un intervalo de confianza dado. (Damodaran, 2009)

Si bien el VaR puede ser utilizado por cualquier entidad para medir su exposición al riesgo, en general es utilizado con mayor frecuencia en el sector financiero para capturar la pérdida potencial en el valor de los portafolios negociados en el mercado durante un período especificado. (Castagna, 2009)

El VaR está definido por la siguiente fórmula:

$$VaR_\alpha = \inf\{x \mid F_L(x) \geq \alpha\} = F_L^{-1}(\alpha) \quad (1.12)$$

Esta medida de riesgo corresponde a un cuantil en la función de distribución empírica (F_L), dejando un área de cola de $1 - \alpha$ a su derecha. Para el presente trabajo este valor representa la pérdida que asumiría el Seguro de Depósitos como máximo el $(1 - \alpha)\%$ de las veces.

1.8.2 Valor en riesgo condicional (TVaR)

Como se menciona en la sección 1.8.1, en gestión del riesgo financiero, el VaR es una medida utilizada para medir el riesgo de pérdida de una cartera específica de los activos financieros. Sin embargo, una desventaja del VaR es que no aporta ninguna información sobre la severidad de las pérdidas más allá del nivel para el cual se calcula. Incluso podría suceder que el VaR de dos variables podría ser exactamente el mismo y sin embargo, una de ellas podría tener una cola más densa (más riesgo) que la otra.

Para mitigar este efecto surge una medida de riesgo conocida como el TVaR, que toma en cuenta la severidad del incumplimiento y no sólo la probabilidad de éste. En general se considera que el TVaR proporciona una mejor medida de riesgo, ya que se ve afectado en los valores extremos de la distribución y se define de la siguiente manera:

$$TVaR_{\alpha}(L) = \frac{1}{1 - \alpha} \int_{\alpha}^1 VaR_{\lambda}(L) d\lambda = E(L | L > VaR_{\alpha}(L)) \quad (1.13)$$

El TVaR por tanto, mide la pérdida esperada promedio de una cartera en un horizonte de tiempo determinado, considerando únicamente los casos en los que las pérdidas son mayores que el VaR. En otras palabras, representa el valor promedio de las pérdidas esperadas en los peores casos. (Peng, 2008)

Como se observa en la presente sección, es posible estimar tanto las probabilidades de riesgo alto como la pérdida promedio y el monto que el Seguro de Depósitos asumiría como pérdida dado un nivel de confianza determinado.

1.9 Metodologías para la definición del porcentaje de recuperación

En los apartados anteriores se ha asumido que el valor de las pérdidas dado el incumplimiento (riesgo alto) es igual a 100%. Sin embargo, a lo largo del proceso de liquidación de las entidades se pueden dar eventuales recuperaciones de acuerdo al orden de prelación.

Al tener en cuenta los ingresos esperados de las recuperaciones se puede concluir que el asegurador de depósitos no necesariamente requiere financiar el 100% de los recursos con la recaudación de las contribuciones, puesto que una parte de esos recursos van a ser recuperados.

Considerando esto, los recursos recaudados a través de las contribuciones y que constituyen la reserva del Seguro de Depósitos (financiación ex-ante) deben ser suficientes para cubrir el porcentaje de recursos que no se espera recuperar con el proceso de liquidación de las entidades.

De forma complementaria, el porcentaje de recursos que se espera recuperar en los procesos de liquidación puede financiarse con alternativas diferentes, como líneas contingentes de crédito. Este esquema de financiamiento busca solucionar un problema de iliquidez temporal del asegurador de depósitos, asociado a un desajuste en términos de flujos de caja.

En principio, la pérdida dado el incumplimiento se puede calcular con datos históricos. Sin embargo, si el seguro de depósitos no cuenta con las bases de datos de las recuperaciones para el cálculo de la pérdida dado el incumplimiento, como se menciona en la sección 1.6, otra opción es considerar el porcentaje de las recuperaciones según los estándares planteados (25 %) por el Comité de Supervisión Bancaria de Basilea (2006), dado que en el caso de un Seguro de Depósitos su exposición es precisamente a entidades financieras.

Capítulo 2

Datos y resultados

En esta sección se definen las variables consideradas en el estudio, así como el conjunto de entidades financieras sobre el cual se analizan e interpretan los resultados obtenidos, tanto en la estimación de los modelos econométricos como en la simulación de las pérdidas para la estimación de la función de distribución empírica.

En el presente trabajo se utilizó información referente a bancos privados, sociedades financieras y mutualistas¹, con periodicidad mensual desde enero de 2007 hasta junio de 2016, con un total de 114 observaciones por cada entidad.

El conjunto de entidades considerado en el análisis incluye aquellas que han obtenido la mayor cantidad de veces una calificación de riesgo *alto* para el período de estudio, según la metodología *CAMEL*, utilizada por la Corporación del Seguro de Depósitos, Fondo de Liquidez y Fondo de Seguros Privados.

Dentro de los bancos se incluyó tres entidades definidas como X_1 , X_2 y X_3 , tres mutualistas definidas como X_4 , X_5 y X_6 , y dos sociedades financieras definidas como X_7 y X_8 .

¹Se ha considerado a las Mutualistas de Ahorro y Crédito para la Vivienda dentro del Sector Financiero Privado, puesto que según lo dispuesto en la Disposición General Vigésimo Novena del Código Orgánico Monetario y Financiero y la Resolución No. 219-2016-F expedida por la Junta de Política y Regulación Monetaria y Financiera, estas entidades deberán resolver su permanencia en el Sector Financiero Popular y Solidario en un plazo de 18 meses contados a partir del 9 de marzo de 2016. Además, estas entidades seguirán siendo reguladas por la Superintendencia de Bancos, hasta que se cumpla dicho plazo.

Para la primera parte de la investigación, referente a la estimación de las probabilidades de riesgo alto, se desarrolló un modelo de regresión logística de series de tiempo para cada entidad. De esta manera, se obtuvo la serie histórica mensual de las probabilidades de riesgo alto de cada entidad financiera.

A partir de ello se construyó la matriz de correlaciones de las probabilidades de pasar a una calificación de riesgo alto (en adelante, probabilidades de riesgo alto), ya que es parte de la información que se necesitó para la aproximación de la función de distribución de pérdidas.

2.1 Definición de variables

2.1.1 Variables independientes

El conjunto de variables independientes se puede dividir en dos grupos: 1) ratios financieros e 2) información macroeconómica del país.

2.1.1.1 Variables financieras

Las variables financieras seleccionadas en el modelo responden a los criterios de: solvencia, liquidez, calidad de activos y rentabilidad. Estas son:

Criterio	Variable	Fórmula
Calidad de activos	Morosidad de cartera (MC)	$\text{Cartera improductiva} / \text{Cartera bruta}$
	Improductividad de activos (IA)	$\text{Activos improductivos} / \text{Activo total}$
	Cobertura de cartera (CC)	$\text{Provisiones de cartera} / \text{Cartera improductiva}$
Liquidez	Liquidez corriente (LC)	$\text{Activo corriente} / \text{Pasivo corriente}$
Solvencia	Apalancamiento del activo (AA)	$\text{Pasivo total} / \text{Activo total}$
	Patrimonio técnico constituido sobre Activos y contingentes ponderados por riesgo (PTAC)	$\text{Patrimonio técnico constituido} / \text{Activos y contingentes ponderados por riesgo}$
Rentabilidad	Rentabilidad (R)	$\text{Utilidad neta} / \text{Total ingresos}$
	Rentabilidad del capital (RC)	$\text{Utilidad neta} / \text{Patrimonio promedio}$
	Margen operativo (MO)	$\text{Margen operacional} / \text{Total ingresos}$

Cuadro 2.1: Variables explicativas financieras

En este estudio se decidió trabajar con las variables en términos de ratio en lugar de porcentajes y en este sentido es importante señalar que la determinación de la escala a utilizar para las variables explicativas es arbitraria, ya que las conclusiones del modelo no se ven alteradas. En muchas ocasiones la escala se suele cambiar por razones didácticas de ilustración de los resultados obtenidos, o cuando el modelo tiene problemas para converger. Sin embargo, en general este aspecto no resulta ser un inconveniente.

Es importante señalar que si bien el grupo de variables seleccionado para el análisis comparte ciertos criterios con los utilizados en la metodología *CAMEL*, los indicadores financieros considerados en el estudio no son expresamente los empleados para el cálculo del nivel de riesgo.

Por otro lado, a diferencia de la regresión lineal, como se mencionó antes, la cual usa el método de estimación por Mínimos Cuadrados Ordinarios (MCO), en la regresión logística se emplea el método de Máxima Verosimilitud (MV) para llevar a cabo la estimación de los parámetros del modelo.

Estos modelos de regresión con variable endógena categórica, también llamados modelos de variable endógena cualitativa o modelos de elección discreta, se suelen emplear para explicar la decisión (Y) que toma un individuo de entre un número limitado de posibles opciones a partir de un conjunto de variables explicativas (X_1, X_2, \dots, X_k). Con lo cual a diferencia de la regresión lineal, estos modelos econométricos pretenden estimar la probabilidad de que ocurra un evento en particular.

2.1.1.2 Variables macroeconómicas

La inclusión de variables macroeconómicas en el análisis de riesgo ha tomado gran importancia en los últimos años. Si bien, en un año específico los factores macroeconómicos no cambian el rango de riesgo de un solo préstamo dentro de una cartera; en un marco más amplio se ha evidenciado que la omisión de datos macroeconómicos generalmente significa deshacerse de información valiosa, lo que se da especialmente cuando el modelo de calificación se emplea en la estimación de probabilidades de que una entidad pase a tener riesgo alto.

En el presente trabajo se incluyó el entorno macroeconómico con el fin de complementar el análisis financiero y evitar errores de especificación del modelo. En este sentido, luego del análisis de seis variables se seleccionó tres variables macroeconómicas que son:

- Media móvil de la variación mensual del Índice de Actividad Económica Coyuntural (IDEAC).²
- Tasa de interés activa mensual (TI).³
- Variación mensual del precio del petróleo (PP).⁴

Es importante señalar que dentro del análisis se exploró la inclusión de estas variables retrasadas un trimestre ($t - 3$), un semestre ($t - 6$) y un año ($t - 12$) considerando su significancia estadística.

2.1.2 Variable dependiente

La variable dependiente considerada en el presente estudio es el nivel de riesgo otorgado a las entidades financieras mediante la metodología *CAMEL*. Esta variable es binaria en el caso de que la serie correspondiente a los niveles de riesgo tenga únicamente dos categorías, si esta variable tiene más de dos categorías para el período de análisis se considera multinomial.

Esta metodología otorga una puntuación que va de 1 hasta 6 de acuerdo a los criterios de: solvencia, liquidez, gestión, resultados y calidad de activos. De esta forma el nivel de riesgo se asigna según el siguiente cuadro:

²En virtud de que el PIB real del Ecuador se calcula de manera trimestral, se prefiere evitar problemas de autocorrelación a través de la imputación de datos faltantes con la técnica del Filtro del Kalman.

³Banco Central del Ecuador, Publicaciones estadísticas, Sector monetario y financiero, Julio de 2016, www.bce.fin.ec/index.php/component/k2/item/754.

⁴Banco Central del Ecuador, Indicadores económicos, Julio de 2016, www.contenido.bce.fin.ec/resumen_ticker.php?ticker_value=petroleo.

Categoría	Nivel de riesgo	Calificación
1	Bajo	$0 > y \leq 2$
2	Medio bajo	$2 > y \leq 3$
3	Medio	$3 > y \leq 4$
4	Medio alto	$4 > y \leq 5$
5	Alto	$5 > y \leq 6$

Cuadro 2.2: Niveles de riesgo

Previo a la estimación de los modelos probabilísticos, a continuación se presenta la exploración de los datos para la entidad X_1 , ya que sobre ésta se describirá toda la metodología del presente estudio. Sin embargo, es necesario señalar que el análisis se realizó para todas las entidades mencionadas al inicio de este capítulo (Ver Apéndices).

2.2 Análisis exploratorio de los datos

Como se indicó anteriormente, para cada entidad se realizó el análisis exploratorio así como la estimación del modelo econométrico. Sin embargo, a manera de ejemplo se describe todo el análisis con datos de una entidad financiera. A continuación se presenta un resumen estadístico de las variables:

	Calificación de riesgo	MC ¹	CC ²	IA ³	LC ⁴	MO ⁵
Mínimo	1.72	0.04	0.30	0.11	0.13	-0.15
1er cuartil	2.97	0.06	0.48	0.18	0.18	0.00
Mediana	3.56	0.07	0.70	0.21	0.23	0.06
3er cuartil	3.46	0.08	0.65	0.21	0.26	0.06
Media	4.13	0.10	0.81	0.23	0.31	0.12
Máximo	5.21	0.14	0.97	0.35	0.54	0.21

Cuadro 2.3: Resumen de estadísticas - parte 1

Nota: ¹ Morosidad de cartera ² Cobertura de cartera

³ Improductividad de activos ⁴ Liquidez corriente

⁵ Margen operativo

Para esta entidad se observa que las calificaciones de riesgo fluctúan entre 1.72 y 5.21, con una media de 4.13 que implica un nivel de riesgo definido como *medio alto*. Por otro lado, las medias de las variables *morosidad de cartera*, *cobertura de cartera*, *improductividad de activos*, *liquidez corriente* y *margen operativo* son 0.10, 0.81, 0.23, 0.31 y 0.12 respectivamente.

Considerando que la media y la mediana de estas variables no presentan diferencias importantes, se asume que la presencia de valores extremos es baja, ya que esto indica que tanto el percentil 50 de la distribución de los datos como el valor promedio de estos, es relativamente cercano. Dicha afirmación se cumple en todos los casos, a excepción de *cobertura de cartera* y *margen operativo* en donde la media es dos veces el valor de la mediana.

	AA ¹	R ²	RC ³	PTAC ⁴	TI ⁵	PP ⁶	IDEAC ⁷
Mínimo	0.79	-0.01	-0.01	0.13	0.07	-0.28	-0.01
1er cuartil	0.83	0.02	0.01	0.15	0.08	-0.05	0.00
Mediana	0.86	0.04	0.02	0.18	0.09	0.01	0.01
3er cuartil	0.85	0.06	0.03	0.18	0.09	0.01	0.01
Media	0.88	0.10	0.04	0.19	0.09	0.06	0.01
Máximo	0.90	0.19	0.13	0.30	0.11	0.24	0.02

Cuadro 2.4: Resumen de estadísticas - parte 2

Nota: ¹ Apalancamiento activo ² Rentabilidad ³ Rentabilidad del capital ⁴ Patrimonio técnico sobre activos y contingentes ponderados por riesgo ⁵ Tasa de interés ⁶ Variación del precio del petróleo ⁷ Variación del Índice de Actividad Económica Conyuntural

Respecto al cuadro 2.4 se observa mayores diferencias entre la media y mediana, especialmente para las variables *variación del precio del petróleo*, y *rentabilidad*.

Para la variable *apalancamiento del activo* se aprecia muy poca diferencia, incluso entre la media y el valor máximo de la distribución de los datos. Mientras que, por ejemplo, la *variación del precio del petróleo* sí muestra gran diferencia entre el valor promedio (0.06) y el valor máximo (0.24), lo

que sugiere que el precio del petróleo ha tenido grandes fluctuaciones durante el período de análisis.

A continuación se presenta la matriz de correlaciones de las variables independientes para la entidad X_1 .

Variables	MC	IA	CC	LC	MO	PTAC	R	RC	PP	IDEAC	TI	AA
MC	1.00	0.66	-0.87	0.16	0.06	0.33	0.26	0.27	-0.51	-0.64	0.62	-0.50
IA		1.00	-0.33	-0.29	-0.46	-0.23	-0.32	-0.18	-0.49	-0.77	0.05	0.15
CC			1.00	-0.36	-0.22	-0.57	-0.46	-0.42	0.39	0.41	-0.75	0.72
LC				1.00	0.39	0.85	0.31	0.25	-0.03	-0.16	0.36	-0.65
MO					1.00	0.57	0.88	0.72	-0.06	0.26	0.30	-0.61
PTAC						1.00	0.61	0.47	-0.21	-0.16	0.49	-0.92
R							1.00	0.75	-0.16	0.22	0.49	-0.73
RC								1.00	0.08	0.16	0.45	-0.56
PP									1.00	0.54	-0.03	0.24
IDEAC										1.00	-0.07	0.13
TI											1.00	-0.70
AA												1.00

Cuadro 2.5: Matriz de correlaciones de las variables independientes

El cuadro anterior permite identificar la existencia de altas correlaciones entre las variables predictoras, valores de correlación altos podrían indicar problemas de multicolinealidad, los cuales se dan cuando dos o más variables independientes en el modelo se determinan aproximadamente por una combinación lineal de otras variables independientes del mismo modelo, lo que podría conducir a estimaciones poco fiables.

En este caso se observa que la mayoría de las correlaciones entre las variables independientes son relativamente bajas, a excepción de las correlaciones entre: *rentabilidad*, y *margen operativo y patrimonio técnico constituido sobre activos y contingentes ponderados por riesgo y apalancamiento del activo*.

Frente a esto, existen varias alternativas para controlar este problema en la estimación del modelo, como pueden ser: eliminar una de las variables que está altamente correlacionada, o usar métodos de estimación que se basen en la selección de variables independientes, con el fin de considerar aquellas con las que se pueda construir un modelo que se ajuste al conjunto de datos.

Para efectos del presente trabajo se empleó el método *stepwise* para la selección de variables en la estimación del modelo econométrico *logit*. A partir de esta selección se probó que de las variables que presentaron

una correlación alta, la metodología admite utilizar a una de ellas para la estimación del modelo, con lo cual se descartan problemas de multicolinealidad; aspecto que también se comprueba estadísticamente a través de los estadísticos pseudo- R^2 .

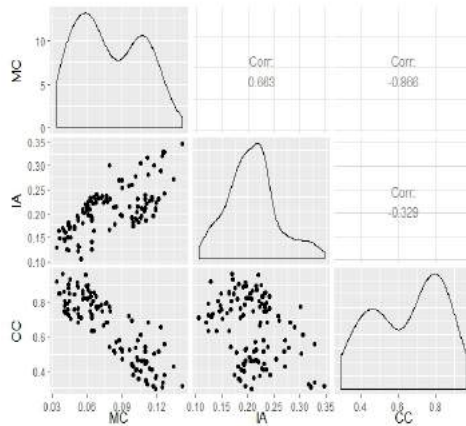


Figura 2.1: Morosidad, improductividad y cobertura de cartera

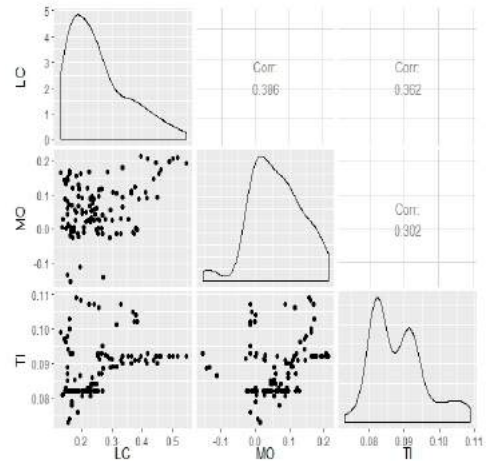


Figura 2.2: Liquidez, margen operativo y tasa de interés

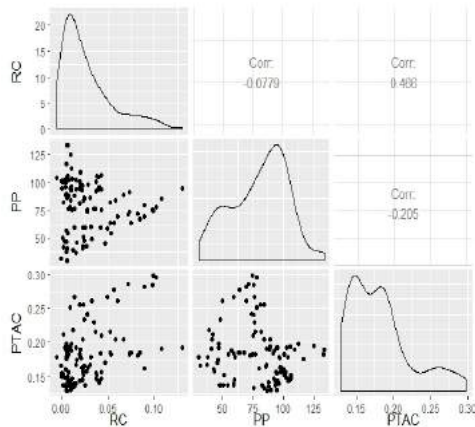


Figura 2.3: Rentabilidad del capital, variación del precio del petróleo, y patrimonio técnico constituido / activos y contingentes ponderados por riesgo

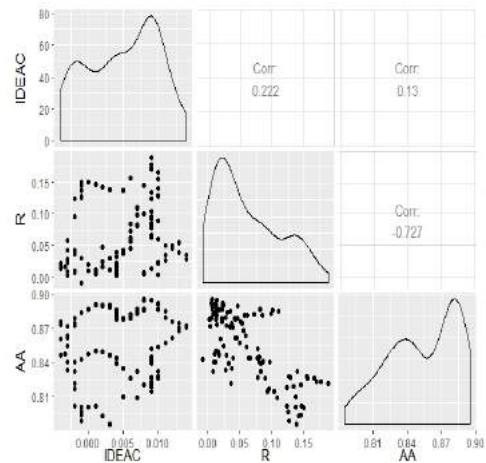


Figura 2.4: Índice de Actividad Económica Coyuntural, rentabilidad y apalancamiento del activo

2.3. ESTIMACIÓN DE LAS PROBABILIDADES DE RIESGO ALTO 43

En las figuras anteriores se visualiza gráficamente la relación entre las variables predictoras.⁵ Por ejemplo, en el cuadro 2.5 la correlación entre *morosidad de cartera* y *cobertura de cartera* es negativa y bastante alta (-0.87). Dicha relación se confirma gráficamente en la figura 2.1, en donde los puntos forman una línea recta (mientras más claramente se aprecie la línea recta, más alta es la correlación lineal entre las variables) y se observa que a niveles bajos de *morosidad de cartera*, la *cobertura de cartera* es alta y viceversa, esto indica que la relación entre ambas es negativa.

Es importante mencionar que, como se indicó anteriormente, el uso de la metodología *stepwise* permite seleccionar las variables que logran un mayor ajuste del modelo al conjunto de datos, lo que se ilustra en la siguiente sección.

2.3 Estimación de las probabilidades de riesgo alto

Para la estimación del modelo se empleó la metodología *stepwise* partiendo del siguiente modelo teórico:

$$\begin{aligned} \log\left(\frac{p}{1-p}\right) = & \beta_0 + \beta_1 MC_t + \beta_2 IA_t + \beta_3 CC_t + \beta_4 MO_t + \beta_5 PTAC_t \\ & + \beta_6 LC_t + \beta_7 AA_t + \beta_8 R_t + \beta_9 RC_t + \beta_{10} IDEAC_t \\ & + \beta_{11} TI_t + \beta_{12} PP_t + \mu_t \end{aligned} \quad (2.1)$$

Para determinar la significancia de las variables se realiza un contraste de hipótesis, en donde, la hipótesis nula (H_0) señala que los coeficientes (β 's) de la regresión son cero, es decir, no aportan explicación al modelo y por tanto no son significativos. Mientras la hipótesis alternativa (H_1) es la negación de la hipótesis nula ($H_1: \neg H_0$).

El criterio utilizado para determinar el resultado de la prueba, es la medida de significancia estadística conocida como *valor p* (valor plausible), este valor es la probabilidad que mide la evidencia en contra de la hipótesis nula. Se rechaza H_0 siempre que el valor p sea inferior al nivel de significancia, valores más bajos representan evidencia fuerte para rechazar la

⁵No se han representado todas las relaciones entre las variables independientes por efectos de visualización. El objetivo es interpretar el concepto de correlación desde la perspectiva de un gráfico.

hipótesis nula en favor de la alternativa.

En el caso de la entidad X_1 la variable dependiente es multinomial, ya que para el período de análisis el nivel de riesgo obtenido ha fluctuado entre 1 y 5 según la metodología *CAMEL*. Con lo cual se estimó un modelo *logit multinomial ordenado*.

<i>Variable dependiente:</i>	
Nivel de riesgo	
IA	54.936*** (9.505) [†]
PTAC	-96.596*** (17.286)
CC	-9.971*** (2.566)
R	-33.325*** (8.813)
Observaciones	114

Nota: *p<0.1; **p<0.05; ***p<0.01

[†] Estos valores corresponden al error estándar del coeficiente

Cuadro 2.6: X_1 : modelo *logit* con variable dependiente multinomial

Según el cuadro 2.6, el modelo evidencia la relación entre las variables independientes respecto a la variable respuesta. A continuación se muestra los estadísticos de las variables explicativas con su respectivo valor plausible.

2.3. ESTIMACIÓN DE LAS PROBABILIDADES DE RIESGO ALTO⁴⁵

	Value	Std. Error	t value	pval
Improductividad de activos	54.94	9.51	5.78	0.00
Patrimonio técnico constituido sobre activos y contingentes ponderados por riesgo	-96.60	17.29	-5.59	0.00
Cobertura de cartera	-9.97	2.57	-3.89	0.00
Rentabilidad	-33.32	8.82	-3.78	0.00
1 2	-22.93	5.50	-4.17	0.00
2 3	-17.53	4.65	-3.77	0.00
3 4	-10.77	4.29	-2.51	0.01
4 5 ⁶	-4.81	4.25	-1.13	0.26

Cuadro 2.7: Estadísticos de las variables independientes

Para este caso todas las variables muestran una relación negativa y estadísticamente significativa con la variable respuesta al 95 % de confianza, a excepción de la variable *improductividad de activos* que tiene relación positiva. Esto quiere decir que aumentos de estas variables están relacionados con una disminución de la probabilidad de tener riesgo alto.

El modelo evidencia que una disminución en estas variables (*cobertura de cartera*, *rentabilidad*, y *patrimonio técnico constituido sobre activos y contingentes ponderados por riesgo*) está asociada con un incremento de la probabilidad de riesgo alto.

La variable *improductividad de activos* muestra una relación positiva y estadísticamente significativa respecto a la probabilidad de tener riesgo alto, con lo cual incrementos en esta variable están asociados con aumentos de dicha probabilidad.

Respecto a los coeficientes de la regresión, se puede interpretar que un incremento de una unidad en *cobertura de cartera*, reduce el *log odds*⁷ ($\log \frac{p}{1-p}$) en 9.97 unidades.

⁶Según el valor plausible, pasar del nivel de riesgo 4 (medio alto) al 5 (alto) no resulta estadísticamente significativo. Sin embargo, es necesario tener en cuenta que el conjunto de datos presenta pocas observaciones en donde el nivel de riesgo sea la categoría 5, en comparación con las demás categorías de riesgo.

⁷El *log odds* es una transformación monótona del *odds ratio*, que a su vez es el ratio de la probabilidad de éxito sobre la probabilidad de fracaso. Cuanto mayores sean las probabilidades, mayores serán los *odds ratio* y *log odds*.

	1	2	3	4	5
1	0.14	0.01	0.00	0.00	0.00
2	0.03	0.05	0.06	0.00	0.00
3	0.00	0.05	0.33	0.03	0.00
4	0.00	0.00	0.03	0.22	0.01
5	0.00	0.00	0.00	0.01	0.04

Cuadro 2.8: Validación cruzada: nivel de riesgo CAMEL vs. nivel de riesgo predicho por el modelo

La validación cruzada es un método de evaluación del modelo que no utiliza medidas estadísticas convencionales (por ejemplo: R^2). Este método se centra en la capacidad de predicción del modelo. Los elementos de la diagonal del cuadro 2.8 representan el porcentaje de aciertos que tuvo el modelo para cada nivel de riesgo, mientras que los demás porcentajes del cuadro corresponden a los errores de clasificación.

En este caso, de un total de 114 observaciones en 88 de ellas (78%) el nivel de riesgo que arrojó el modelo es el mismo que el nivel de riesgo otorgado por la metodología *CAMEL*. Mientras que en 26 observaciones se obtuvo errores de predicción. Con lo cual, el error de clasificación (suma de los elementos fuera de la diagonal sobre el total) es de 22%.

Indicador	Valor
Log-likelihood fitted model	-56.33
Log-likelihood intercept-only	-160.74
G2 ⁸	208.83
McFadden	0.65
Maximum likelihood	0.84
Cragg y Uhler	0.89

Cuadro 2.9: Pseudo R^2

Respecto a los indicadores del cuadro 2.9 se observa que el indicador Mc Fadden es de 0.65, mientras que el de Cragg y Uhler⁹ es 0.89. Como se

⁸Menos dos veces la diferencia en el logaritmo de las verosimilitudes

⁹Para mayor comprensión se prefiere interpretar con estos dos criterios, ya que según su forma de cálculo, es más sencillo interpretar que el “ajuste perfecto” de los datos se

2.3. ESTIMACIÓN DE LAS PROBABILIDADES DE RIESGO ALTO⁴⁷

menciona en la sección 1.4, las medidas pseudo- R^2 pueden arrojar valores bastante distintos entre sí. Sin embargo, en conjunto estas medidas muestran un buen ajuste del modelo a los datos.

Predicciones del modelo

En la figura 2.5 se ven las fluctuaciones de las probabilidades de riesgo alto. Se observa que existe alta variabilidad en las predicciones. Por un lado se tiene valores muy bajos y cercanos a cero, mientras en otros períodos estas probabilidades superan el 15 %.

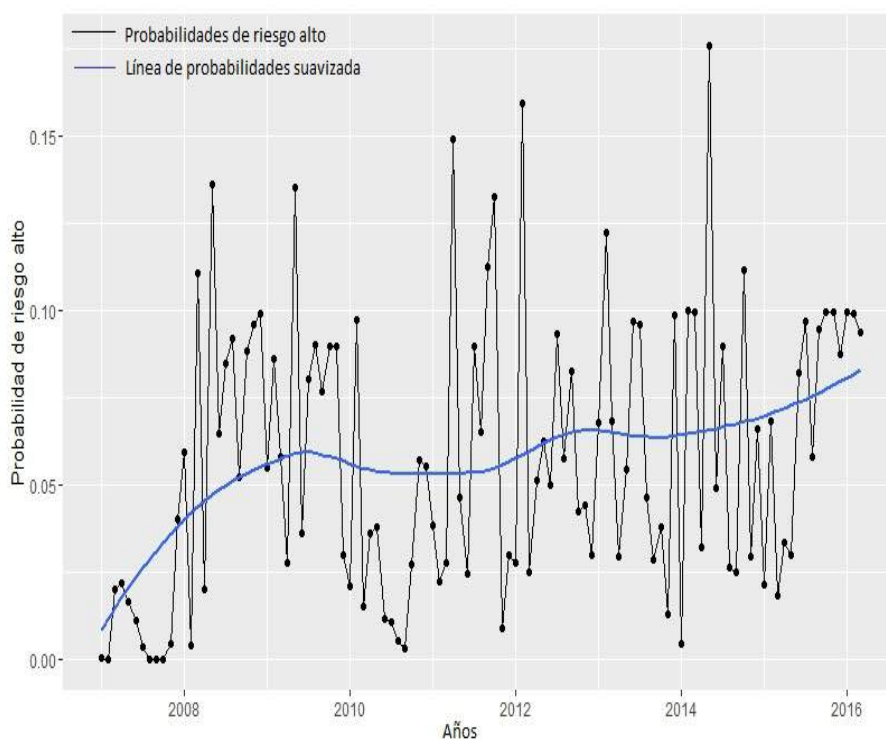


Figura 2.5: Probabilidades de riesgo alto

El cuadro 2.10 muestra un resumen de las estadísticas más relevantes de las predicciones.

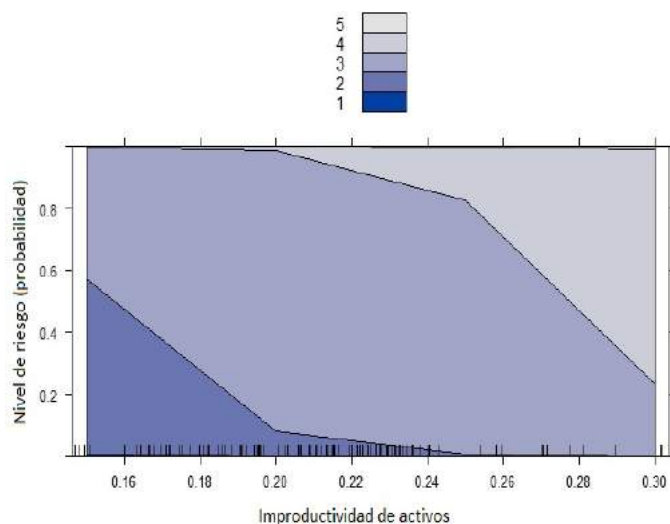
da mientras el indicador más se acerque a la unidad, que en el caso de Cragg y Uhler, este el máximo valor que puede tomar esta medida.

Riesgo alto		
1	Mínimo:	$4.24e^{-5}$
2	1er Cuantil:	0.026
3	Mediana:	0.052
4	Media:	0.058
5	3er Cuantil:	0.090
6	Máximo :	0.176

Cuadro 2.10: X_1 : resumen de las probabilidades de riesgo alto

Las probabilidades de riesgo alto para esta entidad durante el período de análisis tienen una media de 6 % y un máximo de 18 %.

A continuación se muestra los gráficos bivariantes entre las variables predictoras y el nivel de riesgo asociado.

Figura 2.6: *Effect plot*: improductividad de activos vs. nivel de riesgo

En la figura 2.6 se puede apreciar que a niveles más bajos de improductividad de activos, el nivel de riesgo asociado es bajo también. De este modo, la probabilidad de tener riesgo alto también es menor. Mientras que a niveles más altos de este ratio, el riesgo asociado es mayor, de modo que la probabilidad también aumenta.

2.3. ESTIMACIÓN DE LAS PROBABILIDADES DE RIESGO ALTO 49

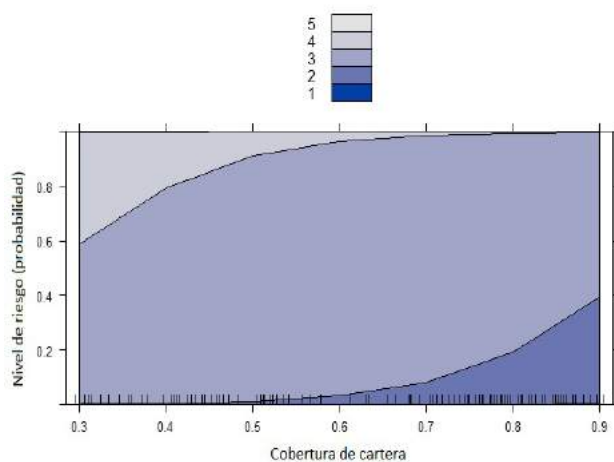


Figura 2.7: *Effect plot*: cobertura de cartera vs. nivel de riesgo

Para la variable *cobertura de cartera* se observa que a mayor cobertura, el riesgo asociado es bajo, con lo cual la probabilidad de tener un riesgo alto es menor. A niveles bajos de *cobertura de cartera*, los niveles de riesgo asociados son más altos, incluso de categoría 5.

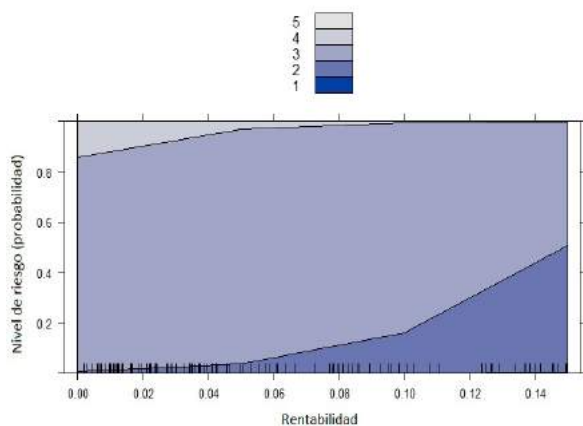


Figura 2.8: *Effect plot*: rentabilidad vs. nivel de riesgo

Para el ratio de *rentabilidad* se observa una forma similar a la figura anterior, a niveles más altos de *rentabilidad* el riesgo asociado es menor. Mientras que a medida que la *rentabilidad* disminuye, los niveles de riesgo relacionados son más altos.

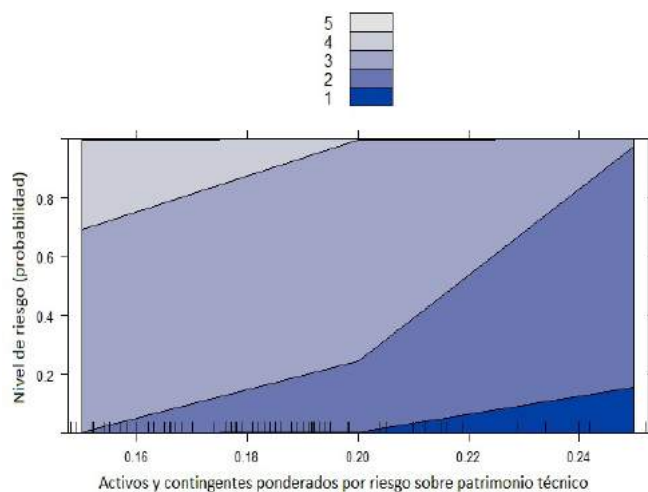


Figura 2.9: *Effect plot*: PTAC vs. nivel de riesgo

Finalmente, la figura 2.9 muestra que a niveles bajos, el nivel de riesgo asociado es alto, a medida que este indicador (PTAC) aumenta, el riesgo disminuye alcanzando niveles de 1.

2.3.1 Matrices de correlación de las entidades financieras

2.3.1.1 Matriz de correlaciones: Sector Financiero Privado

Entidades	X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8
X_1	1.000	0.548	0.298	-0.041	-0.206	0.569	-0.094	0.298
X_2		1.000	0.158	-0.387	-0.401	0.756	-0.315	0.417
X_3			1.000	0.518	0.402	0.308	0.378	0.062
X_4				1.000	0.726	-0.189	0.710	-0.167
X_5					1.000	-0.171	0.764	-0.310
X_6						1.000	-0.080	0.380
X_7							1.000	-0.257
X_8								1.000

Cuadro 2.11: Matriz de correlaciones: entidades con probabilidad de riesgo alto

En el cuadro 2.11 se observa la matriz de correlaciones para el conjunto de entidades con probabilidad de riesgo alto.

2.3. ESTIMACIÓN DE LAS PROBABILIDADES DE RIESGO ALTO 51

Analizando las correlaciones entre bancos, sociedades financieras y mutualistas, se observa que la correlación entre X_1 y X_7 es negativa y baja (-0.094). Mientras la correlación entre X_1 y X_8 es más alta, positiva e igual a 0.30. Por otro lado, la correlación entre X_1 y X_6 es 0.57, entre X_1 y X_5 es -0.21 y entre X_1 y X_4 es -0.04. Estos valores son muy distintos entre sí, lo que sugiere que la relación lineal entre bancos, mutualistas y sociedades financieras es muy diversa y depende de aspectos propios de cada entidad. Esta matriz se desagrega para cada subsistema en los siguientes apartados.

2.3.1.2 Matriz de correlaciones: bancos

Bancos	X_1	X_2	X_3
X_1	1.000	0.548	0.298
X_2		1.000	0.158
X_3			1.000

Cuadro 2.12: Matriz de correlaciones: bancos

Como se puede observar, las correlaciones con la entidad X_1 son positivas e iguales a 0.55 y 0.30 respecto a X_2 y X_3 respectivamente. Lo que significa que para la entidad X_1 , incrementos en la probabilidad de riesgo alto están relacionados con aumentos de esta probabilidad en las otras dos entidades. Por otro lado, la correlación entre X_2 y X_3 es positiva, pero baja e igual a 0.16.

2.3.1.3 Matriz de correlaciones: mutualistas

Mutualistas	X_4	X_5	X_6
X_4	1.000	0.726	-0.189
X_5		1.000	-0.171
X_6			1.000

Cuadro 2.13: Matriz de correlaciones: mutualistas

La correlación entre las entidades X_5 y X_6 es negativa e igual a -0.17. Así también, la correlación entre X_4 y X_6 es -0.19. En ambos casos esto indica que incrementos en la probabilidad de riesgo alto en una de ellas, están

asociados con disminuciones de esta probabilidad para la otra entidad. Sin embargo, esta correlación es bastante baja.

Finalmente, la correlación entre las entidades X_4 y X_5 es positiva y alta, con un valor igual a 0.73.

2.3.1.4 Matriz de correlaciones: sociedades financieras

Sociedades Financieras	X_7	X_8
X_7	1.000	-0.257
X_8		1.000

Cuadro 2.14: Matriz de correlaciones: sociedades financieras

En este caso, la correlación entre sociedades financieras es negativa e igual a -0.26, con lo cual disminuciones en la probabilidad de riesgo alto en una de ellas se relaciona con incrementos de esta probabilidad para la otra entidad.

2.4 Definición de los parámetros para la estimación de la función de pérdidas

La definición de los parámetros se realizó según lo mencionado en la sección 1.6.

- Probabilidades de riesgo alto: p_i , son las probabilidades estimadas por los modelos *logit*.
- Exposición al riesgo: EXP_i , corresponde al monto de depósitos cubiertos de las entidades incluidas en el estudio.
- Pérdida dado el incumplimiento: LGD , corresponde al porcentaje no recuperable de las pérdidas (75 %), de tal modo que el RR conocido como el porcentaje de recuperaciones es 25 %.
- Riesgo sistémico: Y_t definido como un número aleatorio para cada iteración, proveniente de una distribución normal estándar.

- Riesgo idiosincrático: $\epsilon_{n,t}$ definido como números aleatorios distintos para cada entidad dentro de cada iteración, provenientes de una distribución normal estándar.

2.5 Construcción del proceso de simulación

A partir de la definición de los parámetros se realiza el proceso de simulación para $N = 1,000$ con $M = 2,000$ iteraciones.

Para cada uno de los 1,000 eventos simulados se estima X_n , el cual depende del riesgo sistémico y del riesgo idiosincrático, como se indica en la sección 1.7.

Si dicho valor es menor al límite, definido como la función cuantil de la distribución normal (*threshold*), se calcula la pérdida esperada de la cartera según la ecuación 1.9. Además, se calcula la tasa de transición a calificación de riesgo alto (*default rate*), definida como el número de eventos que caen en incumplimiento (riesgo alto) sobre el total de eventos simulados.

Cálculo de la pérdida esperada total

Una vez realizada la simulación se obtiene la pérdida esperada de la cartera como la multiplicación de los siguientes tres factores:

$$\text{Pérdida esperada} = \sum_{i=1}^n EXP_i * p_i * LGD_i \quad (2.2)$$

$$\text{Pérdida esperada} = \text{US\$ } 928,748 \text{ miles}$$

Con lo cual, la pérdida esperada de la cartera es US\$ 928,748 miles y corresponde a 3.52 % del total de depósitos asegurados y 12.10 % del total de depósitos cubiertos.

Si se calcula la media de las simulaciones realizadas, el valor debería ser muy similar al de la ecuación 2.2, es así que de dicho cálculo se obtiene un valor igual a US\$ 923,864 miles.

2.6 Funciones probabilísticas de las pérdidas

2.6.1 Función de densidad de las pérdidas

Como se observa en la figura 2.10, a valores de pérdidas muy elevados la probabilidad de ocurrencia es prácticamente cero. Mientras que pérdidas menores están asociadas a probabilidades de ocurrencia más altas. Se observa la asimetría hacia la derecha de la distribución y la presencia de colas densas.

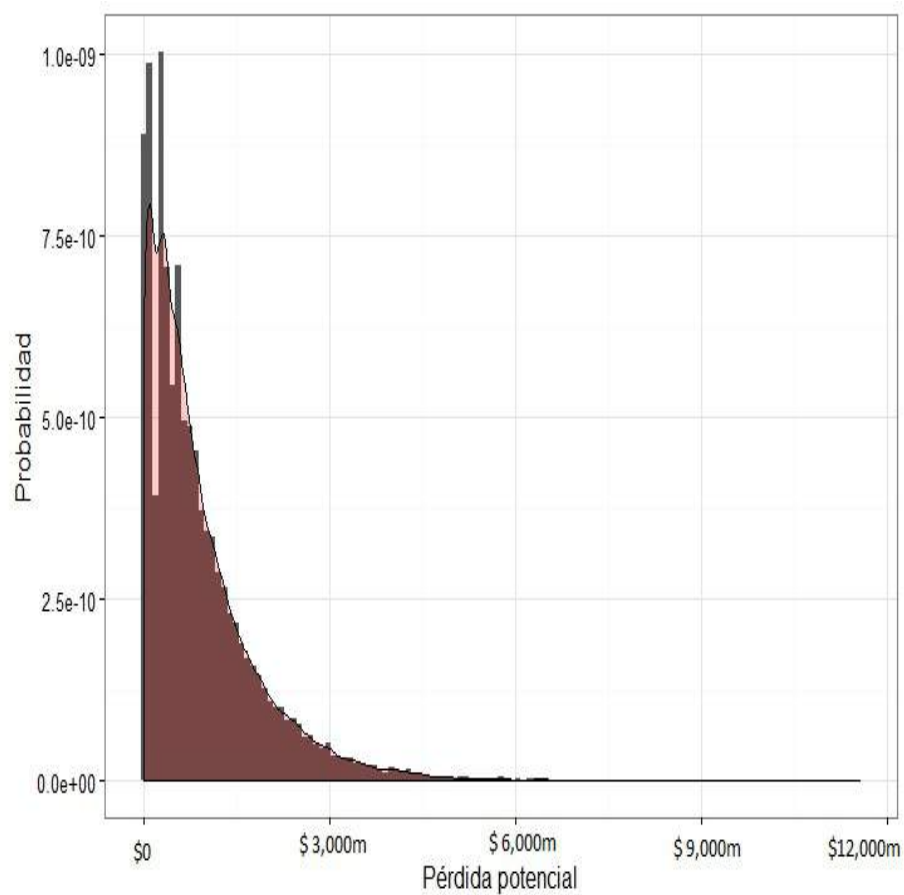


Figura 2.10: Función de densidad empírica de las pérdidas

2.6.2 Función de distribución empírica de las pérdidas

En la figura 2.11, se observa que pérdidas potenciales bajas están asociadas a valores que alcanzan una probabilidad acumulada cercana a 75%. A partir de este nivel, las pérdidas son significativamente mayores. Mientras menor pendiente tenga la curva de la distribución acumulada, mayor es la probabilidad de valores de pérdida bajos.

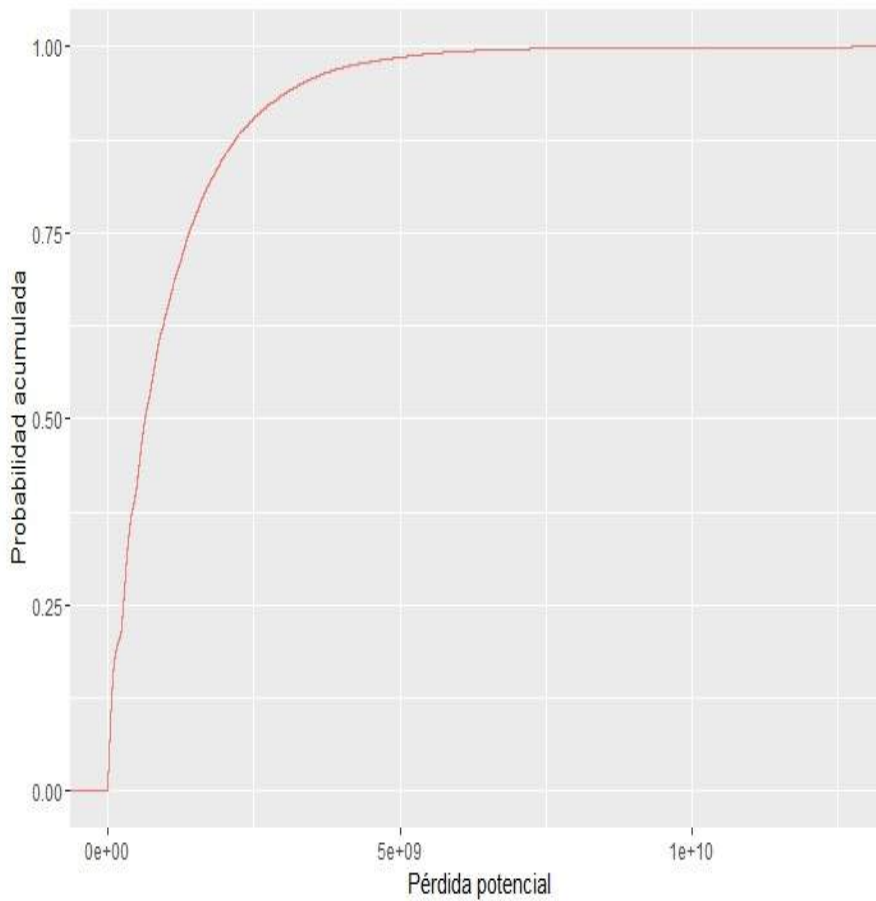


Figura 2.11: Función de distribución empírica de las pérdidas

2.7 Tasa de transición a calificación de riesgo alto (incumplimiento)

La tasa esperada de transición de que una entidad financiera pase a una calificación de riesgo alto (en adelante, tasa de riesgo alto) de las pérdidas simuladas es 2.3 %. Esta tasa es conocida en la literatura internacional como “*default rate*”, la cual se refiere a la tasa de incumplimiento, que como se señala en la sección 1.1, se entiende como la probabilidad de que una entidad financiera pase a una calificación de riesgo alto. En el siguiente cuadro se presentan los estadísticos descriptivos más relevantes:

Riesgo alto		
1	Mínimo:	0.000
2	1er cuartil:	0.009
3	Mediana:	0.020
4	Media:	0.028
5	3er cuartil:	0.037
6	Máximo :	0.314

Cuadro 2.15: Resumen de las tasas de riesgo alto

El percentil 75 de la distribución de tasas de riesgo alto es de 3.7 %, mientras el máximo valor es de 31.4 %, lo cual indica que al menos el 75 % de los datos corresponde a tasas de riesgo alto menores a 4 %.

2.7. TASA DE TRANSICIÓN A CALIFICACIÓN DE RIESGO ALTO (INCUMPLIMIENTO)

A continuación se muestra la función de densidad empírica:

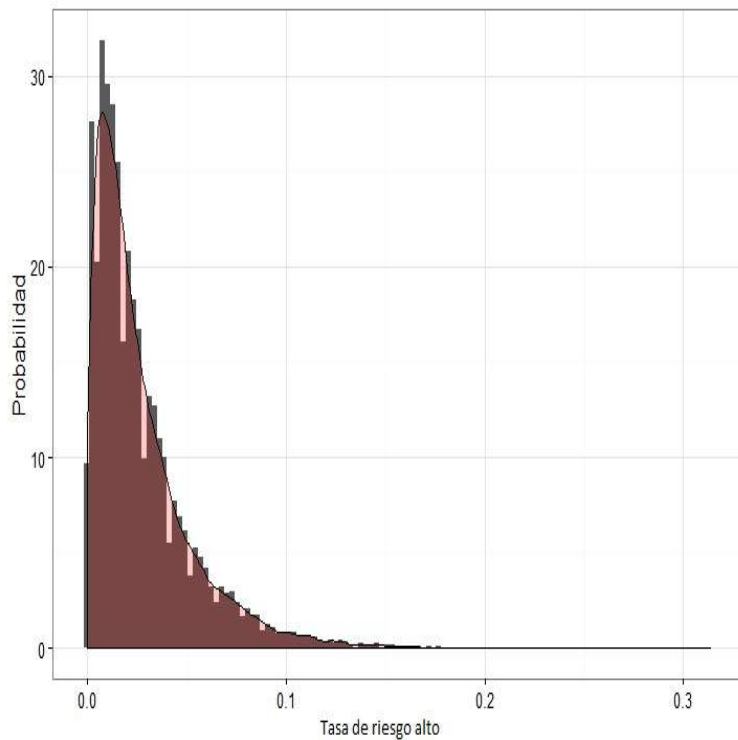


Figura 2.12: Función de densidad empírica de las tasas de riesgo alto

Se observa la misma asimetría hacia la derecha y el alargamiento de la cola de la función de densidad de las tasas de riesgo alto. Sin embargo, la probabilidad de que la tasa de riesgo alto sea cercana a cero llega hasta 31%. Mientras que para tasas de riesgo alto superiores a 10% la probabilidad de ocurrencia se reduce notablemente y se aproxima a cero.

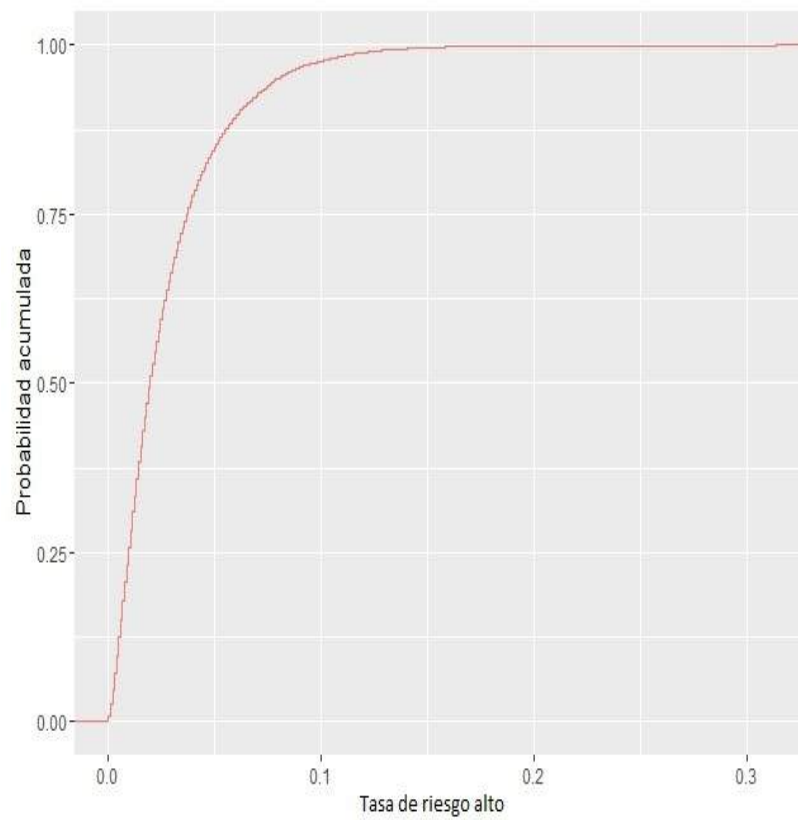


Figura 2.13: Función de distribución empírica de las tasas de riesgo alto

Finalmente, la figura 2.13 muestra que tasas de riesgo alto cercanas a cero alcanzan una probabilidad acumulada cercana a 75%. A partir de este punto, para niveles más altos de estas tasas la probabilidad acumulada se acerca asintóticamente a la unidad.

2.8 Determinación del valor en riesgo y valor en riesgo condicional

Nivel de confianza	Value at Risk			Tail Value at Risk		
	Valor ^a	%DA ^b	%DC ^c	Valor ^a	%DA ^b	%DC ^c
90 %	2,150,474	8.14	28.01	3,121,742	11.82	40.66
95 %	2,799,206	10.60	36.46	3,804,103	14.41	49.54
97.5 %	3,489,618	13.22	45.45	4,506,531	17.07	58.69
99 %	4,401,552	16.67	57.32	5,438,366	20.60	70.83

^aLos valores están en miles de US\$

^bComo porcentaje del total de depósitos asegurados del Sector Financiero Privado

^cComo porcentaje del total de depósitos cubiertos del Sector Financiero Privado

Cuadro 2.16: Medidas de riesgo VaR y TVaR

De acuerdo a las medidas de riesgo calculadas en el cuadro 2.16, se observa que a mayor nivel de confianza la pérdida estimada es mayor. Asimismo, el TVaR es superior al VaR para los diferentes niveles de confianza calculados, lo cual tiene sentido dado que el TVaR es un promedio de las pérdidas superiores al nivel del VaR.

Al 90 % de confianza el VaR representa cerca de 8 % del total de depósitos asegurados y 28 % del total de depósitos cubiertos. Mientras que en un escenario mucho más conservador, con 99 % de confianza el VaR representaría 21 % de los depósitos asegurados y 57 % de los depósitos cubiertos.

Por otro lado, si se considera como medida de riesgo el TVaR (promedio de las pérdidas por encima del nivel del VaR), con 90 % de confianza este valor corresponde a 12 % de los depósitos asegurados y 41 % de los depósitos cubiertos. Mientras que a un nivel de confianza de 95 % el TVaR representa aproximadamente 14 % de depósitos asegurados y 50 % de depósitos cubiertos.

2.9 Determinación del porcentaje de recuperación de cartera

Como se mencionó en la sección 1.9, a lo largo del proceso de liquidación de las entidades se pueden dar eventuales recuperaciones, este porcentaje de recursos que se espera recuperar puede financiarse con alternativas diferentes, por ejemplo, líneas contingentes de crédito. Además, se puede calcular con base en datos históricos o utilizando el porcentaje de recuperaciones (25 %) según los estándares planteados por el Comité de Supervisión Bancaria de Basilea (2006).

Asumiendo un porcentaje de recuperación igual a 25 %, el financiamiento por medio de otras fuentes debería ser 25 % del nivel objetivo del financiamiento total. Para el caso del Seguro de Depósitos del Sector Financiero Privado, por ejemplo, al 90 % de confianza, la pérdida esperada representa 8.14 % de los depósitos asegurados, con lo cual el porcentaje que se espera recuperar es el 25 % de dicho valor, que equivale a 2.04 %, y respecto a los depósitos cubiertos representa 7 %. Mientras que al 95 % de confianza este porcentaje corresponde a 2.65 % de los depósitos asegurados y 9.11 % de los depósitos cubiertos.

2.10 Provisión de pérdidas para entidades con menor nivel de riesgo

Una vez que se ha analizado las entidades que representan un nivel de riesgo alto para la cartera del Seguro de Depósitos se complementa estos resultados con la inclusión de los porcentajes de provisión para las demás entidades financieras, con el fin de mantener un criterio más conservador que abarque al conjunto de entidades que conforman el Sector Financiero Privado.

Para definir el nivel de provisión para las entidades financieras que no se encuentran con riesgo alto se utilizó el nivel de provisiones vigente para los créditos comerciales de las entidades del sector financiero privado. Los porcentajes de provisiones referidos se encuentran detallados en la Resolución No. 209-2016-F, denominada “Norma para la calificación de activos de riesgo y constitución de provisiones por parte de las entidades de los sectores financieros público y privado bajo el control de la Superintendencia de Bancos”, expedida por la Junta de Política y Regulación Monetaria y

2.10. PROVISIÓN DE PÉRDIDAS PARA ENTIDADES CON MENOR NIVEL DE RIESGO

61

Financiera el 12 de febrero de 2016.

A continuación el cuadro 2.17 resume las categorías de riesgo con sus respectivos rangos de pérdida y su relación con los niveles de riesgo otorgados por la metodología *CAMEL*.

Categorías	Calificación	Rango de pérdida esperada	CAMEL	Nivel de riesgo
Con riesgo normal	A-1	1 %	$1 > y \leq 1.99$	Bajo
	A-2	2 %		
	A-3	3 %-5 %		
Con riesgo potencial	B-1	6 %-9 %	$2 > y \leq 2.99$	Medio Bajo
	B-2	10 %-19 %		
Deficientes	C-1	20 %-39 %	$3 > y \leq 3.99$	Medio
	C-2	40 %-59 %		
De dudoso recaudo	D	60 %-99 %	$4 > y \leq 4.99$	Medio Alto
Pérdida	E	100 %	$5 > y \leq 5.99$	Alto

Cuadro 2.17: Provisiones de pérdidas según el nivel de riesgo

En el cuadro 2.18 se muestra el valor total provisionado por tipo de entidad. El monto provisionado para bancos corresponde a 88.80 % del total de provisiones. Este mismo monto representa 3.16 % del total de depósitos cubiertos y 0.92 % del total de depósitos asegurados del Sector Financiero Privado.

Por otro lado, el valor provisionado para sociedades financieras representa 0.30 % de los depósitos cubiertos y 0.09 % del total de depósitos asegurados del Sector Financiero Privado.

Finalmente, el valor provisionado para mutualistas corresponde a 0.09 % del total de depósitos cubiertos y 0.03 % del total de depósitos asegurados del Sector Financiero Privado.

Entidad financiera	Provisión (en miles de US\$)	% total provisiones	% depósitos cubiertos	% depósitos asegurados
Bancos	242,428.18	88.80	3.16	0.92
Sociedades Financieras	23,358.98	8.56	0.30	0.09
Mutualistas	7,209.98	2.64	0.09	0.03
TOTAL	272,997.14	100.00	3.56	1.03

Cuadro 2.18: Provisiones por tipo de entidad

De este modo, la pérdida esperada asciende a US\$ 1,201,745 miles, lo que corresponde a un incremento de 29.4% considerando el valor calculado en la sección 2.5.

En este capítulo se desarrolló toda la metodología sobre la estimación de las probabilidades con los datos de la entidad X_1 , dicha estimación se realizó para todas las entidades incluidas en la sección 2 (ver Apéndice), a partir de aquí, se estimó las correlaciones, pérdida esperada y medidas de riesgo. Finalmente, se calculó las provisiones para el conjunto de entidades que no fueron parte de la estimación de los modelos.

2.11 Resumen de resultados

Los modelos logísticos se estimaron para las ocho entidades incluidas en el análisis y sus resultados pueden ser interpretados según se presenta a continuación para la entidad X_1 , la cual se ha tomado como ejemplo para ilustrar la metodología de la estimación de las probabilidades de riesgo alto. Los resultados para el resto de entidades se presentan desde el apéndice 2 en adelante.

- Las variables *cobertura de cartera*, *rentabilidad*, y *patrimonio técnico constituido sobre activos y contingentes ponderados por riesgo* tienen una relación negativa con la probabilidad de riesgo alto, es decir, incrementos de estas variables se asocian con reducciones de la probabilidad de riesgo alto. Mientras que *improductividad de activos* muestra una relación positiva con la probabilidad de riesgo alto, es decir, incrementos de esta variable se asocian con aumentos de la probabilidad de riesgo alto.

- Respecto a la bondad de ajuste, el modelo arrojó un porcentaje de correcta clasificación igual a 78 %. Mientras que las medidas conocidas como pseudo- R^2 arrojaron valores de 0.65 según el criterio de Mc Fadden y 0.89 según Cragg y Uhler.
- Las predicciones del modelo muestran gran variabilidad que a lo largo del tiempo no cambia significativamente (valor promedio de 5.8 % y desviación estándar de 3.6 %), a excepción del pico que se observa en el año 2014 (17.6 %).
- La matriz de correlaciones, que muestra la relación existente entre las probabilidades de riesgo alto de las distintas entidades financieras, indica que la relación lineal más alta es entre X_5 y X_7 (0.76), seguido de la correlación entre las entidades X_4 y X_5 (0.73). Mientras que las correlaciones más bajas son entre las entidades X_1 y X_4 (-0.04), y X_3 y X_8 (0.062).

Una vez que se estiman las probabilidades de riesgo alto para todas las entidades incluidas en el estudio se realiza el proceso de simulación, del cual se obtuvo los siguientes resultados:

- El valor de la pérdida esperada del Seguro de Depósitos en el plazo de un año es US\$ 928,748 miles y corresponde a 3.52 % de los depósitos asegurados y 12.10 % de los depósitos cubiertos.
- La tasa de riesgo alto esperada de las pérdidas es 2.3 %, dicha tasa corresponde al número de eventos que pasan a riesgo alto sobre el total de eventos simulados.
- La pérdida asumida por el Seguro de Depósitos al 90 % de confianza según la medida de riesgo VaR equivale a 8 % y 28 % de los depósitos asegurados y cubiertos respectivamente. Mientras que si se toma el TVaR como medida de riesgo, estos porcentajes son evidentemente mayores, 12 % de los depósitos asegurados y 41 % de los depósitos cubiertos respectivamente.
- Según el VaR al 95 % de confianza, la pérdida representa 11 % de los depósitos asegurados y 37 % de los depósitos cubiertos. Mientras que el TVaR obtenido para este nivel de confianza, corresponde a 14 % y

50 % de los depósitos asegurados y cubiertos respectivamente.

- Con 99 % de confianza, la pérdida asumida por el Seguro de Depósitos equivale a 17 % y 57 % de los depósitos asegurados y cubiertos respectivamente, de acuerdo con el VaR. Mientras que el TVaR es de 21 % de los depósitos asegurados y 71 % de los depósitos cubiertos.

Capítulo 3

Conclusiones y recomendaciones

3.1 Conclusiones

- Del análisis de las relaciones entre las variables explicativas respecto al nivel de riesgo de las entidades financieras se concluye que del conjunto de variables macroeconómicas, la variable *tasa de interés*, en todos los casos en los que resultó significativa, tiene una relación positiva con el nivel de riesgo, con lo cual aumentos de esta variable se relaciona con aumentos de la probabilidad de que una entidad financiera pase a una calificación de riesgo alto. Por otro lado, es interesante constatar que la *variación del precio del petróleo* no resultó estadísticamente significativa al 95% de confianza, incluso considerando rezagos de la variable, durante el período enero de 2007 hasta junio de 2016, en $t - 3$, $t - 6$, y $t - 12$, lo que quizás se podría explicar en razón de políticas de liquidez contracíclicas o políticas internas de las entidades financieras empleadas durante el período de reducción del precio del petróleo. Finalmente, la *variación del Índice de Actividad Económica Coyuntural* sí resultó significativa para seis de las ocho entidades analizadas. Sin embargo, la relación con la variable dependiente es distinta en cada modelo.
- En lo que respecta a las variables independientes de carácter financiero, se tiene que las variables *cobertura de cartera*, *rentabilidad*, y *liquidez corriente*, en todos los casos en los que resultaron significativas, se relacionan negativamente con la probabilidad de riesgo alto es decir, a mayores niveles de *cobertura de cartera*, *rentabilidad*, y *liqui-*

dez corriente, la probabilidad de riesgo alto es más baja. De su parte, las demás variables no presentan la misma relación en todos los casos en los que resultaron estadísticamente significativas, como son: *margen operativo* y *apalancamiento del activo*. Finalmente, las variables *rentabilidad del capital*, *patrimonio técnico constituido sobre activos y contingentes ponderados por riesgo*, e *improductividad de activos* resultaron significativas solo para una de las ocho entidades analizadas.

- La estimación de los modelos *logit* muestra que las probabilidades de riesgo alto para el conjunto de entidades estudiadas, oscilan entre 0 % y 17.6 % (valor máximo para X_1). Sin embargo, a excepción de la probabilidad máxima de la entidad X_1 , los demás valores llegan a niveles entre 10 % y 12 %.
- Respecto a la capacidad de predicción de los modelos probabilísticos, en general presentan una buena capacidad predictiva, con un error de clasificación promedio de 20 %, un mínimo de 5 % y máximo de 30 %. En lo que se refiere a la bondad de ajuste, según el criterio de Mc Fadden, los modelos de las entidades X_1 , X_5 y X_8 tienen una bondad de ajuste (pseudo- R^2) bastante alta (entre 65 % y 75 %), mientras que el ajuste más bajo es el de la entidad X_2 con 28 %. Sin embargo, se debe tener presente que estas medidas están basadas en la razón de verosimilitud y por tanto no se las puede interpretar como el porcentaje de la varianza total que es explicada por el modelo.
- La función de densidad de las pérdidas esperadas muestra un comportamiento que se alinea a la teoría sobre la que se basan las ciencias actuariales. Es decir, así como en el caso de la contratación de un seguro, la probabilidad de ocurrencia de siniestros es decreciente conforme aumenta el número de siniestros, en este caso la probabilidad de ocurrencia de liquidación forzosa de las entidades financieras también es decreciente. Con lo cual, la probabilidad de valores de pérdida altos tiende a cero, lo que se apreció en la figura 2.10, misma que muestra la asimetría hacia la derecha de la distribución y la presencia de colas densas.
- La tasa de riesgo alto muestra un comportamiento bastante similar al de la función de pérdidas, respecto a la asimetría hacia la derecha y

presencia de colas densas. En este caso se observa que para tasas de riesgo alto cercanas a cero la probabilidad de ocurrencia es alta (alcanza hasta un 31 %). Dicha probabilidad disminuye asintóticamente hacia cero a medida que la tasa de riesgo alto es mayor, lo cual nuevamente se relaciona con lo que la teoría actuarial señala.

- Se calculó los diferentes niveles óptimos del Seguro de Depósitos, es decir tanto el valor esperado de las pérdidas como las medidas de riesgo VaR y TVaR para diferentes niveles de confianza. La pérdida esperada (valor promedio) corresponde a 4 % del total de depósitos asegurados y 12 % de los depósitos cubiertos. Dichos porcentajes son los que se estima requeriría en promedio el Seguro de Depósitos en el plazo de un año.
- En lo referente a los valores extremos de la distribución de pérdidas, el nivel óptimo del Seguro de Depósitos al 90 % de confianza equivale a una reserva de 28 % de los depósitos cubiertos y 8 % de los depósitos asegurados según la medida de riesgo VaR.
- El nivel óptimo del Seguro de Depósitos con 99 % de confianza corresponde a 17 % de los depósitos asegurados y 57 % de los depósitos cubiertos según el VaR. Mientras que si se consideran los peores casos (TVaR), dicho nivel óptimo al 99 % de confianza es de 71 % de los depósitos cubiertos y 21 % de los depósitos asegurados. Es decir, a este nivel de confianza, esta es la reserva que cubriría los desembolsos requeridos al Seguro de Depósitos en el plazo de un año.
- Para complementar el análisis se incluyó el apartado de provisiones, con el fin de mantener un enfoque más conservador y no limitar el estudio exclusivamente a aquellas entidades con riesgo alto. Por lo que se consideró al conjunto de entidades del Sector Financiero Privado que representa un nivel de riesgo más bajo para el Seguro de Depósitos.
- En este sentido, según los datos del cuadro 2.18, el valor provisionado para bancos representa alrededor de 89 % del total de provisiones, mientras sociedades financieras y mutualistas representan 9 % y 3 %

respectivamente, lo que incrementa la pérdida esperada, basada en el modelo probabilístico y en la simulación de los datos en 29 %.

- Si se agrega este valor provisionado a la pérdida esperada, se obtiene que el Seguro de Depósitos requeriría 4.55 % de sus depósitos asegurados y 15.65 % de sus depósitos cubiertos.

3.2 Recomendaciones

- Como se indicó en las conclusiones, la pérdida esperada representa 3.52 % de los depósitos asegurados y 12.10 % de los depósitos cubiertos, si a este nivel óptimo se adiciona el valor de las provisiones, este porcentaje ascendería a 4.55 % de los depósitos asegurados y 15.65 % de los depósitos cubiertos del Seguro de Depósitos en el plazo de un año. Sin embargo, este nivel óptimo varía de acuerdo a los diferentes niveles de confianza para los cuales se calculó las medidas de riesgo (VaR y TVaR). En este estudio, estos porcentajes van desde 8 % hasta 21 % del total de depósitos asegurados y desde 28 % hasta 71 % del total de depósitos cubiertos. En este sentido, es importante que los administradores del Seguro de Depósitos conozcan esta gama de posibles niveles óptimos, con el fin de seleccionar aquel que se ajuste a sus necesidades y la posición frente al riesgo que se desea asumir.
- La metodología desarrollada estima una reserva que cubriría los desembolsos requeridos al Seguro de Depósitos en el plazo de un año. Con lo cual, se sugiere actualizar los resultados y revisar la metodología con periodicidad anual, ya que a partir de un año tanto las entidades consideradas en el presente análisis como las variables explicativas pueden variar en cuanto a la significancia de los modelos y por tanto cambiarían las estimaciones de las probabilidades de riesgo alto.
- Los modelos de espacio-estado, específicamente el *filtro de Kalman*, son útiles para predecir valores de una serie de tiempo, con las técnicas de: predicción, filtrado y alisado. Considerando esto, se recomienda extender la metodología actual y, a través de este método, predecir los valores de las variables explicativas para momentos del tiempo mayores a t ($t+1, t+2, \dots, t+h$). Una vez que se predicen estos valores, con base en los datos hasta junio de 2016, se sugiere incorporarlos al conjunto inicial de datos, para la posterior estimación de las probabilidades de riesgo alto. De este modo, la presente metodología no solo permitiría estimar la pérdida esperada dentro del plazo de un año, sino también predecir el valor de esta pérdida para tiempos mayores a este plazo.¹

¹A medida que se estiman predicciones para momentos del tiempo mayores a t , la probabilidad de que estos valores sean los estimados, es cada vez menor. Es decir, la probabilidad de que la variable x_i tome el valor de j en $t+1$ es mayor que la probabilidad de que x_i tome el valor k en $t+h$.

- En el presente estudio se realizó el análisis para las entidades que conforman el Sector Financiero Privado. Sin embargo, en próximas actualizaciones se recomienda aplicar esta metodología para el Sector Financiero Popular y Solidario, considerando sus particularidades e incluyendo a las mutualistas como parte de este sector (en este estudio fueron incluidas en el Sector Financiero Privado).
- Se ha considerado el porcentaje de recuperaciones según los estándares planteados por el Comité de Supervisión Bancaria de Basilea (2006). Sin embargo, en la medida en la que se disponga de más información sobre estos valores se sugiere incluirlos con base en los datos de la Corporación del Seguro de Depósitos Fondo de Liquidez y Fondos de Seguros Privados para el cálculo de la pérdida esperada ($EL_p = \sum_{i=1}^n EXP_i * p_i * LGD_i$ ²), con el objetivo de precisar aún más los resultados.

² $LGD_i = 1 - RR$; RR: recovery rate (% de recuperación)

Referencias

- Ahangarani, P. M. (2007). A new structural approach to the default risk of companies. *University of Southern California*.
- Allen, D., & Powell, R. (2011). Credit risk measurement methodologies. *Cowan University*.
- Analytics, G. R. . (2014). Credit risk: estimation techniques. *Crisil*.
- Brereton, T., Chan, J. C., & Kroese, D. P. (2012). Montecarlo methods for portfolio credit risk. *University of Queensland and Australian National University*.
- Castagna, A. (2009). Analytical credit VaR with stochastic probabilities of default and recoveries. *Fabio Mercurio of Bloomberg & Iason*.
- Charpentier, A. (2015). *Computational actuarial science using R* (C. . H. Book, Ed.). University of Québec.
- Cosma, R. S. (2012). Advanced data analysis. *Carnegie Mellon University, Dietrich College of Humanities & Social Sciences*.
- Damodaran, A. (2009). Value at risk - VaR. *NYU Stern School of Business*.
- Duffie, D. (2005). Credit risk modeling with affine processes. *Stanford University*.
- Elizalde, A. (2006). Credit risk models IV: understanding and pricing cdos. *Centro de Estudios Monetarios y Financieros*.
- Fox, J. (2010). Logit and probit models. *McMaster University, Department of Sociology*.
- Harvey, A. C. (1990). *Forecasting, structural time series models and the Kalman filter* (C. U. Press, Ed.). The Press Syndicate of the University of Cambridge.
- Hernández, C. (2013). A methodology for determining the target funding level of a deposit insurer. *Fogafin*.
- Hull, J., Nelken, I., & White, A. (2005). Merton's model, credit risk and volatility skews. *University of Toronto, Super Computer Consulting Incorporated*.
- Ioannidou, V. P., & Penas, M. F. (2008). Deposit insurance and bank risk-taking: evidence from internal loan ratings. *Tilburg University and Lancaster University*.
- Jeanblanc, M., & Rutkowski, M. (2001). *Mathematical finance* (Springer, Ed.). Université d'Évry Val d'Essonne, Warsaw University of Technology.
- Kalman, R. E. (1960). A new approach to linear filtering and prediction problems. *Transactions of the ASME-Journal of Basic Engineering*.
- Karoui, N. E., Jeanblanc, M., Jiao, Y., & Zargari, B. (2011). Conditional default probability and density. *Centre de Mathématiques Appliquées*.

- Kochendörfer, A. (2011). Vasicek single factor model. *Technische Universität Kaisers Lautern*.
- Kroese, D. P., Taimre, T., & Botev, Z. I. (2011). *Handbook of monte carlo methods* (J. W. . Sons, Ed.). University of Queensland.
- Kunt, A. D., & Kane, E. (2008). Deposit insurance around the world. *MIT Press*.
- Leo, T. W. (2014). Limited dependent variable. *St. Francis Xavier University*.
- Maccario, A., Sironi, A., & Zazzara, C. (2003). Applying credit risk models to deposit insurance pricing: empirical evidence from the italian banking system. *Unicredit Banca Mobiliare and Università "LUISS-Guido Carli", Università "Luigi Bocconi"*.
- Matjaž, V. (2014). Estimating probability of default and comparing it to credit rating classification by banks. *Bank of Slovenia*.
- Ortiz-Gracia, L., & Masdemont, J. J. (2014). Credit risk contributions under the Vasicek one-factor model: a fast wavelet expansion approximation. *Centre de Recerca Matemàtica*.
- Park, H. M. (2009). Regression models for binary dependent variables. *The University Information Technology Services (UITS) Center for Statistical and Mathematical Computing, Indiana University*.
- Parodi, P. (2015). *Pricing and general insurance* (C. . H. B. C. Press, Ed.).
- Peng, J. (2008). Value at Risk and Tail Value at Risk in uncertain environment. *Institute of Uncertain Systems, Huanggang Normal University*.
- Rodríguez, G. (2007). Generalized linear models. *Princeton University*.
- Rubinstein, R. Y., & Kroese, D. P. (2008). *Simulation and the monte carlo method* (J. W. . Sons, Ed.). University of Queensland.
- Vasicek, O. (1987). Probability of loss on loan portfolio. *KMV*.
- Vasicek, O. (2002). The distribution of loan portfolio value. *KMV*.
- Wilson, T. C. (1998). Portfolio credit risk. *McKinsey & Co. Inc.*
- Zhang, A. (2009). Statistical methods in credit risk modeling. *University of Michigan*.

Apéndice 1

Marco teórico

1.1 Teorema: probabilidades de incumplimiento

La probabilidad de incumplimiento de cada entidad n de un grupo de $n = 1, \dots, N$ entidades en el tiempo t es denotada como $p_{n,t}$ y dada por:

$$\begin{aligned} p_{n,t} &= \Phi(K_{n,t}), \\ K_{n,t} &= \frac{\ln D_n - \ln A_{n,0} - \left(\mu_n - \frac{\sigma_n^2}{2}\right)t}{\sigma_n \sqrt{t}} \end{aligned} \quad (1.1)$$

Donde $A_{n,0}$ corresponde al valor de la firma de cada deudor $n \in (1, \dots, N)$, y el deudor n cae en incumplimiento si el valor del activo cae por debajo de un límite, de modo que $D_n = \mathbf{1}_{(A_n < C_n)}$. Donde Φ corresponde a la función de distribución normal estándar.

Condicionado a la realización del factor sistemático (Y) descrito en la ecuación 1.10, la probabilidad de incumplimiento en el tiempo t de cada entidad financiera n se denota por $p_n(Y)$ y viene dado por:

$$\begin{aligned} p_n(Y) &= P[X_n < K_{n,t} \mid Y] \\ &= P\left[\sqrt{\rho}Y + \sqrt{1-\rho}\epsilon_n < K_n \mid Y\right] \\ &= P\left[\epsilon_n < \frac{K_n - \sqrt{\rho}Y}{\sqrt{1-\rho}} \mid Y\right] \\ &= \Phi\left(\frac{K_n - \sqrt{\rho}Y}{\sqrt{1-\rho}}\right) \end{aligned} \quad (1.2)$$

Además, las variables aleatorias (X_1, \dots, X_N) y las probabilidades de incumplimiento $(p_1(Y), \dots, p_n(Y))$ son independientes.

Partiendo de la ecuación 1.1, donde el valor de $K_{n,t}$ para cada entidad financiera es $K_{n,t} = \Phi^{-1}(p_{n,t})$ y sustituyendo en la ecuación anterior se obtiene que la probabilidad condicional de incumplimiento es:

$$p_n(Y_t) = \Phi \left(\frac{\Phi^{-1}(p_{n,t}) - \sqrt{\rho_t} Y_t}{\sqrt{1 - \rho_t}} \right) \quad (1.3)$$

La variable aleatoria L_t definida como la suma de las variables $L_{1,t}, \dots, L_{N,t}$, representa el número de incumplimientos en el portafolio, misma que tiene la siguiente distribución:

$$P(L \leq m) = \sum_{k=0}^m \binom{N}{k} \int_{-\infty}^{+\infty} p(y)^k (1 - p(y))^{N-k} \phi(y) dy \quad (1.4)$$

Prueba del teorema

La probabilidad de k incumplimientos es:

$$P(L = k) = E(P([L = k] | Y)) = \int_{-\infty}^{+\infty} P(L = k | Y = y) \phi(y) dy \quad (1.5)$$

Donde ϕ es la densidad de Y . Los incumplimientos son condicionales e independientes en Y , por tanto:

$$P(L = k | Y = y) = \binom{N}{k} p(y)^k (1 - p(y))^{N-k} \quad (1.6)$$

Así, para $m \in [1, \dots, N]$ se tiene:

$$P(L \leq m) = \sum_{k=0}^m \binom{N}{k} \int_{-\infty}^{+\infty} p(y)^k (1 - p(y))^{N-k} \phi(y) dy \quad (1.7)$$

1.2 Portafolio homogéneo grande (*LHP*)

Como Schönbucher (2000) y Vasicek (2002) explican, dado que los incumplimientos son independientes cuando se condicionan a la realización de un factor sistemático común Y_t , el supuesto de un portafolio grande y

de igual tamaño implica que aplicando la ley de los grandes números, la fracción de créditos incumplidos en el portafolio converge a la probabilidad individual de incumplimiento para cada crédito $p(Y_t)$.

El portafolio es ponderado con pesos $\omega_1^N, \dots, \omega_n^N, \sum_{n=1}^N \omega_n^N = 1$ tal que:

$$\lim_{N \rightarrow \infty} \sum_{n=1}^N (\omega_n^N)^2 = 0$$

Así, cuando N tiende a infinito los pesos de los créditos tienden a cero.

1.2.1 Lema: tasa de incumplimiento

La tasa de incumplimiento de la cartera se define por la siguiente fórmula:

$$L^N = \sum_{n=1}^N \omega_n^N D_n \in [0, 1] \quad (1.8)$$

Sobre la base de *LHP* se define el siguiente teorema:

$$E(L^N | Y) = p(Y) = \Phi \left(\frac{\Phi^{-1}(p) - \sqrt{\rho}Y}{\sqrt{1-\rho}} \right) \quad (1.9)$$

$$Var(L^N | Y) = \sum_{n=1}^N (\omega_n^N)^2 p(Y)(1-p(Y)) \quad (1.10)$$

1.2.1.1 Prueba del lema

Linealidad de la expectativa condicional de los rendimientos:

$$\begin{aligned} E(L^{(N)} | Y) &= \sum_{n=1}^N \omega_n^N E(D_n | Y) \\ &= \sum_{n=1}^N \omega_n^{(N)} P(D_n | Y) \\ &= p(Y) \sum_{n=1}^N \omega_n^{(N)} \\ &= p(Y) \end{aligned} \quad (1.11)$$

D_n son condicionales e independientes en Y , por tanto:

$$\begin{aligned}
\text{Var}(L^N | Y) &= \sum_{n=1}^N (\omega_n^N)^2 \text{Var}(D_n | Y) \\
&= \sum_{n=1}^N (\omega_n^N)^2 p(Y)(1 - p(Y))
\end{aligned} \tag{1.12}$$

1.2.1.2 Teorema sobre la tasa de incumplimiento

La tasa de incumplimiento en portafolios homogéneos grandes (LHP) converge en probabilidad cuando $N \rightarrow \infty$

$$L^N \rightarrow p(Y) = \Phi \left(\frac{\Phi^{-1}(p) - \sqrt{\rho}Y}{\sqrt{1 - \rho}} \right) \tag{1.13}$$

1.2.1.3 Prueba del teorema

Para un portafolio grande, la variación de la tasa de incumplimiento dado Y tiende a 0.

$$\begin{aligned}
\text{Var}(L^N | Y) &= \sum_{n=1}^N (\omega_n^N)^2 p(Y)(1 - P(y)) \\
&= \leq \frac{1}{4} \sum_{n=1}^N (\omega_n^N)^2 \xrightarrow{N \rightarrow \infty} 0
\end{aligned} \tag{1.14}$$

1.3 Modelos espacio-estado: filtro de Kalman

El *filtro de Kalman* resulta útil en múltiples ámbitos, tanto para medir la serie real (no observable) a partir de una variable observada, como para predecir valores de una serie de tiempo, con lo cual podría resultar de interés realizar predicciones sobre las variables consideradas y por tanto, sobre las probabilidades de riesgo alto. La presente investigación se enfoca en definir un modelo que permita realizar las estimaciones dentro del plazo de un año. Sin embargo, podría resultar de interés realizar predicciones sobre las variables consideradas y por tanto, sobre las probabilidades de riesgo alto.

Los modelos espacio-estado son ampliamente utilizados en una serie de ámbitos. En economía particularmente, son de gran utilidad para predecir el comportamiento de variables cuyas fuerzas impulsoras pueden resultar no observables o medibles. Por ejemplo, el nivel de ingreso de una persona puede depender de factores no observables como su inteligencia, capacida-

des, habilidades, entre otros.

A nivel agregado, la teoría económica sugiere que las variables macroeconómicas e incluso las variables de tipo financiero son impulsadas por factores no observables, como pueden ser: el cambio tecnológico, la acumulación de capital humano, el precio de los activos, entre otros.

Cuando las variables explicativas no son observables, ciertos modelos ya no resultan útiles para estudiar la evolución de las mismas. Sin embargo, con la aplicación de modelos espacio-estado se amplía el marco de modelos econométricos y es posible construir modelos tanto explicativos como predictivos para estas variables.

Estos modelos permiten al investigador modelar una serie de tiempo observable, $(y_t)_{t=1}^T$, que se explica por un vector de variables estado, $(z_t)_{t=1}^T$, el cual es conducido por un proceso estocástico.

En general, estos modelos pueden resultar muy útiles cuando el objetivo es eliminar el ruido de una variable aleatoria. Por tanto, a partir de la información observada se busca estimar el estado $(z_t)_{t=1}^T$ aislando el término de error de $(y_t)_{t=1}^T$.

Estimación de los modelos espacio-estado

El modelo se construye de la siguiente forma:

$$y_t = Hz_t + v_t \quad v_t \sim N(0, \Sigma_v) \quad \text{Ecuación de observación} \quad (1.15)$$

$$z_t = Bz_{t-1} + w_t \quad w_t \sim N(0, \Sigma_w) \quad \text{Ecuación de transición} \quad (1.16)$$

La primera ecuación describe la relación entre la serie de tiempo observada y_t y el estado z_t . En general, se asume que los datos de la variable observada y_t incluyen un término de error que se mide con v_t expresado en la ecuación 1.15. El enfoque tradicional es modelar v_t como un error gaussiano, $v_t \sim N(0, \Sigma_v)$. La segunda ecuación describe la evolución de la variable estado como un proceso estocástico de innovaciones w_t , las cuales también se asume que se distribuyen como una normal $w_t \sim N(0, \Sigma_w)$. (Kalman, 1960)

Los modelos espacio-estado pueden ser formulados de maneras mucho más generales que las ecuaciones 1.15 y 1.16.

En aplicaciones prácticas, las varianzas Σ_w y Σ_v y el sistema de matrices H y B es desconocido y debe ser estimado, para lo cual se aplica la inferencia basada en la verosimilitud, dado que el *filtro de Kalman* permite construir la función de verosimilitud asociada a un modelo espacio-estado.

Filtro de Kalman

Por una parte, se asume que se conoce los datos de $(y_t)_{t=1}^T$. Por otra, se asume que se dispone de valores iniciales razonables para el conjunto de parámetros $\delta = \{H^*, B^*, \Sigma_{*v}, \Sigma_{*w}\}$. Se denota como $f(y_1, y_2, \dots, y_T; \delta)$ a la función de verosimilitud del modelo espacio-estado para un conjunto de parámetros δ . Por el teorema de Bayes se puede factorizar la función de verosimilitud como:

$$\begin{aligned} f(y_1, y_2, \dots, y_T; \delta) &= f(y_1, \delta) f(y_2 | y_1, \delta) f(y_3 | y_2, y_1, \delta) \dots f(y_T | y_{T-1}, \dots, y_1, \delta) \\ &= \prod_{t=1}^T f(y_t | y_{t-1}, \delta) \end{aligned} \quad (1.17)$$

Donde $y^0 = \phi$ y $y^{t-1} = (y_1, y_2, \dots, y_{t-1})$ para $t \geq 2$. El logaritmo de la función de verosimilitud viene dado por:

$$\ln L(y^T, \delta) = \sum_{t=1}^T \ln f(y_t | y^{t-1}, \delta) \quad (1.18)$$

Para poder construir la función de verosimilitud se necesita derivar las funciones de densidad.

$$f(y_t | y^{t-1}, \delta), t = 1, 2, \dots, T \quad (1.19)$$

Para lo cual se pueden usar las técnicas de filtrado del *filtro de Kalman*, puesto que los errores se consideran gaussianos.

El *filtro de Kalman* es un procedimiento recursivo definido en cuatro etapas: inicio, predicción, corrección y construcción de la función de verosimilitud.

Condiciones iniciales

Esta técnica empieza derivando el mejor predictor del estado inicial $z_{0|0}$ y un estimador de la matriz de varianzas y covarianzas $\Sigma_{0|0}^z = E[(z_0 - z_{0|0})(z_0 - z_{0|0})']$. Si el proceso es estacionario, se puede definir $z_{0|0} = z^*$ y $\Sigma_{0|0}^z = \Sigma^*$ tal que:

- $z^* = Bz^*$
- $\Sigma^* = B\Sigma^*B' + \Sigma_w$

Previo a continuar con la predicción, se establece $t = 1$ de modo que $z_{t-1|t-1} = z_{0|0}$ y $\Sigma_{t-1|t-1}^z = \Sigma_{0|0}^z$.

Predicción

En el tiempo t se puede usar $z_{t-1|t-1}$ y $\Sigma_{t-1|t-1}^z$ en la ecuación de transición para obtener:

$$z_{t|t-1} = Bz_{t-1|t-1} \quad (1.20)$$

$$\Sigma_{t|t-1}^z = B\Sigma_{t-1|t-1}^zB' + \Sigma_w \quad (1.21)$$

Ahora se puede usar $z_{t|t-1}$ para construir la predicción $y_{t|t-1} = Hz_{t|t-1}$. Dado que se observa y_t , se puede predecir el error:

$$u_t = y_t - y_{t|t-1} = y_t - Hz_{t|t-1} = v_t + H(z_t - z_{t|t-1}) \quad (1.22)$$

Dado que los errores son gaussianos se cumple que $u_t \sim N(0, \Sigma_v + H\Sigma_{t|t-1}^zH')$. Además, dado que $y_t = u_t + y_{t|t-1}$ se cumple que $f(y_t | y^{t-1}; \delta) = f(u_t, \delta)$.

Así, para construir la función de verosimilitud se necesita $f(y_t | y^{t-1}; \delta)$ para todo $t = 1, 2, \dots, T$; y dado $z_{t-1|t-1}$ y $\Sigma_{t-1|t-1}^z$ se puede calcular $f(y_t | y^{t-1}, \delta)$ de una distribución normal estándar.

$$f(y_t | y^{t-1}, \delta) = \frac{1}{\sqrt{(2\pi)^{n_y} |\Sigma_v + H\Sigma_{t|t-1}^zH'|}} \exp\left(-\frac{u_t'(\Sigma_v + H\Sigma_{t|t-1}^zH')^{-1}u_t}{2}\right) \quad (1.23)$$

Consecuentemente, para calcular $f(y_{t+1} | y^t, \delta)$ se necesita $z_{t|t}$ y $\Sigma_{t|t}^z$. En otras palabras, se necesita corregir las predicciones de los estados usando la nueva información disponible en el tiempo t .

Corrección

Habiendo observado y_t se pueden actualizar las predicciones $z_{t|t-1}$ y $\Sigma_{t|t-1}^z$ según la fórmula del *filtro de Kalman*:

$$z_{t|t} = z_{t|t-1} + K_t(y_t - y_{t|t-1}) = z_{t|t-1} + K_t(y_t - H z_{t|t-1}), \quad (1.24)$$

$$\Sigma_{t|t}^z = \Sigma_{t|t-1}^z - K_t(\Sigma_v + H \Sigma_{t|t-1}^z H') K_t', \quad (1.25)$$

Donde:

$$K_t = \Sigma_{t|t-1}^z H' (H \Sigma_{t|t-1}^z H' + \Sigma_v)^{-1} \quad (1.26)$$

De esta forma se asume que la predicción corregida es una combinación lineal entre la predicción anterior $z_{t|t-1}$ y el error en el tiempo t ($y_t - y_{t|t-1}$).

Dada la forma lineal, K_t se selecciona tal que minimice la varianza del error de predicción. Finalmente, a menos que $t = T$ se incrementa t y se regresa a la predicción. De no ser así, se pasa a la construcción de la función de verosimilitud.

Construcción de la función de verosimilitud

Los dos pasos anteriores calculan recursivamente la función de densidad $f(y_t | y^{t-1}, \delta)$ para $t = 1, 2, \dots, T$, la cual se utiliza para construir:

$$L(y^T, \delta) = \prod_{t=1}^T f(y_t | y^{t-1}, \delta) \quad (1.27)$$

Para la maximización de la función de verosimilitud, al ser una función no lineal se resuelve numéricamente y no de forma analítica. Para ello existen varios métodos, por ejemplo, si la función es continua los métodos basados en el gradiente suelen ser los más utilizados, como el método de Newton Rapson. También se puede utilizar el algoritmo EM (Expected Maximization).

Una vez que se han obtenido las estimaciones que maximizan la función de verosimilitud, es posible utilizar los modelos espacio-estado para hacer predicciones sobre los valores de las variables observables. En particular, se puede usar el estado predictor $z_{T|T}$ implicado en los estimadores máximo verosímiles δ_{ML} , con la ecuación de transición para construir $y_{T+h|T}$ para $h = 1, 2, \dots$, según:

$$y_{T+h|T} = HB^h z_{T|T} \quad (1.28)$$

Como una extensión al presente estudio¹, la utilización de los modelos espacio-estado permitiría predecir los valores de las variables independientes para períodos mayores a t . A partir de aquí, se podría incluir estos valores al conjunto de datos inicial para estimar las probabilidades de riesgo alto a través de los modelos *logit* para períodos $t + 1, t + 2, \dots, t + h$.

¹Incluir valores de predicción está fuera del alcance del presente análisis.

Apéndice 2

Estimación de los modelos *logit* para las entidades financieras

2.1 Entidad financiera X_2

Para esta entidad se estimó un modelo *logit* con variable respuesta de tipo binaria. A continuación se presenta los resultados obtenidos de la estimación del modelo:

<i>Variable dependiente:</i>	
Nivel de riesgo	
IDEAC	-128.449** (53.682)
AA	-46.129** (18.678)
RC	-28.043** (12.745)
Constante	39.093** (15.585)
Observaciones	
	114
Log Likelihood	
	-54.837
Akaike Inf. Crit.	
	117.675
Nota:	
	*p<0.1; **p<0.05; ***p<0.01

Cuadro 2.1: X_2 : modelo *logit* con variable dependiente binaria

De la metodología *stepwise* se obtiene que las variables: *variación del Índice de Actividad Económica Coyuntural*, *rentabilidad del capital*, y *apalancamiento del activo*, tienen una relación estadísticamente significativa y negativa respecto a la variable dependiente binaria, con lo cual, incrementos marginales de estas variables están relacionados con disminuciones de las probabilidades de riesgo alto.

En la siguiente tabla cruzada se valida la capacidad de predicción del modelo, en donde los elementos de la diagonal corresponden al porcentaje de aciertos (correcta clasificación) resultantes de la estimación:

	0	1
0	0.53	0.08
1	0.17	0.22

Cuadro 2.2: Validación cruzada: nivel de riesgo CAMEL vs. nivel de riesgo predicho por el modelo

En este modelo el error de predicción es de 25 % y corresponde al número de veces sobre el total de datos en los cuales el nivel de riesgo predicho por el modelo es distinto del nivel de riesgo otorgado por la calificación CAMEL.

En el cuadro 2.3 se muestra los indicadores de bondad de ajuste del modelo *logit* estimado. En este caso, los valores pseudo- R^2 de Mc Fadden y Cragg y Uhler son 0.28 y 0.42 respectivamente. Es importante destacar que aunque el ajuste para este modelo no es tan alto, estos indicadores presentan niveles satisfactorios en términos de la utilidad del modelo para explicar las relaciones entre las variables, además de su capacidad predictiva que, a través de la validación cruzada, arrojó 75 % de aciertos en la predicción de los niveles de riesgo.

Indicador	Valor
Log-likelihood fitted model	-54.84
Log-likelihood intercept-only	-76.03
G2 ¹	42.38
McFadden	0.28
Maximum likelihood	0.31
Cragg y Uhler	0.42

Cuadro 2.3: Pseudo R^2

A continuación se presenta el gráfico de las predicciones del modelo:

¹Menos dos veces la diferencia en el logaritmo de las verosimilitudes.

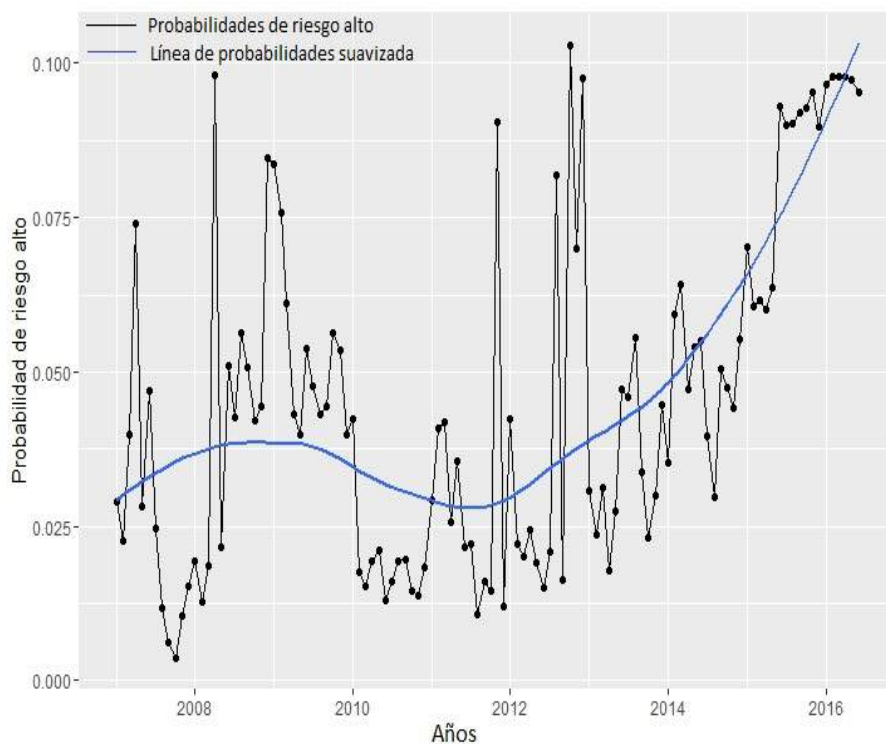


Figura 2.1: X_2 : probabilidades de riesgo alto

Las probabilidades fluctúan entre 0% y 10.3% durante el período de análisis. Se observa una tendencia creciente más pronunciada a partir del año 2013. Este gráfico se resume en el siguiente cuadro:

Riesgo alto		
1	Mínimo:	0,004
2	1er cuartil:	0,021
3	Mediana:	0,042
4	Media:	0,045
5	3er cuartil:	0,060
6	Máximo:	0,103

Cuadro 2.4: X_2 : resumen de las probabilidades de riesgo alto

Los siguientes gráficos bivariantes muestran la relación entre las variables explicativas y la variable respuesta asociada a las probabilidades de riesgo alto²:

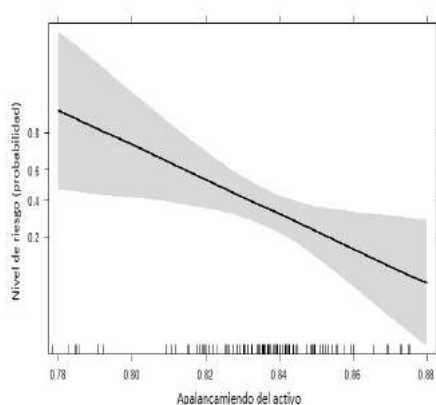


Figura 2.2: X_2 : apalancamiento del activo vs. nivel de riesgo

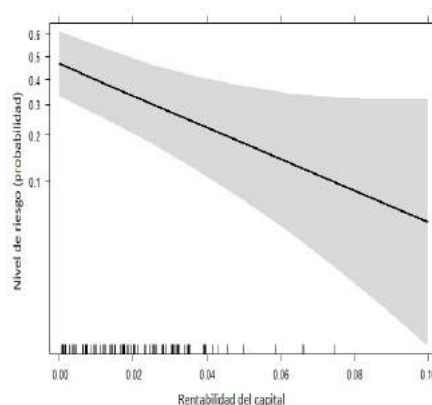


Figura 2.3: X_2 : rentabilidad del capital vs. nivel de riesgo

En el caso de las figuras 2.2 y 2.3 la relación es evidentemente negativa, con lo cual para valores altos de estas variables, la calificación de riesgo es baja y viceversa. Por otro lado, mientras más datos se encuentren en algún intervalo de valores determinado, el ancho de las bandas sombreadas en gris será menor (mayor precisión), mientras que si las observaciones son muy pocas, el área sombreada en gris será más ancha.

²El tipo de gráfico cambia respecto a los presentados en la sección 2.3, ya que depende del número de categorías que tenga la variable dependiente, en este caso esta entidad únicamente presenta dos niveles.

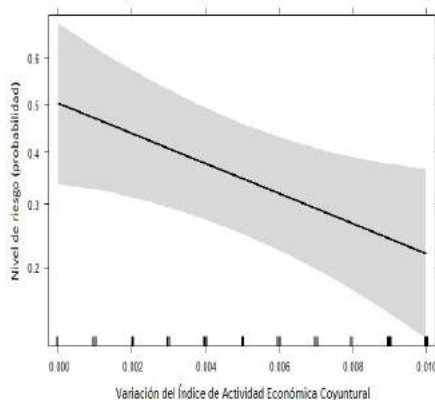


Figura 2.4: X_2 : variación del Índice de Actividad Económica Coyuntural vs. nivel de riesgo

Finalmente, respecto a la *variación del Índice de Actividad Económica Coyuntural* la relación también es negativa, a mayor *variación del Índice de Actividad Económica Coyuntural*, la calificación de riesgo obtenida por esta entidad es menor.

2.2 Entidad financiera X_3

Para esta entidad se estimó un modelo *logit* con variable respuesta de tipo multinomial debido a que las calificaciones de riesgo obtenidas con la metodología CAMEL fueron más de dos durante el periodo de estudio. A continuación se presentan los resultados obtenidos de la estimación del modelo:

<i>Variable dependiente:</i>	
Nivel de riesgo	
IA	22.214** (10.433)
TI	218.130*** (50.465)
IDEAC	-106.776 (83.132)
LC	-21.029*** (7.654)
R	-8.096** (4.284)
Observaciones	114

Nota: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

Cuadro 2.5: X_3 : modelo *logit* con variable dependiente multinomial

En este caso todas las variables resultaron estadísticamente significativas a excepción de la *variación del Índice de Actividad Económica Coyuntural*. Por un lado, *improductividad de activos*, y *tasa de interés* tienen una relación positiva con la variable respuesta. Mientras que *liquidez corriente*, y *rentabilidad* muestran una relación negativa respecto a la variable dependiente.

La siguiente tabla cruzada valida la capacidad de predicción del modelo:

	2	3	4	5
2	0.00	0.02	0.00	0.00
3	0.00	0.16	0.06	0.00
4	0.00	0.04	0.62	0.01
5	0.00	0.00	0.06	0.03

Cuadro 2.6: Validación cruzada: nivel de riesgo CAMEL vs. nivel de riesgo predicho por el modelo

En 81 % de los casos el modelo predijo el nivel de riesgo otorgado por la metodología *CAMEL*, siendo el error de predicción de 19 %.

Indicador	Valor
Log-likelihood fitted model	-61.63
Log-likelihood intercept-only	-102.56
G ² ³	81.86
McFadden	0.40
Maximum likelihood	0.51
Cragg y Uhler	0.61

Cuadro 2.7: Pseudo R^2

Los indicadores de bondad de ajuste del modelo *logit* estimado muestran valores pseudo- R^2 , según el criterio de Mc Fadden y Cragg y Uhler, iguales a 0.40 y 0.61, con lo cual se puede asumir que el modelo predice bien las probabilidades de riesgo alto.

³Menos dos veces la diferencia en el logaritmo de las verosimilitudes

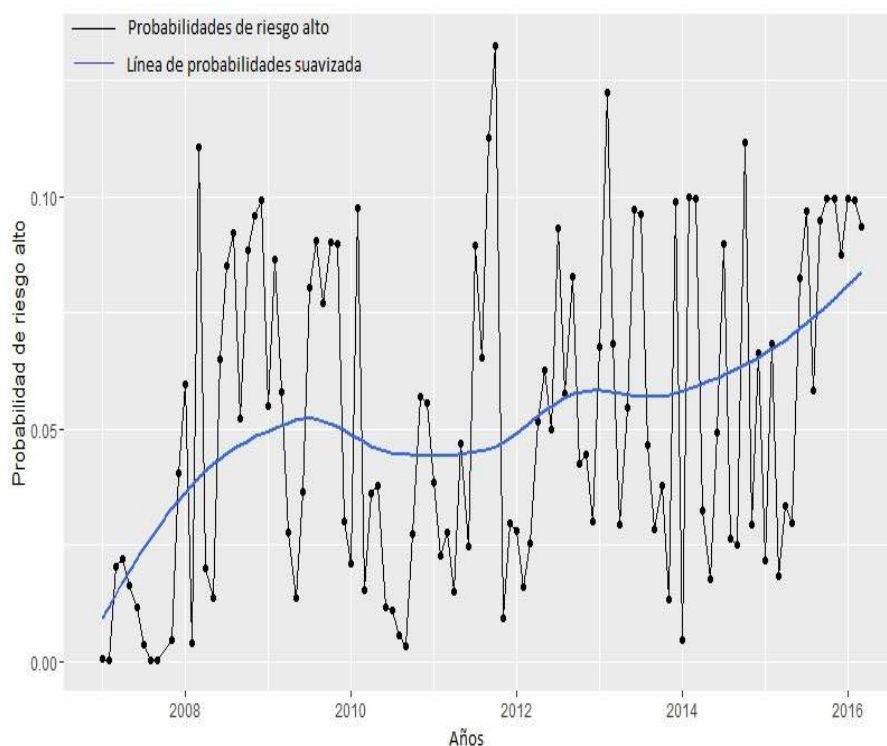


Figura 2.5: X_3 : probabilidades de riesgo alto

Se observa fluctuaciones altas, mostrándose una tendencia creciente para todo período de análisis. En el siguiente cuadro se resumen los estadísticos de las predicciones:

Riesgo alto		
1	Mínimo.:	0,001
2	1er cuartil:	0,02
3	Mediana:	0,034
4	Media:	0,056
5	3er cuartil:	0,078
6	Máximo :	0,094

Cuadro 2.8: X_3 : resumen de las probabilidades de riesgo alto

La probabilidad media de riesgo alto es de 6%, con un máximo de 9%.

APÉNDICE 2. ESTIMACIÓN DE LOS MODELOS LOGIT PARA LAS ENTIDADES FINANCIERAS

A continuación se muestra los gráficos bivariantes entre las variables explicativas y la variable respuesta asociada a las probabilidades de riesgo alto estimadas por el modelo:

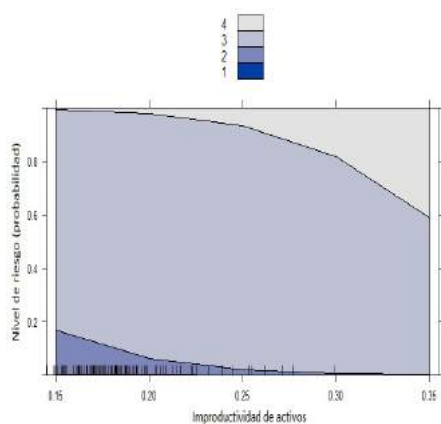


Figura 2.6: X_3 : improductividad activos vs. nivel de riesgo

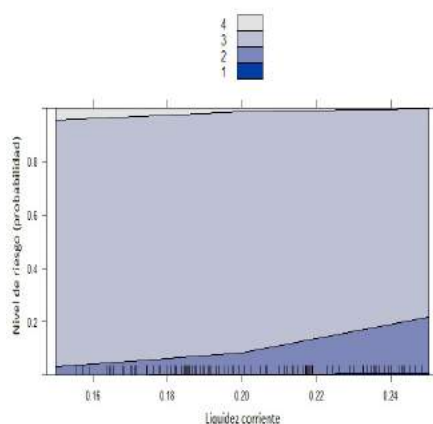


Figura 2.7: X_3 : liquidez corriente vs. nivel de riesgo

Por una parte se observa que a mayores niveles de *improductividad de activos* el riesgo asociado es más alto. Mientras que en el caso de *liquidez corriente*, el nivel de riesgo es prácticamente el mismo para diversos valores de este indicador.

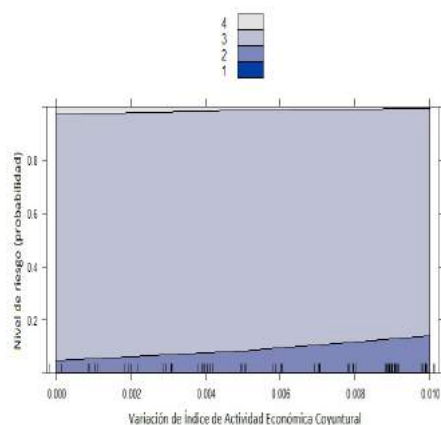


Figura 2.8: X_3 : variación del Índice de Actividad Económica Coyuntural vs. nivel de riesgo

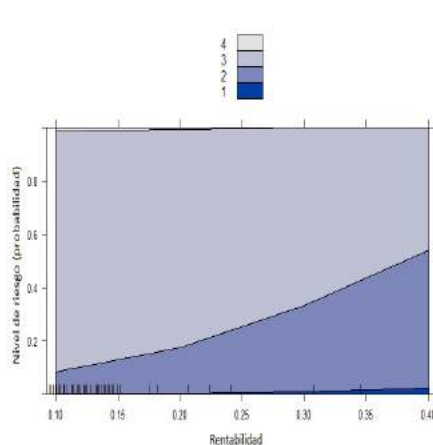


Figura 2.9: X_3 : rentabilidad vs. nivel de riesgo

Para esta entidad la relación de las variables *variación del Índice de Actividad Económica Coyuntural*, y *rentabilidad* indica que a mayores niveles de estas variables el riesgo asociado es más bajo y por tanto la probabilidad también.

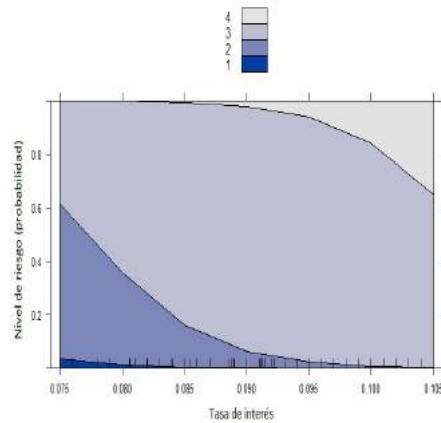


Figura 2.10: X_3 : tasa de interés vs. nivel de riesgo

Respecto a la *tasa de interés* se observa que la relación es positiva con los niveles de riesgo.

2.3 Entidad financiera X_4

Para esta entidad se estimó un modelo *logit* con variable respuesta de tipo multinomial. A continuación se presenta los resultados obtenidos de la estimación del modelo:

<i>Variable dependiente:</i>	
Nivel de riesgo	
CC	-7.397*** (1.499)
MO	-9.659*** (2.778)
LC	-24.846** (10.703)
IDEAC	151.577*** (46.141)
TI	180.597*** (53.679)
PP	-3.590 (2.416)
Observations	114

Nota: *p<0.1; **p<0.05; ***p<0.01

Cuadro 2.9: X_4 : modelo *logit* con variable dependiente multinomial

Para esta regresión las variables *cobertura de cartera*, *margen operativo*, *liquidez corriente* y *variación del precio del petróleo* muestran una relación negativa respecto a la probabilidad de riesgo alto. Sin embargo, esta última no resulta estadísticamente significativa al 5%. De su parte, la *variación del Índice de Actividad Económica Coyuntural* y *tasa de interés* tienen una relación positiva respecto a la probabilidad de riesgo alto y son estadísticamente significativas.

En la siguiente tabla cruzada se valida la capacidad de predicción del modelo⁴:

	2	3	4	5
2	0.13	0.05	0.00	0.00
3	0.04	0.17	0.08	0.01
4	0.00	0.04	0.30	0.01
5	0.00	0.00	0.03	0.14

Cuadro 2.10: Validación cruzada: nivel de riesgo CAMEL vs. nivel de riesgo predicho por el modelo

Del cuadro 2.10 se extrae que la estimación del modelo fue correcta en 74 % de los casos.

Indicador	Valor
Log-likelihood fitted model	-75.98
Log-likelihood intercept-only	-152.65
G ² ⁵	153.32
McFadden	0.50
Maximum likelihood	0.74
Cragg y Uhler	0.79

Cuadro 2.11: Pseudo R^2

En el cuadro 2.11 las medidas de bondad de ajuste del modelo *logit* estimado muestran valores de 0.50 y 0.79 según los criterios de Mc Fadden, y Cragg y Uhler respectivamente, lo que significa que el modelo predice bastante bien los valores de las probabilidades de riesgo alto.

⁴En la tabla cruzada se presenta todas las categorías de riesgo que ha tenido la serie durante el período de análisis. En ciertos casos es posible que la entidad no haya presentado alguna de las categorías de riesgo, en este por ejemplo la entidad no ha obtenido la calificación de riesgo *bajo* a lo largo del período de estudio.

⁵Menos dos veces la diferencia en el logaritmo de las verosimilitudes

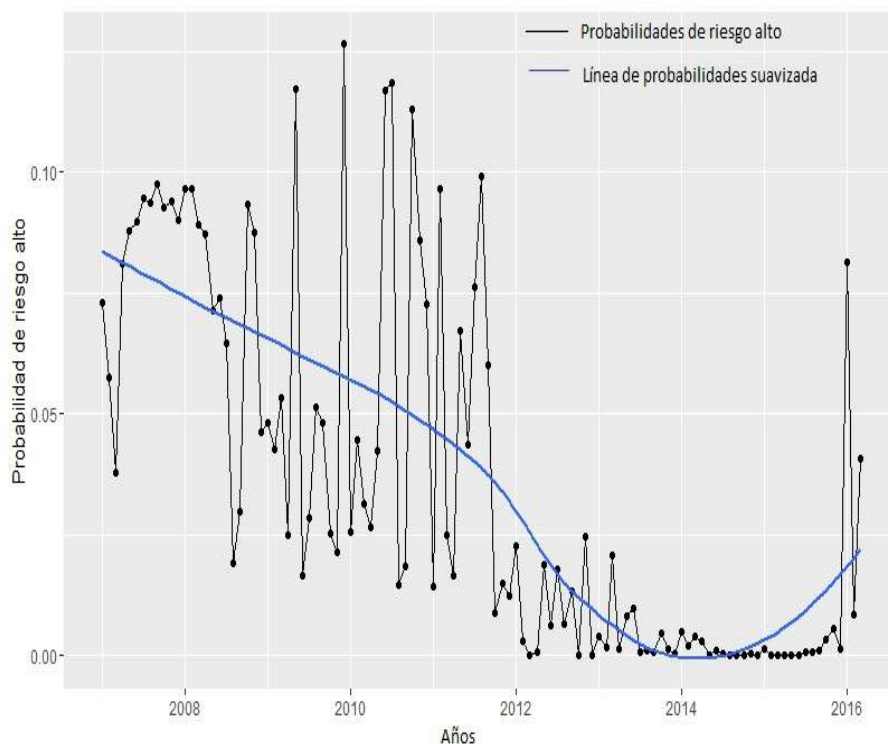


Figura 2.11: X_4 : probabilidades de riesgo alto

En este caso se evidencia que la serie presenta gran varianza para los primeros períodos y valores cercanos y superiores a 10 %. A partir de aquí, las probabilidades se reducen notablemente. Sin embargo, al final del período nuevamente se observa que las probabilidades empiezan a incrementarse.

Riesgo alto		
1	Mínimo:	$3,64e^{-5}$
2	1er cuartil:	0,002
3	Mediana:	0,029
4	Media:	0,036
5	3er cuartil:	0,072
6	Máximo :	0,127

Cuadro 2.12: X_4 : resumen de las probabilidades de riesgo alto

Para esta mutualista la media de las probabilidades es de 4 %, alcanzando un máximo de 13 %.

A continuación se muestra los gráficos bivalentes entre las variables explicativas y la variable respuesta asociada a las probabilidades de riesgo alto estimadas por el modelo:

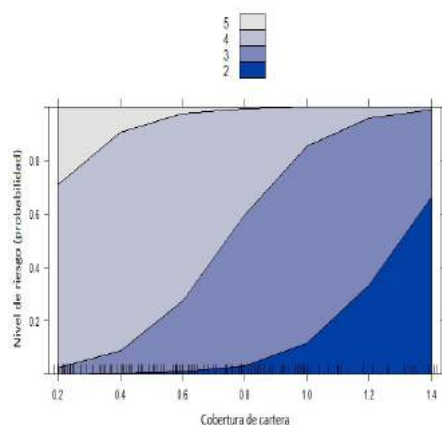


Figura 2.12: X_4 : cobertura de la cartera vs. nivel de riesgo

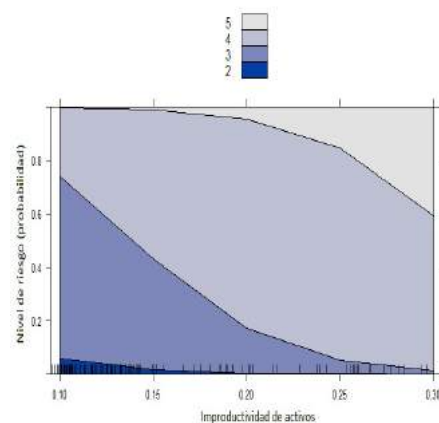


Figura 2.13: X_4 : improductividad de activos vs. nivel de riesgo

Según las figuras 2.12 y 2.13, a valores altos de *cobertura de cartera* menor es el nivel de riesgo asociado, mientras que a mayor nivel de *improductividad de activos* el nivel de riesgo aumenta.

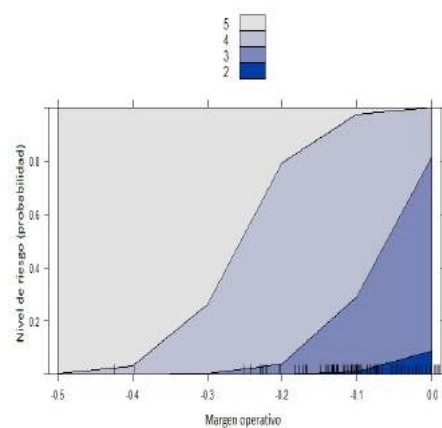


Figura 2.14: X_4 : margen operativo vs. nivel de riesgo

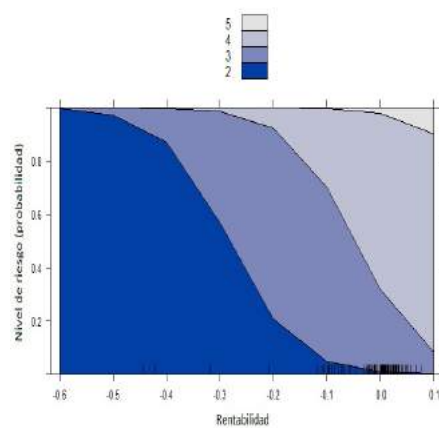


Figura 2.15: X_4 : rentabilidad vs. nivel de riesgo

Respecto a las figuras 2.14 y 2.15 se aprecia que a mayor *margen operativo* menor nivel de riesgo, por tanto, menor probabilidad de riesgo alto, mientras que a menor *rentabilidad* menor nivel de riesgo asociado y probabilidad baja de riesgo alto.

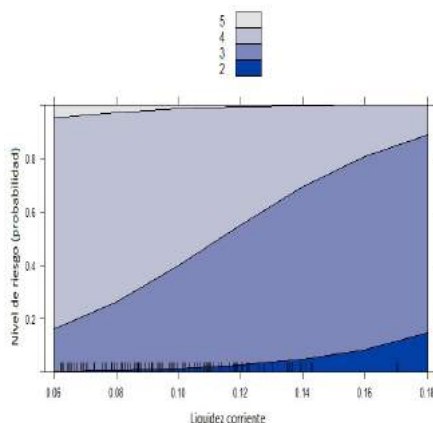


Figura 2.16: X_4 : liquidez corriente vs. nivel de riesgo

Finalmente, para la variable *liquidez corriente* se observa que a valores altos el nivel de riesgo asociado es principalmente medio y medio alto.

2.4 Entidad financiera X_5

Para esta entidad se estimó un modelo *logit* con variable respuesta de tipo binaria. A continuación se presenta los resultados obtenidos de la estimación del modelo:

<i>Variable dependiente:</i>	
Nivel de riesgo	
AA	299.005** (116.186)
CC	-15.414** (7.777)
LC	-100.628*** (34.353)
TI	409.059** (191.381)
MO	17.542* (9.621)
Constante	-289.937*** (99.444)
Observaciones	114
Log Likelihood	-19.053
Akaike Inf. Crit.	50.107
<i>Nota:</i>	*p<0.1; **p<0.05; ***p<0.01

Cuadro 2.13: X_5 : modelo *logit* con variable dependiente binaria

En el caso de la entidad X_5 las variables *apalancamiento del activo*, *tasa de interés* y *margen operativo* tienen una relación positiva respecto a la probabilidad de riesgo alto. Mientras *liquidez corriente* y *cobertura de cartera* tienen una relación negativa respecto a la variable respuesta.

A continuación se presenta la tabla cruzada para validar la capacidad de predicción del modelo:

	0	1
0	0.666	0.026
1	0.061	0.245

Cuadro 2.14: Validación cruzada: nivel de riesgo CAMEL vs. nivel de riesgo predicho por el modelo

En 91 % de los casos predijo correctamente, con lo que el error de predicción es de 9 %.

Indicador	Valor
Log-likelihood fitted model	-19.05
Log-likelihood intercept-only	-70.30
G ² ⁶	102.49
McFadden	0.73
Maximum likelihood	0.59
Cragg y Uhler	0.84

Cuadro 2.15: Pseudo R^2

Para esta entidad los indicadores de bondad de ajuste del modelo *logit* estimado arrojan valores pseudo- R^2 iguales a 0.73 y 0.84 según Mc Fadden y Cragg y Uhler respectivamente. Esto implica que el modelo ajusta bastante bien el conjunto de datos. Además, de la tabla cruzada se obtuvo que el porcentaje de aciertos de clasificación fue de 91 %.

En el siguiente cuadro se resume las estadísticas de las predicciones del modelo estimado:

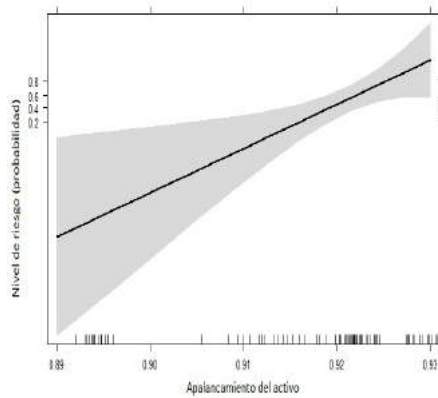
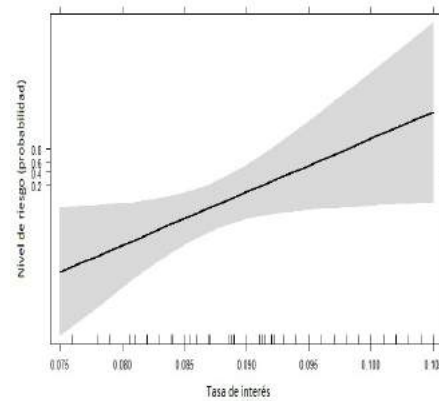
⁶Menos dos veces la diferencia en el logaritmo de las verosimilitudes

Riesgo alto		
1	Mínimo.:	0,000
2	1er cuartil:	0,000
3	Mediana:	0,021
4	Media:	0,036
5	3er cuartil:	0,070
6	Máximo :	0,101

Cuadro 2.16: X_5 : resumen de las probabilidades de riesgo alto

En las estadísticas descriptivas que se muestran en el cuadro anterior se observa que la probabilidad media de riesgo alto es de 4 %, con un máximo de 10 %.

A continuación se muestra los gráficos bivariantes entre las variables explicativas y la variable respuesta asociada a las probabilidades de riesgo alto estimadas por el modelo:

Figura 2.17: X_5 : apalancamiento del activo vs. nivel de riesgoFigura 2.18: X_5 : tasa de interés vs. nivel de riesgo

Según las figuras 2.17 y 2.18 se observa que existe una relación positiva entre el nivel de riesgo y las variables *tasa de interés* y *apalancamiento del activo*, a medida que incrementan los valores de estas dos variables la probabilidad de riesgo alto también aumenta.

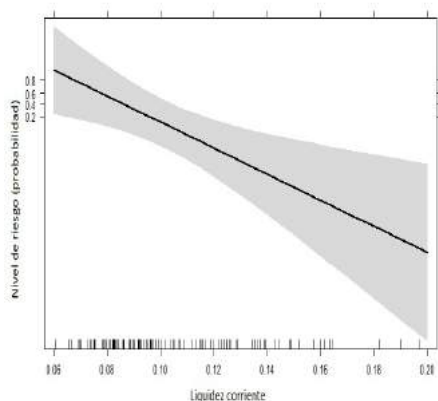


Figura 2.19: X_5 : liquidez corriente vs. nivel de riesgo

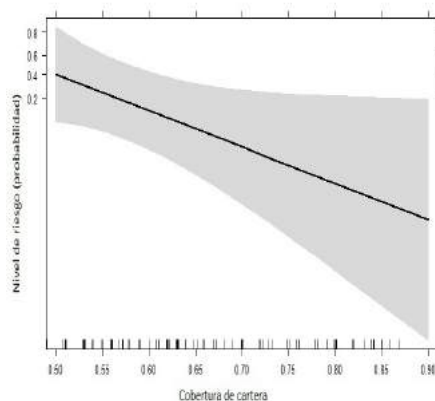


Figura 2.20: X_5 : cobertura de la cartera vs. nivel de riesgo

La relación para estas variables es negativa respecto a los niveles de riesgo, es decir, se observa que a mayores niveles de *liquidez corriente* y *cobertura de cartera* la probabilidad de riesgo alto es más baja.

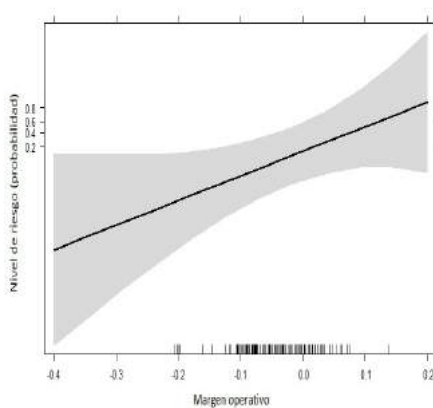


Figura 2.21: X_5 : margen operativo vs. nivel de riesgo

Finalmente, para la entidad X_5 se evidencia una relación positiva entre la variable *margen operativo* y los niveles de riesgo, incrementos en el *margen operativo* están relacionados con aumentos en los niveles de riesgo para esta entidad.

2.5 Entidad financiera X_6

Para esta entidad se estimó un modelo *logit* con variable respuesta de tipo multinomial. A continuación se presenta los resultados obtenidos de la estimación del modelo:

<i>Variable dependiente:</i>	
Nivel de riesgo	
IDEAC	-524.486*** (97.135)
R	-23.198*** (6.277)
TI	110.683*** (41.343)
Observaciones	114
<i>Nota:</i>	*p<0.1; **p<0.05; ***p<0.01

Cuadro 2.17: X_6 : modelo *logit* con variable dependiente multinomial

Una vez empleada la metodología *stepwise*, las variables que resultaron significativas son la *variación del Índice de Actividad Económica Coyuntural*, *rentabilidad*, y la *tasa de interés*. Las dos primeras con una relación negativa respecto a la probabilidad de riesgo alto. Mientras que la *tasa de interés* evidencia una relación positiva.

A continuación se presenta la tabla cruzada para validar la capacidad de predicción del modelo:

	3	4	5
3	0.10	0.04	0.00
4	0.03	0.52	0.06
5	0.00	0.09	0.16

Cuadro 2.18: Validación cruzada: nivel de riesgo CAMEL vs. nivel de riesgo predicho por el modelo

El error de predicción para este modelo es de 22%.

Indicador	Valor
Log-likelihood fitted model	-65.74
Log-likelihood intercept-only	-106.30
G ² ⁷	81.13
McFadden	0.38
Maximum likelihood	0.51
Cragg y Uhler	0.60

Cuadro 2.19: Pseudo R^2

El cuadro 2.19 arroja indicadores de 0.38 y 0.60 de acuerdo a los criterios de Mc Fadden y Cragg y Uhler respectivamente. Con lo cual la bondad de ajuste es buena, lo que también se corrobora con la validación a través de la tabla cruzada, donde el porcentaje de aciertos de clasificación fue de 78%.

⁷Menos dos veces la diferencia en el logaritmo de las verosimilitudes

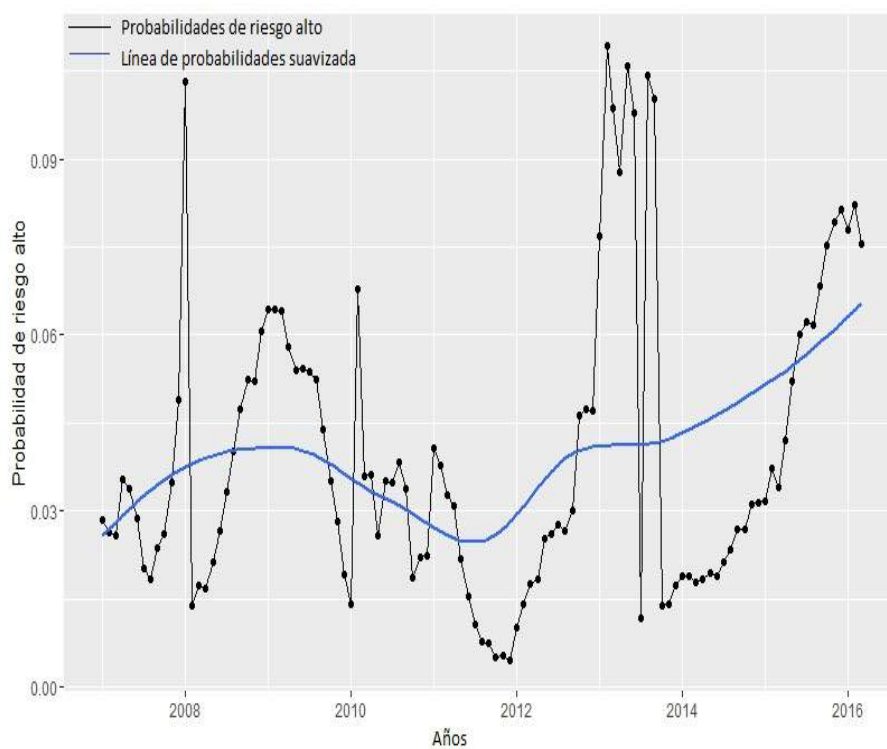


Figura 2.22: X_6 : probabilidades de riesgo alto

En la figura 2.22 se observa que las probabilidades de riesgo alto presentan volatilidad a lo largo de todo el período de estudio, alcanzando niveles superiores a 10 % y probabilidades muy bajas cercanas a cero. La tendencia de la serie es claramente creciente desde finales de 2013.

Riesgo alto		
1	Mínimo:	0,005
2	1er cuartil:	0,023
3	Mediana:	0,031
4	Media:	0,034
5	3er cuartil:	0,047
6	Máximo :	0,106

Cuadro 2.20: X_6 : resumen de las probabilidades de riesgo alto

La probabilidad media es de 3 % con un máximo de 11 %.

APÉNDICE 2. ESTIMACIÓN DE LOS MODELOS LOGIT PARA LAS ENTIDADES FINANCIERAS
106

A continuación se muestra los gráficos bivariantes entre las variables explicativas y la variable respuesta asociada a las probabilidades de riesgo alto estimadas por el modelo:

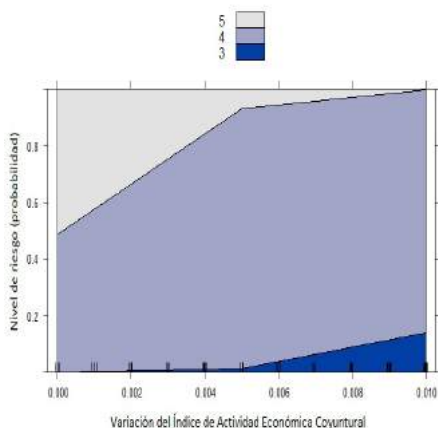


Figura 2.23: X_6 : variación del Índice de Actividad Económica Coyuntural vs. nivel de riesgo

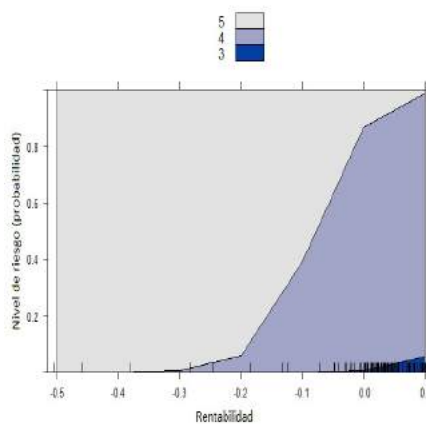
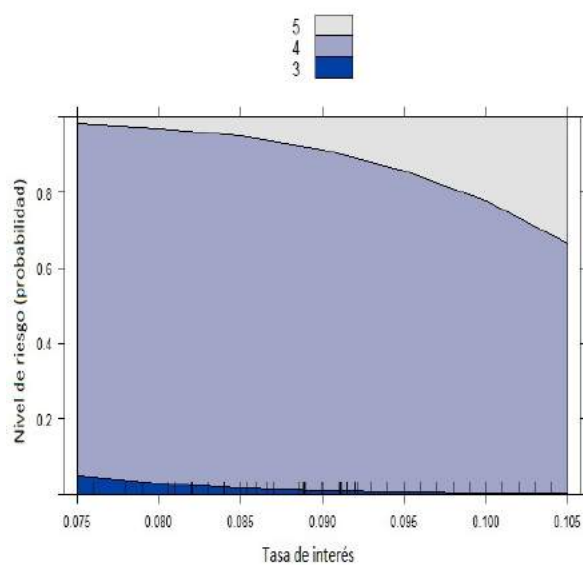


Figura 2.24: X_6 : rentabilidad vs. nivel de riesgo

En la figura 2.23 se muestra que a mayor valor de la variable independiente, el nivel de riesgo asociado es más bajo. Por otro lado, en la figura 2.24 se observa un comportamiento similar. Sin embargo, se observa un nivel de riesgo alto asociado a los niveles bajos y medios de la variable *rentabilidad*, a partir de aquí el riesgo es menor.

Figura 2.25: X_6 : tasa de interés vs. nivel de riesgo

En el caso de la figura 2.25 se observa que para casi todos los niveles de *tasa de interés* el nivel de riesgo asociado es medio alto. Sin embargo, a niveles altos de este indicador el riesgo aumenta.

2.6 Entidad financiera X_7

Para esta entidad se estimó un modelo *logit* con variable respuesta de tipo multinomial. A continuación se presenta los resultados obtenidos de la estimación del modelo:

<i>Variable dependiente:</i>	
Nivel de riesgo	
CC	-6.788*** (1.341)
AA	46.390*** (10.685)
IDEAC	-160.382** (67.696)
Observaciones	114
<i>Nota:</i>	*p<0.1; **p<0.05; ***p<0.01

Cuadro 2.21: X_7 : modelo *logit* con variable dependiente multinomial

Para la entidad X_7 la variable que muestra una relación positiva es el *apalancamiento del activo*. Mientras que *cobertura de cartera* y *variación del Índice de Actividad Económica Coyuntural* tienen una relación negativa con la probabilidad de riesgo alto.

A continuación se presenta la tabla cruzada para validar la capacidad de predicción del modelo:

	2	3	4	5
2	0.13	0.04	0.00	0.00
3	0.04	0.05	0.09	0.00
4	0.00	0.10	0.34	0.05
5	0.00	0.00	0.03	0.13

Cuadro 2.22: Validación cruzada: nivel de riesgo CAMEL vs. nivel de riesgo predicho por el modelo

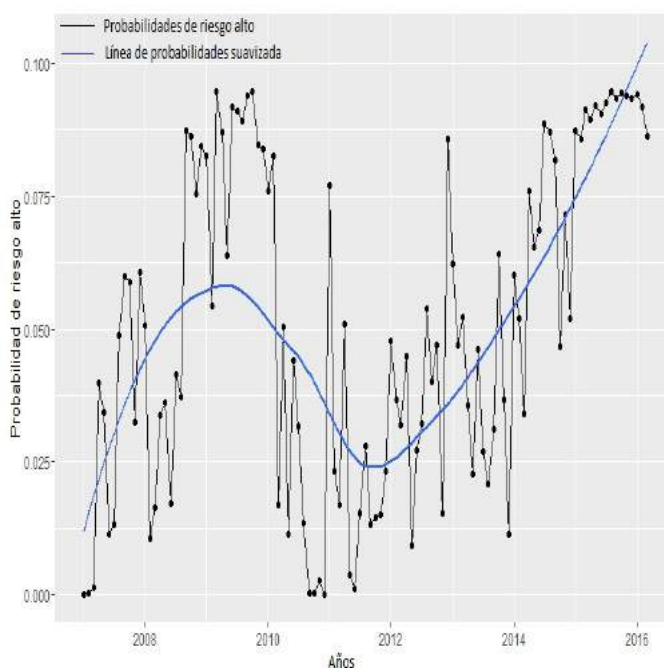
El error de predicción de este modelo es de 30 %.

⁸Menos dos veces la diferencia en el logaritmo de las verosimilitudes

Indicador	Valor
Log-likelihood fitted model	-71.87
Log-likelihood intercept-only	-144.60
G^2	145.45
McFadden	0.50
Maximum likelihood	0.72
Cragg y Uhler	0.78

Cuadro 2.23: Pseudo R^2

Para este modelo, el cuadro 2.23 muestra valores pseudo- R^2 bastante altos. El indicador de Mc Fadden es de 0.50, mientras el de Cragg y Uhler es igual a 0.78. La capacidad predictiva es ligeramente alta según la validación cruzada, ya que el porcentaje de aciertos en la clasificación de los niveles de riesgo es de 70%. En general el modelo resulta útil para fines tanto explicativos como predictivos.

Figura 2.26: X_7 : probabilidades de riesgo alto

En este caso se observa que las probabilidades de riesgo alto tienen una

gran varianza para todo el período de análisis. Desde el año 2007 al 2010 se observa una tendencia creciente, a partir de aquí caen las probabilidades, para volver a mantener la tendencia creciente desde el año 2012 en adelante.

A continuación se presenta las estadísticas más relevantes de las predicciones del modelo:

Riesgo alto		
1	Mínimo:	$2,94e^{-5}$
2	1er cuartil:	0,023
3	Mediana:	0,049
4	Media:	0,051
5	3er cuartil:	0,085
6	Máximo :	0,095

Cuadro 2.24: X_7 : resumen de las probabilidades de riesgo alto

La probabilidad media de riesgo alto es de 5%, la probabilidad del percentil 75 es de 0.09 y un máximo de 10%.

A continuación se muestra los gráficos bivariantes entre las variables explicativas y la variable respuesta asociada a las probabilidades de riesgo alto estimadas por el modelo:

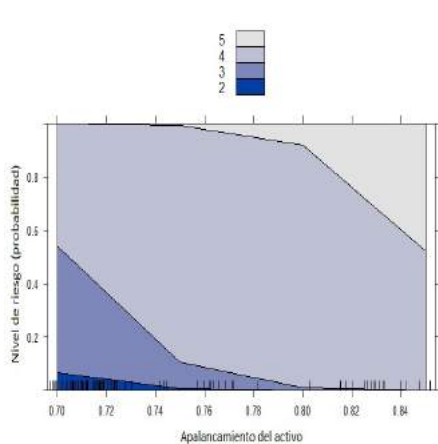


Figura 2.27: X_7 : apalancamiento del activo vs. nivel de riesgo

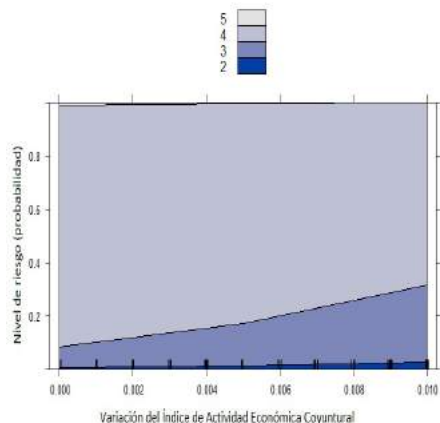


Figura 2.28: X_7 : variación del Índice de Actividad Económica Coyuntural vs. nivel de riesgo

Según la figura 2.27 a medida que el valor de este indicador es más alto, el riesgo asociado también lo es. Mientras que en el caso de la figura 2.28, la *variación del Índice de Actividad Económica Coyuntural* no muestra una relación muy evidente respecto a la probabilidad de riesgo alto, como se puede apreciar, prácticamente para todos los niveles de este indicador el riesgo asociado es medio alto.

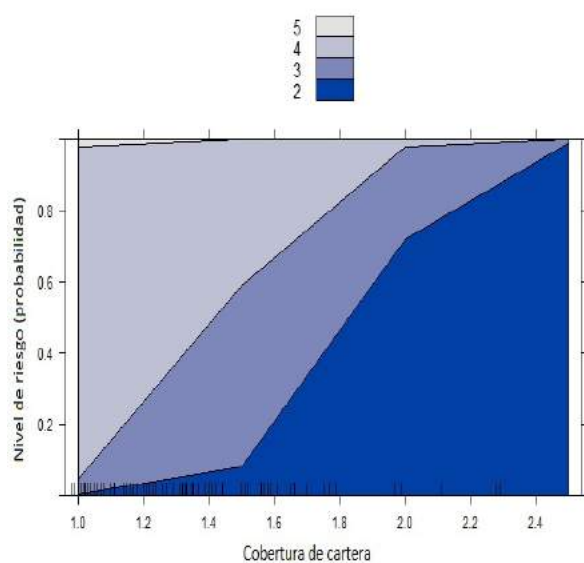


Figura 2.29: X_7 : cobertura de cartera vs. nivel de riesgo

La figura 2.29 muestra diferencias en los niveles de riesgo asociados a los valores de *cobertura de cartera*, se observa que a mayor cobertura menor es el riesgo asociado y por tanto la probabilidad.

2.7 Entidad financiera X_8

Para esta entidad se estimó un modelo *logit* con variable respuesta de tipo binaria.

A continuación se presenta los resultados obtenidos de la estimación del modelo:

Variable dependiente:	
Nivel de riesgo	
CC	-7.351*** (2.117)
PTAC	-51.665*** (13.979)
R	-15.032*** (5.763)
IDEAC	152.139 (99.781)
Constante	13.787*** (3.250)
Observaciones	
	114
Log Likelihood	
	-19.006
Akaike Inf. Crit.	
	48.012
Nota:	
	*p<0.1; **p<0.05; ***p<0.01

Cuadro 2.25: X_8 : modelo *logit* con variable dependiente binaria

Para esta entidad las variables estadísticamente significativas que muestran una relación negativa con la variable respuesta son: *cobertura de cartera*, y *patrimonio técnico constituido sobre activos y contingentes ponderados por riesgo y rentabilidad*. Por tanto, incrementos de estas variables están relacionados con disminuciones de la probabilidad de riesgo alto. Mientras *variación del Índice de Actividad Económica Coyuntural* muestra una relación positiva con la probabilidad de riesgo alto. Sin embargo, no es estadísticamente significativa.

En la siguiente tabla cruzada se valida la capacidad de predicción del modelo:

	0	1
0	0.60	0.04
1	0.01	0.35

Cuadro 2.26: Validación cruzada: nivel de riesgo CAMEL vs. nivel de riesgo predicho por el modelo

El error de predicción es de 5% para X_8 .

Indicador	Valor
Log-likelihood fitted model	-19.00
Log-likelihood intercept-only	-75.02
G2 ⁹	112.04
McFadden	0.75
Maximum likelihood	0.63
Cragg y Uhler	0.85

Cuadro 2.27: Pseudo R^2

Para esta entidad los indicadores de bondad de ajuste del modelo *logit* estimado muestran un pseudo- R^2 de Mc Fadden igual a 0.75, mientras el valor de Cragg y Uhler es de 0.85, esto quiere decir que el modelo predice bastante bien las probabilidades de riesgo alto.

A continuación se resume las estadísticas más relevantes de la serie de predicciones:

⁹Menos dos veces la diferencia en el logaritmo de las verosimilitudes

Riesgo alto		
1	Mínimo:	0,000
2	1er cuartil:	$2,30e^{-7}$
3	Mediana:	0,017
4	Media:	0,039
5	3er cuartil:	0,096
6	Máximo :	0,112

Cuadro 2.28: X_8 : resumen de las probabilidades de riesgo alto

Las probabilidades fluctúan entre 0 % y 11.2 % durante el período de análisis.

Los siguientes gráficos bivariantes muestran la relación entre las variables explicativas y la variable respuesta asociada a las probabilidades de riesgo alto estimadas por el modelo:

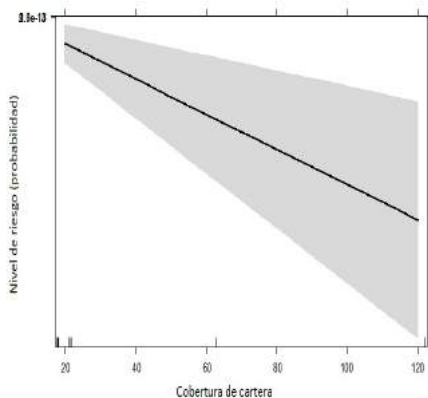


Figura 2.30: X_8 : cobertura de cartera vs. nivel de riesgo

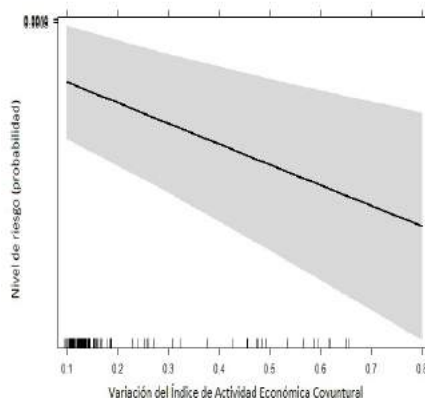


Figura 2.31: X_8 : patrimonio técnico constituido / activos y contingentes ponderados por riesgo vs. nivel de riesgo

Estas dos variables muestran la misma relación con la variable respuesta, a mayores ratios de *cobertura de cartera*, y *patrimonio técnico sobre activos y contingentes ponderados por riesgo*, la probabilidad de riesgo alto es menor.

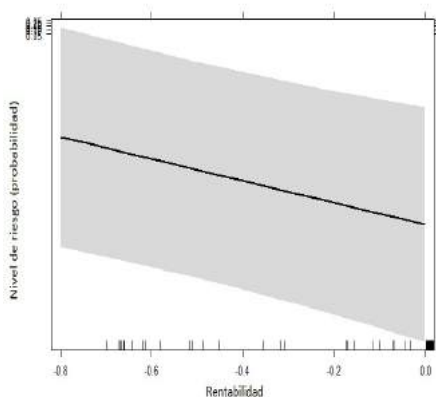


Figura 2.32: X_8 : rentabilidad vs. nivel de riesgo

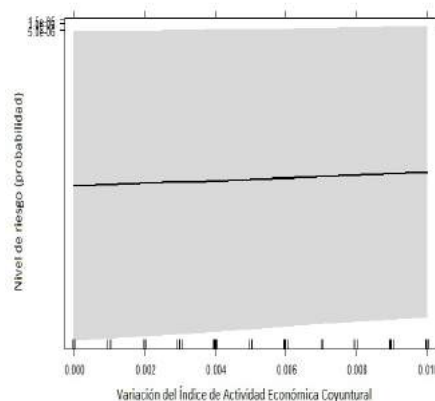


Figura 2.33: X_8 : variación del Índice de Actividad Económica Coyuntural vs. nivel de riesgo

Finalmente, para las variables *rentabilidad*, y *variación del Índice de Actividad Económica Coyuntural*, se observa que el nivel de riesgo prácticamente es el mismo para el rango de valores de estos indicadores, principalmente para la *variación del Índice de Actividad Económica Coyuntural*.

2.8 Provisiones bancos

En el cuadro 2.29 se muestra el total de depósitos cubiertos para cada entidad financiera (en adelante, EFI) y el valor de la provisión que se ha considerado de acuerdo a su nivel de riesgo. Dicho nivel de riesgo se calculó como el promedio de la calificación *CAMEL* para el período comprendido entre enero de 2007 hasta junio de 2016¹⁰.

El porcentaje considerado para la provisión se calculó como el promedio del rango de pérdida esperada para cada categoría según se presenta en el cuadro 2.17.

El total provisionado asciende a 242,428 miles de US\$, lo que corresponde a 3.33% de los depósitos cubiertos y 0.98% del total de depósitos asegurados de los bancos. Mientras representa 3.16% del total de depósitos

¹⁰Este período corresponde al mismo que se consideró para la estimación de los modelos econométricos.

APÉNDICE 2. ESTIMACIÓN DE LOS MODELOS LOGIT PARA LAS
116 ENTIDADES FINANCIERAS

cubiertos y 0.92 % del total de depósitos asegurados del Sector Financiero Privado.

Bancos ^a	Total depósitos ^b cubiertos	Nivel de riesgo	Provisión ^b
EFI 1	2,637,286	Bajo	79,119
EFI 2	825,364	Bajo	24,761
EFI 3	789,798	Bajo	23,694
EFI 4	780,193	Bajo	23,406
EFI 5	497,857	Bajo	14,936
EFI 6	406,837	Bajo	12,205
EFI 7	392,352	Bajo	11,771
EFI 8	199,811	Bajo	5,994
EFI 9	194,213	Bajo	5,826
EFI 10	124,701	Bajo	3,741
EFI 11	102,371	Medio Bajo	12,796
EFI 12	65,655	Medio Bajo	8,207
EFI 13	41,860	Medio Bajo	5,232
EFI 14	15,890	Medio Bajo	2,258
EFI 15	8,541	Medio	3,374
EFI 16	7,330	Medio	2,895
EFI 17	4,628	Medio	1,828
EFI 18	12,826	Bajo	385
TOTAL			242,428

^aConstan todas las entidades a excepción de las consideradas en el análisis previo.

^bLos valores están en miles de US\$

Cuadro 2.29: Provisiones bancos

2.9 Provisiones mutualistas

Respecto a las mutualistas, el cuadro 2.30 presenta valores para una sola mutualista, entidad que por su nivel de riesgo fue excluida del análisis previo.

El monto asciende a 7,210 en miles de US\$, valor que representa 2.59 % del total de depósitos cubiertos de mutualistas y 1.24 % de sus depósitos

asegurados. Lo que corresponde a 0.09 % del total de depósitos cubiertos y 0.03 % del total de depósitos asegurados del Sector Financiero Privado.

Mutualistas	Total depósitos ^a cubiertos	Nivel de riesgo	Provisión ^b
EFI 19	57,680	Medio Bajo	7,210
TOTAL			7,210

^aLos valores están en miles de US\$

Cuadro 2.30: Provisiones mutualistas

2.10 Provisiones sociedades financieras

Finalmente, se incluye dentro de las provisiones las sociedades financieras que no forman parte del análisis econométrico previo.

El monto asciende a US\$ 23,359 miles, lo cual corresponde a 18.70 % de los depósitos cubiertos y 2.41 % de los depósitos asegurados de las sociedades financieras. Mientras que representa 0.30 % de los depósitos cubiertos y 0.09 % del total de depósitos asegurados del Sector Financiero Privado.

Sociedades Financieras ^a	Total depósitos ^b cubiertos	Nivel de riesgo	Provisión ^b
EFI 20	93,218	Medio Bajo	11,652
EFI 21	21,764	Medio	8,597
EFI 22	6,181	Medio	2,441
EFI 23	1,414	Medio Bajo	177
EFI 24	1,227	Medio	485
EFI 25	248	Bajo	7
TOTAL			23,359

^aConstan todas las entidades a excepción de las consideradas en el análisis previo.

^bLos valores están en miles de US\$

Cuadro 2.31: Provisiones sociedades financieras

Determinación de los niveles óptimos del Seguro de Depósitos del Ecuador a partir de su función de distribución de pérdidas esperadas y de la estimación de las probabilidades de riesgo alto de las entidades financieras

Autor: Andrea Bayancela Espinel

Coordinación y edición: Eugenio Paladines y David Villamar

Corporación del Seguro de Depósitos,
Fondo de Liquidez y Fondo de Seguros Privados

Diciembre 2016

Determination of the target levels of the Deposit Insurance in Ecuador from its expected loss distribution function and the estimation of the probabilities of high risk rating of the financial entities

Resumen

El presente estudio determina los niveles óptimos del Seguro de Depósitos del Sector Financiero Privado en el Ecuador, con datos desde enero de 2007 hasta junio de 2016, a través de modelos de riesgo crediticio para la estimación tanto de las probabilidades de que las entidades financieras pasen a una calificación de riesgo alto de liquidación forzosa (incumplimiento), como de la función de distribución de pérdidas esperadas, medidas de riesgo y provisiones del Seguro de Depósitos.

Los modelos de riesgo crediticio utilizan indicadores macroeconómicos y propios de las entidades financieras para estimar las probabilidades de las variaciones en las calificaciones de riesgo de las entidades financieras. Del análisis de los indicadores macroeconómicos se desprende que incrementos en la tasa de interés activa están relacionados con una mayor probabilidad de que una entidad financiera pase a una calificación de riesgo alto. Mientras que de los indicadores propios de las entidades, se observa que aumentos en *rentabilidad*, *liquidez corriente*, y *cobertura de cartera* se asocian con una menor probabilidad de que una entidad financiera pase a una calificación de riesgo alto.

A partir de las probabilidades obtenidas de que una entidad financiera pase a una calificación de riesgo alto, se estimó la función de distribución de pérdidas esperadas del Seguro de Depósitos. De esta última, se determinó que los niveles óptimos del mismo equivalen a una pérdida esperada de 4 % de los depósitos asegurados y 12 % de los depósitos cubiertos; mientras que, si se consideran los valores extremos, dichos niveles de pérdida van desde 8 % de los depósitos asegurados y 28 % de los depósitos cubiertos con 90 % de confianza; hasta 17 % de los depósitos asegurados y 71 % de los depósitos cubiertos al 99 % de confianza.

Con el objetivo de mantener un enfoque más conservador, a los valores de pérdida esperados se debería añadir el valor de provisiones correspondiente a las entidades financieras con menor nivel de riesgo. Dicho valor representa 1.03 % de los depósitos asegurados y 3.56 % de los depósitos cubiertos.

Palabras clave: seguro de depósitos, versión asintótica modelo Vasicek, probabilidad de riesgo de incumplimiento, función de distribución de pérdidas, Value at Risk, Tail Value at Risk.

Código JEL: C15, C53, G21, G28, G33, F34.

Abstract

This study determines the target levels for the Deposit Insurance of the Private Financial Sector in Ecuador, with data from January 2007 to June 2016, through credit risk models for estimating the likelihood that financial institutions move to a high-risk rating of compulsory liquidation (default), and its distribution function of expected losses, risk measures and provisions of the Deposit Insurance.

Credit risk models use macroeconomic and financial entities' indicators to estimate probabilities of variations in the entities' risk ratings. Analysis of macroeconomic indicators shows how increases in the interest rates are related to a greater probability that a financial entity moves to a high-risk rating, whereas the entities' indicators show how increases in profitability, current liquidity and portfolio coverage ratio are associated with a lower probability that a financial entity moves to a high-risk rating.

Based on the calculated probabilities that a financial entity moves to a high-risk rating, the distribution function of expected losses of the Deposit Insurance was estimated. Using this function, it was determined that the optimum levels of the Deposit Insurance are equal to an expected loss of 4% of preferred deposits and 12% of covered deposits; whereas if the extreme values of the loss distribution are considered, these optimum levels range from 28% of covered deposits and 8% of preferred deposits with 90% of confidence; to 71% of covered deposits and 17% of preferred deposits with 99% confidence.

To take a more conservative approach, to the expected loss values should be added the value of the provisions regarding the financial entities with lower risk ratings. That value represents 1.03% of preferred deposits and 3.56% of covered deposits.

Keywords: deposit insurance, asymptotic version of Vasicek model, default probability, loss distribution function, Value at Risk, Tail Value at Risk.

Código JEL: C15, C53, G21, G28, G33, F34.

Índice general

Resumen	2
Índice general	4
Índice de figuras	6
Índice de cuadros	7
1 Marco teórico	17
1.1 Modelos de riesgo crediticio	17
1.2 Modelo asintótico unifactorial de Vasicek	20
1.3 Modelos de variable dependiente binaria <i>logit</i>	22
1.4 Bondad de ajuste y especificación del modelo	26
1.5 Metodología <i>stepwise</i>	28
1.6 Estimación de la pérdida esperada	29
1.7 Generación de la función de distribución empírica	30
1.8 Medidas de riesgo: valor en riesgo y valor en riesgo condicional	32
1.8.1 Valor en riesgo (VaR)	32
1.8.2 Valor en riesgo condicional (TVaR)	32
1.9 Metodologías para la definición del porcentaje de recuperación	33
2 Datos y resultados	35
2.1 Definición de variables	36
2.1.1 Variables independientes	36
2.1.1.1 Variables financieras	36
2.1.1.2 Variables macroeconómicas	37
2.1.2 Variable dependiente	38
2.2 Análisis exploratorio de los datos	39
2.3 Estimación de las probabilidades de riesgo alto	43
2.3.1 Matrices de correlación de las entidades financieras .	50
2.3.1.1 Matriz de correlaciones: Sector Financiero Privado	50

2.3.1.2	Matriz de correlaciones: bancos	51
2.3.1.3	Matriz de correlaciones: mutualistas	51
2.3.1.4	Matriz de correlaciones: sociedades financieras	52
2.4	Definición de los parámetros para la estimación de la función de pérdidas	52
2.5	Construcción del proceso de simulación	53
2.6	Funciones probabilísticas de las pérdidas	54
2.6.1	Función de densidad de las pérdidas	54
2.6.2	Función de distribución empírica de las pérdidas	55
2.7	Tasa de transición a calificación de riesgo alto (incumplimiento)	56
2.8	Determinación del valor en riesgo y valor en riesgo condicional	59
2.9	Determinación del porcentaje de recuperación de cartera	60
2.10	Provisión de pérdidas para entidades con menor nivel de riesgo	60
2.11	Resumen de resultados	62
3	Conclusiones y recomendaciones	65
3.1	Conclusiones	65
3.2	Recomendaciones	69
	Referencias	71
Apéndice 1	Marco teórico	73
1.1	Teorema: probabilidades de incumplimiento	73
1.2	Portafolio homogéneo grande (<i>LHP</i>)	74
1.2.1	Lema: tasa de incumplimiento	75
1.2.1.1	Prueba del lema	75
1.2.1.2	Teorema sobre la tasa de incumplimiento	76
1.2.1.3	Prueba del teorema	76
1.3	Modelos espacio-estado: filtro de Kalman	76
Apéndice 2	Estimación de los modelos <i>logit</i> para las entidades financieras	83
2.1	Entidad financiera X_2	83
2.2	Entidad financiera X_3	88
2.3	Entidad financiera X_4	93
2.4	Entidad financiera X_5	98
2.5	Entidad financiera X_6	103
2.6	Entidad financiera X_7	107
2.7	Entidad financiera X_8	111
2.8	Provisiones bancos	115
2.9	Provisiones mutualistas	116

2.10 Provisiones sociedades financieras	117
---	-----

Índice de figuras

1.1 Función de densidad de las pérdidas	20
1.2 Funciones de densidad de las distribuciones normal y logística	23
1.3 Funciones de distribución acumulada normal y logística	23
2.1 Morosidad, improductividad y cobertura de cartera	42
2.2 Liquidez, margen operativo y tasa de interés	42
2.3 Rentabilidad del capital, variación del precio del petróleo, y patrimonio técnico constituido / activos y contingentes ponderados por riesgo	42
2.4 Índice de Actividad Económica Coyuntural, rentabilidad y apalancamiento del activo	42
2.5 Probabilidades de riesgo alto	47
2.6 <i>Effect plot</i> : improductividad de activos vs. nivel de riesgo	48
2.7 <i>Effect plot</i> : cobertura de cartera vs. nivel de riesgo	49
2.8 <i>Effect plot</i> : rentabilidad vs. nivel de riesgo	49
2.9 <i>Effect plot</i> : PTAC vs. nivel de riesgo	50
2.10 Función de densidad empírica de las pérdidas	54
2.11 Función de distribución empírica de las pérdidas	55
2.12 Función de densidad empírica de las tasas de riesgo alto	57
2.13 Función de distribución empírica de las tasas de riesgo alto	58
2.1 X_2 : probabilidades de riesgo alto	86
2.2 X_2 : apalancamiento del activo vs. nivel de riesgo	87
2.3 X_2 : rentabilidad del capital vs. nivel de riesgo	87
2.4 X_2 : variación del Índice de Actividad Económica Coyuntural vs. nivel de riesgo	88
2.5 X_3 : probabilidades de riesgo alto	91
2.6 X_3 : improductividad activos vs. nivel de riesgo	92
2.7 X_3 : liquidez corriente vs. nivel de riesgo	92
2.8 X_3 : variación del Índice de Actividad Económica Coyuntural vs. nivel de riesgo	92

2.9	X_3 : rentabilidad vs. nivel de riesgo	92
2.10	X_3 : tasa de interés vs. nivel de riesgo	93
2.11	X_4 : probabilidades de riesgo alto	96
2.12	X_4 : cobertura de la cartera vs. nivel de riesgo	97
2.13	X_4 : improductividad de activos vs. nivel de riesgo	97
2.14	X_4 : margen operativo vs. nivel de riesgo	97
2.15	X_4 : rentabilidad vs. nivel de riesgo	97
2.16	X_4 : liquidez corriente vs. nivel de riesgo	98
2.17	X_5 : apalancamiento del activo vs. nivel de riesgo	101
2.18	X_5 : tasa de interés vs. nivel de riesgo	101
2.19	X_5 : liquidez corriente vs. nivel de riesgo	102
2.20	X_5 : cobertura de la cartera vs. nivel de riesgo	102
2.21	X_5 : margen operativo vs. nivel de riesgo	102
2.22	X_6 : probabilidades de riesgo alto	105
2.23	X_6 : variación del Índice de Actividad Económica Coyuntural vs. nivel de riesgo	106
2.24	X_6 : rentabilidad vs. nivel de riesgo	106
2.25	X_6 : tasa de interés vs. nivel de riesgo	107
2.26	X_7 : probabilidades de riesgo alto	109
2.27	X_7 : apalancamiento del activo vs. nivel de riesgo	110
2.28	X_7 : variación del Índice de Actividad Económica Coyuntural vs. nivel de riesgo	110
2.29	X_7 : cobertura de cartera vs. nivel de riesgo	111
2.30	X_8 : cobertura de cartera vs. nivel de riesgo	114
2.31	X_8 : patrimonio técnico constituido / activos y contingentes pon- derados por riesgo vs. nivel de riesgo	114
2.32	X_8 : rentabilidad vs. nivel de riesgo	115
2.33	X_8 : variación del Índice de Actividad Económica Coyuntural vs. nivel de riesgo	115

Índice de cuadros

2.1	VARIABLES explicativas financieras	36
2.2	Niveles de riesgo	39
2.3	Resumen de estadísticas - parte 1	39

2.4	Resumen de estadísticas - parte 2	40
2.5	Matriz de correlaciones de las variables independientes	41
2.6	X_1 : modelo <i>logit</i> con variable dependiente multinomial	44
2.7	Estadísticos de las variables independientes	45
2.8	Validación cruzada: nivel de riesgo CAMEL vs. nivel de riesgo predicho por el modelo	46
2.9	Pseudo R^2	46
2.10	X_1 : resumen de las probabilidades de riesgo alto	48
2.11	Matriz de correlaciones: entidades con probabilidad de riesgo alto	50
2.12	Matriz de correlaciones: bancos	51
2.13	Matriz de correlaciones: mutualistas	51
2.14	Matriz de correlaciones: sociedades financieras	52
2.15	Resumen de las tasas de riesgo alto	56
2.16	Medidas de riesgo VaR y TVaR	59
2.17	Provisiones de pérdidas según el nivel de riesgo	61
2.18	Provisiones por tipo de entidad	62
2.1	X_2 : modelo <i>logit</i> con variable dependiente binaria	84
2.2	Validación cruzada: nivel de riesgo CAMEL vs. nivel de riesgo predicho por el modelo	85
2.3	Pseudo R^2	85
2.4	X_2 : resumen de las probabilidades de riesgo alto	86
2.5	X_3 : modelo <i>logit</i> con variable dependiente multinomial	89
2.6	Validación cruzada: nivel de riesgo CAMEL vs. nivel de riesgo predicho por el modelo	90
2.7	Pseudo R^2	90
2.8	X_3 : resumen de las probabilidades de riesgo alto	91
2.9	X_4 : modelo <i>logit</i> con variable dependiente multinomial	94
2.10	Validación cruzada: nivel de riesgo CAMEL vs. nivel de riesgo predicho por el modelo	95
2.11	Pseudo R^2	95
2.12	X_4 : resumen de las probabilidades de riesgo alto	96
2.13	X_5 : modelo <i>logit</i> con variable dependiente binaria	99
2.14	Validación cruzada: nivel de riesgo CAMEL vs. nivel de riesgo predicho por el modelo	100
2.15	Pseudo R^2	100
2.16	X_5 : resumen de las probabilidades de riesgo alto	101
2.17	X_6 : modelo <i>logit</i> con variable dependiente multinomial	103
2.18	Validación cruzada: nivel de riesgo CAMEL vs. nivel de riesgo predicho por el modelo	104
2.19	Pseudo R^2	104

2.20	X_6 : resumen de las probabilidades de riesgo alto	105
2.21	X_7 : modelo <i>logit</i> con variable dependiente multinomial	108
2.22	Validación cruzada: nivel de riesgo CAMEL vs. nivel de riesgo predicho por el modelo	108
2.23	Pseudo R^2	109
2.24	X_7 : resumen de las probabilidades de riesgo alto	110
2.25	X_8 : modelo <i>logit</i> con variable dependiente binaria	112
2.26	Validación cruzada: nivel de riesgo CAMEL vs. nivel de riesgo predicho por el modelo	113
2.27	Pseudo R^2	113
2.28	X_8 : resumen de las probabilidades de riesgo alto	114
2.29	Provisiones bancos	116
2.30	Provisiones mutualistas	117
2.31	Provisiones sociedades financieras	117

Glosario

AIC: criterio de información de Akaike (*Akaike Information Criterion*).

Algoritmo EM: algoritmo esperanza-maximización (*Expected Maximization Algorithm*).

BIC: criterio de información de Bayes (*Bayesian Information Criterion*).

BLUE: mejor estimador lineal insesgado (*Best Linear Unbiased Estimator*).

CAMEL rating: calificación de riesgo calculada con base en 4 aspectos: adecuación de capital, activos, gestión, rentabilidad y liquidez. (*Capital adequacy, Assets, Management Capability, Earnings, Liquidity*).

Depósitos asegurados: depósitos a la vista o a plazo fijo efectuados por personas naturales o jurídicas en las entidades financieras de los Sectores Financieros Privado y Popular y Solidario, bajo la forma de cuentas corrientes, de ahorros, depósitos a plazo fijo u otras modalidades legalmente aceptadas, de acuerdo con la ley vigente en el Ecuador para el pago del seguro.

Depósitos cubiertos: depósitos que el Seguro de Depósitos cubre para cada persona natural o jurídica, los cuales son diferenciados por cada uno de los sectores financieros asegurados; para el caso de las entidades financieras privadas, los depósitos cubiertos son iguales a dos veces la fracción básica exenta vigente del impuesto a la renta, pero en ningún caso inferior a US\$ 32.000,00 (treinta y dos mil dólares de los Estados Unidos de América).

IADI: Asociación Internacional de Aseguradores de Depósitos (*International Association of Deposit Insurers*).

LGD: pérdida dado el incumplimiento (*Loss Given Default*).

LHP: portafolio homogéneo grande (*Large Homogeneous Portfolio*).

PTAC: patrimonio técnico constituido respecto a activos y contingentes ponderados por riesgo.

ML: máxima verosimilitud (*Maximum Likelihood*) .

Metodología stepwise: metodología paso a paso.

OLS: mínimos cuadrados ordinarios (*Ordinary least squares*).

RR: tasa de recuperación (*Recovery Rate*).

RSS: suma de cuadrados de residuos (*Residual Sum of Squares*).

Seguro de Depósitos: sistema establecido para proteger (de forma limitada) los depósitos efectuados por personas naturales y jurídicas en las entidades de los Sectores Financieros Privado y Popular y Solidario autorizadas por los respectivos organismos de control. El Seguro de Depósitos se compone de los siguientes recursos, en lo que a cada sector financiero corresponde: las contribuciones que realizan las entidades de los Sectores Financieros Privado y Popular y Solidario, de conformidad con lo previsto en la ley vigente; el rendimiento de las inversiones y las utilidades líquidas de cada ejercicio anual; las donaciones que reciba; los provenientes de préstamos o líneas contingentes obtenidos para el financiamiento de sus actividades; los provenientes de préstamos entre los fideicomisos del Seguro de Depósitos; y, los pagos derivados de la liquidación forzosa de una entidad financiera.

TVaR: valor en riesgo condicional (*Tail Value at Risk*).

VaR: valor en riesgo (*Value at Risk*).

Introducción

Un tema ampliamente discutido en la actualidad es la gestión de los niveles de riesgo que enfrentan las entidades financieras, para lo cual las autoridades emplean mecanismos de supervisión y control, con el objetivo de procurar la estabilidad y sostenibilidad de los sistemas financieros. Para controlar los costos asociados con posibles problemas financieros, las autoridades implementan redes de seguridad financiera que buscan reducir la probabilidad de que ocurra la liquidación forzosa de una entidad financiera y minimizar su impacto en la economía.

Los principales mecanismos de seguridad financiera que estas redes suelen incluir son: supervisión financiera, fondo de liquidez, fortalecimiento patrimonial, resolución financiera, seguro de depósitos, entre otros.

En este sentido, el establecimiento de un Seguro de Depósitos explícito ha llegado a ser considerado como uno de los pilares fundamentales sobre el cual debe construirse un sistema financiero moderno. Es así que el número de jurisdicciones que ofrecen un Seguro de Depósitos explícito ha aumentado rápidamente, pasando de 12 en 1974 a 113 en el año 2014, según datos de la Asociación Internacional de Aseguradores de Depósitos (IADI). Hoy en día la mayoría de los países de la Organización para la Cooperación y el Desarrollo Económico (OCDE) y un creciente número de países emergentes cuentan con un sistema de Seguro de Depósitos explícito.

Dada la popularidad que ha tenido el Seguro de Depósitos explícito, se podría pensar que el diseño y la operación de un sistema eficiente es relativamente sencillo. Sin embargo, en la práctica los administradores de este mecanismo de seguridad financiera se enfrentan a una constante disyuntiva que hace su tarea bastante compleja ya que no solo deben procurar la estabilidad y sostenibilidad financiera, sino también evitar que las entidades financieras asuman mayores riesgos y otras prácticas que podrían resultar imprudentes.

Uno de los principales desafíos a los que se enfrentan los gestores de los seguros de depósitos consiste en establecer un adecuado equilibrio entre la prevención de una crisis y el control respecto al nivel de riesgo asumido por parte de las entidades financieras.

Con el fin de que las entidades financieras mantengan niveles de riesgo adecuados se han desarrollado diversos instrumentos como: requerimientos de calidad de activos, niveles mínimos de patrimonio, control de pasivos, entre otros.

Además, un enfoque alternativo más reciente consiste en utilizar modelos de riesgo de crédito en la administración de los Seguros de Depósitos, con el objeto de probar si estos modelos permiten cuantificar el riesgo de una cartera compuesta por las entidades financieras que asegura.

Este nuevo instrumento sugiere que “si el seguro de depósitos es visto como un servicio que utilizan las entidades financieras, la pregunta es ¿cuál es el precio que debe tener este servicio?”. (Maccario, Sironi, & Zazzara, 2003, p. 2)

Una posible respuesta consiste en que el precio debe reflejar el riesgo que la entidad financiera representa para el sistema de Seguro de Depósitos. Este enfoque de pérdida esperada de los precios se ajusta con las mejores prácticas que se han desarrollado en la industria financiera en los últimos años, esencialmente los principios de Basilea.

Adicionalmente conviene destacar que en el caso de Ecuador, la crisis financiera acaecida en 1999 develaría la necesidad de un esquema de supervisión estricto por parte de los organismos de control sobre las entidades financieras del país, con el objetivo de proteger los recursos de los depositantes a través de la regulación de las operaciones de colocación de créditos. En tal sentido a partir del año 2007 se diseña una nueva arquitectura financiera nacional, desarrollándose varios mecanismos de seguridad financiera, que junto a un oportuno y efectivo control estatal de carácter preventivo (basado en un enfoque de riesgos), permita definir el conjunto de procedimientos y medidas para resolver situaciones de liquidación de entidades financieras, beneficiando primordialmente a los depositantes.

Como parte de la nueva arquitectura financiera nacional, en el año 2008 se expide la Ley de Creación de la Red de Seguridad Financiera, la cual parte del esquema de dolarización adoptado a finales del siglo XX y se ampara en el artículo 308 de la Constitución de la República, el cual determina que las actividades financieras son un servicio de orden público que deberá tener como finalidad fundamental la preservación de los depósitos y atender los requerimientos de financiamiento para la consecución de los objetivos de desarrollo del país.

Dicha ley dictó la creación de la Corporación del Seguro de Depósitos; que más adelante, en el año 2014, con la expedición del Código Orgánico Monetario y Financiero, se transformó en la Corporación del Seguro de Depósitos, Fondo de Liquidez y Fondo de Seguros Privados (COSEDE), siendo ahora esta institución la encargada de la administración general de algunos de los principales mecanismos de seguridad financiera.

Considerando lo expuesto, la presente investigación tiene como objetivo determinar los niveles óptimos del patrimonio del Seguro de Depósitos del Sector Financiero Privado. Para ello se emplea modelos de riesgo crediticio que estiman las probabilidades de que las entidades financieras presenten un riesgo alto de liquidación forzosa (incumplimiento) y la función de distribución de pérdidas esperadas del Seguro de Depósitos.

El presente trabajo se divide en tres capítulos. El capítulo 1 describe el marco teórico mediante la explicación de los modelos econométricos y las técnicas estadísticas utilizadas. El capítulo 2 describe la base de datos, define las variables y presenta los resultados de estimación, tanto de los modelos como de la pérdida esperada y medidas de riesgo. Finalmente, el capítulo 3 presenta las conclusiones del estudio y recomendaciones para futuros trabajos de investigación.

Capítulo 1

Marco teórico

1.1 Modelos de riesgo crediticio

En esta sección se describe las diferentes metodologías de los modelos de riesgo de crédito y su aplicabilidad dentro de un sistema de Seguro de Depósitos.

El riesgo de crédito se puede definir como la probabilidad de que una entidad financiera incurra en pérdidas debido a la falta de pago de un deudor, sea de un préstamo u otra línea de financiamiento. Este evento se produce cuando el deudor es incapaz de cumplir con su obligación legal de acuerdo con el contrato de financiamiento. Por ejemplo, falta de pago de títulos, quiebra corporativa, falta de pago de la tarjeta de crédito o la ejecución de una hipoteca.

Los modelos de medición del riesgo de crédito se pueden agrupar en dos categorías principales, cuya diferencia radica en la disponibilidad de los datos: mediciones directas de crédito y precios observados en el mercado de crédito.

Las dos categorías de modelos de medición de riesgo de crédito pueden ser caracterizadas por los tipos de probabilidad de incumplimiento, siendo el primer tipo de probabilidad el observado y el otro implícito. Dicho de otro modo, la primera categoría de riesgo de crédito corresponde a las observaciones directas de los valores de incumplimiento (riesgo explícito). Mientras que la segunda categoría de riesgo de crédito se refiere a la probabilidad de incumplimiento tomada de datos financieros e información macroeconómica disponible (riesgo implícito). (Zhang, 2009)

Actualmente la literatura académica de riesgo de crédito se ha inclinado mayoritariamente al estudio de las probabilidades de incumplimiento, es decir, al riesgo implícito de las entidades financieras a través de la estimación de modelos estadísticos.

El principal resultado de un modelo de riesgo de crédito es la función de densidad de las pérdidas esperadas¹ del portafolio. A partir del análisis de dicha distribución de pérdidas, una entidad financiera puede estimar la pérdida esperada de su cartera de crédito. La misma es igual a la media (incondicional) de la distribución de pérdidas, que representa la cantidad que la entidad puede esperar perder dentro de un período específico de tiempo (generalmente un año).

Las entidades financieras aplican modelos de riesgo de crédito para evaluar el capital económico necesario para hacer frente a los riesgos asociados con sus carteras de crédito. En este sentido, las provisiones para pérdidas de crédito deben cubrir las pérdidas esperadas. Del mismo modo se puede aplicar esta lógica a las exposiciones de crédito de los sistemas de Seguro de Depósitos explícitos. En el caso de un portafolio de crédito bancario, la pérdida se produce cuando un prestatario no paga, mientras que en el caso de un Seguro de Depósitos la pérdida se da cuando una entidad asegurada cae en liquidación forzosa, generándose un pago por parte del Seguro de Depósitos.

Siguiendo este enfoque, un Seguro de Depósitos puede ser comparado con una entidad financiera en la medición de riesgo de incumplimiento de sus contrapartes. Mientras para las entidades financieras las contrapartes son los prestatarios, para las entidades de Seguro de Depósitos son las propias entidades financieras, con exposiciones equivalentes a los depósitos cubiertos. El Seguro de Depósitos puede entenderse como el administrador de una cartera de exposiciones de crédito contingente para las entidades que asegura. (Maccario et al., 2003, p. 4-6)

Por otra parte, las distribuciones empíricas de las pérdidas de las carteras de crédito son generalmente asimétricas con mayor densidad en los valores más bajos de pérdidas. Es decir, la probabilidad de incurrir en pérdidas

¹Una función de densidad es una función matemática de una variable aleatoria continua cuyo área bajo la curva representa la probabilidad de que el valor de esa variable caiga en un intervalo determinado.

extremadamente altas tiende a cero. Sin embargo, es mayor que la probabilidad que se obtendría de una distribución normal estándar en estos niveles.

Por tanto, al enfrentarse el Seguro de Depósitos a una situación similar a la distribución de pérdidas de una cartera de crédito, la probabilidad de ocurrencia de pérdidas con valores bajos es relativamente alta debido al incumplimiento de entidades financieras pequeñas, mientras que a medida que las pérdidas son mayores se reduce la probabilidad asintóticamente hacia cero, lo cual se daría en caso de que el incumplimiento sea por parte de las entidades financieras grandes.

En esta investigación, por incumplimiento se entiende la probabilidad de que una entidad financiera pase a una calificación de riesgo alto, esta definición difiere del concepto tradicional de incumplimiento que se entiende como la probabilidad de liquidación de una entidad.

La probabilidad de incumplimiento es mayor a aquella que se generaría de una distribución normal, lo que se ilustra en la figura 1.1, en la cual se observa la asimetría de la distribución hacia la derecha, siendo la cola evidentemente más densa que la de una distribución normal estándar². Con lo cual, en la medida en que la distribución sea asimétrica, determinar los cuantiles extremos de la distribución se torna un aspecto clave.

²Datos simulados de una distribución empírica.

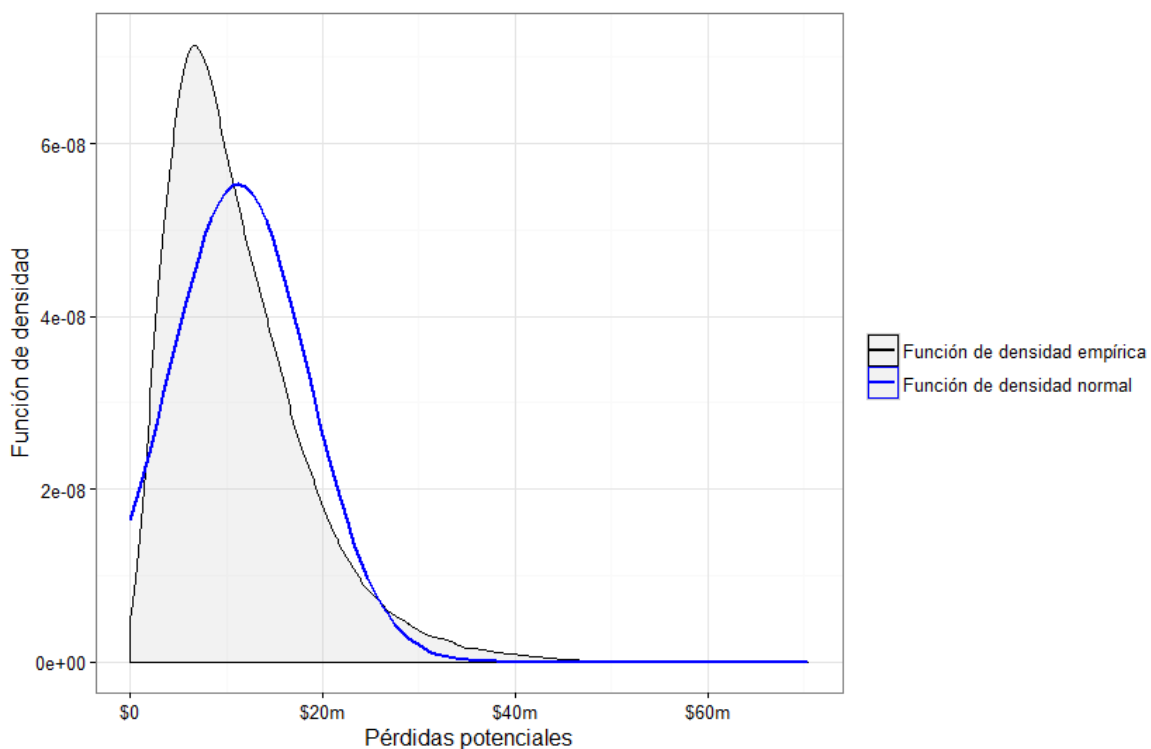


Figura 1.1: Función de densidad de las pérdidas

El objetivo del presente trabajo es aproximar la distribución de pérdidas del Seguro de Depósitos para el caso de Ecuador, a partir de la estimación de las probabilidades de que las entidades pasen a riesgo alto y de la simulación de las pérdidas esperadas del Seguro de Depósitos empleando la versión asintótica del modelo de Vasicek.

Esta distribución resulta útil para identificar el nivel apropiado de los recursos necesarios para hacer frente a las posibles pérdidas potenciales del Seguro de Depósitos.

1.2 Modelo asintótico unifactorial de Vasicek

El modelo asintótico unifactorial de Vasicek resulta atractivo dentro del campo de la estadística actuarial, en razón de que su objetivo es ofrecer una forma práctica para obtener la función de distribución tanto de la morosidad de la cartera como de las tasas de pérdida de una entidad

financiera. (Analytics, 2014)

Para obtener la tasa de morosidad de la cartera no solo es necesario conocer las probabilidades de incumplimiento $(p_{1,t}, \dots, p_{N,t})$, sino también la estructura de correlación de estas probabilidades, dado que la única variable aleatoria que afecta el estado de cada entidad financiera n en el tiempo t (incumplimiento o no incumplimiento) es $X_{n,t}$, la cual representa las pérdidas. De esta manera, la estructura de correlación entre las probabilidades de incumplimiento de las entidades se debe introducir a través de la variable aleatoria normal $X_{1,t}, \dots, X_{N,t}$. Se asume que el coeficiente de correlación para cada par de variables aleatorias $X_{n,t}$ y $X_{m,t}$ es $\rho_{n,m,t}$. (Vasicek, 1987)

Por otro lado, se puede pensar que existe un factor aleatorio o fuente de incertidumbre que afecta a todas las entidades financieras exactamente de la misma manera. Así, se pueden expresar las variables aleatorias $X_{n,t}, \dots, X_{m,t}$ de la siguiente forma:

$$X_{n,t} = \sqrt{\rho_t} * Y_t + \sqrt{1 - \rho_t} * \epsilon_{n,t} \quad (1.1)$$

Para cada $n = 1, \dots, N$, donde Y_t y $\epsilon_{1,t}, \dots, \epsilon_{N,t}$ son independientes e idénticamente distribuidas como una distribución normal estándar.

La fórmula 1.1 se puede interpretar tal que la variable aleatoria que representa las pérdidas ($X_{n,t}$) se expresa como la suma de dos factores de riesgo, por un lado el riesgo común (Y_t) y por otro el riesgo idiosincrático ($\epsilon_{n,t}$) que es particular para cada entidad financiera.

Por tanto:

- Y_t es el riesgo sistemático común que afecta a todas las entidades financieras (estado de la economía).
- $\epsilon_{n,t}$ es el factor idiosincrático, particular para cada entidad financiera.
- ρ_t definido como $Corr(X_{n,t}, X_{m,t})$ es el coeficiente de correlación entre cada par de entidades financieras, que para este modelo se asume igual en todo n y m .

1.3 Modelos de variable dependiente binaria *logit*

Los modelos econométricos *logit* y *probit* son apropiados para modelar una variable dependiente dicotómica o multinomial (más de dos categorías de respuesta) en relación a un conjunto de variables categóricas o numéricas (escalas de intervalo o ratios). La necesidad de utilizar estos modelos surge en virtud de que la aproximación clásica a través del modelo de regresión lineal presenta varios problemas en cuanto a la estimación.³

Cuando la variable dependiente es categórica, la estimación por *Mínimos Cuadrados Ordinarios* ya no produce el mejor estimador lineal insesgado (*BLUE*). Por consiguiente, se han desarrollado varios modelos de regresión para trabajar con este tipo de variables. (Leo, 2014)

Los modelos *logit* y *probit* pretenden resolver estos problemas mediante el ajuste de una función no lineal, mientras que las estimaciones con el modelo clásico de regresión se obtienen a partir de la siguiente función:

$$\hat{Y} = X\beta + \mu \quad (1.2)$$

Además estos modelos asumen que la relación entre las variables no es lineal y por tanto la función adopta otra forma.

$$\hat{Y} = f(X\beta + \mu) \quad (1.3)$$

Dicha función no es una línea recta como en el caso del modelo de regresión clásico, sino una curva en forma de *S* que: 1) respeta los límites de la variable dependiente; 2) admite diferentes tasas de variación en los extremos inferior y superior; y 3) elimina el problema de heterocedasticidad, asumiendo especificación correcta de las variables independientes.

La principal diferencia entre los modelos *logit* y *probit* consiste en la forma de la distribución de los errores (perturbaciones) que asumen ambos modelos. Por un lado, el modelo *logit* utiliza la función de distribución acumulada de la distribución logística, cuya función de densidad viene dada

³La forma funcional de la técnica conocida como *OLS* puede conducir a predicciones que no correspondan a probabilidades y por tanto se encuentren fuera del intervalo $[0, 1]$, asumiendo el mismo efecto marginal sobre la variable dependiente. Además, esta técnica suele revelar problemas de heterocedasticidad con este tipo de variables.

por $\lambda(\epsilon) = e^\epsilon / (1 + e^\epsilon)^2$, con media 0 y varianza $\frac{\pi^2}{3}$. Mientras que el modelo *probit* utiliza la función de distribución acumulada de la distribución normal estándar, $\Phi(\epsilon) = \frac{1}{\sqrt{2\pi}} e^{-\frac{\epsilon^2}{2}}$, con varianza 1. (Park, 2009)

Ambas funciones transforman los valores de tal manera que se encuentren en el intervalo $[0,1]$. Por lo tanto, cualquiera que sea el valor de $X\beta + \mu$, se puede transformar para obtener una probabilidad.

Como se puede apreciar en las siguientes figuras, un punto importante a destacar sobre la forma de estas funciones consiste en que la función de densidad (PDF) de la distribución normal estándar tiene un pico más alto y colas más delgadas que la distribución de probabilidad logística estándar. En la función de distribución logística estándar el pico es más bajo, pero las colas son más densas (*heavy tails*). Como resultado, la función de densidad acumulada (CDF) de la distribución normal estándar es más pronunciada para los valores centrales que la CDF de la distribución logística estándar y se acerca más rápidamente a cero en ambas colas. (Fox, 2010)

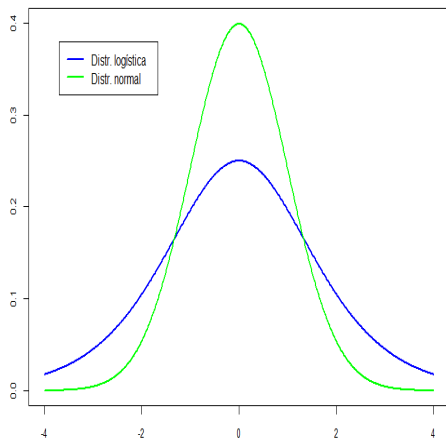


Figura 1.2: Funciones de densidad de las distribuciones normal y logística

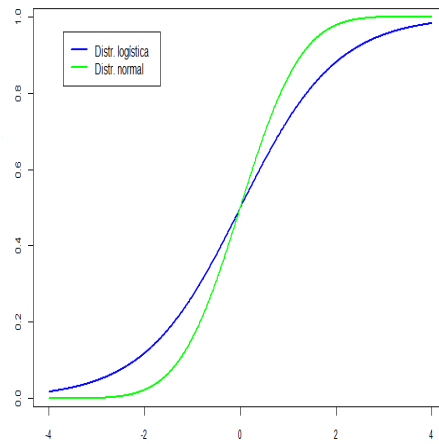


Figura 1.3: Funciones de distribución acumulada normal y logística

Los dos modelos producen diferentes estimaciones de parámetros. En los modelos de respuesta discreta, las estimaciones de un modelo *logit* son aproximadamente $\pi/\sqrt{3}$ veces más grandes que las del modelo *probit*. Sin embargo, los estimadores de las variables independientes tienen los mismos efectos estandarizados en ambos modelos.

La elección entre modelos *logit* o *probit* está relacionada con la estimación y la familiaridad con cada modelo, más que a aspectos teóricos o de interpretación. En general, los modelos *logit* alcanzan convergencia bastante bien. Actualmente, la potencia de cálculo ha mejorado y se han desarrollado nuevos algoritmos que hacen que la importancia de esta decisión sea cada vez menor.⁴

Estimación del modelo econométrico *logit*

Asumiendo que se tienen k ensayos de Bernoulli, y_1, \dots, y_k , y que la observación i -ésima puede ser tratada como la realización de una variable aleatoria Y_i . Se supone que Y_i tiene una distribución binomial. (Rodríguez, 2007)

$$Y_i \sim \text{Binomial}(n_i, \pi_i) \quad (1.4)$$

Con denominador binomial n_i , probabilidad de éxito π_i y datos individuales $n_i = 1$ para todo i . Además, se supone que el *logit* de la probabilidad subyacente π_i es una función lineal de los predictores.

$$\text{logit}(\pi_i) = x_i' \beta, \quad (1.5)$$

Donde x_i es un vector de covariables y β es el vector de coeficientes de la regresión.

El modelo definido en las ecuaciones 1.4 y 1.5 es un modelo lineal generalizado con respuesta de tipo *binomial* y función de enlace *logit*.

Los coeficientes de regresión pueden interpretarse de la misma manera que en los modelos lineales, considerando que la variable dependiente es un *logit* y no una media. Por lo tanto, β_j representa el cambio en el *logit* de la probabilidad asociado a una unidad de cambio en la j -ésima variable predictora, manteniendo todo lo demás constante. Así, también es posible expresar los resultados en la escala *logit*, lo cual resulta más sencillo de interpretar.

Si se obtiene la exponencial de la ecuación 1.5 se tendría que la probabilidad para la i -ésima unidad está dada por:

⁴Los modelos *probit* y *logit* son estimados por el método de Máxima Verosimilitud (*ML*), en lugar del método de *Mínimos Cuadrados Ordinarios* (*MCO*).

$$\frac{\pi_i}{1 - \pi_i} = \exp(x'_i\beta) \quad (1.6)$$

Esta expresión define un modelo multiplicativo de las probabilidades. Por ejemplo, si se tuviera que cambiar el predictor i -ésimo en una unidad manteniendo las demás variables constantes, se multiplicarían las probabilidades por $\exp(\beta_j)$.

Bajo el supuesto de que el predictor $x'_i\beta$ es lineal y se incrementa x_j en una unidad, se obtiene $x'_i\beta + \beta_j$. Si se calcula su exponencial, esta expresión sería $\exp(x'_i\beta) * \exp(\beta_j)$. Por lo tanto, el coeficiente exponencial $\exp(\beta_j)$ es lo que se conoce como *odds ratio* o razón de probabilidades. Transformar los resultados en un efecto multiplicativo de los *odds ratio*, es a menudo bastante útil ya que los resultados se expresan en una escala más sencilla.

Despejando la probabilidad π_i en el modelo *logit* de la ecuación 1.6 se obtiene lo siguiente:

$$\pi_i = \frac{\exp(x'_i\beta)}{1 + \exp(x'_i\beta)} \quad (1.7)$$

Mientras que el lado izquierdo de la igualdad está expresado en términos de probabilidad, el lado derecho es una función no lineal de los predictores. De esta forma, se expresa el efecto en la probabilidad dado un aumento de una unidad en la variable predictora. Es posible obtener una respuesta aproximada tomando la derivada con respecto a x_j , que evidentemente tiene sentido únicamente para variables predictoras continuas.

Para ello se utiliza la regla del cociente:

$$\frac{d\pi_i}{dx_{ij}} = \beta_j\pi_i(1 - \pi_i) \quad (1.8)$$

Por lo tanto, el efecto del predictor j -ésimo en la probabilidad π_i depende del coeficiente β_j .

1.4 Bondad de ajuste y especificación del modelo

En este apartado se estudia los principales criterios estadísticos empleados para medir tanto la bondad de ajuste como la especificación de los modelos vistos en la sección 1.1.

Para los modelos econométricos basados en variables dependientes categóricas existen dos enfoques para determinar si el modelo estimado se ajusta al conjunto de datos. El primer enfoque consiste en emplear medidas que permitan evaluar la capacidad predictiva del modelo. El segundo enfoque prueba si el modelo debe ser más complejo (por ejemplo, adición o diferenciación de variables explicativas), si la especificación es correcta o se necesita incluir interacciones para representar satisfactoriamente los datos.

La capacidad predictiva del modelo se evalúa a través de una tabla cruzada que contrasta los valores observados de la variable respuesta, con los valores estimados por el modelo para cada categoría.

Respecto a la bondad de ajuste, en modelos *OLS* es posible obtener una medida conocida como R^2 , la cual representa: el porcentaje de la varianza explicada por los regresores del modelo, el cuadrado del coeficiente de correlación entre los datos ajustados y predichos y la transformación del estadístico F . Sin embargo, al estimar modelos con variable respuesta de tipo categórica, por ejemplo los modelos *logit*, no existe una medida estadística equivalente al R^2 , ya que las estimaciones de los modelos de regresión logística se realizan por máxima verosimilitud a través de un proceso iterativo, es decir, no se calculan con el objetivo de minimizar la varianza de los errores. Sin embargo, hay una serie de medidas conocidas como pseudo- R^2 que se han desarrollado con el fin de evaluar el ajuste de estos modelos.

Entre ellas se encuentran:

- R^2 **McFadden**: se conoce como el índice de la razón de verosimilitud. Esta medida compara la verosimilitud del modelo incluyendo todos los predictores versus el modelo considerando únicamente el intercepto. Para un modelo de regresión logística, el logaritmo de la verosimilitud es siempre negativo (el logaritmo de una probabilidad será un valor igual o menor a cero). Si el modelo que incluye tanto el intercepto como las variables regresoras no predice la variable respuesta mejor que el modelo que considera únicamente el intercepto, el logaritmo

del modelo completo no será mucho más alto que el logaritmo del modelo sin variables independientes, con lo cual el ratio entre estos dos será de aproximadamente uno. De esta manera, según la forma de cálculo ($1 -$ razón de verosimilitud), el resultado será cercano a cero. Este indicador tiende a ser más pequeño que el R^2 , valores entre 0.2 y 0.4 se consideran satisfactorios.

$$R_{McFadden}^2 = 1 - \frac{L_\beta}{L_\alpha}$$

- **R^2 Verosimilitud máxima:** expresa el ajuste del modelo como una transformación de la razón de verosimilitud χ^2 , de forma análoga a la de R^2 en regresión por mínimos cuadrados ordinarios, que se expresa como una transformación del estadístico F .

$$R_{VM}^2 = 1 - \left(\frac{L_\alpha}{L_\beta} \right)^{2/N}$$

$$R_{VM}^2 = 1 - \exp(-G^2/N), \text{ donde:}$$

$$G^2 = -2\ln(L_\alpha/L_\beta)$$

- **R^2 Cragg-Uhler:** Cragg-Uhler propusieron un índice relativo que puede alcanzar el valor de uno. Si el modelo predice perfectamente y tiene probabilidad de uno, el indicador de Cragg y Uhler será igual a uno, si la razón de verosimilitudes es uno, el valor del R^2 será igual a cero.

$$R_{Cragg-Uhler}^2 = \frac{1 - (L_\alpha/L_\beta)^{2/N}}{1 - L_\alpha^{2/N}}$$

Se debe tener presente que no existe un criterio uniforme relacionado con el valor de R^2 y pseudo- R^2 a partir del cual se puede considerar un “buen ajuste”, eso depende principalmente del objetivo del investigador y del tipo de modelo, el cual puede ser explicativo o predictivo. En el presente estudio interesan tanto las relaciones entre las variables (explicativo) como la capacidad predictiva del modelo, con lo que valores pseudo- R^2 muy bajos no serían adecuados para estimar las probabilidades de riesgo alto. Por otro lado, al calcular cada una de estas medidas (pseudo- R^2) con base en criterios distintos, en muchas ocasiones resulta conveniente evaluar varios modelos para comparar los valores, de este modo, mientras más alto sea el valor del pseudo R^2 , mejor será la predicción del modelo respecto a la

variable respuesta.

En el presente estudio se empleó los dos enfoques, es decir, se realizó la validación cruzada del modelo para evaluar su capacidad de predicción y se calculó las medidas conocidas como pseudo- R^2 para determinar la bondad de ajuste del modelo.

1.5 Metodología *stepwise*

En este apartado se revisará el uso de la metodología *stepwise*, la misma que será empleada para la selección del modelo.

La metodología *stepwise* suele ser utilizada cuando existe un número relativamente grande de posibles covariables. Con n covariables existen 2^n modelos posibles. De esta forma, esta regresión se construye a partir de un proceso iterativo que implica la elección de las variables de predicción mediante un procedimiento automático basado en un criterio específico. Esta metodología tiene dos enfoques:

- **Selección hacia adelante**, implica iniciar sin variables en el modelo. Luego, se prueba la adición de cada variable y se va seleccionando la variable que más contribuya a mejorar el modelo. Este proceso se repite hasta que no se produzca alguna mejora significativa con base en el criterio específico.
- **Selección hacia atrás**, implica iniciar con todas las variables en el modelo. Luego, se prueba la eliminación de cada variable; la variable se elimina siempre que su omisión mejore el modelo, dicho proceso se repite hasta que no se produzca mejoría en el modelo por la eliminación de alguna variable ⁵.

Esta metodología suele utilizar el “Criterio de información Akaike” (AIC)⁶ ($AIC = -2 \ln L + 2k$), el cual es un poco más flexible respecto a la inclusión de una variable en el modelo, en comparación con el “Criterio de Información de Bayes” (BIC) ⁷ ($BIC = -2 \ln L + k \ln N$), este criterio es más

⁵En el presente documento se aplicó la regresión paso a paso utilizando ambos métodos.

⁶Sea L el valor máximo de la función de verosimilitud y k el número de parámetros estimados en el modelo.

⁷Sea L el valor máximo de la función de verosimilitud, k el número de parámetros libres que se estima y N el número de datos observados.

riguroso en la selección de una variable. En general la elección de uno u otro criterio depende de la finalidad para la cual se desarrolle el modelo.

Como se mencionó, esta metodología es ampliamente utilizada cuando se tiene un gran número de posibles variables explicativas. El procedimiento se utiliza principalmente en el análisis de regresión. Sin embargo, el enfoque básico es aplicable en muchas formas de selección del modelo.

Con esta sección concluye el marco teórico referente a la primera parte de la presente investigación, la cual corresponde a los modelos econométricos empleados para la estimación de las probabilidades de incumplimiento (riesgo alto), la segunda parte comprende la estimación de la pérdida esperada basada en la técnica de simulación.

1.6 Estimación de la pérdida esperada

En esta sección se describe la segunda parte del marco teórico del presente trabajo, que abarca el proceso de simulación sobre el cual se estimará tanto la pérdida esperada como las diversas medidas de riesgo para el Seguro de Depósitos en el plazo de un año.

En un modelo de incumplimiento (*default-mode model*) la pérdida esperada de la cartera (EL_p) es igual al producto de las exposiciones al riesgo individuales para cada una de las n entidades (EXP_i), las probabilidades de incumplimiento (p_i) de cada entidad, es decir, la proporción de deudores que podrían caer en incumplimiento dentro de un marco de tiempo dado (generalmente de un año) y las pérdidas dado el incumplimiento (LGD_i). Tal que: (Matjaž, 2014)

$$EL_p = \sum_{i=1}^n EXP_i * p_i * LGD_i \quad (1.9)$$

- EXP : la exposición del sistema de Seguro de Depósitos para cada entidad se define como el monto de sus depósitos cubiertos.
- LGD : $1 - RR$, donde RR corresponde a la tasa de recuperación, (*Recovery Rate*), la cual se puede calcular con la base de datos de recuperaciones. Otra alternativa para el cálculo de dicho porcentaje es utilizar el método establecido por el Comité de Supervisión de Basilea (2006) en el contexto de los requisitos de capital. La Comisión

propone un LGD^8 de 75 % de todos los préstamos, dicha aproximación tiene sentido dado que la mayor exposición del sistema de Seguro de Depósitos se da en el sector bancario.

- p : estas probabilidades pueden ser estimadas como se menciona en la sección 1.3, a través de modelos econométricos *logit*, en donde es necesario definir una variable respuesta categórica que represente el evento de incurrir en incumplimiento.

1.7 Generación de la función de distribución empírica

Una vez que se ha estimado las probabilidades de incumplimiento se debe construir el proceso iterativo a partir de la ecuación 1.10, correspondiente a la versión asintótica del modelo de Vasicek, la cual se expresa a continuación:

$$X_{n,t} = \sqrt{\rho_t} * Y_t + \sqrt{1 - \rho_t} * \epsilon_{n,t} \quad (1.10)$$

Donde el crédito i se asume en incumplimiento cuando la variable latente X_i toma un valor menor al límite de crédito especificado. Para efectos de la simulación se generan m iteraciones, con m valores aleatorios para la variable *Threshold*, que se define como un número aleatorio que representa el límite para determinar si la entidad cae o no en incumplimiento. (Hernández, 2013)

$$X_i < Threshold \quad (1.11)$$

El principal rol de una entidad financiera es brindar el servicio de intermediación. Cuando los activos de la entidad no permiten hacer frente a los pasivos exigibles de la misma, se le dificulta ofrecer el servicio referido y podría incurrir en incumplimiento. De esta manera, se generan suficientes eventos que simulen esta situación y así permitan aproximar la función de densidad y de distribución de las pérdidas.

⁸Según el Comité de Supervisión de Basilea, cuando no existan medidas para el cálculo del LGD se sugiere considerar 45 % en créditos preferentes y 75 % en créditos subordinados. En este caso se decidió tomar 75 %, con el objetivo de realizar el análisis bajo el escenario más conservador, es decir, aquel que asume el mayor porcentaje de pérdidas.

A nivel macroeconómico se conoce que hay un aspecto que afecta a todas las entidades financieras independientemente de su tamaño, aunque en diferente magnitud, que es el estado de la economía Y_t . Para reflejar este hecho, se define la variable Y_t como un número aleatorio proveniente de una distribución normal estándar (media 0 y desviación típica 1), tal que corresponda al estado actual de la economía. Un número alto implicaría una buena situación económica, un número bajo implicaría un mal estado económico.

Con el fin de reflejar la idiosincrasia de una entidad en particular, se define $\epsilon_{n,t}$ también como un número aleatorio específico para la entidad i , proveniente de una función de densidad normal estándar.

Para ponderar el efecto macroeconómico se multiplica Y_t por las correlaciones de la cartera (es decir, la fuerza de la relación entre la cartera de créditos) y para ponderar la “singularidad” de la entidad se multiplica por $1 - \rho$ (es decir, lo contrario de su relación con otras entidades financieras).

Para cada iteración se establece un nuevo Y_t estado económico, que todas las entidades compartirán para esa iteración. Un nuevo $\epsilon_{n,t}$ se genera para cada entidad dentro de cada iteración. Una vez que se ha calculado suficientes simulaciones los escenarios de pérdida que se obtienen finalmente convergen en la pérdida total esperada de la cartera, lo cual da lugar a la generación de la función de distribución empírica y las probabilidades asociadas para toda la gama de pérdidas. (Charpentier, 2015)

1.8 Medidas de riesgo: valor en riesgo y valor en riesgo condicional

1.8.1 Valor en riesgo (VaR)

En su forma más general, el VaR mide la pérdida potencial en el valor de un activo de riesgo o de un portafolio durante un período definido para un intervalo de confianza dado. (Damodaran, 2009)

Si bien el VaR puede ser utilizado por cualquier entidad para medir su exposición al riesgo, en general es utilizado con mayor frecuencia en el sector financiero para capturar la pérdida potencial en el valor de los portafolios negociados en el mercado durante un período especificado. (Castagna, 2009)

El VaR está definido por la siguiente fórmula:

$$VaR_\alpha = \inf\{x \mid F_L(x) \geq \alpha\} = F_L^{-1}(\alpha) \quad (1.12)$$

Esta medida de riesgo corresponde a un cuantil en la función de distribución empírica (F_L), dejando un área de cola de $1 - \alpha$ a su derecha. Para el presente trabajo este valor representa la pérdida que asumiría el Seguro de Depósitos como máximo el $(1 - \alpha)\%$ de las veces.

1.8.2 Valor en riesgo condicional (TVaR)

Como se menciona en la sección 1.8.1, en gestión del riesgo financiero, el VaR es una medida utilizada para medir el riesgo de pérdida de una cartera específica de los activos financieros. Sin embargo, una desventaja del VaR es que no aporta ninguna información sobre la severidad de las pérdidas más allá del nivel para el cual se calcula. Incluso podría suceder que el VaR de dos variables podría ser exactamente el mismo y sin embargo, una de ellas podría tener una cola más densa (más riesgo) que la otra.

Para mitigar este efecto surge una medida de riesgo conocida como el TVaR, que toma en cuenta la severidad del incumplimiento y no sólo la probabilidad de éste. En general se considera que el TVaR proporciona una mejor medida de riesgo, ya que se ve afectado en los valores extremos de la distribución y se define de la siguiente manera:

$$TVaR_\alpha(L) = \frac{1}{1-\alpha} \int_\alpha^1 VaR_\lambda(L) d\lambda = E(L \mid L > VaR_\alpha(L)) \quad (1.13)$$

El TVaR por tanto, mide la pérdida esperada promedio de una cartera en un horizonte de tiempo determinado, considerando únicamente los casos en los que las pérdidas son mayores que el VaR. En otras palabras, representa el valor promedio de las pérdidas esperadas en los peores casos. (Peng, 2008)

Como se observa en la presente sección, es posible estimar tanto las probabilidades de riesgo alto como la pérdida promedio y el monto que el Seguro de Depósitos asumiría como pérdida dado un nivel de confianza determinado.

1.9 Metodologías para la definición del porcentaje de recuperación

En los apartados anteriores se ha asumido que el valor de las pérdidas dado el incumplimiento (riesgo alto) es igual a 100%. Sin embargo, a lo largo del proceso de liquidación de las entidades se pueden dar eventuales recuperaciones de acuerdo al orden de prelación.

Al tener en cuenta los ingresos esperados de las recuperaciones se puede concluir que el asegurador de depósitos no necesariamente requiere financiar el 100% de los recursos con la recaudación de las contribuciones, puesto que una parte de esos recursos van a ser recuperados.

Considerando esto, los recursos recaudados a través de las contribuciones y que constituyen la reserva del Seguro de Depósitos (financiación ex-ante) deben ser suficientes para cubrir el porcentaje de recursos que no se espera recuperar con el proceso de liquidación de las entidades.

De forma complementaria, el porcentaje de recursos que se espera recuperar en los procesos de liquidación puede financiarse con alternativas diferentes, como líneas contingentes de crédito. Este esquema de financiamiento busca solucionar un problema de iliquidez temporal del asegurador de depósitos, asociado a un desajuste en términos de flujos de caja.

En principio, la pérdida dado el incumplimiento se puede calcular con datos históricos. Sin embargo, si el seguro de depósitos no cuenta con las bases de datos de las recuperaciones para el cálculo de la pérdida dado el incumplimiento, como se menciona en la sección 1.6, otra opción es considerar el porcentaje de las recuperaciones según los estándares planteados (25 %) por el Comité de Supervisión Bancaria de Basilea (2006), dado que en el caso de un Seguro de Depósitos su exposición es precisamente a entidades financieras.

Capítulo 2

Datos y resultados

En esta sección se definen las variables consideradas en el estudio, así como el conjunto de entidades financieras sobre el cual se analizan e interpretan los resultados obtenidos, tanto en la estimación de los modelos econométricos como en la simulación de las pérdidas para la estimación de la función de distribución empírica.

En el presente trabajo se utilizó información referente a bancos privados, sociedades financieras y mutualistas¹, con periodicidad mensual desde enero de 2007 hasta junio de 2016, con un total de 114 observaciones por cada entidad.

El conjunto de entidades considerado en el análisis incluye aquellas que han obtenido la mayor cantidad de veces una calificación de riesgo *alto* para el período de estudio, según la metodología *CAMEL*, utilizada por la Corporación del Seguro de Depósitos, Fondo de Liquidez y Fondo de Seguros Privados.

Dentro de los bancos se incluyó tres entidades definidas como X_1 , X_2 y X_3 , tres mutualistas definidas como X_4 , X_5 y X_6 , y dos sociedades financieras definidas como X_7 y X_8 .

¹Se ha considerado a las Mutualistas de Ahorro y Crédito para la Vivienda dentro del Sector Financiero Privado, puesto que según lo dispuesto en la Disposición General Vigésimo Novena del Código Orgánico Monetario y Financiero y la Resolución No. 219-2016-F expedida por la Junta de Política y Regulación Monetaria y Financiera, estas entidades deberán resolver su permanencia en el Sector Financiero Popular y Solidario en un plazo de 18 meses contados a partir del 9 de marzo de 2016. Además, estas entidades seguirán siendo reguladas por la Superintendencia de Bancos, hasta que se cumpla dicho plazo.

Para la primera parte de la investigación, referente a la estimación de las probabilidades de riesgo alto, se desarrolló un modelo de regresión logística de series de tiempo para cada entidad. De esta manera, se obtuvo la serie histórica mensual de las probabilidades de riesgo alto de cada entidad financiera.

A partir de ello se construyó la matriz de correlaciones de las probabilidades de pasar a una calificación de riesgo alto (en adelante, probabilidades de riesgo alto), ya que es parte de la información que se necesitó para la aproximación de la función de distribución de pérdidas.

2.1 Definición de variables

2.1.1 Variables independientes

El conjunto de variables independientes se puede dividir en dos grupos: 1) ratios financieros e 2) información macroeconómica del país.

2.1.1.1 Variables financieras

Las variables financieras seleccionadas en el modelo responden a los criterios de: solvencia, liquidez, calidad de activos y rentabilidad. Estas son:

Criterio	Variable	Fórmula
Calidad de activos	Morosidad de cartera (MC)	$\text{Cartera improductiva} / \text{Cartera bruta}$
	Improductividad de activos (IA)	$\text{Activos improductivos} / \text{Activo total}$
	Cobertura de cartera (CC)	$\text{Provisiones de cartera} / \text{Cartera improductiva}$
Liquidez	Liquidez corriente (LC)	$\text{Activo corriente} / \text{Pasivo corriente}$
Solvencia	Apalancamiento del activo (AA)	$\text{Pasivo total} / \text{Activo total}$
	Patrimonio técnico constituido sobre Activos y contingentes ponderados por riesgo (PTAC)	$\text{Patrimonio técnico constituido} / \text{Activos y contingentes ponderados por riesgo}$
Rentabilidad	Rentabilidad (R)	$\text{Utilidad neta} / \text{Total ingresos}$
	Rentabilidad del capital (RC)	$\text{Utilidad neta} / \text{Patrimonio promedio}$
	Margen operativo (MO)	$\text{Margen operacional} / \text{Total ingresos}$

Cuadro 2.1: Variables explicativas financieras

En este estudio se decidió trabajar con las variables en términos de ratio en lugar de porcentajes y en este sentido es importante señalar que la determinación de la escala a utilizar para las variables explicativas es arbitraria, ya que las conclusiones del modelo no se ven alteradas. En muchas ocasiones la escala se suele cambiar por razones didácticas de ilustración de los resultados obtenidos, o cuando el modelo tiene problemas para converger. Sin embargo, en general este aspecto no resulta ser un inconveniente.

Es importante señalar que si bien el grupo de variables seleccionado para el análisis comparte ciertos criterios con los utilizados en la metodología *CAMEL*, los indicadores financieros considerados en el estudio no son expresamente los empleados para el cálculo del nivel de riesgo.

Por otro lado, a diferencia de la regresión lineal, como se mencionó antes, la cual usa el método de estimación por Mínimos Cuadrados Ordinarios (MCO), en la regresión logística se emplea el método de Máxima Verosimilitud (MV) para llevar a cabo la estimación de los parámetros del modelo.

Estos modelos de regresión con variable endógena categórica, también llamados modelos de variable endógena cualitativa o modelos de elección discreta, se suelen emplear para explicar la decisión (Y) que toma un individuo de entre un número limitado de posibles opciones a partir de un conjunto de variables explicativas (X_1, X_2, \dots, X_k). Con lo cual a diferencia de la regresión lineal, estos modelos econométricos pretenden estimar la probabilidad de que ocurra un evento en particular.

2.1.1.2 Variables macroeconómicas

La inclusión de variables macroeconómicas en el análisis de riesgo ha tomado gran importancia en los últimos años. Si bien, en un año específico los factores macroeconómicos no cambian el rango de riesgo de un solo préstamo dentro de una cartera; en un marco más amplio se ha evidenciado que la omisión de datos macroeconómicos generalmente significa deshacerse de información valiosa, lo que se da especialmente cuando el modelo de calificación se emplea en la estimación de probabilidades de que una entidad pase a tener riesgo alto.

En el presente trabajo se incluyó el entorno macroeconómico con el fin de complementar el análisis financiero y evitar errores de especificación del modelo. En este sentido, luego del análisis de seis variables se seleccionó tres variables macroeconómicas que son:

- Media móvil de la variación mensual del Índice de Actividad Económica Coyuntural (IDEAC).²
- Tasa de interés activa mensual (TI).³
- Variación mensual del precio del petróleo (PP).⁴

Es importante señalar que dentro del análisis se exploró la inclusión de estas variables retrasadas un trimestre ($t - 3$), un semestre ($t - 6$) y un año ($t - 12$) considerando su significancia estadística.

2.1.2 Variable dependiente

La variable dependiente considerada en el presente estudio es el nivel de riesgo otorgado a las entidades financieras mediante la metodología *CAMEL*. Esta variable es binaria en el caso de que la serie correspondiente a los niveles de riesgo tenga únicamente dos categorías, si esta variable tiene más de dos categorías para el período de análisis se considera multinomial.

Esta metodología otorga una puntuación que va de 1 hasta 6 de acuerdo a los criterios de: solvencia, liquidez, gestión, resultados y calidad de activos. De esta forma el nivel de riesgo se asigna según el siguiente cuadro:

²En virtud de que el PIB real del Ecuador se calcula de manera trimestral, se prefiere evitar problemas de autocorrelación a través de la imputación de datos faltantes con la técnica del Filtro del Kalman.

³Banco Central del Ecuador, Publicaciones estadísticas, Sector monetario y financiero, Julio de 2016, www.bce.fin.ec/index.php/component/k2/item/754.

⁴Banco Central del Ecuador, Indicadores económicos, Julio de 2016, www.contenido.bce.fin.ec/resumen_ticker.php?ticker_value=petroleo.

Categoría	Nivel de riesgo	Calificación
1	Bajo	$0 > y \leq 2$
2	Medio bajo	$2 > y \leq 3$
3	Medio	$3 > y \leq 4$
4	Medio alto	$4 > y \leq 5$
5	Alto	$5 > y \leq 6$

Cuadro 2.2: Niveles de riesgo

Previo a la estimación de los modelos probabilísticos, a continuación se presenta la exploración de los datos para la entidad X_1 , ya que sobre ésta se describirá toda la metodología del presente estudio. Sin embargo, es necesario señalar que el análisis se realizó para todas las entidades mencionadas al inicio de este capítulo (Ver Apéndices).

2.2 Análisis exploratorio de los datos

Como se indicó anteriormente, para cada entidad se realizó el análisis exploratorio así como la estimación del modelo econométrico. Sin embargo, a manera de ejemplo se describe todo el análisis con datos de una entidad financiera. A continuación se presenta un resumen estadístico de las variables:

	Calificación de riesgo	MC ¹	CC ²	IA ³	LC ⁴	MO ⁵
Mínimo	1.72	0.04	0.30	0.11	0.13	-0.15
1er cuartil	2.97	0.06	0.48	0.18	0.18	0.00
Mediana	3.56	0.07	0.70	0.21	0.23	0.06
3er cuartil	3.46	0.08	0.65	0.21	0.26	0.06
Media	4.13	0.10	0.81	0.23	0.31	0.12
Máximo	5.21	0.14	0.97	0.35	0.54	0.21

Cuadro 2.3: Resumen de estadísticas - parte 1

Nota: ¹ Morosidad de cartera ² Cobertura de cartera

³ Improductividad de activos ⁴ Liquidez corriente

⁵ Margen operativo

Para esta entidad se observa que las calificaciones de riesgo fluctúan entre 1.72 y 5.21, con una media de 4.13 que implica un nivel de riesgo definido como *medio alto*. Por otro lado, las medias de las variables *morosidad de cartera*, *cobertura de cartera*, *improductividad de activos*, *liquidez corriente* y *margen operativo* son 0.10, 0.81, 0.23, 0.31 y 0.12 respectivamente.

Considerando que la media y la mediana de estas variables no presentan diferencias importantes, se asume que la presencia de valores extremos es baja, ya que esto indica que tanto el percentil 50 de la distribución de los datos como el valor promedio de estos, es relativamente cercano. Dicha afirmación se cumple en todos los casos, a excepción de *cobertura de cartera* y *margen operativo* en donde la media es dos veces el valor de la mediana.

	AA ¹	R ²	RC ³	PTAC ⁴	TI ⁵	PP ⁶	IDEAC ⁷
Mínimo	0.79	-0.01	-0.01	0.13	0.07	-0.28	-0.01
1er cuartil	0.83	0.02	0.01	0.15	0.08	-0.05	0.00
Mediana	0.86	0.04	0.02	0.18	0.09	0.01	0.01
3er cuartil	0.85	0.06	0.03	0.18	0.09	0.01	0.01
Media	0.88	0.10	0.04	0.19	0.09	0.06	0.01
Máximo	0.90	0.19	0.13	0.30	0.11	0.24	0.02

Cuadro 2.4: Resumen de estadísticas - parte 2

Nota: ¹ Apalancamiento activo ² Rentabilidad ³ Rentabilidad del capital ⁴ Patrimonio técnico sobre activos y contingentes ponderados por riesgo ⁵ Tasa de interés ⁶ Variación del precio del petróleo ⁷ Variación del Índice de Actividad Económica Conyuntural

Respecto al cuadro 2.4 se observa mayores diferencias entre la media y mediana, especialmente para las variables *variación del precio del petróleo*, y *rentabilidad*.

Para la variable *apalancamiento del activo* se aprecia muy poca diferencia, incluso entre la media y el valor máximo de la distribución de los datos. Mientras que, por ejemplo, la *variación del precio del petróleo* sí muestra gran diferencia entre el valor promedio (0.06) y el valor máximo (0.24), lo

que sugiere que el precio del petróleo ha tenido grandes fluctuaciones durante el período de análisis.

A continuación se presenta la matriz de correlaciones de las variables independientes para la entidad X_1 .

Variables	MC	IA	CC	LC	MO	PTAC	R	RC	PP	IDEAC	TI	AA
MC	1.00	0.66	-0.87	0.16	0.06	0.33	0.26	0.27	-0.51	-0.64	0.62	-0.50
IA		1.00	-0.33	-0.29	-0.46	-0.23	-0.32	-0.18	-0.49	-0.77	0.05	0.15
CC			1.00	-0.36	-0.22	-0.57	-0.46	-0.42	0.39	0.41	-0.75	0.72
LC				1.00	0.39	0.85	0.31	0.25	-0.03	-0.16	0.36	-0.65
MO					1.00	0.57	0.88	0.72	-0.06	0.26	0.30	-0.61
PTAC						1.00	0.61	0.47	-0.21	-0.16	0.49	-0.92
R							1.00	0.75	-0.16	0.22	0.49	-0.73
RC								1.00	0.08	0.16	0.45	-0.56
PP									1.00	0.54	-0.03	0.24
IDEAC										1.00	-0.07	0.13
TI											1.00	-0.70
AA												1.00

Cuadro 2.5: Matriz de correlaciones de las variables independientes

El cuadro anterior permite identificar la existencia de altas correlaciones entre las variables predictoras, valores de correlación altos podrían indicar problemas de multicolinealidad, los cuales se dan cuando dos o más variables independientes en el modelo se determinan aproximadamente por una combinación lineal de otras variables independientes del mismo modelo, lo que podría conducir a estimaciones poco fiables.

En este caso se observa que la mayoría de las correlaciones entre las variables independientes son relativamente bajas, a excepción de las correlaciones entre: *rentabilidad*, y *margen operativo* y *patrimonio técnico constituido sobre activos y contingentes ponderados por riesgo* y *apalancamiento del activo*.

Frente a esto, existen varias alternativas para controlar este problema en la estimación del modelo, como pueden ser: eliminar una de las variables que está altamente correlacionada, o usar métodos de estimación que se basen en la selección de variables independientes, con el fin de considerar aquellas con las que se pueda construir un modelo que se ajuste al conjunto de datos.

Para efectos del presente trabajo se empleó el método *stepwise* para la selección de variables en la estimación del modelo econométrico *logit*. A partir de esta selección se probó que de las variables que presentaron

una correlación alta, la metodología admite utilizar a una de ellas para la estimación del modelo, con lo cual se descartan problemas de multicolinealidad; aspecto que también se comprueba estadísticamente a través de los estadísticos pseudo- R^2 .

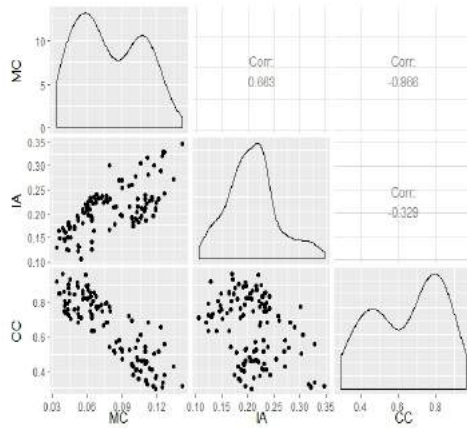


Figura 2.1: Morosidad, improductividad y cobertura de cartera

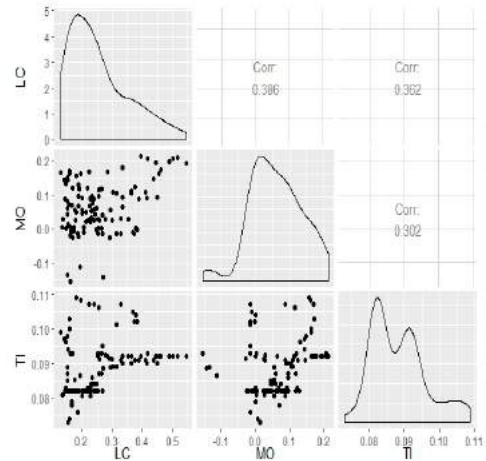


Figura 2.2: Liquidez, margen operativo y tasa de interés

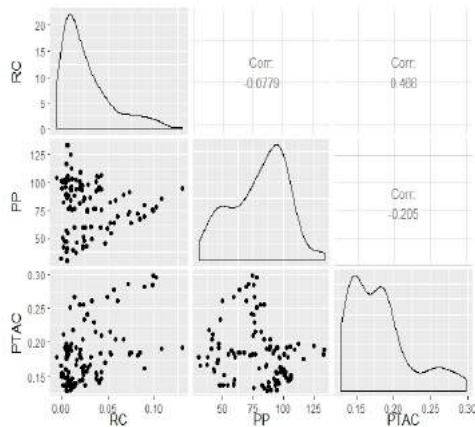


Figura 2.3: Rentabilidad del capital, variación del precio del petróleo, y patrimonio técnico constituido / activos y contingentes ponderados por riesgo

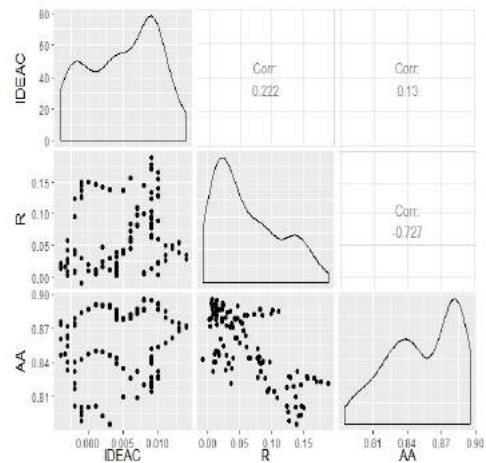


Figura 2.4: Índice de Actividad Económica Coyuntural, rentabilidad y apalancamiento del activo

2.3. ESTIMACIÓN DE LAS PROBABILIDADES DE RIESGO ALTO 43

En las figuras anteriores se visualiza gráficamente la relación entre las variables predictoras.⁵ Por ejemplo, en el cuadro 2.5 la correlación entre *morosidad de cartera* y *cobertura de cartera* es negativa y bastante alta (-0.87). Dicha relación se confirma gráficamente en la figura 2.1, en donde los puntos forman una línea recta (mientras más claramente se aprecie la línea recta, más alta es la correlación lineal entre las variables) y se observa que a niveles bajos de *morosidad de cartera*, la *cobertura de cartera* es alta y viceversa, esto indica que la relación entre ambas es negativa.

Es importante mencionar que, como se indicó anteriormente, el uso de la metodología *stepwise* permite seleccionar las variables que logran un mayor ajuste del modelo al conjunto de datos, lo que se ilustra en la siguiente sección.

2.3 Estimación de las probabilidades de riesgo alto

Para la estimación del modelo se empleó la metodología *stepwise* partiendo del siguiente modelo teórico:

$$\begin{aligned} \log\left(\frac{p}{1-p}\right) = & \beta_0 + \beta_1 MC_t + \beta_2 IA_t + \beta_3 CC_t + \beta_4 MO_t + \beta_5 PTAC_t \\ & + \beta_6 LC_t + \beta_7 AA_t + \beta_8 R_t + \beta_9 RC_t + \beta_{10} IDEAC_t \\ & + \beta_{11} TI_t + \beta_{12} PP_t + \mu_t \end{aligned} \quad (2.1)$$

Para determinar la significancia de las variables se realiza un contraste de hipótesis, en donde, la hipótesis nula (H_0) señala que los coeficientes (β 's) de la regresión son cero, es decir, no aportan explicación al modelo y por tanto no son significativos. Mientras la hipótesis alternativa (H_1) es la negación de la hipótesis nula ($H_1: \neg H_0$).

El criterio utilizado para determinar el resultado de la prueba, es la medida de significancia estadística conocida como *valor p* (valor plausible), este valor es la probabilidad que mide la evidencia en contra de la hipótesis nula. Se rechaza H_0 siempre que el valor p sea inferior al nivel de significancia, valores más bajos representan evidencia fuerte para rechazar la

⁵No se han representado todas las relaciones entre las variables independientes por efectos de visualización. El objetivo es interpretar el concepto de correlación desde la perspectiva de un gráfico.

hipótesis nula en favor de la alternativa.

En el caso de la entidad X_1 la variable dependiente es multinomial, ya que para el período de análisis el nivel de riesgo obtenido ha fluctuado entre 1 y 5 según la metodología *CAMEL*. Con lo cual se estimó un modelo *logit multinomial ordenado*.

<i>Variable dependiente:</i>	
Nivel de riesgo	
IA	54.936*** (9.505) [†]
PTAC	-96.596*** (17.286)
CC	-9.971*** (2.566)
R	-33.325*** (8.813)
Observaciones	114

Nota: *p<0.1; **p<0.05; ***p<0.01

[†] Estos valores corresponden al error estándar del coeficiente

Cuadro 2.6: X_1 : modelo *logit* con variable dependiente multinomial

Según el cuadro 2.6, el modelo evidencia la relación entre las variables independientes respecto a la variable respuesta. A continuación se muestra los estadísticos de las variables explicativas con su respectivo valor plausible.

2.3. ESTIMACIÓN DE LAS PROBABILIDADES DE RIESGO ALTO⁴⁵

	Value	Std. Error	t value	pval
Improductividad de activos	54.94	9.51	5.78	0.00
Patrimonio técnico constituido sobre activos y contingentes ponderados por riesgo	-96.60	17.29	-5.59	0.00
Cobertura de cartera	-9.97	2.57	-3.89	0.00
Rentabilidad	-33.32	8.82	-3.78	0.00
1 2	-22.93	5.50	-4.17	0.00
2 3	-17.53	4.65	-3.77	0.00
3 4	-10.77	4.29	-2.51	0.01
4 5 ⁶	-4.81	4.25	-1.13	0.26

Cuadro 2.7: Estadísticos de las variables independientes

Para este caso todas las variables muestran una relación negativa y estadísticamente significativa con la variable respuesta al 95 % de confianza, a excepción de la variable *improductividad de activos* que tiene relación positiva. Esto quiere decir que aumentos de estas variables están relacionados con una disminución de la probabilidad de tener riesgo alto.

El modelo evidencia que una disminución en estas variables (*cobertura de cartera*, *rentabilidad*, y *patrimonio técnico constituido sobre activos y contingentes ponderados por riesgo*) está asociada con un incremento de la probabilidad de riesgo alto.

La variable *improductividad de activos* muestra una relación positiva y estadísticamente significativa respecto a la probabilidad de tener riesgo alto, con lo cual incrementos en esta variable están asociados con aumentos de dicha probabilidad.

Respecto a los coeficientes de la regresión, se puede interpretar que un incremento de una unidad en *cobertura de cartera*, reduce el *log odds*⁷ ($\log \frac{p}{1-p}$) en 9.97 unidades.

⁶Según el valor plausible, pasar del nivel de riesgo 4 (medio alto) al 5 (alto) no resulta estadísticamente significativo. Sin embargo, es necesario tener en cuenta que el conjunto de datos presenta pocas observaciones en donde el nivel de riesgo sea la categoría 5, en comparación con las demás categorías de riesgo.

⁷El *log odds* es una transformación monótona del *odds ratio*, que a su vez es el ratio de la probabilidad de éxito sobre la probabilidad de fracaso. Cuanto mayores sean las probabilidades, mayores serán los *odds ratio* y *log odds*.

	1	2	3	4	5
1	0.14	0.01	0.00	0.00	0.00
2	0.03	0.05	0.06	0.00	0.00
3	0.00	0.05	0.33	0.03	0.00
4	0.00	0.00	0.03	0.22	0.01
5	0.00	0.00	0.00	0.01	0.04

Cuadro 2.8: Validación cruzada: nivel de riesgo CAMEL vs. nivel de riesgo predicho por el modelo

La validación cruzada es un método de evaluación del modelo que no utiliza medidas estadísticas convencionales (por ejemplo: R^2). Este método se centra en la capacidad de predicción del modelo. Los elementos de la diagonal del cuadro 2.8 representan el porcentaje de aciertos que tuvo el modelo para cada nivel de riesgo, mientras que los demás porcentajes del cuadro corresponden a los errores de clasificación.

En este caso, de un total de 114 observaciones en 88 de ellas (78%) el nivel de riesgo que arrojó el modelo es el mismo que el nivel de riesgo otorgado por la metodología *CAMEL*. Mientras que en 26 observaciones se obtuvo errores de predicción. Con lo cual, el error de clasificación (suma de los elementos fuera de la diagonal sobre el total) es de 22%.

Indicador	Valor
Log-likelihood fitted model	-56.33
Log-likelihood intercept-only	-160.74
G2 ⁸	208.83
McFadden	0.65
Maximum likelihood	0.84
Cragg y Uhler	0.89

Cuadro 2.9: Pseudo R^2

Respecto a los indicadores del cuadro 2.9 se observa que el indicador Mc Fadden es de 0.65, mientras que el de Cragg y Uhler⁹ es 0.89. Como se

⁸Menos dos veces la diferencia en el logaritmo de las verosimilitudes

⁹Para mayor comprensión se prefiere interpretar con estos dos criterios, ya que según su forma de cálculo, es más sencillo interpretar que el “ajuste perfecto” de los datos se

2.3. ESTIMACIÓN DE LAS PROBABILIDADES DE RIESGO ALTO⁴⁷

menciona en la sección 1.4, las medidas pseudo- R^2 pueden arrojar valores bastante distintos entre sí. Sin embargo, en conjunto estas medidas muestran un buen ajuste del modelo a los datos.

Predicciones del modelo

En la figura 2.5 se ven las fluctuaciones de las probabilidades de riesgo alto. Se observa que existe alta variabilidad en las predicciones. Por un lado se tiene valores muy bajos y cercanos a cero, mientras en otros períodos estas probabilidades superan el 15 %.

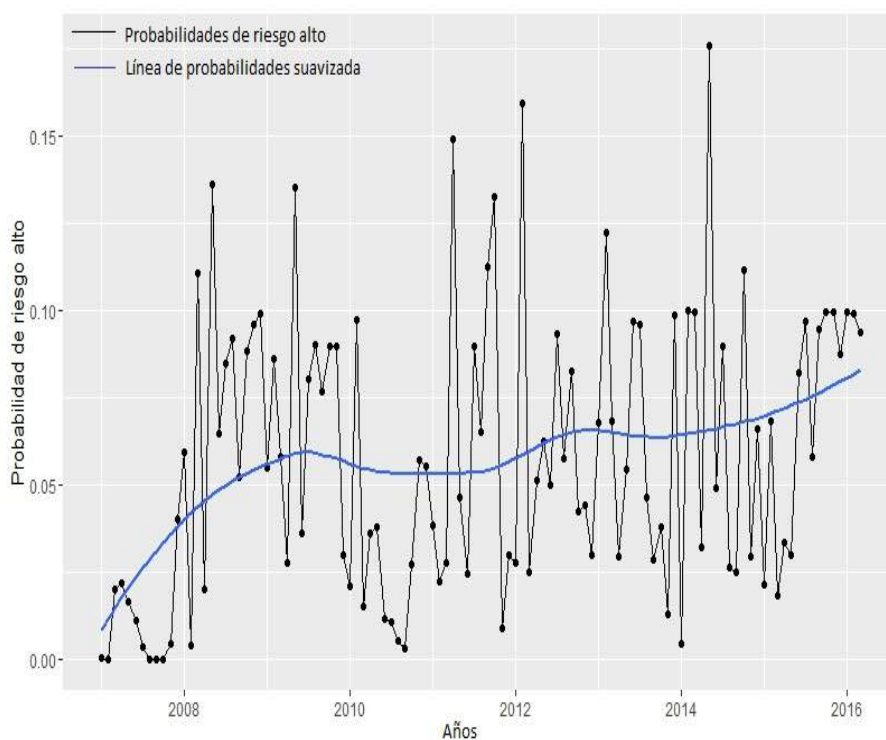


Figura 2.5: Probabilidades de riesgo alto

El cuadro 2.10 muestra un resumen de las estadísticas más relevantes de las predicciones.

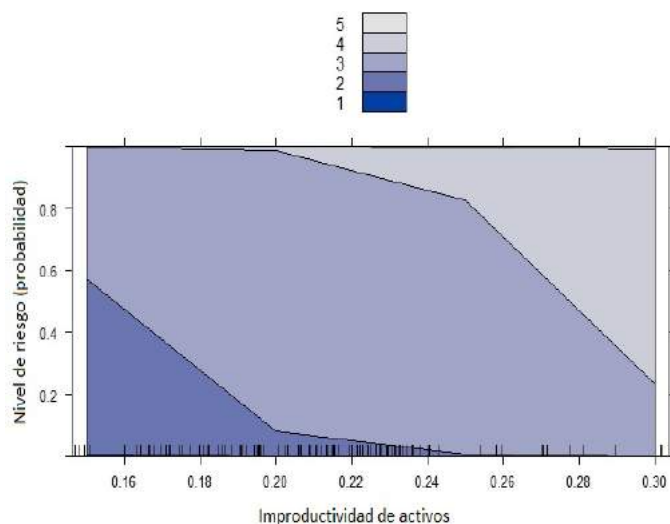
da mientras el indicador más se acerque a la unidad, que en el caso de Cragg y Uhler, este el máximo valor que puede tomar esta medida.

Riesgo alto		
1	Mínimo:	$4.24e^{-5}$
2	1er Cuantil:	0.026
3	Mediana:	0.052
4	Media:	0.058
5	3er Cuantil:	0.090
6	Máximo :	0.176

Cuadro 2.10: X_1 : resumen de las probabilidades de riesgo alto

Las probabilidades de riesgo alto para esta entidad durante el período de análisis tienen una media de 6 % y un máximo de 18 %.

A continuación se muestra los gráficos bivariantes entre las variables predictoras y el nivel de riesgo asociado.

Figura 2.6: *Effect plot*: improductividad de activos vs. nivel de riesgo

En la figura 2.6 se puede apreciar que a niveles más bajos de improductividad de activos, el nivel de riesgo asociado es bajo también. De este modo, la probabilidad de tener riesgo alto también es menor. Mientras que a niveles más altos de este ratio, el riesgo asociado es mayor, de modo que la probabilidad también aumenta.

2.3. ESTIMACIÓN DE LAS PROBABILIDADES DE RIESGO ALTO 49

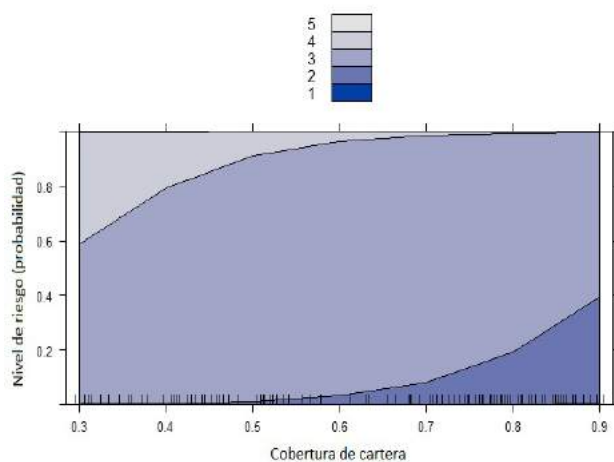


Figura 2.7: *Effect plot*: cobertura de cartera vs. nivel de riesgo

Para la variable *cobertura de cartera* se observa que a mayor cobertura, el riesgo asociado es bajo, con lo cual la probabilidad de tener un riesgo alto es menor. A niveles bajos de *cobertura de cartera*, los niveles de riesgo asociados son más altos, incluso de categoría 5.

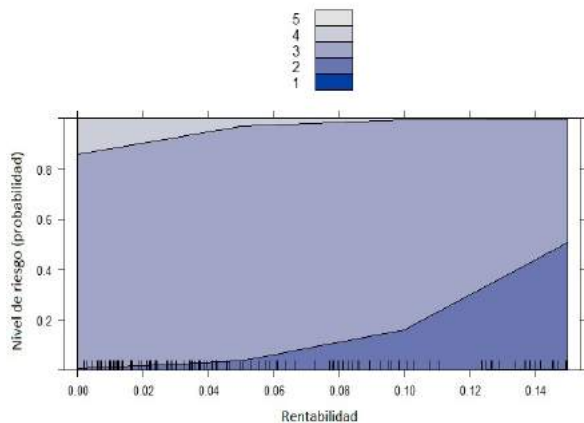


Figura 2.8: *Effect plot*: rentabilidad vs. nivel de riesgo

Para el ratio de *rentabilidad* se observa una forma similar a la figura anterior, a niveles más altos de *rentabilidad* el riesgo asociado es menor. Mientras que a medida que la *rentabilidad* disminuye, los niveles de riesgo relacionados son más altos.

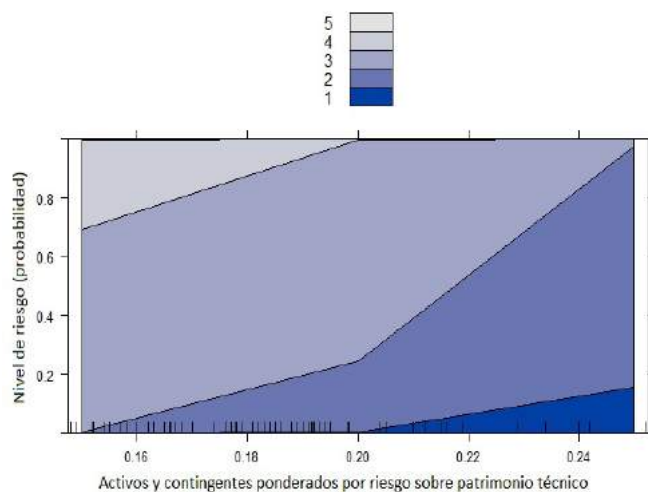


Figura 2.9: *Effect plot*: PTAC vs. nivel de riesgo

Finalmente, la figura 2.9 muestra que a niveles bajos, el nivel de riesgo asociado es alto, a medida que este indicador (PTAC) aumenta, el riesgo disminuye alcanzando niveles de 1.

2.3.1 Matrices de correlación de las entidades financieras

2.3.1.1 Matriz de correlaciones: Sector Financiero Privado

Entidades	X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8
X_1	1.000	0.548	0.298	-0.041	-0.206	0.569	-0.094	0.298
X_2		1.000	0.158	-0.387	-0.401	0.756	-0.315	0.417
X_3			1.000	0.518	0.402	0.308	0.378	0.062
X_4				1.000	0.726	-0.189	0.710	-0.167
X_5					1.000	-0.171	0.764	-0.310
X_6						1.000	-0.080	0.380
X_7							1.000	-0.257
X_8								1.000

Cuadro 2.11: Matriz de correlaciones: entidades con probabilidad de riesgo alto

En el cuadro 2.11 se observa la matriz de correlaciones para el conjunto de entidades con probabilidad de riesgo alto.

2.3. ESTIMACIÓN DE LAS PROBABILIDADES DE RIESGO ALTO 51

Analizando las correlaciones entre bancos, sociedades financieras y mutualistas, se observa que la correlación entre X_1 y X_7 es negativa y baja (-0.094). Mientras la correlación entre X_1 y X_8 es más alta, positiva e igual a 0.30. Por otro lado, la correlación entre X_1 y X_6 es 0.57, entre X_1 y X_5 es -0.21 y entre X_1 y X_4 es -0.04. Estos valores son muy distintos entre sí, lo que sugiere que la relación lineal entre bancos, mutualistas y sociedades financieras es muy diversa y depende de aspectos propios de cada entidad. Esta matriz se desagrega para cada subsistema en los siguientes apartados.

2.3.1.2 Matriz de correlaciones: bancos

Bancos	X_1	X_2	X_3
X_1	1.000	0.548	0.298
X_2		1.000	0.158
X_3			1.000

Cuadro 2.12: Matriz de correlaciones: bancos

Como se puede observar, las correlaciones con la entidad X_1 son positivas e iguales a 0.55 y 0.30 respecto a X_2 y X_3 respectivamente. Lo que significa que para la entidad X_1 , incrementos en la probabilidad de riesgo alto están relacionados con aumentos de esta probabilidad en las otras dos entidades. Por otro lado, la correlación entre X_2 y X_3 es positiva, pero baja e igual a 0.16.

2.3.1.3 Matriz de correlaciones: mutualistas

Mutualistas	X_4	X_5	X_6
X_4	1.000	0.726	-0.189
X_5		1.000	-0.171
X_6			1.000

Cuadro 2.13: Matriz de correlaciones: mutualistas

La correlación entre las entidades X_5 y X_6 es negativa e igual a -0.17. Así también, la correlación entre X_4 y X_6 es -0.19. En ambos casos esto indica que incrementos en la probabilidad de riesgo alto en una de ellas, están

asociados con disminuciones de esta probabilidad para la otra entidad. Sin embargo, esta correlación es bastante baja.

Finalmente, la correlación entre las entidades X_4 y X_5 es positiva y alta, con un valor igual a 0.73.

2.3.1.4 Matriz de correlaciones: sociedades financieras

Sociedades Financieras	X_7	X_8
X_7	1.000	-0.257
X_8		1.000

Cuadro 2.14: Matriz de correlaciones: sociedades financieras

En este caso, la correlación entre sociedades financieras es negativa e igual a -0.26, con lo cual disminuciones en la probabilidad de riesgo alto en una de ellas se relaciona con incrementos de esta probabilidad para la otra entidad.

2.4 Definición de los parámetros para la estimación de la función de pérdidas

La definición de los parámetros se realizó según lo mencionado en la sección 1.6.

- Probabilidades de riesgo alto: p_i , son las probabilidades estimadas por los modelos *logit*.
- Exposición al riesgo: EXP_i , corresponde al monto de depósitos cubiertos de las entidades incluidas en el estudio.
- Pérdida dado el incumplimiento: LGD , corresponde al porcentaje no recuperable de las pérdidas (75 %), de tal modo que el RR conocido como el porcentaje de recuperaciones es 25 %.
- Riesgo sistémico: Y_t definido como un número aleatorio para cada iteración, proveniente de una distribución normal estándar.

- Riesgo idiosincrático: $\epsilon_{n,t}$ definido como números aleatorios distintos para cada entidad dentro de cada iteración, provenientes de una distribución normal estándar.

2.5 Construcción del proceso de simulación

A partir de la definición de los parámetros se realiza el proceso de simulación para $N = 1,000$ con $M = 2,000$ iteraciones.

Para cada uno de los 1,000 eventos simulados se estima X_n , el cual depende del riesgo sistémico y del riesgo idiosincrático, como se indica en la sección 1.7.

Si dicho valor es menor al límite, definido como la función cuantil de la distribución normal (*threshold*), se calcula la pérdida esperada de la cartera según la ecuación 1.9. Además, se calcula la tasa de transición a calificación de riesgo alto (*default rate*), definida como el número de eventos que caen en incumplimiento (riesgo alto) sobre el total de eventos simulados.

Cálculo de la pérdida esperada total

Una vez realizada la simulación se obtiene la pérdida esperada de la cartera como la multiplicación de los siguientes tres factores:

$$\text{Pérdida esperada} = \sum_{i=1}^n EXP_i * p_i * LGD_i \quad (2.2)$$

$$\text{Pérdida esperada} = \text{US\$ } 928,748 \text{ miles}$$

Con lo cual, la pérdida esperada de la cartera es US\$ 928,748 miles y corresponde a 3.52 % del total de depósitos asegurados y 12.10 % del total de depósitos cubiertos.

Si se calcula la media de las simulaciones realizadas, el valor debería ser muy similar al de la ecuación 2.2, es así que de dicho cálculo se obtiene un valor igual a US\$ 923,864 miles.

2.6 Funciones probabilísticas de las pérdidas

2.6.1 Función de densidad de las pérdidas

Como se observa en la figura 2.10, a valores de pérdidas muy elevados la probabilidad de ocurrencia es prácticamente cero. Mientras que pérdidas menores están asociadas a probabilidades de ocurrencia más altas. Se observa la asimetría hacia la derecha de la distribución y la presencia de colas densas.

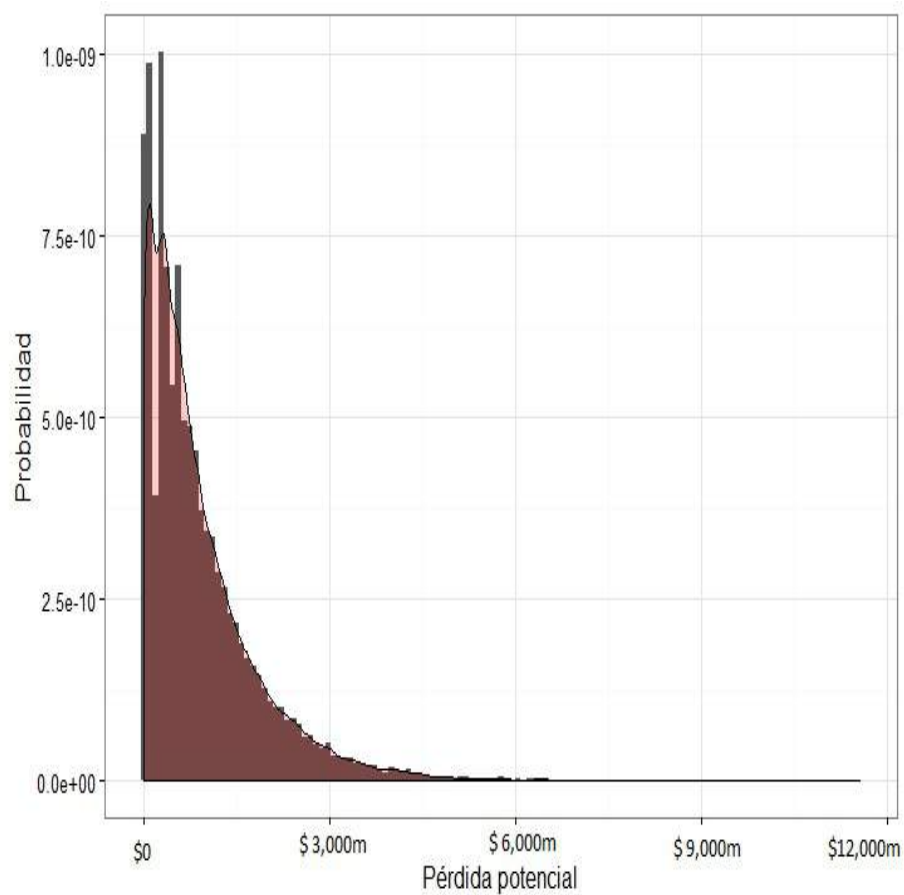


Figura 2.10: Función de densidad empírica de las pérdidas

2.6.2 Función de distribución empírica de las pérdidas

En la figura 2.11, se observa que pérdidas potenciales bajas están asociadas a valores que alcanzan una probabilidad acumulada cercana a 75%. A partir de este nivel, las pérdidas son significativamente mayores. Mientras menor pendiente tenga la curva de la distribución acumulada, mayor es la probabilidad de valores de pérdida bajos.

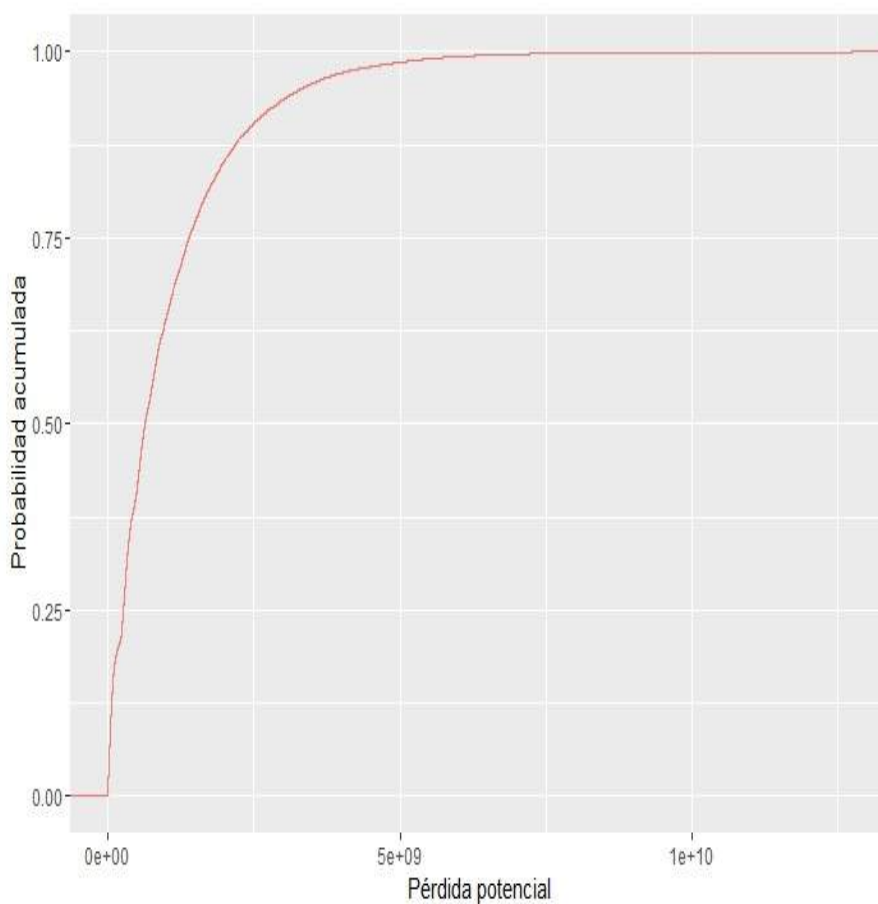


Figura 2.11: Función de distribución empírica de las pérdidas

2.7 Tasa de transición a calificación de riesgo alto (incumplimiento)

La tasa esperada de transición de que una entidad financiera pase a una calificación de riesgo alto (en adelante, tasa de riesgo alto) de las pérdidas simuladas es 2.3 %. Esta tasa es conocida en la literatura internacional como “*default rate*”, la cual se refiere a la tasa de incumplimiento, que como se señala en la sección 1.1, se entiende como la probabilidad de que una entidad financiera pase a una calificación de riesgo alto. En el siguiente cuadro se presentan los estadísticos descriptivos más relevantes:

Riesgo alto		
1	Mínimo:	0.000
2	1er cuartil:	0.009
3	Mediana:	0.020
4	Media:	0.028
5	3er cuartil:	0.037
6	Máximo :	0.314

Cuadro 2.15: Resumen de las tasas de riesgo alto

El percentil 75 de la distribución de tasas de riesgo alto es de 3.7 %, mientras el máximo valor es de 31.4 %, lo cual indica que al menos el 75 % de los datos corresponde a tasas de riesgo alto menores a 4 %.

2.7. TASA DE TRANSICIÓN A CALIFICACIÓN DE RIESGO ALTO (INCUMPLIMIENTO)

A continuación se muestra la función de densidad empírica:

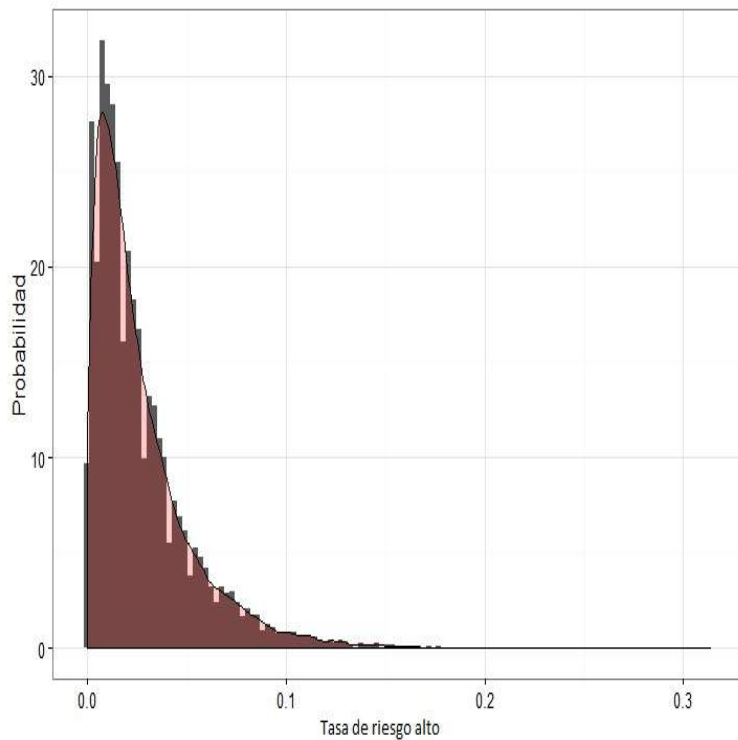


Figura 2.12: Función de densidad empírica de las tasas de riesgo alto

Se observa la misma asimetría hacia la derecha y el alargamiento de la cola de la función de densidad de las tasas de riesgo alto. Sin embargo, la probabilidad de que la tasa de riesgo alto sea cercana a cero llega hasta 31%. Mientras que para tasas de riesgo alto superiores a 10% la probabilidad de ocurrencia se reduce notablemente y se aproxima a cero.

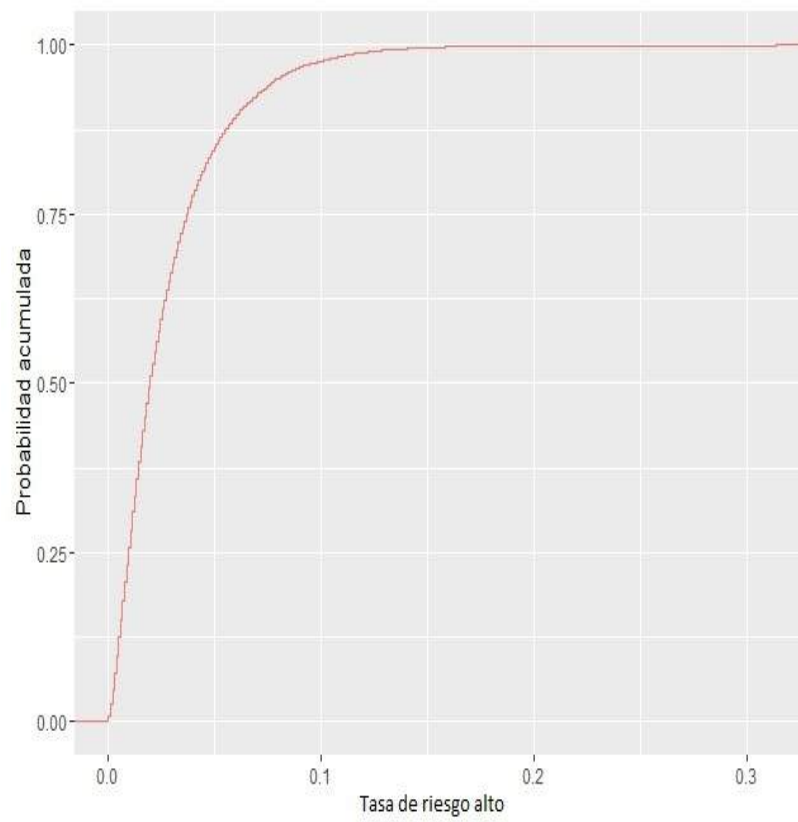


Figura 2.13: Función de distribución empírica de las tasas de riesgo alto

Finalmente, la figura 2.13 muestra que tasas de riesgo alto cercanas a cero alcanzan una probabilidad acumulada cercana a 75%. A partir de este punto, para niveles más altos de estas tasas la probabilidad acumulada se acerca asintóticamente a la unidad.

2.8 Determinación del valor en riesgo y valor en riesgo condicional

Nivel de confianza	Value at Risk			Tail Value at Risk		
	Valor ^a	%DA ^b	%DC ^c	Valor ^a	%DA ^b	%DC ^c
90 %	2,150,474	8.14	28.01	3,121,742	11.82	40.66
95 %	2,799,206	10.60	36.46	3,804,103	14.41	49.54
97.5 %	3,489,618	13.22	45.45	4,506,531	17.07	58.69
99 %	4,401,552	16.67	57.32	5,438,366	20.60	70.83

^aLos valores están en miles de US\$

^bComo porcentaje del total de depósitos asegurados del Sector Financiero Privado

^cComo porcentaje del total de depósitos cubiertos del Sector Financiero Privado

Cuadro 2.16: Medidas de riesgo VaR y TVaR

De acuerdo a las medidas de riesgo calculadas en el cuadro 2.16, se observa que a mayor nivel de confianza la pérdida estimada es mayor. Asimismo, el TVaR es superior al VaR para los diferentes niveles de confianza calculados, lo cual tiene sentido dado que el TVaR es un promedio de las pérdidas superiores al nivel del VaR.

Al 90 % de confianza el VaR representa cerca de 8 % del total de depósitos asegurados y 28 % del total de depósitos cubiertos. Mientras que en un escenario mucho más conservador, con 99 % de confianza el VaR representaría 21 % de los depósitos asegurados y 57 % de los depósitos cubiertos.

Por otro lado, si se considera como medida de riesgo el TVaR (promedio de las pérdidas por encima del nivel del VaR), con 90 % de confianza este valor corresponde a 12 % de los depósitos asegurados y 41 % de los depósitos cubiertos. Mientras que a un nivel de confianza de 95 % el TVaR representa aproximadamente 14 % de depósitos asegurados y 50 % de depósitos cubiertos.

2.9 Determinación del porcentaje de recuperación de cartera

Como se mencionó en la sección 1.9, a lo largo del proceso de liquidación de las entidades se pueden dar eventuales recuperaciones, este porcentaje de recursos que se espera recuperar puede financiarse con alternativas diferentes, por ejemplo, líneas contingentes de crédito. Además, se puede calcular con base en datos históricos o utilizando el porcentaje de recuperaciones (25 %) según los estándares planteados por el Comité de Supervisión Bancaria de Basilea (2006).

Asumiendo un porcentaje de recuperación igual a 25 %, el financiamiento por medio de otras fuentes debería ser 25 % del nivel objetivo del financiamiento total. Para el caso del Seguro de Depósitos del Sector Financiero Privado, por ejemplo, al 90 % de confianza, la pérdida esperada representa 8.14 % de los depósitos asegurados, con lo cual el porcentaje que se espera recuperar es el 25 % de dicho valor, que equivale a 2.04 %, y respecto a los depósitos cubiertos representa 7 %. Mientras que al 95 % de confianza este porcentaje corresponde a 2.65 % de los depósitos asegurados y 9.11 % de los depósitos cubiertos.

2.10 Provisión de pérdidas para entidades con menor nivel de riesgo

Una vez que se ha analizado las entidades que representan un nivel de riesgo alto para la cartera del Seguro de Depósitos se complementa estos resultados con la inclusión de los porcentajes de provisión para las demás entidades financieras, con el fin de mantener un criterio más conservador que abarque al conjunto de entidades que conforman el Sector Financiero Privado.

Para definir el nivel de provisión para las entidades financieras que no se encuentran con riesgo alto se utilizó el nivel de provisiones vigente para los créditos comerciales de las entidades del sector financiero privado. Los porcentajes de provisiones referidos se encuentran detallados en la Resolución No. 209-2016-F, denominada “Norma para la calificación de activos de riesgo y constitución de provisiones por parte de las entidades de los sectores financieros público y privado bajo el control de la Superintendencia de Bancos”, expedida por la Junta de Política y Regulación Monetaria y

2.10. PROVISIÓN DE PÉRDIDAS PARA ENTIDADES CON MENOR NIVEL DE RIESGO

61

Financiera el 12 de febrero de 2016.

A continuación el cuadro 2.17 resume las categorías de riesgo con sus respectivos rangos de pérdida y su relación con los niveles de riesgo otorgados por la metodología *CAMEL*.

Categorías	Calificación	Rango de pérdida esperada	CAMEL	Nivel de riesgo
Con riesgo normal	A-1	1 %	1 > y ≤ 1.99	Bajo
	A-2	2 %		
	A-3	3 %-5 %		
Con riesgo potencial	B-1	6 %-9 %	2 > y ≤ 2.99	Medio Bajo
	B-2	10 %-19 %		
Deficientes	C-1	20 %-39 %	3 > y ≤ 3.99	Medio
	C-2	40 %-59 %		
De dudoso recaudo	D	60 %-99 %	4 > y ≤ 4.99	Medio Alto
Pérdida	E	100 %	5 > y ≤ 5.99	Alto

Cuadro 2.17: Provisiones de pérdidas según el nivel de riesgo

En el cuadro 2.18 se muestra el valor total provisionado por tipo de entidad. El monto provisionado para bancos corresponde a 88.80 % del total de provisiones. Este mismo monto representa 3.16 % del total de depósitos cubiertos y 0.92 % del total de depósitos asegurados del Sector Financiero Privado.

Por otro lado, el valor provisionado para sociedades financieras representa 0.30 % de los depósitos cubiertos y 0.09 % del total de depósitos asegurados del Sector Financiero Privado.

Finalmente, el valor provisionado para mutualistas corresponde a 0.09 % del total de depósitos cubiertos y 0.03 % del total de depósitos asegurados del Sector Financiero Privado.

Entidad financiera	Provisión (en miles de US\$)	% total provisiones	% depósitos cubiertos	% depósitos asegurados
Bancos	242,428.18	88.80	3.16	0.92
Sociedades Financieras	23,358.98	8.56	0.30	0.09
Mutualistas	7,209.98	2.64	0.09	0.03
TOTAL	272,997.14	100.00	3.56	1.03

Cuadro 2.18: Provisiones por tipo de entidad

De este modo, la pérdida esperada asciende a US\$ 1,201,745 miles, lo que corresponde a un incremento de 29.4% considerando el valor calculado en la sección 2.5.

En este capítulo se desarrolló toda la metodología sobre la estimación de las probabilidades con los datos de la entidad X_1 , dicha estimación se realizó para todas las entidades incluidas en la sección 2 (ver Apéndice), a partir de aquí, se estimó las correlaciones, pérdida esperada y medidas de riesgo. Finalmente, se calculó las provisiones para el conjunto de entidades que no fueron parte de la estimación de los modelos.

2.11 Resumen de resultados

Los modelos logísticos se estimaron para las ocho entidades incluidas en el análisis y sus resultados pueden ser interpretados según se presenta a continuación para la entidad X_1 , la cual se ha tomado como ejemplo para ilustrar la metodología de la estimación de las probabilidades de riesgo alto. Los resultados para el resto de entidades se presentan desde el apéndice 2 en adelante.

- Las variables *cobertura de cartera*, *rentabilidad*, y *patrimonio técnico constituido sobre activos y contingentes ponderados por riesgo* tienen una relación negativa con la probabilidad de riesgo alto, es decir, incrementos de estas variables se asocian con reducciones de la probabilidad de riesgo alto. Mientras que *improductividad de activos* muestra una relación positiva con la probabilidad de riesgo alto, es decir, incrementos de esta variable se asocian con aumentos de la probabilidad de riesgo alto.

- Respecto a la bondad de ajuste, el modelo arrojó un porcentaje de correcta clasificación igual a 78 %. Mientras que las medidas conocidas como pseudo- R^2 arrojaron valores de 0.65 según el criterio de Mc Fadden y 0.89 según Cragg y Uhler.
- Las predicciones del modelo muestran gran variabilidad que a lo largo del tiempo no cambia significativamente (valor promedio de 5.8 % y desviación estándar de 3.6 %), a excepción del pico que se observa en el año 2014 (17.6 %).
- La matriz de correlaciones, que muestra la relación existente entre las probabilidades de riesgo alto de las distintas entidades financieras, indica que la relación lineal más alta es entre X_5 y X_7 (0.76), seguido de la correlación entre las entidades X_4 y X_5 (0.73). Mientras que las correlaciones más bajas son entre las entidades X_1 y X_4 (-0.04), y X_3 y X_8 (0.062).

Una vez que se estiman las probabilidades de riesgo alto para todas las entidades incluidas en el estudio se realiza el proceso de simulación, del cual se obtuvo los siguientes resultados:

- El valor de la pérdida esperada del Seguro de Depósitos en el plazo de un año es US\$ 928,748 miles y corresponde a 3.52 % de los depósitos asegurados y 12.10 % de los depósitos cubiertos.
- La tasa de riesgo alto esperada de las pérdidas es 2.3 %, dicha tasa corresponde al número de eventos que pasan a riesgo alto sobre el total de eventos simulados.
- La pérdida asumida por el Seguro de Depósitos al 90 % de confianza según la medida de riesgo VaR equivale a 8 % y 28 % de los depósitos asegurados y cubiertos respectivamente. Mientras que si se toma el TVaR como medida de riesgo, estos porcentajes son evidentemente mayores, 12 % de los depósitos asegurados y 41 % de los depósitos cubiertos respectivamente.
- Según el VaR al 95 % de confianza, la pérdida representa 11 % de los depósitos asegurados y 37 % de los depósitos cubiertos. Mientras que el TVaR obtenido para este nivel de confianza, corresponde a 14 % y

50 % de los depósitos asegurados y cubiertos respectivamente.

- Con 99 % de confianza, la pérdida asumida por el Seguro de Depósitos equivale a 17 % y 57 % de los depósitos asegurados y cubiertos respectivamente, de acuerdo con el VaR. Mientras que el TVaR es de 21 % de los depósitos asegurados y 71 % de los depósitos cubiertos.

Capítulo 3

Conclusiones y recomendaciones

3.1 Conclusiones

- Del análisis de las relaciones entre las variables explicativas respecto al nivel de riesgo de las entidades financieras se concluye que del conjunto de variables macroeconómicas, la variable *tasa de interés*, en todos los casos en los que resultó significativa, tiene una relación positiva con el nivel de riesgo, con lo cual aumentos de esta variable se relaciona con aumentos de la probabilidad de que una entidad financiera pase a una calificación de riesgo alto. Por otro lado, es interesante constatar que la *variación del precio del petróleo* no resultó estadísticamente significativa al 95% de confianza, incluso considerando rezagos de la variable, durante el período enero de 2007 hasta junio de 2016, en $t - 3$, $t - 6$, y $t - 12$, lo que quizás se podría explicar en razón de políticas de liquidez contracíclicas o políticas internas de las entidades financieras empleadas durante el período de reducción del precio del petróleo. Finalmente, la *variación del Índice de Actividad Económica Coyuntural* sí resultó significativa para seis de las ocho entidades analizadas. Sin embargo, la relación con la variable dependiente es distinta en cada modelo.
- En lo que respecta a las variables independientes de carácter financiero, se tiene que las variables *cobertura de cartera*, *rentabilidad*, y *liquidez corriente*, en todos los casos en los que resultaron significativas, se relacionan negativamente con la probabilidad de riesgo alto es decir, a mayores niveles de *cobertura de cartera*, *rentabilidad*, y *liqui-*

dez corriente, la probabilidad de riesgo alto es más baja. De su parte, las demás variables no presentan la misma relación en todos los casos en los que resultaron estadísticamente significativas, como son: *margen operativo* y *apalancamiento del activo*. Finalmente, las variables *rentabilidad del capital*, *patrimonio técnico constituido sobre activos y contingentes ponderados por riesgo*, e *improductividad de activos* resultaron significativas solo para una de las ocho entidades analizadas.

- La estimación de los modelos *logit* muestra que las probabilidades de riesgo alto para el conjunto de entidades estudiadas, oscilan entre 0 % y 17.6 % (valor máximo para X_1). Sin embargo, a excepción de la probabilidad máxima de la entidad X_1 , los demás valores llegan a niveles entre 10 % y 12 %.
- Respecto a la capacidad de predicción de los modelos probabilísticos, en general presentan una buena capacidad predictiva, con un error de clasificación promedio de 20 %, un mínimo de 5 % y máximo de 30 %. En lo que se refiere a la bondad de ajuste, según el criterio de Mc Fadden, los modelos de las entidades X_1 , X_5 y X_8 tienen una bondad de ajuste (pseudo- R^2) bastante alta (entre 65 % y 75 %), mientras que el ajuste más bajo es el de la entidad X_2 con 28 %. Sin embargo, se debe tener presente que estas medidas están basadas en la razón de verosimilitud y por tanto no se las puede interpretar como el porcentaje de la varianza total que es explicada por el modelo.
- La función de densidad de las pérdidas esperadas muestra un comportamiento que se alinea a la teoría sobre la que se basan las ciencias actuariales. Es decir, así como en el caso de la contratación de un seguro, la probabilidad de ocurrencia de siniestros es decreciente conforme aumenta el número de siniestros, en este caso la probabilidad de ocurrencia de liquidación forzosa de las entidades financieras también es decreciente. Con lo cual, la probabilidad de valores de pérdida altos tiende a cero, lo que se apreció en la figura 2.10, misma que muestra la asimetría hacia la derecha de la distribución y la presencia de colas densas.
- La tasa de riesgo alto muestra un comportamiento bastante similar al de la función de pérdidas, respecto a la asimetría hacia la derecha y

presencia de colas densas. En este caso se observa que para tasas de riesgo alto cercanas a cero la probabilidad de ocurrencia es alta (alcanza hasta un 31 %). Dicha probabilidad disminuye asintóticamente hacia cero a medida que la tasa de riesgo alto es mayor, lo cual nuevamente se relaciona con lo que la teoría actuarial señala.

- Se calculó los diferentes niveles óptimos del Seguro de Depósitos, es decir tanto el valor esperado de las pérdidas como las medidas de riesgo VaR y TVaR para diferentes niveles de confianza. La pérdida esperada (valor promedio) corresponde a 4 % del total de depósitos asegurados y 12 % de los depósitos cubiertos. Dichos porcentajes son los que se estima requeriría en promedio el Seguro de Depósitos en el plazo de un año.
- En lo referente a los valores extremos de la distribución de pérdidas, el nivel óptimo del Seguro de Depósitos al 90 % de confianza equivale a una reserva de 28 % de los depósitos cubiertos y 8 % de los depósitos asegurados según la medida de riesgo VaR.
- El nivel óptimo del Seguro de Depósitos con 99 % de confianza corresponde a 17 % de los depósitos asegurados y 57 % de los depósitos cubiertos según el VaR. Mientras que si se consideran los peores casos (TVaR), dicho nivel óptimo al 99 % de confianza es de 71 % de los depósitos cubiertos y 21 % de los depósitos asegurados. Es decir, a este nivel de confianza, esta es la reserva que cubriría los desembolsos requeridos al Seguro de Depósitos en el plazo de un año.
- Para complementar el análisis se incluyó el apartado de provisiones, con el fin de mantener un enfoque más conservador y no limitar el estudio exclusivamente a aquellas entidades con riesgo alto. Por lo que se consideró al conjunto de entidades del Sector Financiero Privado que representa un nivel de riesgo más bajo para el Seguro de Depósitos.
- En este sentido, según los datos del cuadro 2.18, el valor provisionado para bancos representa alrededor de 89 % del total de provisiones, mientras sociedades financieras y mutualistas representan 9 % y 3 %

respectivamente, lo que incrementa la pérdida esperada, basada en el modelo probabilístico y en la simulación de los datos en 29 %.

- Si se agrega este valor provisionado a la pérdida esperada, se obtiene que el Seguro de Depósitos requeriría 4.55 % de sus depósitos asegurados y 15.65 % de sus depósitos cubiertos.

3.2 Recomendaciones

- Como se indicó en las conclusiones, la pérdida esperada representa 3.52 % de los depósitos asegurados y 12.10 % de los depósitos cubiertos, si a este nivel óptimo se adiciona el valor de las provisiones, este porcentaje ascendería a 4.55 % de los depósitos asegurados y 15.65 % de los depósitos cubiertos del Seguro de Depósitos en el plazo de un año. Sin embargo, este nivel óptimo varía de acuerdo a los diferentes niveles de confianza para los cuales se calculó las medidas de riesgo (VaR y TVaR). En este estudio, estos porcentajes van desde 8 % hasta 21 % del total de depósitos asegurados y desde 28 % hasta 71 % del total de depósitos cubiertos. En este sentido, es importante que los administradores del Seguro de Depósitos conozcan esta gama de posibles niveles óptimos, con el fin de seleccionar aquel que se ajuste a sus necesidades y la posición frente al riesgo que se desea asumir.
- La metodología desarrollada estima una reserva que cubriría los desembolsos requeridos al Seguro de Depósitos en el plazo de un año. Con lo cual, se sugiere actualizar los resultados y revisar la metodología con periodicidad anual, ya que a partir de un año tanto las entidades consideradas en el presente análisis como las variables explicativas pueden variar en cuanto a la significancia de los modelos y por tanto cambiarían las estimaciones de las probabilidades de riesgo alto.
- Los modelos de espacio-estado, específicamente el *filtro de Kalman*, son útiles para predecir valores de una serie de tiempo, con las técnicas de: predicción, filtrado y alisado. Considerando esto, se recomienda extender la metodología actual y, a través de este método, predecir los valores de las variables explicativas para momentos del tiempo mayores a t ($t+1, t+2, \dots, t+h$). Una vez que se predicen estos valores, con base en los datos hasta junio de 2016, se sugiere incorporarlos al conjunto inicial de datos, para la posterior estimación de las probabilidades de riesgo alto. De este modo, la presente metodología no solo permitiría estimar la pérdida esperada dentro del plazo de un año, sino también predecir el valor de esta pérdida para tiempos mayores a este plazo.¹

¹A medida que se estiman predicciones para momentos del tiempo mayores a t , la probabilidad de que estos valores sean los estimados, es cada vez menor. Es decir, la probabilidad de que la variable x_i tome el valor de j en $t+1$ es mayor que la probabilidad de que x_i tome el valor k en $t+h$.

- En el presente estudio se realizó el análisis para las entidades que conforman el Sector Financiero Privado. Sin embargo, en próximas actualizaciones se recomienda aplicar esta metodología para el Sector Financiero Popular y Solidario, considerando sus particularidades e incluyendo a las mutualistas como parte de este sector (en este estudio fueron incluidas en el Sector Financiero Privado).
- Se ha considerado el porcentaje de recuperaciones según los estándares planteados por el Comité de Supervisión Bancaria de Basilea (2006). Sin embargo, en la medida en la que se disponga de más información sobre estos valores se sugiere incluirlos con base en los datos de la Corporación del Seguro de Depósitos Fondo de Liquidez y Fondos de Seguros Privados para el cálculo de la pérdida esperada ($EL_p = \sum_{i=1}^n EXP_i * p_i * LGD_i$ ²), con el objetivo de precisar aún más los resultados.

² $LGD_i = 1 - RR$; RR: recovery rate (% de recuperación)

Referencias

- Ahangarani, P. M. (2007). A new structural approach to the default risk of companies. *University of Southern California*.
- Allen, D., & Powell, R. (2011). Credit risk measurement methodologies. *Cowan University*.
- Analytics, G. R. . (2014). Credit risk: estimation techniques. *Crisil*.
- Brereton, T., Chan, J. C., & Kroese, D. P. (2012). Montecarlo methods for portfolio credit risk. *University of Queensland and Australian National University*.
- Castagna, A. (2009). Analytical credit VaR with stochastic probabilities of default and recoveries. *Fabio Mercurio of Bloomberg & Iason*.
- Charpentier, A. (2015). *Computational actuarial science using R* (C. . H. Book, Ed.). University of Québec.
- Cosma, R. S. (2012). Advanced data analysis. *Carnegie Mellon University, Dietrich College of Humanities & Social Sciences*.
- Damodaran, A. (2009). Value at risk - VaR. *NYU Stern School of Business*.
- Duffie, D. (2005). Credit risk modeling with affine processes. *Stanford University*.
- Elizalde, A. (2006). Credit risk models IV: understanding and pricing cdos. *Centro de Estudios Monetarios y Financieros*.
- Fox, J. (2010). Logit and probit models. *McMaster University, Department of Sociology*.
- Harvey, A. C. (1990). *Forecasting, structural time series models and the Kalman filter* (C. U. Press, Ed.). The Press Syndicate of the University of Cambridge.
- Hernández, C. (2013). A methodology for determining the target funding level of a deposit insurer. *Fogafin*.
- Hull, J., Nelken, I., & White, A. (2005). Merton's model, credit risk and volatility skews. *University of Toronto, Super Computer Consulting Incorporated*.
- Ioannidou, V. P., & Penas, M. F. (2008). Deposit insurance and bank risk-taking: evidence from internal loan ratings. *Tilburg University and Lancaster University*.
- Jeanblanc, M., & Rutkowski, M. (2001). *Mathematical finance* (Springer, Ed.). Université d'Évry Val d'Essonne, Warsaw University of Technology.
- Kalman, R. E. (1960). A new approach to linear filtering and prediction problems. *Transactions of the ASME-Journal of Basic Engineering*.
- Karoui, N. E., Jeanblanc, M., Jiao, Y., & Zargari, B. (2011). Conditional default probability and density. *Centre de Mathématiques Appliquées*.

- Kochendörfer, A. (2011). Vasicek single factor model. *Technische Universität Kaisers Lautern*.
- Kroese, D. P., Taimre, T., & Botev, Z. I. (2011). *Handbook of monte carlo methods* (J. W. . Sons, Ed.). University of Queensland.
- Kunt, A. D., & Kane, E. (2008). Deposit insurance around the world. *MIT Press*.
- Leo, T. W. (2014). Limited dependent variable. *St. Francis Xavier University*.
- Maccario, A., Sironi, A., & Zazzara, C. (2003). Applying credit risk models to deposit insurance pricing: empirical evidence from the italian banking system. *Unicredit Banca Mobiliare and Università "LUISS-Guido Carli", Università "Luigi Bocconi"*.
- Matjaž, V. (2014). Estimating probability of default and comparing it to credit rating classification by banks. *Bank of Slovenia*.
- Ortiz-Gracia, L., & Masdemont, J. J. (2014). Credit risk contributions under the Vasicek one-factor model: a fast wavelet expansion approximation. *Centre de Recerca Matemática*.
- Park, H. M. (2009). Regression models for binary dependent variables. *The University Information Technology Services (UITS) Center for Statistical and Mathematical Computing, Indiana University*.
- Parodi, P. (2015). *Pricing and general insurance* (C. . H. B. C. Press, Ed.).
- Peng, J. (2008). Value at Risk and Tail Value at Risk in uncertain environment. *Institute of Uncertain Systems, Huanggang Normal University*.
- Rodríguez, G. (2007). Generalized linear models. *Princeton University*.
- Rubinstein, R. Y., & Kroese, D. P. (2008). *Simulation and the monte carlo method* (J. W. . Sons, Ed.). University of Queensland.
- Vasicek, O. (1987). Probability of loss on loan portfolio. *KMV*.
- Vasicek, O. (2002). The distribution of loan portfolio value. *KMV*.
- Wilson, T. C. (1998). Portfolio credit risk. *McKinsey & Co. Inc.*.
- Zhang, A. (2009). Statistical methods in credit risk modeling. *University of Michigan*.

Apéndice 1

Marco teórico

1.1 Teorema: probabilidades de incumplimiento

La probabilidad de incumplimiento de cada entidad n de un grupo de $n = 1, \dots, N$ entidades en el tiempo t es denotada como $p_{n,t}$ y dada por:

$$\begin{aligned} p_{n,t} &= \Phi(K_{n,t}), \\ K_{n,t} &= \frac{\ln D_n - \ln A_{n,0} - \left(\mu_n - \frac{\sigma_n^2}{2}\right)t}{\sigma_n \sqrt{t}} \end{aligned} \quad (1.1)$$

Donde $A_{n,0}$ corresponde al valor de la firma de cada deudor $n \in (1, \dots, N)$, y el deudor n cae en incumplimiento si el valor del activo cae por debajo de un límite, de modo que $D_n = \mathbf{1}_{(A_n < C_n)}$. Donde Φ corresponde a la función de distribución normal estándar.

Condicionado a la realización del factor sistemático (Y) descrito en la ecuación 1.10, la probabilidad de incumplimiento en el tiempo t de cada entidad financiera n se denota por $p_n(Y)$ y viene dado por:

$$\begin{aligned} p_n(Y) &= P[X_n < K_{n,t} \mid Y] \\ &= P\left[\sqrt{\rho}Y + \sqrt{1-\rho}\epsilon_n < K_n \mid Y\right] \\ &= P\left[\epsilon_n < \frac{K_n - \sqrt{\rho}Y}{\sqrt{1-\rho}} \mid Y\right] \\ &= \Phi\left(\frac{K_n - \sqrt{\rho}Y}{\sqrt{1-\rho}}\right) \end{aligned} \quad (1.2)$$

Además, las variables aleatorias (X_1, \dots, X_N) y las probabilidades de incumplimiento $(p_1(Y), \dots, p_n(Y))$ son independientes.

Partiendo de la ecuación 1.1, donde el valor de $K_{n,t}$ para cada entidad financiera es $K_{n,t} = \Phi^{-1}(p_{n,t})$ y sustituyendo en la ecuación anterior se obtiene que la probabilidad condicional de incumplimiento es:

$$p_n(Y_t) = \Phi \left(\frac{\Phi^{-1}(p_{n,t}) - \sqrt{\rho_t} Y_t}{\sqrt{1 - \rho_t}} \right) \quad (1.3)$$

La variable aleatoria L_t definida como la suma de las variables $L_{1,t}, \dots, L_{N,t}$, representa el número de incumplimientos en el portafolio, misma que tiene la siguiente distribución:

$$P(L \leq m) = \sum_{k=0}^m \binom{N}{k} \int_{-\infty}^{+\infty} p(y)^k (1 - p(y))^{N-k} \phi(y) dy \quad (1.4)$$

Prueba del teorema

La probabilidad de k incumplimientos es:

$$P(L = k) = E(P([L = k] | Y)) = \int_{-\infty}^{+\infty} P(L = k | Y = y) \phi(y) dy \quad (1.5)$$

Donde ϕ es la densidad de Y . Los incumplimientos son condicionales e independientes en Y , por tanto:

$$P(L = k | Y = y) = \binom{N}{k} p(y)^k (1 - p(y))^{N-k} \quad (1.6)$$

Así, para $m \in [1, \dots, N]$ se tiene:

$$P(L \leq m) = \sum_{k=0}^m \binom{N}{k} \int_{-\infty}^{+\infty} p(y)^k (1 - p(y))^{N-k} \phi(y) dy \quad (1.7)$$

1.2 Portafolio homogéneo grande (*LHP*)

Como Schönbucher (2000) y Vasicek (2002) explican, dado que los incumplimientos son independientes cuando se condicionan a la realización de un factor sistemático común Y_t , el supuesto de un portafolio grande y

de igual tamaño implica que aplicando la ley de los grandes números, la fracción de créditos incumplidos en el portafolio converge a la probabilidad individual de incumplimiento para cada crédito $p(Y_t)$.

El portafolio es ponderado con pesos $\omega_1^N, \dots, \omega_n^N, \sum_{n=1}^N \omega_n^N = 1$ tal que:

$$\lim_{N \rightarrow \infty} \sum_{n=1}^N (\omega_n^N)^2 = 0$$

Así, cuando N tiende a infinito los pesos de los créditos tienden a cero.

1.2.1 Lema: tasa de incumplimiento

La tasa de incumplimiento de la cartera se define por la siguiente fórmula:

$$L^N = \sum_{n=1}^N \omega_n^N D_n \in [0, 1] \quad (1.8)$$

Sobre la base de *LHP* se define el siguiente teorema:

$$E(L^N | Y) = p(Y) = \Phi \left(\frac{\Phi^{-1}(p) - \sqrt{\rho}Y}{\sqrt{1-\rho}} \right) \quad (1.9)$$

$$Var(L^N | Y) = \sum_{n=1}^N (\omega_n^N)^2 p(Y)(1-p(Y)) \quad (1.10)$$

1.2.1.1 Prueba del lema

Linealidad de la expectativa condicional de los rendimientos:

$$\begin{aligned} E(L^{(N)} | Y) &= \sum_{n=1}^N \omega_n^N E(D_n | Y) \\ &= \sum_{n=1}^N \omega_n^{(N)} P(D_n | Y) \\ &= p(Y) \sum_{n=1}^N \omega_n^{(N)} \\ &= p(Y) \end{aligned} \quad (1.11)$$

D_n son condicionales e independientes en Y , por tanto:

$$\begin{aligned}
\text{Var}(L^N | Y) &= \sum_{n=1}^N (\omega_n^N)^2 \text{Var}(D_n | Y) \\
&= \sum_{n=1}^N (\omega_n^N)^2 p(Y)(1 - p(Y))
\end{aligned} \tag{1.12}$$

1.2.1.2 Teorema sobre la tasa de incumplimiento

La tasa de incumplimiento en portafolios homogéneos grandes (LHP) converge en probabilidad cuando $N \rightarrow \infty$

$$L^N \rightarrow p(Y) = \Phi \left(\frac{\Phi^{-1}(p) - \sqrt{\rho}Y}{\sqrt{1 - \rho}} \right) \tag{1.13}$$

1.2.1.3 Prueba del teorema

Para un portafolio grande, la variación de la tasa de incumplimiento dado Y tiende a 0.

$$\begin{aligned}
\text{Var}(L^N | Y) &= \sum_{n=1}^N (\omega_n^N)^2 p(Y)(1 - P(y)) \\
&= \leq \frac{1}{4} \sum_{n=1}^N (\omega_n^N)^2 \xrightarrow{N \rightarrow \infty} 0
\end{aligned} \tag{1.14}$$

1.3 Modelos espacio-estado: filtro de Kalman

El *filtro de Kalman* resulta útil en múltiples ámbitos, tanto para medir la serie real (no observable) a partir de una variable observada, como para predecir valores de una serie de tiempo, con lo cual podría resultar de interés realizar predicciones sobre las variables consideradas y por tanto, sobre las probabilidades de riesgo alto. La presente investigación se enfoca en definir un modelo que permita realizar las estimaciones dentro del plazo de un año. Sin embargo, podría resultar de interés realizar predicciones sobre las variables consideradas y por tanto, sobre las probabilidades de riesgo alto.

Los modelos espacio-estado son ampliamente utilizados en una serie de ámbitos. En economía particularmente, son de gran utilidad para predecir el comportamiento de variables cuyas fuerzas impulsoras pueden resultar no observables o medibles. Por ejemplo, el nivel de ingreso de una persona puede depender de factores no observables como su inteligencia, capacida-

des, habilidades, entre otros.

A nivel agregado, la teoría económica sugiere que las variables macroeconómicas e incluso las variables de tipo financiero son impulsadas por factores no observables, como pueden ser: el cambio tecnológico, la acumulación de capital humano, el precio de los activos, entre otros.

Cuando las variables explicativas no son observables, ciertos modelos ya no resultan útiles para estudiar la evolución de las mismas. Sin embargo, con la aplicación de modelos espacio-estado se amplía el marco de modelos econométricos y es posible construir modelos tanto explicativos como predictivos para estas variables.

Estos modelos permiten al investigador modelar una serie de tiempo observable, $(y_t)_{t=1}^T$, que se explica por un vector de variables estado, $(z_t)_{t=1}^T$, el cual es conducido por un proceso estocástico.

En general, estos modelos pueden resultar muy útiles cuando el objetivo es eliminar el ruido de una variable aleatoria. Por tanto, a partir de la información observada se busca estimar el estado $(z_t)_{t=1}^T$ aislando el término de error de $(y_t)_{t=1}^T$.

Estimación de los modelos espacio-estado

El modelo se construye de la siguiente forma:

$$y_t = Hz_t + v_t \quad v_t \sim N(0, \Sigma_v) \quad \text{Ecuación de observación} \quad (1.15)$$

$$z_t = Bz_{t-1} + w_t \quad w_t \sim N(0, \Sigma_w) \quad \text{Ecuación de transición} \quad (1.16)$$

La primera ecuación describe la relación entre la serie de tiempo observada y_t y el estado z_t . En general, se asume que los datos de la variable observada y_t incluyen un término de error que se mide con v_t expresado en la ecuación 1.15. El enfoque tradicional es modelar v_t como un error gaussiano, $v_t \sim N(0, \Sigma_v)$. La segunda ecuación describe la evolución de la variable estado como un proceso estocástico de innovaciones w_t , las cuales también se asume que se distribuyen como una normal $w_t \sim N(0, \Sigma_w)$. (Kalman, 1960)

Los modelos espacio-estado pueden ser formulados de maneras mucho más generales que las ecuaciones 1.15 y 1.16.

En aplicaciones prácticas, las varianzas Σ_w y Σ_v y el sistema de matrices H y B es desconocido y debe ser estimado, para lo cual se aplica la inferencia basada en la verosimilitud, dado que el *filtro de Kalman* permite construir la función de verosimilitud asociada a un modelo espacio-estado.

Filtro de Kalman

Por una parte, se asume que se conoce los datos de $(y_t)_{t=1}^T$. Por otra, se asume que se dispone de valores iniciales razonables para el conjunto de parámetros $\delta = \{H^*, B^*, \Sigma_{*v}, \Sigma_{*w}\}$. Se denota como $f(y_1, y_2, \dots, y_T; \delta)$ a la función de verosimilitud del modelo espacio-estado para un conjunto de parámetros δ . Por el teorema de Bayes se puede factorizar la función de verosimilitud como:

$$\begin{aligned} f(y_1, y_2, \dots, y_T; \delta) &= f(y_1, \delta) f(y_2 | y_1, \delta) f(y_3 | y_2, y_1, \delta) \dots f(y_T | y_{T-1}, \dots, y_1, \delta) \\ &= \prod_{t=1}^T f(y_t | y_{t-1}, \delta) \end{aligned} \quad (1.17)$$

Donde $y^0 = \phi$ y $y^{t-1} = (y_1, y_2, \dots, y_{t-1})$ para $t \geq 2$. El logaritmo de la función de verosimilitud viene dado por:

$$\ln L(y^T, \delta) = \sum_{t=1}^T \ln f(y_t | y^{t-1}, \delta) \quad (1.18)$$

Para poder construir la función de verosimilitud se necesita derivar las funciones de densidad.

$$f(y_t | y^{t-1}, \delta), t = 1, 2, \dots, T \quad (1.19)$$

Para lo cual se pueden usar las técnicas de filtrado del *filtro de Kalman*, puesto que los errores se consideran gaussianos.

El *filtro de Kalman* es un procedimiento recursivo definido en cuatro etapas: inicio, predicción, corrección y construcción de la función de verosimilitud.

Condiciones iniciales

Esta técnica empieza derivando el mejor predictor del estado inicial $z_{0|0}$ y un estimador de la matriz de varianzas y covarianzas $\Sigma_{0|0}^z = E[(z_0 - z_{0|0})(z_0 - z_{0|0})']$. Si el proceso es estacionario, se puede definir $z_{0|0} = z^*$ y $\Sigma_{0|0}^z = \Sigma^*$ tal que:

- $z^* = Bz^*$
- $\Sigma^* = B\Sigma^*B' + \Sigma_w$

Previo a continuar con la predicción, se establece $t = 1$ de modo que $z_{t-1|t-1} = z_{0|0}$ y $\Sigma_{t-1|t-1}^z = \Sigma_{0|0}^z$.

Predicción

En el tiempo t se puede usar $z_{t-1|t-1}$ y $\Sigma_{t-1|t-1}^z$ en la ecuación de transición para obtener:

$$z_{t|t-1} = Bz_{t-1|t-1} \quad (1.20)$$

$$\Sigma_{t|t-1}^z = B\Sigma_{t-1|t-1}^zB' + \Sigma_w \quad (1.21)$$

Ahora se puede usar $z_{t|t-1}$ para construir la predicción $y_{t|t-1} = Hz_{t|t-1}$. Dado que se observa y_t , se puede predecir el error:

$$u_t = y_t - y_{t|t-1} = y_t - Hz_{t|t-1} = v_t + H(z_t - z_{t|t-1}) \quad (1.22)$$

Dado que los errores son gaussianos se cumple que $u_t \sim N(0, \Sigma_v + H\Sigma_{t|t-1}^zH')$. Además, dado que $y_t = u_t + y_{t|t-1}$ se cumple que $f(y_t | y^{t-1}; \delta) = f(u_t, \delta)$.

Así, para construir la función de verosimilitud se necesita $f(y_t | y^{t-1}; \delta)$ para todo $t = 1, 2, \dots, T$; y dado $z_{t-1|t-1}$ y $\Sigma_{t-1|t-1}^z$ se puede calcular $f(y_t | y^{t-1}, \delta)$ de una distribución normal estándar.

$$f(y_t | y^{t-1}, \delta) = \frac{1}{\sqrt{(2\pi)^{n_y} |\Sigma_v + H\Sigma_{t|t-1}^zH'|}} \exp\left(-\frac{u_t'(\Sigma_v + H\Sigma_{t|t-1}^zH')^{-1}u_t}{2}\right) \quad (1.23)$$

Consecuentemente, para calcular $f(y_{t+1} | y^t, \delta)$ se necesita $z_{t|t}$ y $\Sigma_{t|t}^z$. En otras palabras, se necesita corregir las predicciones de los estados usando la nueva información disponible en el tiempo t .

Corrección

Habiendo observado y_t se pueden actualizar las predicciones $z_{t|t-1}$ y $\Sigma_{t|t-1}^z$ según la fórmula del *filtro de Kalman*:

$$z_{t|t} = z_{t|t-1} + K_t(y_t - y_{t|t-1}) = z_{t|t-1} + K_t(y_t - H z_{t|t-1}), \quad (1.24)$$

$$\Sigma_{t|t}^z = \Sigma_{t|t-1}^z - K_t(\Sigma_v + H \Sigma_{t|t-1}^z H') K_t', \quad (1.25)$$

Donde:

$$K_t = \Sigma_{t|t-1}^z H' (H \Sigma_{t|t-1}^z H' + \Sigma_v)^{-1} \quad (1.26)$$

De esta forma se asume que la predicción corregida es una combinación lineal entre la predicción anterior $z_{t|t-1}$ y el error en el tiempo t ($y_t - y_{t|t-1}$).

Dada la forma lineal, K_t se selecciona tal que minimice la varianza del error de predicción. Finalmente, a menos que $t = T$ se incrementa t y se regresa a la predicción. De no ser así, se pasa a la construcción de la función de verosimilitud.

Construcción de la función de verosimilitud

Los dos pasos anteriores calculan recursivamente la función de densidad $f(y_t | y^{t-1}, \delta)$ para $t = 1, 2, \dots, T$, la cual se utiliza para construir:

$$L(y^T, \delta) = \prod_{t=1}^T f(y_t | y^{t-1}, \delta) \quad (1.27)$$

Para la maximización de la función de verosimilitud, al ser una función no lineal se resuelve numéricamente y no de forma analítica. Para ello existen varios métodos, por ejemplo, si la función es continua los métodos basados en el gradiente suelen ser los más utilizados, como el método de Newton Rapson. También se puede utilizar el algoritmo EM (Expected Maximization).

Una vez que se han obtenido las estimaciones que maximizan la función de verosimilitud, es posible utilizar los modelos espacio-estado para hacer predicciones sobre los valores de las variables observables. En particular, se puede usar el estado predictor $z_{T|T}$ implicado en los estimadores máximo verosímiles δ_{ML} , con la ecuación de transición para construir $y_{T+h|T}$ para $h = 1, 2, \dots$, según:

$$y_{T+h|T} = HB^h z_{T|T} \quad (1.28)$$

Como una extensión al presente estudio¹, la utilización de los modelos espacio-estado permitiría predecir los valores de las variables independientes para períodos mayores a t . A partir de aquí, se podría incluir estos valores al conjunto de datos inicial para estimar las probabilidades de riesgo alto a través de los modelos *logit* para períodos $t + 1, t + 2, \dots, t + h$.

¹Incluir valores de predicción está fuera del alcance del presente análisis.

Apéndice 2

Estimación de los modelos *logit* para las entidades financieras

2.1 Entidad financiera X_2

Para esta entidad se estimó un modelo *logit* con variable respuesta de tipo binaria. A continuación se presenta los resultados obtenidos de la estimación del modelo:

<i>Variable dependiente:</i>	
Nivel de riesgo	
IDEAC	-128.449** (53.682)
AA	-46.129** (18.678)
RC	-28.043** (12.745)
Constante	39.093** (15.585)
Observaciones	
	114
Log Likelihood	
	-54.837
Akaike Inf. Crit.	
	117.675
Nota:	
	*p<0.1; **p<0.05; ***p<0.01

Cuadro 2.1: X_2 : modelo *logit* con variable dependiente binaria

De la metodología *stepwise* se obtiene que las variables: *variación del Índice de Actividad Económica Coyuntural*, *rentabilidad del capital*, y *apalancamiento del activo*, tienen una relación estadísticamente significativa y negativa respecto a la variable dependiente binaria, con lo cual, incrementos marginales de estas variables están relacionados con disminuciones de las probabilidades de riesgo alto.

En la siguiente tabla cruzada se valida la capacidad de predicción del modelo, en donde los elementos de la diagonal corresponden al porcentaje de aciertos (correcta clasificación) resultantes de la estimación:

	0	1
0	0.53	0.08
1	0.17	0.22

Cuadro 2.2: Validación cruzada: nivel de riesgo CAMEL vs. nivel de riesgo predicho por el modelo

En este modelo el error de predicción es de 25 % y corresponde al número de veces sobre el total de datos en los cuales el nivel de riesgo predicho por el modelo es distinto del nivel de riesgo otorgado por la calificación CAMEL.

En el cuadro 2.3 se muestra los indicadores de bondad de ajuste del modelo *logit* estimado. En este caso, los valores pseudo- R^2 de Mc Fadden y Cragg y Uhler son 0.28 y 0.42 respectivamente. Es importante destacar que aunque el ajuste para este modelo no es tan alto, estos indicadores presentan niveles satisfactorios en términos de la utilidad del modelo para explicar las relaciones entre las variables, además de su capacidad predictiva que, a través de la validación cruzada, arrojó 75 % de aciertos en la predicción de los niveles de riesgo.

Indicador	Valor
Log-likelihood fitted model	-54.84
Log-likelihood intercept-only	-76.03
G2 ¹	42.38
McFadden	0.28
Maximum likelihood	0.31
Cragg y Uhler	0.42

Cuadro 2.3: Pseudo R^2

A continuación se presenta el gráfico de las predicciones del modelo:

¹Menos dos veces la diferencia en el logaritmo de las verosimilitudes.

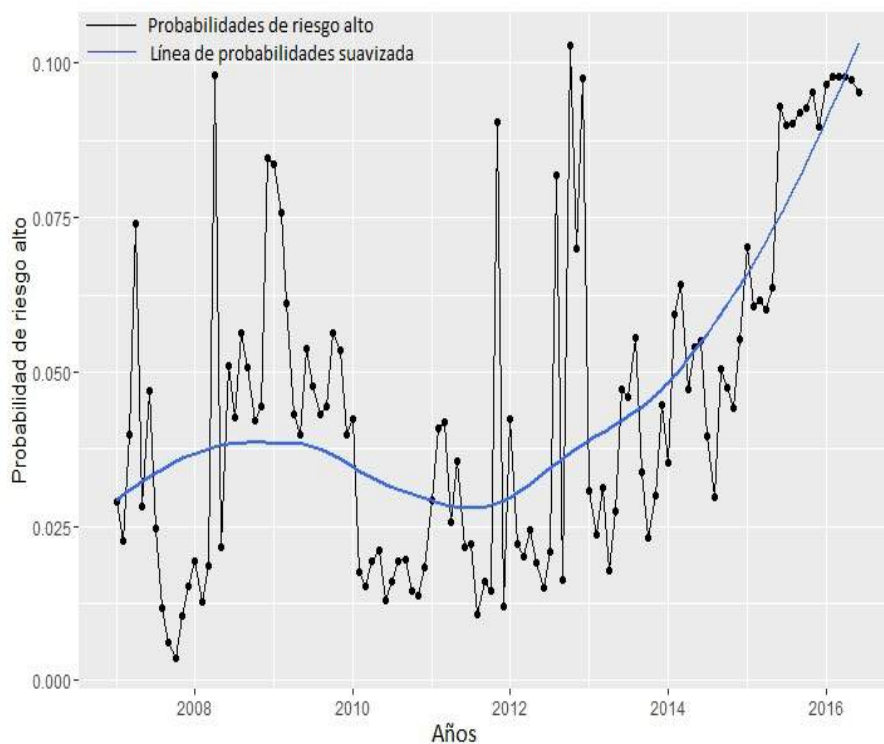


Figura 2.1: X_2 : probabilidades de riesgo alto

Las probabilidades fluctúan entre 0% y 10.3% durante el período de análisis. Se observa una tendencia creciente más pronunciada a partir del año 2013. Este gráfico se resume en el siguiente cuadro:

Riesgo alto		
1	Mínimo:	0,004
2	1er cuartil:	0,021
3	Mediana:	0,042
4	Media:	0,045
5	3er cuartil:	0,060
6	Máximo:	0,103

Cuadro 2.4: X_2 : resumen de las probabilidades de riesgo alto

Los siguientes gráficos bivariantes muestran la relación entre las variables explicativas y la variable respuesta asociada a las probabilidades de riesgo alto²:

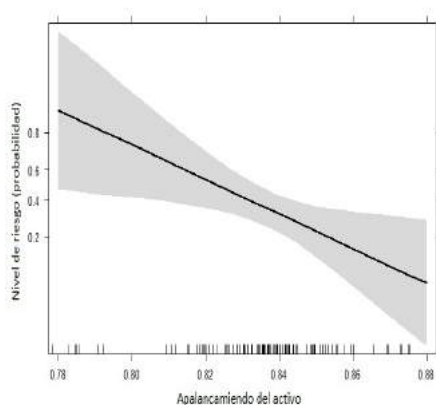


Figura 2.2: X_2 : apalancamiento del activo vs. nivel de riesgo

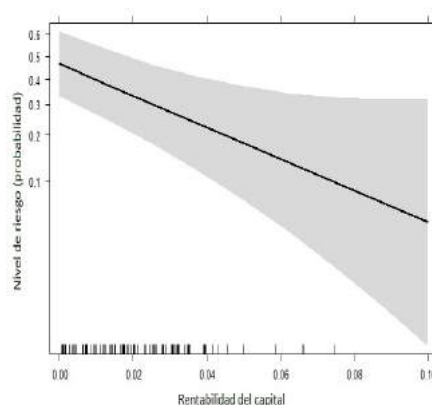


Figura 2.3: X_2 : rentabilidad del capital vs. nivel de riesgo

En el caso de las figuras 2.2 y 2.3 la relación es evidentemente negativa, con lo cual para valores altos de estas variables, la calificación de riesgo es baja y viceversa. Por otro lado, mientras más datos se encuentren en algún intervalo de valores determinado, el ancho de las bandas sombreadas en gris será menor (mayor precisión), mientras que si las observaciones son muy pocas, el área sombreada en gris será más ancha.

²El tipo de gráfico cambia respecto a los presentados en la sección 2.3, ya que depende del número de categorías que tenga la variable dependiente, en este caso esta entidad únicamente presenta dos niveles.

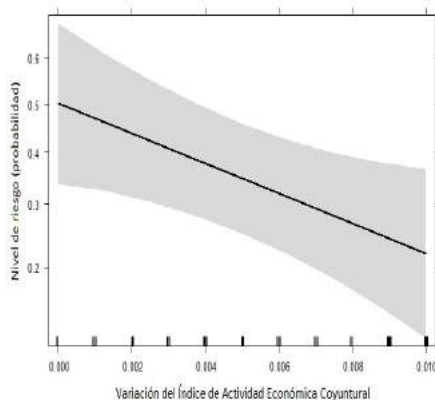


Figura 2.4: X_2 : variación del Índice de Actividad Económica Coyuntural vs. nivel de riesgo

Finalmente, respecto a la *variación del Índice de Actividad Económica Coyuntural* la relación también es negativa, a mayor *variación del Índice de Actividad Económica Coyuntural*, la calificación de riesgo obtenida por esta entidad es menor.

2.2 Entidad financiera X_3

Para esta entidad se estimó un modelo *logit* con variable respuesta de tipo multinomial debido a que las calificaciones de riesgo obtenidas con la metodología CAMEL fueron más de dos durante el periodo de estudio. A continuación se presentan los resultados obtenidos de la estimación del modelo:

<i>Variable dependiente:</i>	
Nivel de riesgo	
IA	22.214** (10.433)
TI	218.130*** (50.465)
IDEAC	-106.776 (83.132)
LC	-21.029*** (7.654)
R	-8.096** (4.284)
Observaciones	114

Nota: *p<0.1; **p<0.05; ***p<0.01

Cuadro 2.5: X_3 : modelo *logit* con variable dependiente multinomial

En este caso todas las variables resultaron estadísticamente significativas a excepción de la *variación del Índice de Actividad Económica Coyuntural*. Por un lado, *improductividad de activos*, y *tasa de interés* tienen una relación positiva con la variable respuesta. Mientras que *liquidez corriente*, y *rentabilidad* muestran una relación negativa respecto a la variable dependiente.

La siguiente tabla cruzada valida la capacidad de predicción del modelo:

	2	3	4	5
2	0.00	0.02	0.00	0.00
3	0.00	0.16	0.06	0.00
4	0.00	0.04	0.62	0.01
5	0.00	0.00	0.06	0.03

Cuadro 2.6: Validación cruzada: nivel de riesgo CAMEL vs. nivel de riesgo predicho por el modelo

En 81 % de los casos el modelo predijo el nivel de riesgo otorgado por la metodología *CAMEL*, siendo el error de predicción de 19 %.

Indicador	Valor
Log-likelihood fitted model	-61.63
Log-likelihood intercept-only	-102.56
G ² ³	81.86
McFadden	0.40
Maximum likelihood	0.51
Cragg y Uhler	0.61

Cuadro 2.7: Pseudo R^2

Los indicadores de bondad de ajuste del modelo *logit* estimado muestran valores pseudo- R^2 , según el criterio de Mc Fadden y Cragg y Uhler, iguales a 0.40 y 0.61, con lo cual se puede asumir que el modelo predice bien las probabilidades de riesgo alto.

³Menos dos veces la diferencia en el logaritmo de las verosimilitudes

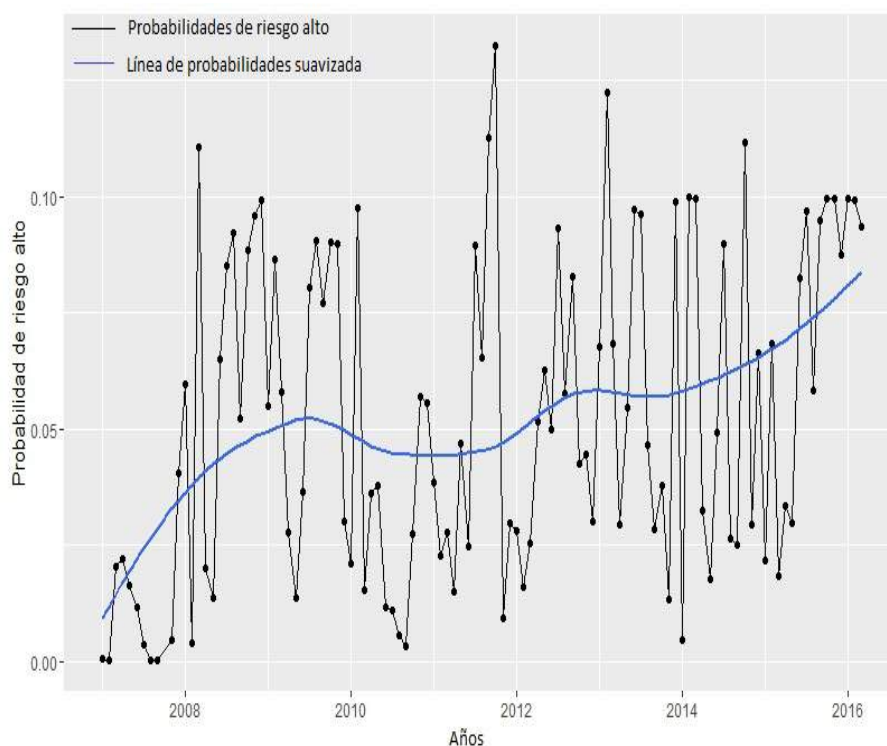


Figura 2.5: X_3 : probabilidades de riesgo alto

Se observa fluctuaciones altas, mostrándose una tendencia creciente para todo período de análisis. En el siguiente cuadro se resumen los estadísticos de las predicciones:

Riesgo alto		
1	Mínimo.:	0,001
2	1er cuartil:	0,02
3	Mediana:	0,034
4	Media:	0,056
5	3er cuartil:	0,078
6	Máximo :	0,094

Cuadro 2.8: X_3 : resumen de las probabilidades de riesgo alto

La probabilidad media de riesgo alto es de 6%, con un máximo de 9%.

APÉNDICE 2. ESTIMACIÓN DE LOS MODELOS LOGIT PARA LAS ENTIDADES FINANCIERAS

A continuación se muestra los gráficos bivariantes entre las variables explicativas y la variable respuesta asociada a las probabilidades de riesgo alto estimadas por el modelo:

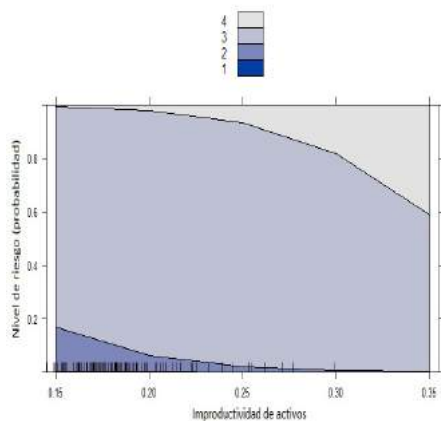


Figura 2.6: X_3 : improductividad activos vs. nivel de riesgo

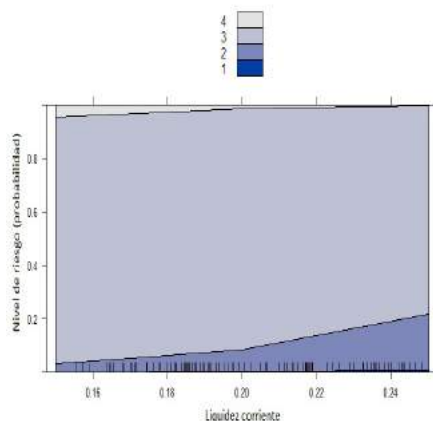


Figura 2.7: X_3 : liquidez corriente vs. nivel de riesgo

Por una parte se observa que a mayores niveles de *improductividad de activos* el riesgo asociado es más alto. Mientras que en el caso de *liquidez corriente*, el nivel de riesgo es prácticamente el mismo para diversos valores de este indicador.

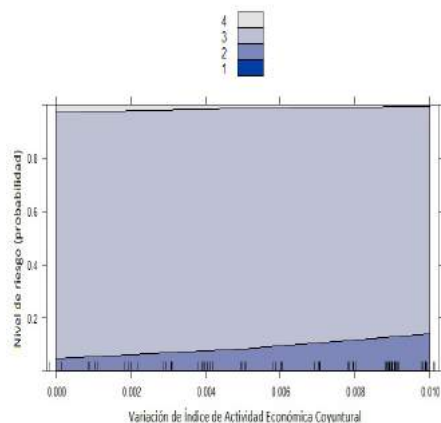


Figura 2.8: X_3 : variación del Índice de Actividad Económica Coyuntural vs. nivel de riesgo

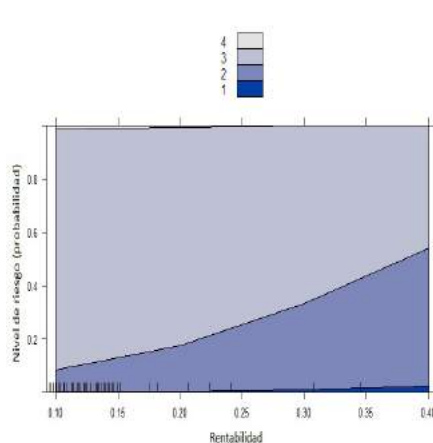


Figura 2.9: X_3 : rentabilidad vs. nivel de riesgo

Para esta entidad la relación de las variables *variación del Índice de Actividad Económica Coyuntural*, y *rentabilidad* indica que a mayores niveles de estas variables el riesgo asociado es más bajo y por tanto la probabilidad también.

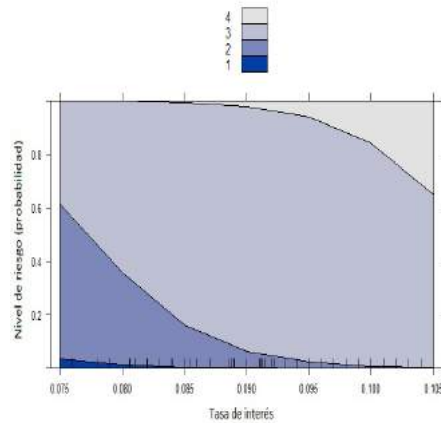


Figura 2.10: X_3 : tasa de interés vs. nivel de riesgo

Respecto a la *tasa de interés* se observa que la relación es positiva con los niveles de riesgo.

2.3 Entidad financiera X_4

Para esta entidad se estimó un modelo *logit* con variable respuesta de tipo multinomial. A continuación se presenta los resultados obtenidos de la estimación del modelo:

<i>Variable dependiente:</i>	
Nivel de riesgo	
CC	-7.397*** (1.499)
MO	-9.659*** (2.778)
LC	-24.846** (10.703)
IDEAC	151.577*** (46.141)
TI	180.597*** (53.679)
PP	-3.590 (2.416)
Observations	114

Nota: *p<0.1; **p<0.05; ***p<0.01

Cuadro 2.9: X_4 : modelo *logit* con variable dependiente multinomial

Para esta regresión las variables *cobertura de cartera*, *margen operativo*, *liquidez corriente* y *variación del precio del petróleo* muestran una relación negativa respecto a la probabilidad de riesgo alto. Sin embargo, esta última no resulta estadísticamente significativa al 5%. De su parte, la *variación del Índice de Actividad Económica Coyuntural* y *tasa de interés* tienen una relación positiva respecto a la probabilidad de riesgo alto y son estadísticamente significativas.

En la siguiente tabla cruzada se valida la capacidad de predicción del modelo⁴:

	2	3	4	5
2	0.13	0.05	0.00	0.00
3	0.04	0.17	0.08	0.01
4	0.00	0.04	0.30	0.01
5	0.00	0.00	0.03	0.14

Cuadro 2.10: Validación cruzada: nivel de riesgo CAMEL vs. nivel de riesgo predicho por el modelo

Del cuadro 2.10 se extrae que la estimación del modelo fue correcta en 74% de los casos.

Indicador	Valor
Log-likelihood fitted model	-75.98
Log-likelihood intercept-only	-152.65
G ² ⁵	153.32
McFadden	0.50
Maximum likelihood	0.74
Cragg y Uhler	0.79

Cuadro 2.11: Pseudo R^2

En el cuadro 2.11 las medidas de bondad de ajuste del modelo *logit* estimado muestran valores de 0.50 y 0.79 según los criterios de Mc Fadden, y Cragg y Uhler respectivamente, lo que significa que el modelo predice bastante bien los valores de las probabilidades de riesgo alto.

⁴En la tabla cruzada se presenta todas las categorías de riesgo que ha tenido la serie durante el período de análisis. En ciertos casos es posible que la entidad no haya presentado alguna de las categorías de riesgo, en este por ejemplo la entidad no ha obtenido la calificación de riesgo *bajo* a lo largo del período de estudio.

⁵Menos dos veces la diferencia en el logaritmo de las verosimilitudes

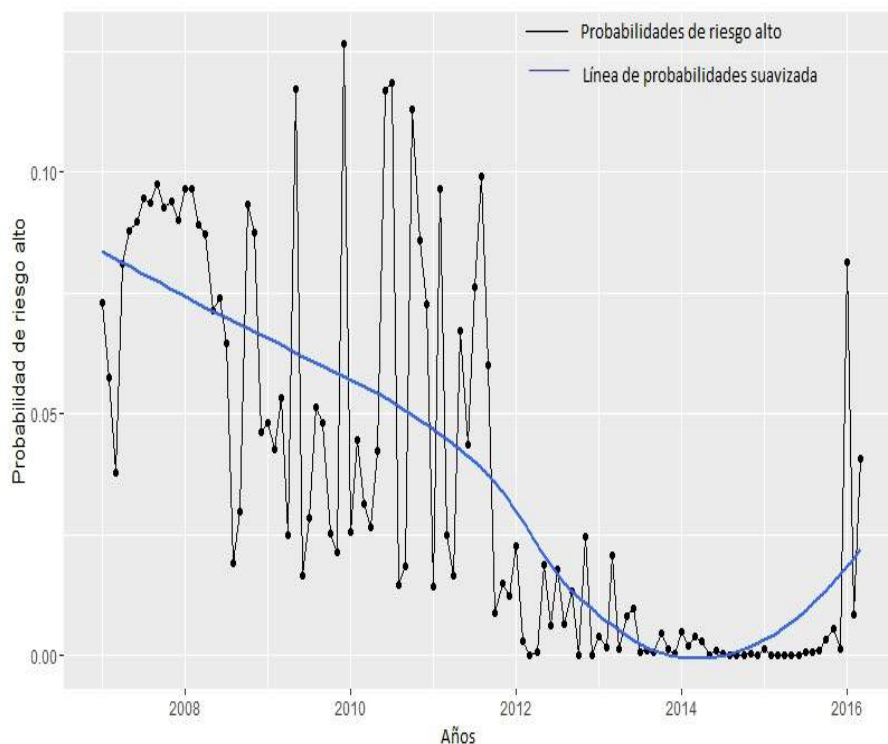


Figura 2.11: X_4 : probabilidades de riesgo alto

En este caso se evidencia que la serie presenta gran varianza para los primeros períodos y valores cercanos y superiores a 10 %. A partir de aquí, las probabilidades se reducen notablemente. Sin embargo, al final del período nuevamente se observa que las probabilidades empiezan a incrementarse.

Riesgo alto		
1	Mínimo:	$3,64e^{-5}$
2	1er cuartil:	0,002
3	Mediana:	0,029
4	Media:	0,036
5	3er cuartil:	0,072
6	Máximo :	0,127

Cuadro 2.12: X_4 : resumen de las probabilidades de riesgo alto

Para esta mutualista la media de las probabilidades es de 4 %, alcanzando un máximo de 13 %.

A continuación se muestra los gráficos bivalentes entre las variables explicativas y la variable respuesta asociada a las probabilidades de riesgo alto estimadas por el modelo:

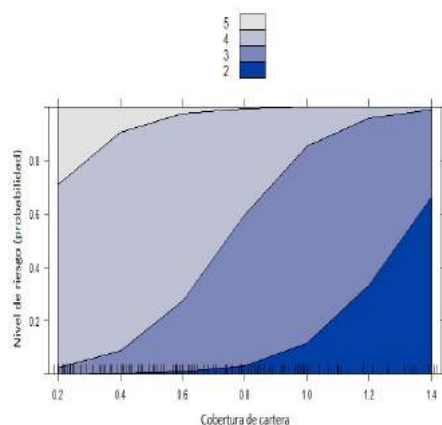


Figura 2.12: X_4 : cobertura de la cartera vs. nivel de riesgo

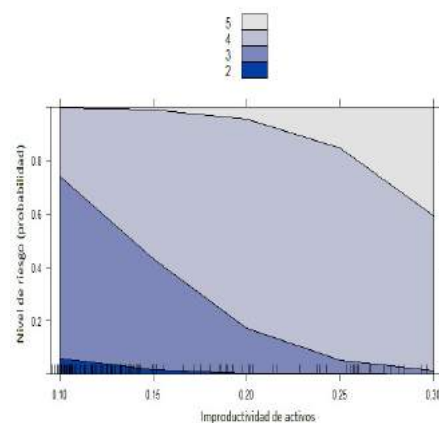


Figura 2.13: X_4 : improductividad de activos vs. nivel de riesgo

Según las figuras 2.12 y 2.13, a valores altos de *cobertura de cartera* menor es el nivel de riesgo asociado, mientras que a mayor nivel de *improductividad de activos* el nivel de riesgo aumenta.

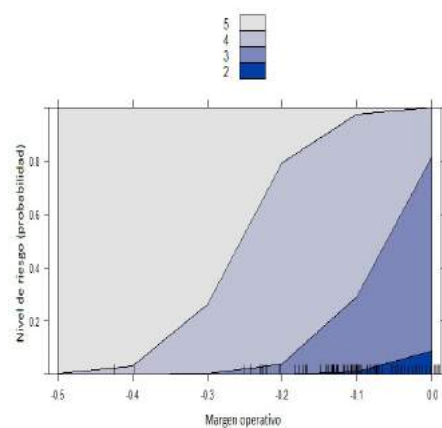


Figura 2.14: X_4 : margen operativo vs. nivel de riesgo

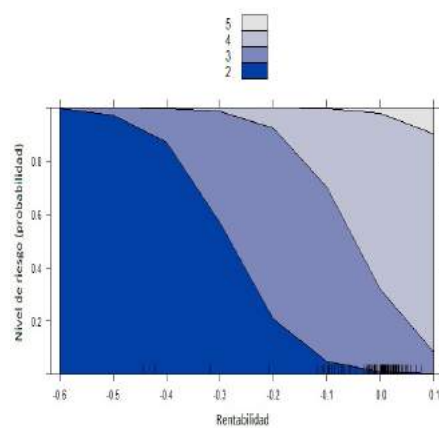


Figura 2.15: X_4 : rentabilidad vs. nivel de riesgo

Respecto a las figuras 2.14 y 2.15 se aprecia que a mayor *margen operativo* menor nivel de riesgo, por tanto, menor probabilidad de riesgo alto, mientras que a menor *rentabilidad* menor nivel de riesgo asociado y probabilidad baja de riesgo alto.

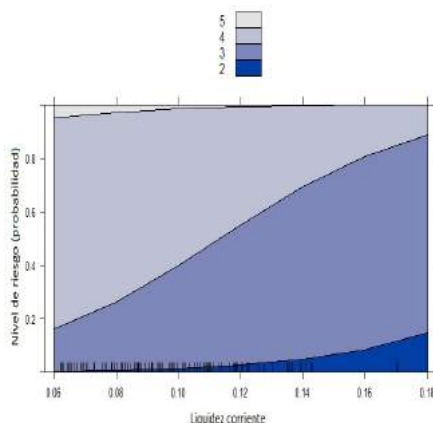


Figura 2.16: X_4 : liquidez corriente vs. nivel de riesgo

Finalmente, para la variable *liquidez corriente* se observa que a valores altos el nivel de riesgo asociado es principalmente medio y medio alto.

2.4 Entidad financiera X_5

Para esta entidad se estimó un modelo *logit* con variable respuesta de tipo binaria. A continuación se presenta los resultados obtenidos de la estimación del modelo:

<i>Variable dependiente:</i>	
Nivel de riesgo	
AA	299.005** (116.186)
CC	-15.414** (7.777)
LC	-100.628*** (34.353)
TI	409.059** (191.381)
MO	17.542* (9.621)
Constante	-289.937*** (99.444)
Observaciones	114
Log Likelihood	-19.053
Akaike Inf. Crit.	50.107
<i>Nota:</i>	*p<0.1; **p<0.05; ***p<0.01

Cuadro 2.13: X_5 : modelo *logit* con variable dependiente binaria

En el caso de la entidad X_5 las variables *apalancamiento del activo*, *tasa de interés* y *margen operativo* tienen una relación positiva respecto a la probabilidad de riesgo alto. Mientras *liquidez corriente* y *cobertura de cartera* tienen una relación negativa respecto a la variable respuesta.

A continuación se presenta la tabla cruzada para validar la capacidad de predicción del modelo:

	0	1
0	0.666	0.026
1	0.061	0.245

Cuadro 2.14: Validación cruzada: nivel de riesgo CAMEL vs. nivel de riesgo predicho por el modelo

En 91 % de los casos predijo correctamente, con lo que el error de predicción es de 9 %.

Indicador	Valor
Log-likelihood fitted model	-19.05
Log-likelihood intercept-only	-70.30
G ² ⁶	102.49
McFadden	0.73
Maximum likelihood	0.59
Cragg y Uhler	0.84

Cuadro 2.15: Pseudo R^2

Para esta entidad los indicadores de bondad de ajuste del modelo *logit* estimado arrojan valores pseudo- R^2 iguales a 0.73 y 0.84 según Mc Fadden y Cragg y Uhler respectivamente. Esto implica que el modelo ajusta bastante bien el conjunto de datos. Además, de la tabla cruzada se obtuvo que el porcentaje de aciertos de clasificación fue de 91 %.

En el siguiente cuadro se resume las estadísticas de las predicciones del modelo estimado:

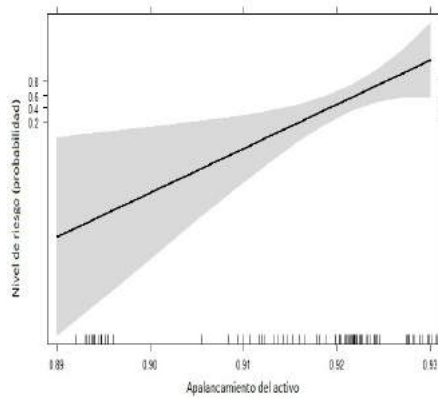
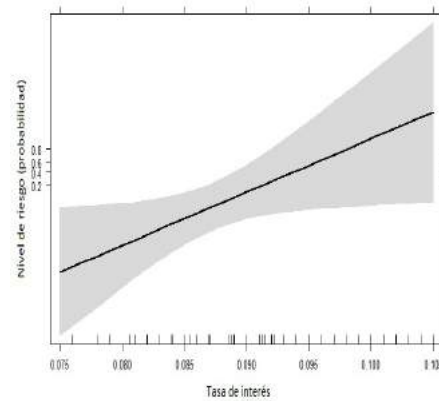
⁶Menos dos veces la diferencia en el logaritmo de las verosimilitudes

Riesgo alto		
1	Mínimo.:	0,000
2	1er cuartil:	0,000
3	Mediana:	0,021
4	Media:	0,036
5	3er cuartil:	0,070
6	Máximo :	0,101

Cuadro 2.16: X_5 : resumen de las probabilidades de riesgo alto

En las estadísticas descriptivas que se muestran en el cuadro anterior se observa que la probabilidad media de riesgo alto es de 4 %, con un máximo de 10 %.

A continuación se muestra los gráficos bivariantes entre las variables explicativas y la variable respuesta asociada a las probabilidades de riesgo alto estimadas por el modelo:

Figura 2.17: X_5 : apalancamiento del activo vs. nivel de riesgoFigura 2.18: X_5 : tasa de interés vs. nivel de riesgo

Según las figuras 2.17 y 2.18 se observa que existe una relación positiva entre el nivel de riesgo y las variables *tasa de interés* y *apalancamiento del activo*, a medida que incrementan los valores de estas dos variables la probabilidad de riesgo alto también aumenta.

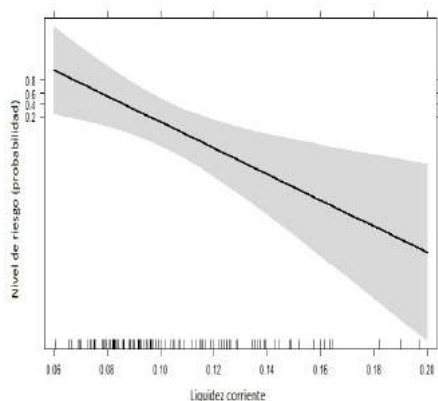


Figura 2.19: X_5 : liquidez corriente vs. nivel de riesgo

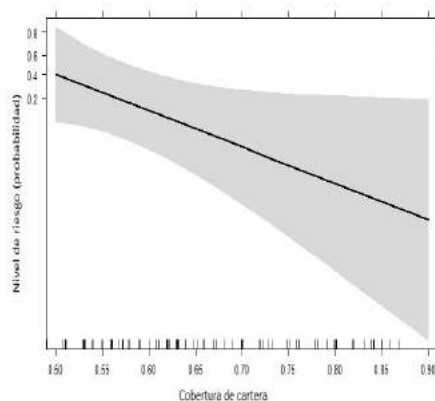


Figura 2.20: X_5 : cobertura de la cartera vs. nivel de riesgo

La relación para estas variables es negativa respecto a los niveles de riesgo, es decir, se observa que a mayores niveles de *liquidez corriente* y *cobertura de cartera* la probabilidad de riesgo alto es más baja.

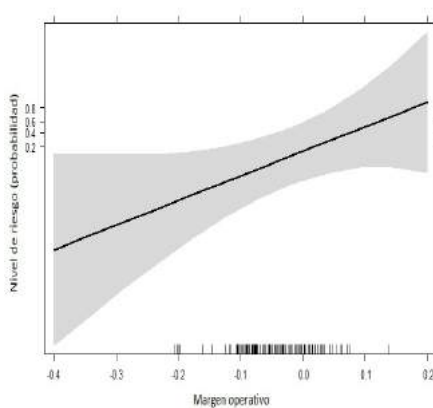


Figura 2.21: X_5 : margen operativo vs. nivel de riesgo

Finalmente, para la entidad X_5 se evidencia una relación positiva entre la variable *margen operativo* y los niveles de riesgo, incrementos en el *margen operativo* están relacionados con aumentos en los niveles de riesgo para esta entidad.

2.5 Entidad financiera X_6

Para esta entidad se estimó un modelo *logit* con variable respuesta de tipo multinomial. A continuación se presenta los resultados obtenidos de la estimación del modelo:

<i>Variable dependiente:</i>	
Nivel de riesgo	
IDEAC	-524.486*** (97.135)
R	-23.198*** (6.277)
TI	110.683*** (41.343)
Observaciones	114
<i>Nota:</i>	* $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

Cuadro 2.17: X_6 : modelo *logit* con variable dependiente multinomial

Una vez empleada la metodología *stepwise*, las variables que resultaron significativas son la *variación del Índice de Actividad Económica Coyuntural*, *rentabilidad*, y la *tasa de interés*. Las dos primeras con una relación negativa respecto a la probabilidad de riesgo alto. Mientras que la *tasa de interés* evidencia una relación positiva.

A continuación se presenta la tabla cruzada para validar la capacidad de predicción del modelo:

	3	4	5
3	0.10	0.04	0.00
4	0.03	0.52	0.06
5	0.00	0.09	0.16

Cuadro 2.18: Validación cruzada: nivel de riesgo CAMEL vs. nivel de riesgo predicho por el modelo

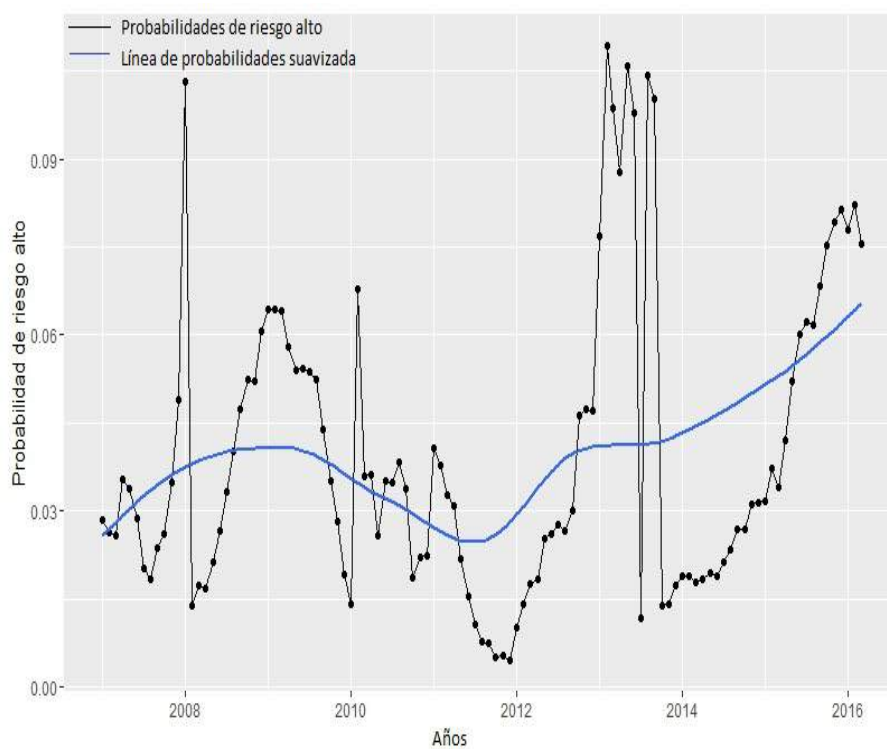
El error de predicción para este modelo es de 22%.

Indicador	Valor
Log-likelihood fitted model	-65.74
Log-likelihood intercept-only	-106.30
G ² ⁷	81.13
McFadden	0.38
Maximum likelihood	0.51
Cragg y Uhler	0.60

Cuadro 2.19: Pseudo R^2

El cuadro 2.19 arroja indicadores de 0.38 y 0.60 de acuerdo a los criterios de Mc Fadden y Cragg y Uhler respectivamente. Con lo cual la bondad de ajuste es buena, lo que también se corrobora con la validación a través de la tabla cruzada, donde el porcentaje de aciertos de clasificación fue de 78%.

⁷Menos dos veces la diferencia en el logaritmo de las verosimilitudes

Figura 2.22: X_6 : probabilidades de riesgo alto

En la figura 2.22 se observa que las probabilidades de riesgo alto presentan volatilidad a lo largo de todo el período de estudio, alcanzando niveles superiores a 10 % y probabilidades muy bajas cercanas a cero. La tendencia de la serie es claramente creciente desde finales de 2013.

Riesgo alto		
1	Mínimo:	0,005
2	1er cuartil:	0,023
3	Mediana:	0,031
4	Media:	0,034
5	3er cuartil:	0,047
6	Máximo :	0,106

Cuadro 2.20: X_6 : resumen de las probabilidades de riesgo alto

La probabilidad media es de 3 % con un máximo de 11 %.

APÉNDICE 2. ESTIMACIÓN DE LOS MODELOS LOGIT PARA LAS ENTIDADES FINANCIERAS
106

A continuación se muestra los gráficos bivariantes entre las variables explicativas y la variable respuesta asociada a las probabilidades de riesgo alto estimadas por el modelo:

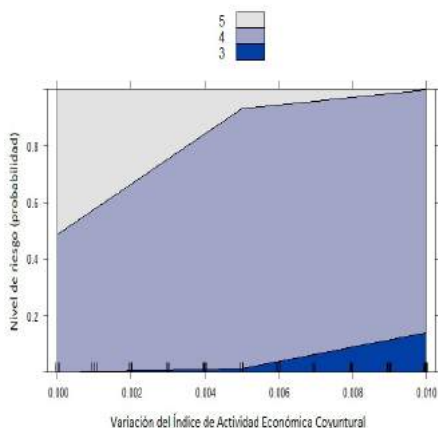


Figura 2.23: X_6 : variación del Índice de Actividad Económica Coyuntural vs. nivel de riesgo

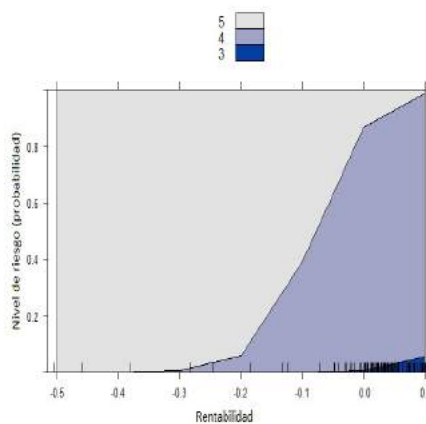
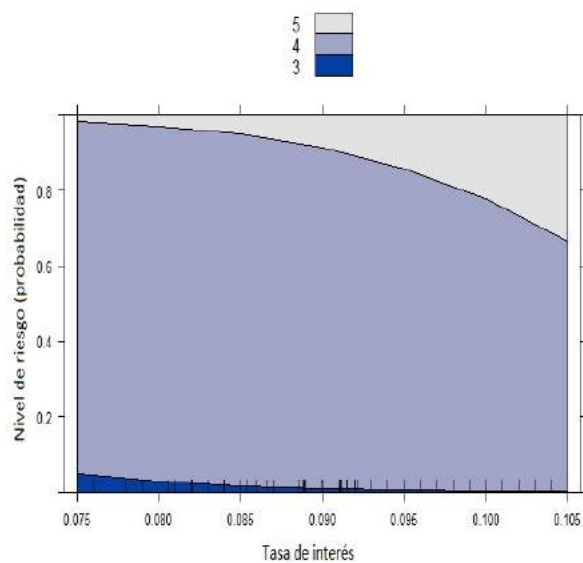


Figura 2.24: X_6 : rentabilidad vs. nivel de riesgo

En la figura 2.23 se muestra que a mayor valor de la variable independiente, el nivel de riesgo asociado es más bajo. Por otro lado, en la figura 2.24 se observa un comportamiento similar. Sin embargo, se observa un nivel de riesgo alto asociado a los niveles bajos y medios de la variable *rentabilidad*, a partir de aquí el riesgo es menor.

Figura 2.25: X_6 : tasa de interés vs. nivel de riesgo

En el caso de la figura 2.25 se observa que para casi todos los niveles de *tasa de interés* el nivel de riesgo asociado es medio alto. Sin embargo, a niveles altos de este indicador el riesgo aumenta.

2.6 Entidad financiera X_7

Para esta entidad se estimó un modelo *logit* con variable respuesta de tipo multinomial. A continuación se presenta los resultados obtenidos de la estimación del modelo:

<i>Variable dependiente:</i>	
Nivel de riesgo	
CC	-6.788*** (1.341)
AA	46.390*** (10.685)
IDEAC	-160.382** (67.696)
Observaciones	114
<i>Nota:</i>	*p<0.1; **p<0.05; ***p<0.01

Cuadro 2.21: X_7 : modelo *logit* con variable dependiente multinomial

Para la entidad X_7 la variable que muestra una relación positiva es el *apalancamiento del activo*. Mientras que *cobertura de cartera* y *variación del Índice de Actividad Económica Coyuntural* tienen una relación negativa con la probabilidad de riesgo alto.

A continuación se presenta la tabla cruzada para validar la capacidad de predicción del modelo:

	2	3	4	5
2	0.13	0.04	0.00	0.00
3	0.04	0.05	0.09	0.00
4	0.00	0.10	0.34	0.05
5	0.00	0.00	0.03	0.13

Cuadro 2.22: Validación cruzada: nivel de riesgo CAMEL vs. nivel de riesgo predicho por el modelo

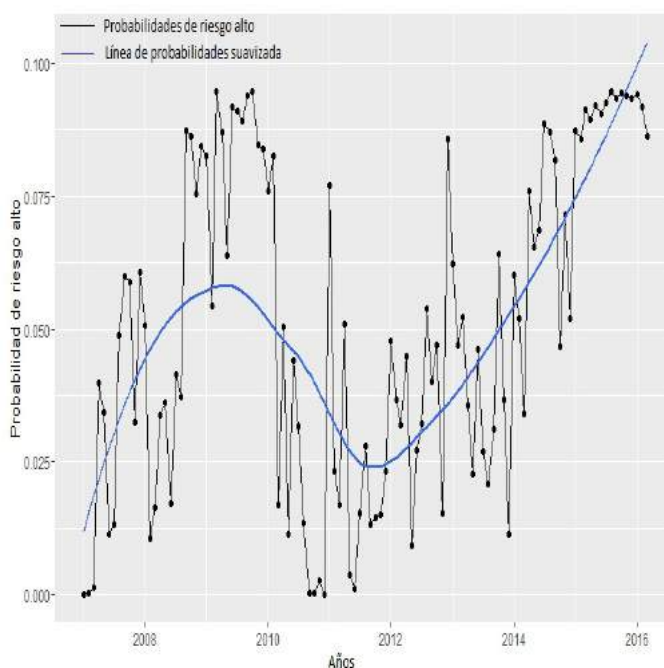
El error de predicción de este modelo es de 30 %.

⁸Menos dos veces la diferencia en el logaritmo de las verosimilitudes

Indicador	Valor
Log-likelihood fitted model	-71.87
Log-likelihood intercept-only	-144.60
G^2	145.45
McFadden	0.50
Maximum likelihood	0.72
Cragg y Uhler	0.78

Cuadro 2.23: Pseudo R^2

Para este modelo, el cuadro 2.23 muestra valores pseudo- R^2 bastante altos. El indicador de Mc Fadden es de 0.50, mientras el de Cragg y Uhler es igual a 0.78. La capacidad predictiva es ligeramente alta según la validación cruzada, ya que el porcentaje de aciertos en la clasificación de los niveles de riesgo es de 70%. En general el modelo resulta útil para fines tanto explicativos como predictivos.

Figura 2.26: X_7 : probabilidades de riesgo alto

En este caso se observa que las probabilidades de riesgo alto tienen una

gran varianza para todo el período de análisis. Desde el año 2007 al 2010 se observa una tendencia creciente, a partir de aquí caen las probabilidades, para volver a mantener la tendencia creciente desde el año 2012 en adelante.

A continuación se presenta las estadísticas más relevantes de las predicciones del modelo:

Riesgo alto		
1	Mínimo:	$2,94e^{-5}$
2	1er cuartil:	0,023
3	Mediana:	0,049
4	Media:	0,051
5	3er cuartil:	0,085
6	Máximo :	0,095

Cuadro 2.24: X_7 : resumen de las probabilidades de riesgo alto

La probabilidad media de riesgo alto es de 5%, la probabilidad del percentil 75 es de 0.09 y un máximo de 10%.

A continuación se muestra los gráficos bivariantes entre las variables explicativas y la variable respuesta asociada a las probabilidades de riesgo alto estimadas por el modelo:

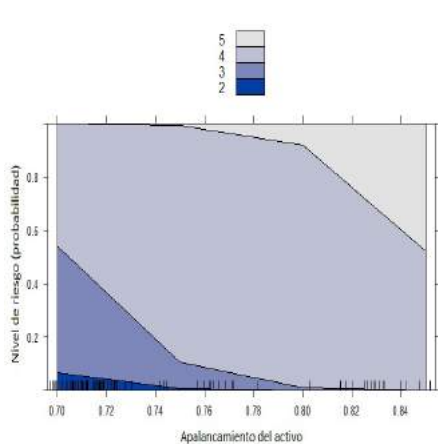


Figura 2.27: X_7 : apalancamiento del activo vs. nivel de riesgo

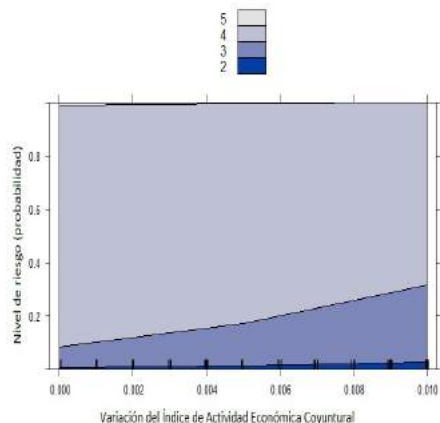


Figura 2.28: X_7 : variación del Índice de Actividad Económica Coyuntural vs. nivel de riesgo

Según la figura 2.27 a medida que el valor de este indicador es más alto, el riesgo asociado también lo es. Mientras que en el caso de la figura 2.28, la *variación del Índice de Actividad Económica Coyuntural* no muestra una relación muy evidente respecto a la probabilidad de riesgo alto, como se puede apreciar, prácticamente para todos los niveles de este indicador el riesgo asociado es medio alto.

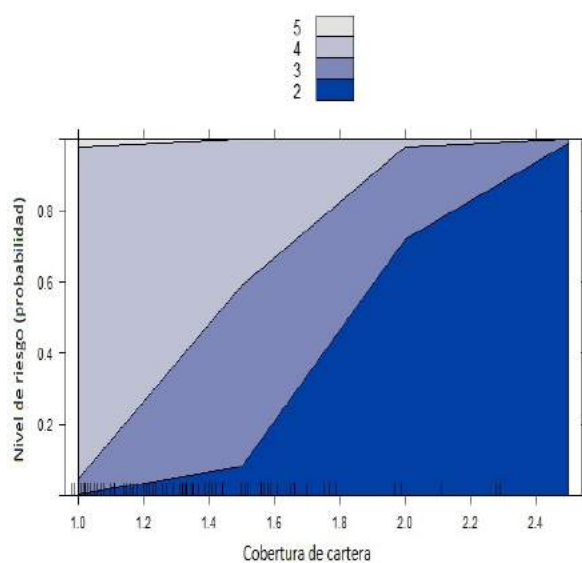


Figura 2.29: X_7 : cobertura de cartera vs. nivel de riesgo

La figura 2.29 muestra diferencias en los niveles de riesgo asociados a los valores de *cobertura de cartera*, se observa que a mayor cobertura menor es el riesgo asociado y por tanto la probabilidad.

2.7 Entidad financiera X_8

Para esta entidad se estimó un modelo *logit* con variable respuesta de tipo binaria.

A continuación se presenta los resultados obtenidos de la estimación del modelo:

<i>Variable dependiente:</i>	
Nivel de riesgo	
CC	-7.351*** (2.117)
PTAC	-51.665*** (13.979)
R	-15.032*** (5.763)
IDEAC	152.139 (99.781)
Constante	13.787*** (3.250)
Observaciones	
	114
Log Likelihood	
	-19.006
Akaike Inf. Crit.	
	48.012
Nota:	
	*p<0.1; **p<0.05; ***p<0.01

Cuadro 2.25: X_8 : modelo *logit* con variable dependiente binaria

Para esta entidad las variables estadísticamente significativas que muestran una relación negativa con la variable respuesta son: *cobertura de cartera*, y *patrimonio técnico constituido sobre activos y contingentes ponderados por riesgo y rentabilidad*. Por tanto, incrementos de estas variables están relacionados con disminuciones de la probabilidad de riesgo alto. Mientras *variación del Índice de Actividad Económica Coyuntural* muestra una relación positiva con la probabilidad de riesgo alto. Sin embargo, no es estadísticamente significativa.

En la siguiente tabla cruzada se valida la capacidad de predicción del modelo:

	0	1
0	0.60	0.04
1	0.01	0.35

Cuadro 2.26: Validación cruzada: nivel de riesgo CAMEL vs. nivel de riesgo predicho por el modelo

El error de predicción es de 5% para X_8 .

Indicador	Valor
Log-likelihood fitted model	-19.00
Log-likelihood intercept-only	-75.02
G2 ⁹	112.04
McFadden	0.75
Maximum likelihood	0.63
Cragg y Uhler	0.85

Cuadro 2.27: Pseudo R^2

Para esta entidad los indicadores de bondad de ajuste del modelo *logit* estimado muestran un pseudo- R^2 de Mc Fadden igual a 0.75, mientras el valor de Cragg y Uhler es de 0.85, esto quiere decir que el modelo predice bastante bien las probabilidades de riesgo alto.

A continuación se resume las estadísticas más relevantes de la serie de predicciones:

⁹Menos dos veces la diferencia en el logaritmo de las verosimilitudes

Riesgo alto		
1	Mínimo:	0,000
2	1er cuartil:	$2,30e^{-7}$
3	Mediana:	0,017
4	Media:	0,039
5	3er cuartil:	0,096
6	Máximo :	0,112

Cuadro 2.28: X_8 : resumen de las probabilidades de riesgo alto

Las probabilidades fluctúan entre 0 % y 11.2 % durante el período de análisis.

Los siguientes gráficos bivariantes muestran la relación entre las variables explicativas y la variable respuesta asociada a las probabilidades de riesgo alto estimadas por el modelo:

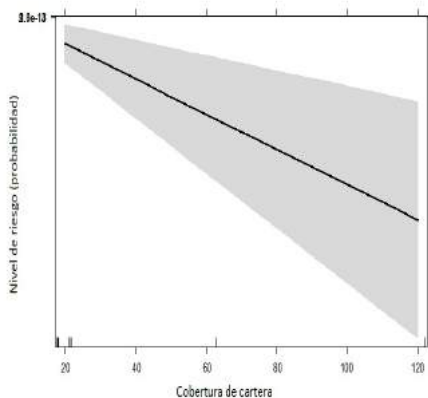


Figura 2.30: X_8 : cobertura de cartera vs. nivel de riesgo

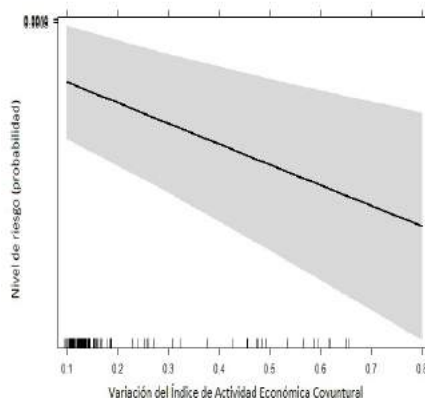


Figura 2.31: X_8 : patrimonio técnico constituido / activos y contingentes ponderados por riesgo vs. nivel de riesgo

Estas dos variables muestran la misma relación con la variable respuesta, a mayores ratios de *cobertura de cartera*, y *patrimonio técnico sobre activos y contingentes ponderados por riesgo*, la probabilidad de riesgo alto es menor.

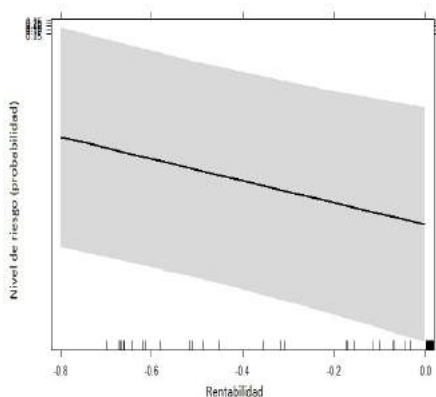


Figura 2.32: X_8 : rentabilidad vs. nivel de riesgo

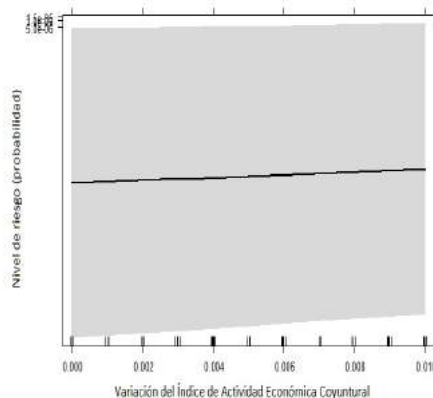


Figura 2.33: X_8 : variación del Índice de Actividad Económica Coyuntural vs. nivel de riesgo

Finalmente, para las variables *rentabilidad*, y *variación del Índice de Actividad Económica Coyuntural*, se observa que el nivel de riesgo prácticamente es el mismo para el rango de valores de estos indicadores, principalmente para la *variación del Índice de Actividad Económica Coyuntural*.

2.8 Provisiones bancos

En el cuadro 2.29 se muestra el total de depósitos cubiertos para cada entidad financiera (en adelante, EFI) y el valor de la provisión que se ha considerado de acuerdo a su nivel de riesgo. Dicho nivel de riesgo se calculó como el promedio de la calificación *CAMEL* para el período comprendido entre enero de 2007 hasta junio de 2016¹⁰.

El porcentaje considerado para la provisión se calculó como el promedio del rango de pérdida esperada para cada categoría según se presenta en el cuadro 2.17.

El total provisionado asciende a 242,428 miles de US\$, lo que corresponde a 3.33% de los depósitos cubiertos y 0.98% del total de depósitos asegurados de los bancos. Mientras representa 3.16% del total de depósitos

¹⁰Este período corresponde al mismo que se consideró para la estimación de los modelos econométricos.

APÉNDICE 2. ESTIMACIÓN DE LOS MODELOS LOGIT PARA LAS
116 ENTIDADES FINANCIERAS

cubiertos y 0.92 % del total de depósitos asegurados del Sector Financiero Privado.

Bancos ^a	Total depósitos ^b cubiertos	Nivel de riesgo	Provisión ^b
EFI 1	2,637,286	Bajo	79,119
EFI 2	825,364	Bajo	24,761
EFI 3	789,798	Bajo	23,694
EFI 4	780,193	Bajo	23,406
EFI 5	497,857	Bajo	14,936
EFI 6	406,837	Bajo	12,205
EFI 7	392,352	Bajo	11,771
EFI 8	199,811	Bajo	5,994
EFI 9	194,213	Bajo	5,826
EFI 10	124,701	Bajo	3,741
EFI 11	102,371	Medio Bajo	12,796
EFI 12	65,655	Medio Bajo	8,207
EFI 13	41,860	Medio Bajo	5,232
EFI 14	15,890	Medio Bajo	2,258
EFI 15	8,541	Medio	3,374
EFI 16	7,330	Medio	2,895
EFI 17	4,628	Medio	1,828
EFI 18	12,826	Bajo	385
TOTAL			242,428

^aConstan todas las entidades a excepción de las consideradas en el análisis previo.

^bLos valores están en miles de US\$

Cuadro 2.29: Provisiones bancos

2.9 Provisiones mutualistas

Respecto a las mutualistas, el cuadro 2.30 presenta valores para una sola mutualista, entidad que por su nivel de riesgo fue excluida del análisis previo.

El monto asciende a 7,210 en miles de US\$, valor que representa 2.59 % del total de depósitos cubiertos de mutualistas y 1.24 % de sus depósitos

asegurados. Lo que corresponde a 0.09 % del total de depósitos cubiertos y 0.03 % del total de depósitos asegurados del Sector Financiero Privado.

Mutualistas	Total depósitos ^a cubiertos	Nivel de riesgo	Provisión ^b
EFI 19	57,680	Medio Bajo	7,210
TOTAL			7,210

^aLos valores están en miles de US\$

Cuadro 2.30: Provisiones mutualistas

2.10 Provisiones sociedades financieras

Finalmente, se incluye dentro de las provisiones las sociedades financieras que no forman parte del análisis econométrico previo.

El monto asciende a US\$ 23,359 miles, lo cual corresponde a 18.70 % de los depósitos cubiertos y 2.41 % de los depósitos asegurados de las sociedades financieras. Mientras que representa 0.30 % de los depósitos cubiertos y 0.09 % del total de depósitos asegurados del Sector Financiero Privado.

Sociedades Financieras ^a	Total depósitos ^b cubiertos	Nivel de riesgo	Provisión ^b
EFI 20	93,218	Medio Bajo	11,652
EFI 21	21,764	Medio	8,597
EFI 22	6,181	Medio	2,441
EFI 23	1,414	Medio Bajo	177
EFI 24	1,227	Medio	485
EFI 25	248	Bajo	7
TOTAL			23,359

^aConstan todas las entidades a excepción de las consideradas en el análisis previo.

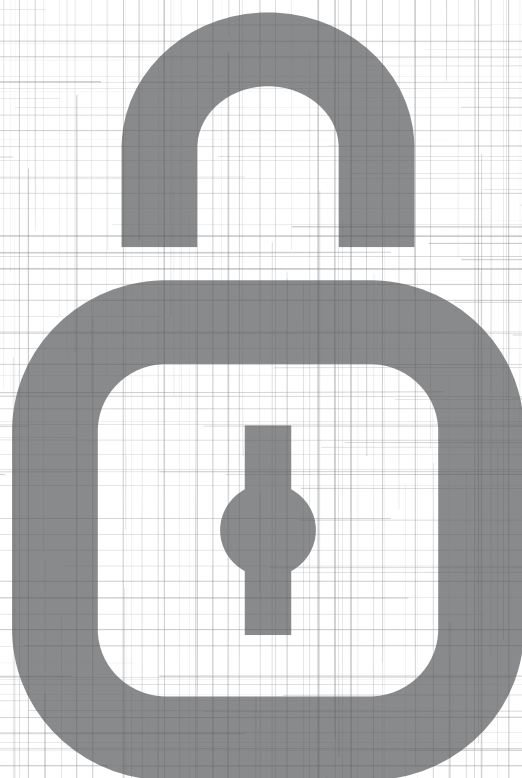
^bLos valores están en miles de US\$

Cuadro 2.31: Provisiones sociedades financieras

Autor: Econ. Andrea Bayancela Espinel

Coordinación y edición: Econ. David Villamar Cabezas
y Econ. Eugenio Paladines Camacho

Diciembre 2016



© Corporación del Seguro de Depósitos, Fondo de Liquidez
y Fondos de Seguros Privados (COSEDE)
Av. 12 de Octubre N26-97 y Abraham Lincoln Edificio 1492 Planta Baja
Telf: 593-2 2986510
www.cosedec.gov.ec