



UNIVERSIDAD DE CHILE  
FACULTAD DE CIENCIAS FÍSICAS Y MATEMÁTICAS  
DEPARTAMENTO DE INGENIERÍA INDUSTRIAL

DESAFÍOS DE LA INDUSTRIA BANCARIA EN LA ESTIMACIÓN DE LOSS GIVEN  
DEFAULT

MEMORIA PARA OPTAR AL TÍTULO DE INGENIERO CIVIL INDUSTRIAL

VÍCTOR DAMIÁN CUEVAS VÉLIZ

PROFESOR GUÍA:  
CARLOS PULGAR ARATA

MIEMBROS DE LA COMISIÓN:  
JAVIER SUAZO SÁEZ  
LUIS LLANOS COLLADO

SANTIAGO DE CHILE  
2019

**RESUMEN DE LA MEMORIA PARA OPTAR  
AL TÍTULO DE INGENIERO CIVIL  
INDUSTRIAL  
POR: VÍCTOR DAMIÁN CUEVAS VÉLIZ  
FECHA: 10/07/2019  
PROFESOR GUÍA: CARLOS PULGAR  
ARATA**

## **DESAFÍOS DE LA INDUSTRIA BANCARIA CHILENA EN LA ESTIMACIÓN DE LOSS GIVEN DEFAULT**

La regulación bancaria chilena, en línea con lo propuesto por estándares regulatorios internacionales, modela el monto que un banco debiese tener provisionado como el producto de tres variables: PD, EAD y LGD. El presente trabajo se centra en el estudio de la LGD, o Loss Given Default, que corresponde a la fracción de un crédito que no será recuperada en caso de que esté en incumplimiento.

Se busca estudiar los distintos modelos que existen en la literatura para la estimación de la LGD. En particular, se busca proponer un modelo y sus posibles variaciones utilizando datos de créditos de consumo chilenos, para así analizar los valores de LGD y su interacción con factores de riesgo del deudor.

Para esto, se segmenta la cartera de créditos según sus valores de LGD, y con un árbol de decisión se modelan tanto las probabilidades de pertenecer a cada grupo, como los valores que tienen dado que están en cada grupo.

Se concluye que deudores con menos renta, con montos castigados altos o sin un crédito hipotecario adicional tienen un valor esperado de LGD mayor, mientras que la relación entre la LGD con el monto y la antigüedad del crédito varía según el modelo utilizado. Los modelos propuestos pueden ser complejizados por cada institución agregando otras variables o incorporando el proceso de cobranza.

La siguiente sección de este trabajo muestra que, pese a la diversidad de modelos a elegir, pueden existir diferencias importantes en los valores de LGD al cambiar la forma en que es definida. La definición de default y cura son fijadas normativamente, pero hay otros factores que son decididos por cada banco. En particular, la LGD muestra una variación de un 1,9% al cambiar la tasa de descuento y de más de un 12% dependiendo de la medición de gastos que se elija.

Se concluye sobre la importancia de estandarizar los supuestos utilizados para el cálculo de LGD, dado que diferencias en los valores estimados tienen un impacto directo en el cálculo de provisiones y capital.

## TABLA DE CONTENIDO

1. INTRODUCCIÓN.....	1
2. LGD TEÓRICA .....	6
2.1. Pérdida Esperada .....	6
2.2. LGD.....	7
3. DATOS .....	9
3.1. Cálculo de una LGD por crédito .....	9
3.2. Gastos.....	10
3.3. Variables explicativas.....	11
4. DESCRIPCIÓN Y COMPARACIÓN DE MODELOS.....	14
4.1. Cálculo de una LGD promedio .....	14
4.2. Árboles de decisión.....	15
4.2.1. Probabilidad de pertenecer al grupo LGD = 0 o LGD > 0.....	17
4.2.2. Probabilidad de pertenecer al grupo LGD entre 0 y 1 o LGD = 1 .....	19
4.2.3. LGD entre 0 y 1 .....	22
4.2.3.1. Regresión lineal .....	23
4.2.3.2. Regresión logística.....	24
4.2.3.3. Regresión beta.....	26
4.2.4. Comentarios generales.....	28
4.2.5. Comparación con otros trabajos de LGD.....	29
4.3. Otros modelos.....	31
5. FACTORES QUE AFECTAN LA ESTIMACIÓN DE LGD .....	32
5.1. Definición de default.....	32
5.2. Definición de cura .....	34
5.3. Tasa de descuento.....	37
5.4. Costos.....	39
5.5. Horizonte de recuperación .....	42
5.6. Recesiones .....	43
6. CONCLUSIÓN.....	45
7. BIBLIOGRAFÍA.....	47

## 1. INTRODUCCIÓN

El Comité de Basilea fue establecido por los bancos centrales de Alemania, Bélgica, Canadá, Estados Unidos, Francia, Italia, Japón, Países Bajos, Reino Unido, Suecia y Suiza en el año 1974, como respuesta a una serie de disturbios en los mercados de divisas y la industria bancaria en dicha época, en particular, la quiebra del banco alemán Herstatt Bank.

Su objetivo era potenciar la estabilidad financiera promoviendo una mayor calidad en la regulación bancaria, sirviendo como un espacio de cooperación entre sus países miembros. Desde su creación, el comité pasó de estar conformado por los países del G10 a ahora contar con 45 instituciones de 28 jurisdicciones.

A partir de este comité surgieron acuerdos que buscaban establecer requerimientos mínimos de capital para que un banco pueda evitar su insolvencia en caso de algún evento crediticio. El primer acuerdo, Basilea I, fue aprobado en el año 1988, Basilea II, en el año 2004 y Basilea III surgió como respuesta a la crisis subprime.

Uno de los puntos que fue propuesto por el comité fue que los bancos provisionaran de acuerdo a la Pérdida Esperada (en inglés Expected Loss) que lograran calcular. Las provisiones que un banco tiene son montos que buscan cubrir pérdidas que éste tiene durante su ejercicio bancario, por ejemplo, cuando ante condiciones normales deudores dejan de pagar sus créditos.

Una de las razones para provisionar de acuerdo al cálculo de una Pérdida Esperada es buscar un valor que se anticipe a potenciales pérdidas futuras, en lugar de solo ser un valor histórico. Según lo propuesto en Basilea, la Pérdida Esperada será el producto de tres cantidades: PD (Probability of Default), EAD (Exposure At Default) y LGD (Loss Given Default).

La idea general es que un deudor puede estar en default con una probabilidad PD, y en dicho caso el banco perdería el monto de su crédito EAD. Pero no necesariamente perderá la totalidad del monto, dado que puede llegar a recuperar una parte, por ejemplo, gracias a acciones de cobranza. El porcentaje del crédito que no pudo recuperar de dicho crédito será la LGD.

El estado de default, también llamado incumplimiento, busca capturar la percepción por parte del banco de que los pagos futuros de dicho deudor son menos probables. No existe una única forma de definir un default, por lo que al cambiar las condiciones necesarias para que un crédito esté en incumplimiento, los valores estimados de Pérdida Esperada pueden cambiar. Dado esto, la Comisión para el Mercado Financiero (CMF), que es la entidad reguladora de la industria bancaria chilena, entrega una definición estandarizada a utilizar por la industria.

Precisando, el primer término, PD es la abreviación de Probability of Default, y representa la probabilidad que tiene un cliente de caer en incumplimiento. Por tratarse de una probabilidad, es una cantidad que puede variar entre 0 y 1. El segundo término es EAD, en inglés Exposure At Default, y representa el monto que un banco estaría expuesto a perder en caso de que un crédito esté en default. Finalmente, la LGD, en inglés Loss Given Default, representa la porción de exposición que un banco efectivamente perdería en caso de un default. Un deudor puede estar en default, pero luego de esto el banco puede recuperar cierta cantidad de dicha exposición. Este último término será la variable de interés para el presente trabajo.

Luego, la LGD representa qué tanto no fue recuperado de la exposición del crédito. Es por eso que también se usa la convención  $LGD = (1 - R)$ , con  $R$  la tasa de recuperación. Para su cálculo se necesita contabilizar todos los flujos que un banco recibe luego de un default, que serán recuperaciones. Por otro lado, para lograr dichas recuperaciones el banco pudo haber incurrido en acciones de cobranza, lo que representa un costo y por lo tanto una pérdida a restar de los montos que había recuperado.

Estos flujos económicos deben ser descontados a la fecha del default con alguna tasa de descuento apropiada, para así hacerlos comparables con el monto EAD que había en el momento de un default, y así obtener una tasa de recuperación. Al igual que con la definición de default, tampoco existe una tasa de descuento correcta a utilizar, por lo que la elección de dicha tasa también afectará el valor de LGD y con ello el cálculo de Pérdida Esperada.

Al tratarse de una proporción, se espera que los valores de LGD estén siempre entre cero y uno, pero éste no es necesariamente el caso. Por ejemplo, un banco pudo haber incurrido en gastos de cobranza, pero no lograr ninguna recuperación, por lo que no solo perdió la totalidad de la exposición de dicho crédito (que sería una  $LGD=100\%$ ), sino que también perdió dichos gastos. Su LGD en ese caso será mayor a 100%.

Chile, al igual que otros países, ha adaptado las directrices normativas entregadas por el Comité de Basilea a sus regulaciones internas. En particular, el 3 de octubre del 2018 se aprobó en la Cámara de Diputados la nueva Ley General de Bancos, que busca adoptar los estándares regulatorios de Basilea III.

De la misma forma que en los Acuerdos de Basilea, la regulación chilena también basa los requerimientos de provisiones que un banco debiese tener en la Pérdida Esperada, y también como el producto de PD, EAD y LGD.

Los valores de estos términos que un banco utilizará, podrán ser estimados a partir de modelos internos, los que bajo aprobación del regulador podrán ser utilizados para el cálculo de provisiones.

Por otro lado, el banco puede utilizar los valores propuestos por la CMF, que se denomina modelo estándar de provisiones para la cartera grupal. En el Compendio de Normas Contables<sup>1</sup> estos valores están definidos para créditos hipotecarios, operaciones de leasing comercial, préstamos estudiantiles y colocaciones comerciales genéricas y de factoraje.

Modelar la PD ha sido el objetivo de sistemas de crédito por más de 50 años, dado que se buscaban modelos de originación que permitieran entender mejor la capacidad de pago de los deudores a los que les otorgarían créditos. Este no ha sido el caso de la LGD, que fue estudiada en mayor detalle con la llegada de los Acuerdos de Basilea<sup>2</sup>.

Proponer y calibrar modelos de LGD supone limitaciones respecto a otros modelos de PD, por ejemplo, asociadas al menor número de datos con los que se cuenta. Al comparar directamente la cantidad de datos, la probabilidad de default representará qué tantos clientes cayeron en default, del total de todos los créditos vigentes, mientras que la LGD representará la proporción de pérdida que se incurrirá solo de aquellos créditos en default. La PD contará con la totalidad de una cartera de créditos para realizar sus estimaciones, mientras que LGD solo contará con aquellos que cayeron en default, que en ciertas carteras podrían incluso no existir.

El horizonte temporal para el que se contabilizarán las recuperaciones y pérdidas luego del incumplimiento también supone una limitación para el número de datos. El contabilizar recuperaciones que ocurran dentro de los 24 meses luego del default, significará que los créditos que estén en incumplimiento hace menos de 24 meses aún estarán sin resolver. Esto significa que no se podrían utilizar todos estos créditos como datos, o al hacerlo se debe tener la consideración de que sus valores pueden cambiar en los próximos meses.

Adicionalmente, la LGD históricamente tiene una distribución que se escapa de modelos clásicos como la distribución normal. En primer lugar, si no se consideran los gastos, los valores de LGD están entre cero y uno, por lo que no se ajusta a distribuciones de probabilidad que toman cualquier número real.

En segundo lugar, la distribución de los datos dentro de este intervalo suele ser bimodal<sup>3</sup>, es decir que recuperaciones muy altas o muy bajas tienen una mayor probabilidad de ocurrencia. En particular, existen casos en que una mayor proporción de los datos tiene  $LGD=0$  y  $LGD=1$ , y una menor cantidad de créditos con pérdidas entre esos valores. Al hablar de una distribución bimodal, nuevamente ciertas distribuciones clásicas como la normal no logran capturar este comportamiento.

Para analizar qué tanto ha sido estudiada la LGD en la industria local, se toma como referencia la Serie de Estudios Normativos de la Comisión para el Mercado Financiero (CMF). Se encuentran tres documentos relativos a modelos estándar para el cálculo de provisiones.

El primero, del año 2014<sup>4</sup>, propone tablas con valores de LGD subdividida con respecto a factores de riesgo para carteras hipotecarias residenciales. El documento de julio de 2018<sup>5</sup> genera una propuesta similar, pero para la cartera comercial, subdividida en leasing, estudiantil y créditos comerciales genéricos. Finalmente, el documento publicado en agosto de 2018<sup>6</sup> propone un marco general para el diseño de un modelo estándar, donde se presenta una fórmula discreta para el cálculo de la LGD por individuo, y menciona que la LGD podría tener una distribución beta o ZOIB.

Considerando todo lo expuesto anteriormente, la oportunidad de este trabajo está en explorar los distintos métodos existentes en la literatura para el modelamiento de la LGD.

Para esto, se estudia en mayor detalle un modelo de árbol de decisión para el modelamiento de la LGD, para luego ser aplicado en una base de créditos de consumo chilenos en incumplimiento.

Dentro de los modelos estándar propuestos por la CMF no existe uno para este tipo de créditos, por lo que este trabajo contribuirá a analizar el comportamiento de la LGD aplicada a dicha cartera. Adicionalmente, al analizar la literatura internacional sobre LGD tampoco se encontraron referencias a la utilización de créditos de consumo chilenos para su cálculo.

Por otro lado, se analizará cómo cambian los valores de LGD cuando cambian sus supuestos al, por ejemplo, usar una definición de default alternativa, cambiar la tasa de descuento a utilizada o al cambiar los cálculos de costos de cada banco.

Este análisis de sensibilidad permitirá entender los factores que afectan los valores de LGD si los bancos eligen realizar modelos internos para su estimación, por lo que estas propuestas permitirán que las áreas de riesgo de conozcan de forma más específica el riesgo asociado a su cartera y los montos a provisionar.

Adicionalmente, estas propuestas también van en el interés del organismo regulador, que adopta un rol macroprudencial para buscar la estabilidad del sistema financiero en su totalidad. El valor de la LGD tiene implicancias directas en el cálculo de provisiones que un banco tiene, y junto a la normativa introducida en la Ley General de Bancos también afecta los cálculos de capital.

Según lo expuesto, el objetivo general de este trabajo es:

Proponer un modelo para la estimación de la LGD usando créditos de consumo chilenos e identificar la sensibilidad a los factores que afectan su cálculo, para así proponer recomendaciones para calcular y modelar la LGD.

Y los objetivos específicos:

- Aplicar un modelo de estimación de LGD a datos reales de créditos de consumo chilenos.
- Describir las ventajas y desventajas del modelo propuesto y sus variaciones y de otros modelos no utilizados.
- Identificar los factores que afectan al cálculo de la LGD y calcular la sensibilidad a cambios de estos factores.
- Proponer recomendaciones para evitar la variabilidad en los valores de LGD entre distintas instituciones financieras.

Según los objetivos descritos, la metodología de este trabajo es:

- Realizar una revisión bibliográfica para identificar los modelos utilizados en la literatura para la estimación de la LGD.
- Elegir un modelo adecuado a los datos que se tiene.

- Aplicar el modelo y distintas variaciones de éste a datos de crédito de consumo de chilenos.
- Analizar los resultados para identificar las ventajas y desventajas de cada variación.
- Describir las ventajas y desventajas de los modelos no utilizados.
- Identificar factores que afectan al cálculo de LGD.
- Calcular numéricamente la sensibilidad de la LGD promedio a los distintos supuestos, según sea posible con los datos que se dispone.
- Generar recomendaciones basadas en una revisión bibliográfica para estandarizar estos supuestos y evitar diferencias entre las estimaciones de distintas instituciones financieras.

El presente informe comienza en la sección 2 explicando el valor teórico que tiene la LGD para cada crédito. En la sección 3 se describe la base de datos a utilizar, detallando cómo fue calculada la LGD y las variables explicativas a utilizar en los modelos. La sección 4 describe el cálculo de una LGD grupal, en particular se muestra el cálculo de una LGD promedio, un modelo de árbol de decisión y se enumeran modelos más complejos que se podrían utilizar. La sección 5 discute la sensibilidad de los valores de LGD según los supuestos con los que se construyó y la sección 6 termina con las conclusiones de esta memoria.



## 2. LGD TEÓRICA

### 2.1. Pérdida Esperada

La Pérdida Esperada (PE), en inglés Expected Loss (EL), es una cantidad que mide las pérdidas que un banco podría tener durante el transcurso de un ejercicio bancario normal, cuando, por ejemplo, una porción de los deudores deja de pagar sus créditos en dicho periodo. Es luego según este monto que los bancos deben calcular sus provisiones para estar preparados a cubrir estas pérdidas.

Como su nombre lo indica, la Pérdida Esperada es el valor esperado de la distribución de la pérdida que puede tener un banco, como se muestra en la figura 1.

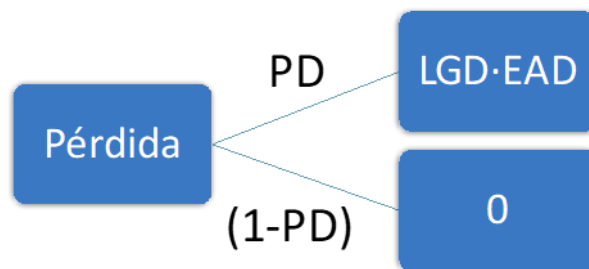


Figura 1. Distribución de pérdida en función de parámetros regulatorios.

Existe una probabilidad  $PD$  de que el deudor esté en default, por lo que el banco perdería la exposición  $EAD$  de dicho crédito. Pero aun cuando el deudor está en default, el banco podría incurrir en acciones de cobranza y recuperar una parte de dicho crédito, por lo que el banco solo perdería una porción  $LGD$  del monto  $EAD$ . Por otro lado, con una probabilidad  $(1 - PD)$  el deudor pagará, por lo que la pérdida para el banco sería cero.

Finalmente, la Pérdida Esperada (PE) según esta notación será:

$$PE = E(\text{Pérdida}) = PD \cdot LGD \cdot EAD + (1 - PD) \cdot 0$$
$$PE = PD \cdot LGD \cdot EAD$$

$PD$  es la Probabilidad de Incumplimiento (en inglés Probability of Default),  $LGD$  es la Pérdida Dado Incumplimiento (en inglés Loss Given Default) y  $EAD$  es la exposición (en inglés Exposure At Default). Estas tres variables tienen implícitamente una definición de default.

El default o incumplimiento corresponde a un punto en que se considera que las recuperaciones de un deudor y sus créditos será remota, dado que muestra una deteriorada o nula capacidad de pago. Regulatoriamente en Chile según el Compendio de Normas Contables<sup>1</sup>, un deudor estará en default si al cierre de un mes cumple alguna de las siguientes condiciones:

- Atraso igual o superior a 90 días en el pago de intereses o capital de algún crédito.

- Se les otorgue un crédito para dejar vigente una operación que presentaba más de 60 días de atraso en su pago.
- Hayan sido objeto de reestructuración forzosa o condonación parcial de una deuda.

Es decir, cuando un deudor cumpla alguna de las tres condiciones anteriores, el 100% del monto del capital de sus créditos será la EAD, y la proporción de dicha exposición que el banco no logrará recuperar será la LGD.

## 2.2. LGD

La LGD, Pérdida Dado el Incumplimiento o severidad de la pérdida (en inglés Loss Given Default), corresponde a la proporción de una deuda que no es recuperada en caso de un default. Alternativamente, si  $R$  es la tasa de recuperación de dicho crédito, la LGD puede describirse como  $LGD = (1 - R)$ .

Para ejemplificar lo anterior, en la figura 2 se muestra la historia de un crédito. El punto 1 corresponde al inicio de un crédito por un monto de 200. Luego de sucesivos pagos, el deudor cae en default en el punto 2, con una exposición (EAD) de 100. En el punto 3 el banco gasta 10 en acciones de cobranza, gracias a lo cual el deudor paga 50 de lo que debía.

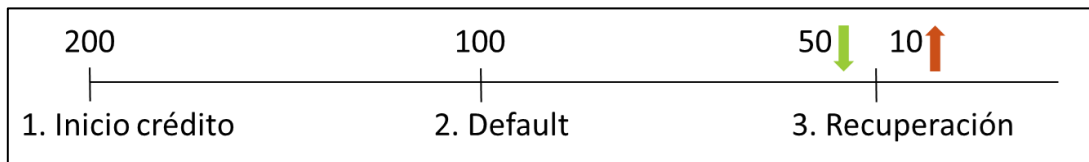


Figura 2. Ejemplo de la historia de un crédito.

El banco logró recuperar  $50 - 10 = 40$ , pero este monto sigue sin ser comparable con los 100 de la exposición al default, por lo que estos flujos se descuentan a la fecha de incumplimiento. Si el factor de descuento es 0,9, la recuperación al momento de default sería  $0,9 \cdot (50 - 10) = 36$  y luego la tasa de recuperación de dicho crédito sería:

$$R = \frac{\text{Monto recuperado}}{\text{Exposición}} = \frac{36}{100} = 36\%$$

Con esto, la LGD sería  $LGD = (1 - R) = (1 - 36\%) = 64\%$ . Es decir que, de una exposición de 100, el banco perdió un 64%.

Matemáticamente, la LGD para un crédito  $i$  en el período  $t$  será:

$$LGD_{it} = \text{Máx} \left\{ 1 - \frac{\sum_{j>\tau_{io}}^{\tau_{io}+\Delta_i} \text{rec}_{ij}(r_i) - \text{gas}_{ij}(r_i)}{EAD_i}, 0 \right\} = \text{Máx} \left\{ 1 - \frac{\text{Rec}_{it}(r_i) - \text{Gas}_{it}(r_i)}{EAD_i}, 0 \right\}$$

Es decir, la suma de los flujos recuperados y gastados desde el período de default  $\tau_{i0}$  durante un horizonte de recuperación  $\Delta_i$ . Estos flujos son descontados con una tasa de descuento  $r_i$ , de acuerdo a sus respectivos plazos.

Cabe destacar que la fórmula es una función máximo, que evita que los valores de LGD puedan ser negativos.

### 3. DATOS

Para el presente trabajo se utilizan datos de deudores innominados del sistema bancario chileno, informados por los bancos a través de la CMF, los cuales fueron analizados y trabajados en la misma institución. Los datos son utilizados únicamente para los modelos y análisis descritos en este informe, de acuerdo a la confidencialidad que requiere su uso.

La base está compuesta por créditos de consumo chilenos que cayeron en default entre julio del 2009 hasta diciembre del 2011.

Del total de créditos, 7.609 fueron castigados. El castigo corresponde a la baja del Estado de Situación Financiera del activo correspondiente a la respectiva operación. En el caso de la regulación chilena, el Compendio de Normas Contables<sup>1</sup> define que un crédito será castigado si alguna de las siguientes condiciones se cumple:

- El banco, basado en toda la información disponible, concluye que no obtendrá ningún flujo de la colocación registrada en el activo.
- Cuando una acreencia sin título ejecutivo cumpla 90 días desde que fue registrada en el activo.
- Al cumplirse el plazo de prescripción de las acciones para demandar el cobro mediante un juicio ejecutivo o al momento del rechazo o abandono de la ejecución del título por resolución judicial ejecutoriada.
- Cuando el tiempo de mora de una operación alcance el plazo para castigar, que en el caso de créditos de consumo corresponde a 6 meses.

Los 11.512 créditos en default restantes no fueron castigados. De dichos créditos, solo se considerarán aquellos que luego de cuatro meses ya no están en incumplimiento. Estos créditos durante el presente trabajo son llamados créditos curados, que son créditos que pese a estar en incumplimiento, no son castigados dado que el deudor logra pagar la deuda.

Con la definición de cura descrita, los créditos no castigados a utilizar que cumplen esta condición serán 4.386. La elección y reducción de la base será ilustrativa de cómo cambiarán los valores de LGD al considerar más o menos casos curados en la sección 5.2.

Dicho lo anterior, al unir los créditos que fueron castigados y los que fueron curados se consolida una base de 11.995 créditos en incumplimiento.

#### 3.1. Cálculo de una LGD por crédito

La LGD para cada crédito es calculada como:

$$LGD_i = \begin{cases} 1 - Rf_i, & \text{Si crédito fue castigado} \\ 0, & \text{Si crédito no fue castigado} \end{cases}$$

Donde  $Rf_i$  corresponde a la tasa de recuperación de dicho crédito sin considerar los gastos que se incurrieron para la recuperación. Los flujos son descontados a la fecha de incumplimiento usando la tasa de consumo promedio de dicho banco. La tasa de recuperación se obtiene de la base de castigos.

Cuando un crédito es curado, es decir, estuvo en default, pero luego retomó sus pagos de forma normal, su pérdida es cero, por lo que su LGD es cero.

Cabe destacar que este cálculo de LGD es un aproximado. La tasa  $Rf_i$  es calculada como:

$$Rf_i = \frac{\sum_{j>C_{i0}}^{C_{i0}+24} rec_{ij}(r_i)}{\text{Monto Castigado}}$$

Que contabiliza todas las recuperaciones que ocurran dentro de los 24 meses luego de la fecha  $C_{i0}$  de castigo. Lo correcto sería dividir las recuperaciones por la exposición al default (EAD) y no el monto castigado. Se discute sobre los efectos de elegir el horizonte de recuperaciones en 24 meses en la sección 5.5.

### 3.2. Gastos

El cálculo anterior no considera los gastos que incurre cada banco en cada operación, dado que no se dispone de esa información para cada crédito. Los bancos reportan las recuperaciones y gastos totales en los que incurrió de forma trimestral.

Dada esta limitación, para poder incorporar los costos, se propone calcular una tasa  $h_{it}$  que se debería incorporar al valor final de LGD, de la siguiente forma:

$$h_{it} = \frac{Rec_{it}(r_i) - Gas_{it}(r_i)}{Rec_{it}(r_i)}$$

Donde  $Rec_{it}(r_i)$  y  $Gas_{it}(r_i)$  es la suma de recuperaciones y la suma de gastos descontados usando la tasa  $r_i$ , respectivamente, para cada trimestre  $t$ . Esta tasa  $h_{it}$  es llamada tasa de recuperación efectiva durante el presente trabajo, y corresponde al porcentaje de la recuperación trimestral que efectivamente se recibe al restarle los gastos.

En un periodo dado, la relación entre  $h_i$ , la tasa de recuperación sin costos  $Rf_i$  usada y la tasa de recuperación real  $R_i$  será:

$$R_i = \frac{Rec_i(r_i) - Gas_i(r_i)}{EAD_i} = \left( \frac{Rec_i(r_i) - Gas_i(r_i)}{Rec_i(r_i)} \right) \cdot \left( \frac{Rec_i(r_i)}{EAD_i} \right) = h_i \cdot Rf_i$$

Por lo que la LGD real sería:

$$LGD_i = (1 - R_i) = (1 - h_i \cdot Rf_i)$$

Se dispone de información de gastos y recuperaciones trimestrales para 12 instituciones financieras entre diciembre del 2013 y marzo del 2017. La tasa de recuperación efectiva promedio para cada institución durante este periodo varía entre 74,12% y 99,87%. Se decide utilizar la tasa promedio de todas las instituciones  $h = 89,011\%$ . Se discute sobre efectos de la tasa promedio  $h$  elegida en el valor de LGD en la sección 5.4.

### 3.3. Variables explicativas

Las variables que se tienen para un deudor en incumplimiento y sus respectivas descripciones se muestran en la tabla 1.

Variable	Descripción
Renta	Renta anual del deudor, en pesos chilenos, que corresponde al promedio de todos los valores reportados por las distintas instituciones financieras.
Colocación	Corresponde al monto por pagar de todos los créditos de consumo que tiene el deudor en dicho banco. En pesos chilenos.
Hipotecario	Variable que vale 1 si el deudor posee, adicionalmente al crédito de consumo, un crédito hipotecario y 0 si no.

Tabla 1. Variables explicativas al momento del default.

Respecto a la variable Renta, la inclusión de ésta para los modelos de la sección 4 radica en el potencial que puede tener en explicar si un deudor es más o menos riesgoso y tiene una mejor o peor capacidad de pago en caso de un default.

Similarmente, en el caso de la variable Colocación se busca explicar si el tamaño del crédito puede o no influir en si un deudor tiene o no capacidad de repagar su deuda en caso de incumplimiento.

La variable Hipotecario muestra si un cliente tiene o no un crédito hipotecario adicional. Al tener un crédito hipotecario, un deudor tiene un inmueble al cual un banco puede recurrir en caso de un default si este inmueble es una garantía general. El proceso podría tener más dificultades si el inmueble fuese una garantía específica para el crédito hipotecario. Los créditos de consumo usados en el presente informe no tienen colateral, por lo que en el caso de un incumplimiento el banco arriesga perder la totalidad del crédito.

Adicionalmente a las cuatro variables anteriores, se tienen variables para aquellos créditos que fueron castigados, por lo que la parte de los créditos que no fue castigada no cuenta con estos valores. Las variables se muestran en la tabla 2.

Variable	Descripción
Castigo	Monto que fue castigado por el banco, es decir, sacado de sus libros contables. En UF.
Antigüedad	Meses que pasaron entre que el crédito fue originado y cuando fue castigado.

Tabla 2. Variables explicativas al momento del castigo.

Al igual que en el caso anterior, se busca explicar el riesgo y la capacidad de pago en caso de default tanto con el tamaño del monto castigado, como con los meses de antigüedad del crédito.

Para analizar la distribución de las variables en los datos, en la tabla 3 se muestran los percentiles 25, 50 y 75.

Percentil	Renta (MM\$)	Colocación (MM\$)	Hipotecario	Castigo (UF)	Antigüedad (meses)
25	2,61	0,18	0	123	6
50	4,48	0,59	0	369	6
75	7,81	1,59	0	789	8

Tabla 3. Percentiles de variables explicativas.

Es decir, la mitad de los deudores de la muestra tienen una renta anual de menos de 4,48 millones de pesos, una colocación menor o igual a \$593.600, o un monto castigado menor o igual a 369 UF.

En el caso de la antigüedad tanto los percentiles 25 y 50 son iguales a 6 meses, lo que no es extraño considerando que un crédito de consumo con más de 6 meses de mora será castigado. Es decir, son deudores que luego del inicio del crédito no pagaron ninguna cuota, por lo que directamente a los 3 meses estuvieron en incumplimiento y a los 6 meses en castigo.

Todos los percentiles de la variable Hipotecario son cero, esto se debe a que solo 670 deudores de la muestra poseen un crédito hipotecario.

En la tabla 4 se muestra la matriz de correlación entre las variables. Se destaca que la correlación más alta es entre la colocación de consumo y el monto castigado, lo que es razonable considerando que mayores montos castigados vengan de créditos de mayor monto. Cabe destacar que una alta correlación podría generar coeficientes no significativos o simplemente distorsionar los valores de los parámetros estimados.

La segunda correlación más alta es entre el monto del crédito y la renta promedio. De forma similar, esta relación nos indica que puede existir una relación donde ciertos deudores con mayores rentas son los que poseen los créditos de mayor monto.

	LGD	Renta	Colocación	Hipotecario	Castigo	Antigüedad
LGD	1	-0,21	-0,13	-0,17	0,056	0,013
Renta	-0,21	1	0,36	0,14	0,17	-0,029
Colocación	-0,13	0,36	1	0,14	0,56	0,11
Hipotecario	-0,17	0,14	0,14	1	0,06	0,029
Castigo	0,056	0,17	0,56	0,06	1	0,19
Antigüedad	0,013	-0,029	0,11	0,029	0,19	1

Tabla 4. Matriz de correlación de variables explicativas.



## 4. DESCRIPCIÓN Y COMPARACIÓN DE MODELOS

### 4.1. Cálculo de una LGD promedio

La forma más básica del cálculo de LGD según Til Schuermann (2014)<sup>3</sup> es el cálculo de una LGD promedio para una cartera de créditos. Teniendo la variable  $LGD_i$  para cada individuo, la LGD promedio simple de la muestra será:

$$LGD = \frac{\sum_{i=1}^N LGD_i}{N}$$

Con N el número de créditos, en este caso 11.995.

La LGD promedio es 51,702% y al incorporar la tasa de recuperación efectiva promedio  $h = 89,011\%$  de la sección 3.2, la LGD promedio es 57,009%. Lo anterior nos dice que, del monto total de los créditos en default, los bancos lograron recuperar en promedio el 57,009%.

En este caso se tomó el promedio sobre la muestra completa, pero se podría subdividir la cartera de acuerdo a ciertas características de cada individuo, como por ejemplo el tipo de crédito o la antigüedad del cliente en el banco, obteniendo así un valor promedio de LGD en cada segmento.

Esta metodología de crear tablas con valores de LGD promedio para cada segmento es la que se puede encontrar en el Modelo Estándar del Compendio de Normas Contables de la CMF para créditos hipotecarios, préstamos estudiantiles y operaciones de leasing comercial<sup>1</sup>.

Cabe destacar la dependencia de los valores de LGD según la segmentación elegida. Al realizar evaluaciones de provisiones grupales como es este caso, el Compendio de Normas Contables requiere la conformación de grupos de créditos con características homogéneas en cuanto a tipo de deudores y condiciones pactadas.

Al calcular estos valores, también existe sensibilidad a los periodos de estimación, siendo posible calcular dicho promedio sobre un periodo largo (incluso todos los datos, como es el caso de este trabajo) o sobre una ventana móvil de algún largo definido.

La ventaja de calcular este valor promedio, según Schuermann, es que es relativamente fácil de calcular, dado que no se necesitan sofisticadas habilidades para su modelamiento y que es a la vez fácil de implementar.

Sin embargo, este enfoque se ve limitado por la cantidad de datos que se pueden obtener de LGD como se explicó anteriormente, lo que al segmentar la cartera reduce aún más los datos en cada segmento.

## 4.2. Árboles de decisión

Thomas et al<sup>2</sup> plantea otro modelamiento para créditos a personas sin garantías usando árboles de decisión. Un árbol de decisión para modelar LGD busca representar en un diagrama las distintas formas en que podría ser calculada la LGD. Por ejemplo, la LGD no es solo el resultado de la habilidad para pagar o no del cliente, sino también de otros factores, tales como cuáles son las acciones de un banco para cobrar dicha deuda: cobranza hecha por el banco, servicios de cobranza externos o venta de la deuda a un tercero.

Los distintos resultados del proceso de cobranza permiten entender a priori los posibles valores que puede tomar la LGD, independiente de la capacidad de pago del deudor. Por ejemplo:

- Cobranza hecha por el banco (sin considerar costos):  $0\% \leq LGD \leq 100\%$ .
- Cobranza por agente externo con comisión del 40%:  $40\% \leq LGD \leq 100\%$ .
- Venta de deuda a un 5% del capital:  $LGD = 95\%$ .

Otros criterios para clasificar el cálculo de LGD pueden ser si el banco tiene o no registros para encontrar al deudor. Las formas de realizar estos diagramas pueden detallarse aún más, incluyendo, por ejemplo, las posibilidades más precisas en cuanto al envío de correos, y qué acciones se toman ante una respuesta exitosa o no. Un ejemplo de modelamiento las decisiones que puede tomar se muestra en la figura 3.

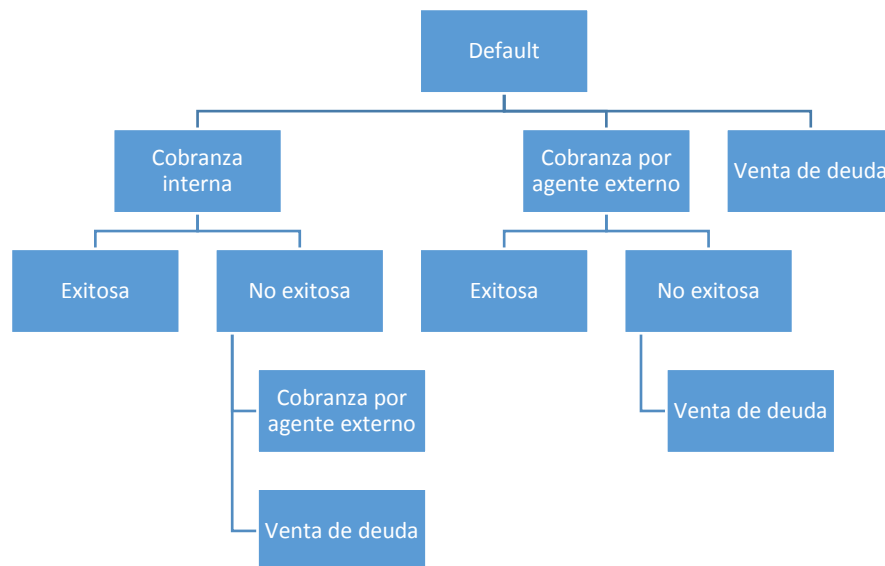


Figura 3. Árbol de decisión para un deudor en default.

Todo lo anterior representa el espectro de posibilidades de acciones que un banco puede tomar para buscar recuperaciones. A continuación, se presenta un modelo para describir las posibilidades que tiene un deudor para pagar dado que está en incumplimiento. Con esto, lo que se modela es el comportamiento de pago de un deudor de manera independiente a las acciones que el banco por sí mismo pueda tomar.

La primera pregunta que se busca responder es, dentro del total de defaults en la base de datos, cuántos tipos de clientes existen. En la figura 4 se muestra un histograma de cuánta fue la pérdida dado el incumplimiento de todos los deudores de la base de datos.

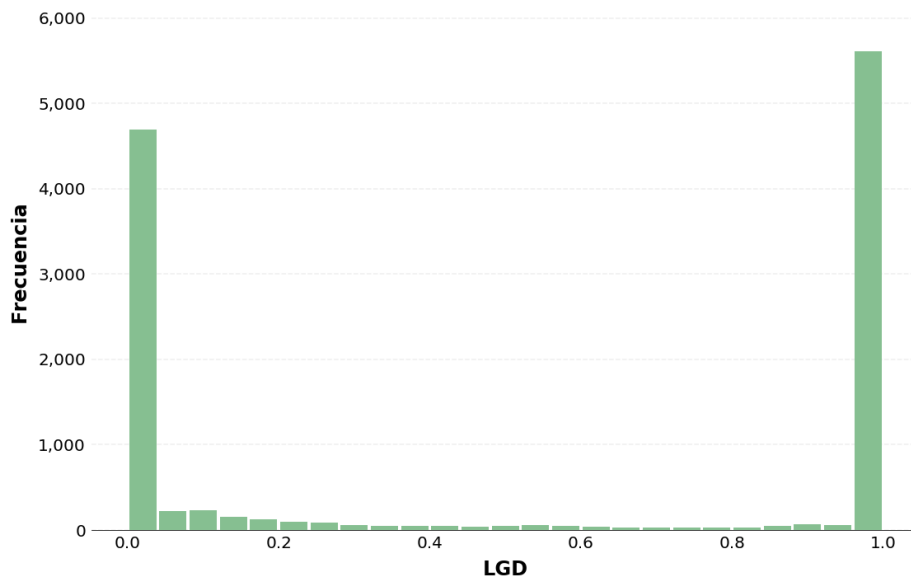


Figura 4. Histograma de LGD para todos los deudores.

Se observa que se pueden distinguir a priori tres tipos de créditos, aquellos para los cuales no hubo recuperación y que por lo tanto su LGD es uno, aquellos que tuvieron una pérdida total, por lo que su LGD es 1 y aquellos cuya LGD está entre cero y uno. Con esto, se genera un árbol simplificado que modela el comportamiento de estos tres tipos de deudores, que se muestra en la figura 5.

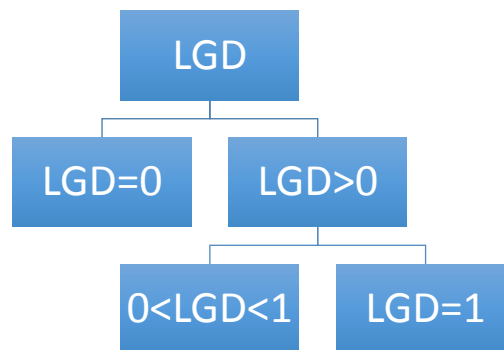


Figura 5. Árbol de decisión para un deudor en default

Es decir que, o bien su LGD es cero, o bien es mayor que cero. Luego, para aquellos créditos cuya LGD es mayor a cero, tendremos los casos de LGD igual a uno, o entre cero y uno.

El objetivo de realizar la primera separación es decir que existen diferencias en el comportamiento de pago entre el grupo de LGD=0 y LGD>0. El primer grupo cayó en

default, pero luego pagó toda su deuda y salió de dicha condición, por lo que no hubo ninguna pérdida. Este grupo corresponde a créditos curados, para los cuales se retomaron sus pagos sin necesidad de llegar al castigo

El segundo grupo son deudores que efectivamente generaron una pérdida para el banco. Estos deudores cayeron en default, luego fueron castigados, y finalmente el banco tuvo que realizar acciones de cobranza, gracias a las cuales pudieron recuperar una fracción o nada de la deuda.

Dado que un cliente estaba en el segundo grupo y tuvo una pérdida, el objetivo de realizar la segunda separación es decir que existen diferencias entre el grupo  $LGD \in (0,1)$  y  $LGD=1$ . El primer grupo cayó en default y solo pagó su deuda de forma parcial, mientras que el segundo no pagó nada de su deuda. Este último caso sería el de un deudor que cayó en default y al que el banco no pudo volver a contactar, o deudores para los cuales el banco decidió que no realizaría acciones de cobranza.

La LGD para un cliente del grupo  $LGD=0$ , dado que no tuvo pérdida, será cero, y la LGD para un cliente del grupo  $LGD=1$  que tuvo una pérdida total será uno. El modelamiento del árbol está compuesto de tres partes, las que serán descritas en el resto del capítulo. Cabe destacar que estas tres partes corresponden a tres procesos distintos, por lo que la relación entre las variables y la LGD no tiene por qué ser necesariamente la misma.

#### **4.2.1. Probabilidad de pertenecer al grupo $LGD = 0$ o $LGD > 0$**

Se busca modelar la probabilidad de que el cliente pertenezca a uno de los dos grupos, y al tratarse de una probabilidad, sus valores tienen que estar entre 0 y 1. Un modelo adecuado para responder a esta pregunta es una regresión logística.

La regresión logística requiere que la variable dependiente sea binaria, cuyos dos valores representan a ambos grupos. En este caso se define la variable  $LGD\_1$  como:

$$LGD\_1_i = \begin{cases} 0, & \text{si } LGD_i = 0 \\ 1, & \text{si } LGD_i > 0 \end{cases}$$

La distribución de los datos de esta variable se muestra en la figura 6. Un 31,2% de los créditos tiene una LGD igual a cero y un 61,8% tiene una LGD mayor a cero.

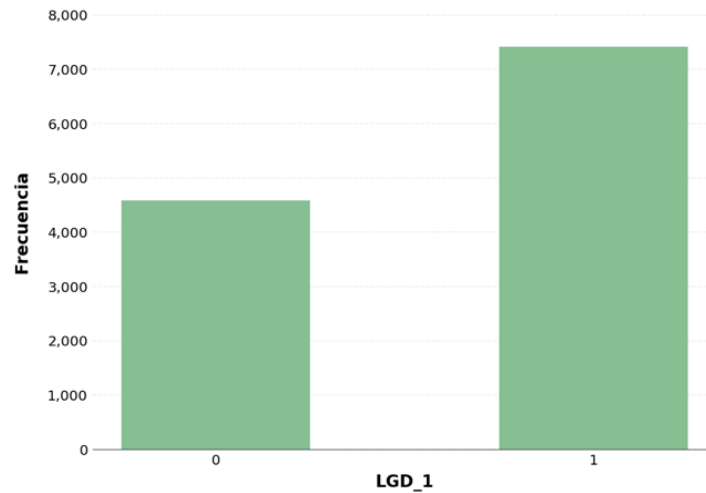


Figura 6. Histograma variable dependiente de regresión logística para primera probabilidad.

De la tabla 2, se utilizan las variables disponibles al momento del default, pero no las de castigo, dado que la mayoría de los créditos con LGD=0 en incumplimiento no fueron castigados. Estas variables son:

- Renta: Renta anual del deudor en pesos.
- Colocación: Colocaciones totales de consumo en dicho banco.
- Hipotecario: 1 si tiene crédito hipotecario, 0 si no.

Donde se toma el logaritmo natural de Renta y Colocación. Los resultados de la calibración de este modelo se muestran en la tabla 5.

Variable	Coeficiente	P-value
Constante	8,66	$5 \cdot 10^{-110}$
Renta	-0,55	$3,9 \cdot 10^{-102}$
Colocación	0,02	0,028
Hipotecario	-1,06	$2 \cdot 10^{-33}$

Tabla 5. Resultados regresión logística primera probabilidad.

Todas las variables y la constante del modelo son significativas al 5%.

Dada esta calibración, para una regresión logística la probabilidad  $p_i$  de que un deudor  $i$  pertenezca al grupo LGD>0 es:

$$p_i = \frac{e^{(\beta_0 + \beta_1 \cdot Renta + \beta_2 \cdot Colocación + \beta_3 \cdot Hipotecario)}}{1 + e^{(\beta_0 + \beta_1 \cdot Renta + \beta_2 \cdot Colocación + \beta_3 \cdot Hipotecario)}}$$

Y equivalentemente, la probabilidad  $(1 - p_i)$  de que un deudor  $i$  pertenezca al grupo LGD=0 es:

$$(1 - p_i) = \frac{1}{1 + e^{(\beta_0 + \beta_1 \cdot Renta + \beta_2 \cdot Colocación + \beta_3 \cdot Hipotecario)}}$$

Interpretando los signos de los coeficientes según la fórmula de  $(1 - p_i)$ , la probabilidad de que pertenezca al grupo LGD=0, aumentará:

- Si Renta aumenta. Lo que esto nos dice es que aquellos deudores con rentas mayores tienen una mayor probabilidad de pagar toda su deuda en caso de un default, sin que el banco incurra en ningún tipo de pérdida por dicho crédito. Este resultado concuerda con la intuición de que un deudor que tenga un ingreso mayor tendrá una mejor capacidad de pago.
- Si Colocación disminuye. Con esto, aquellos deudores con créditos de menor monto tendrán una mayor probabilidad de repagar toda su deuda si están en incumplimiento y en el mayor de los casos sin que los créditos sean castigados. Una potencial hipótesis a este resultado es que créditos con menor monto serán más fáciles de pagar. Usando un ejemplo, un deudor tendrá menos problemas pagando una deuda de \$200.000 que una de \$1.000.000.
- Si la variable Hipotecario aumenta. Esto muestra que deudores que adicionalmente posean un crédito hipotecario tendrán una mayor probabilidad de pagar la totalidad de su deuda en caso de default. En particular, evaluando el modelo en la mediana de la renta y en la mediana de la colocación, al pasar de Hipotecario = 0 a Hipotecario = 1, la probabilidad de pertenecer al grupo LGD = 0 aumenta de 0,35 a 0,61.

Una potencial hipótesis, directa de este resultado, es que el comportamiento de pago de deudores con otro crédito hipotecario adicional es mejor.

Por otro lado, deudores con un crédito hipotecario tienen una garantía para el banco. En caso de que el deudor deje de pagar su crédito hipotecario el banco podrá exigir la vivienda como pago. El hecho de que el banco tenga esta garantía permite que la posibilidad de repactación del crédito sea más probable que para un deudor sin garantía.

Adicionalmente, se calibra una versión alternativa de este modelo, donde se utilizan dummies para segmentar las variables Renta y Colocación, pero no se observa monotonía en sus coeficientes por lo que este modelo no es usado.

#### **4.2.2. Probabilidad de pertenecer al grupo LGD entre 0 y 1 o LGD = 1**

Dado que ya no se consideran los casos con LGD igual a cero, la distribución de los casos restantes se muestra en la figura 7.

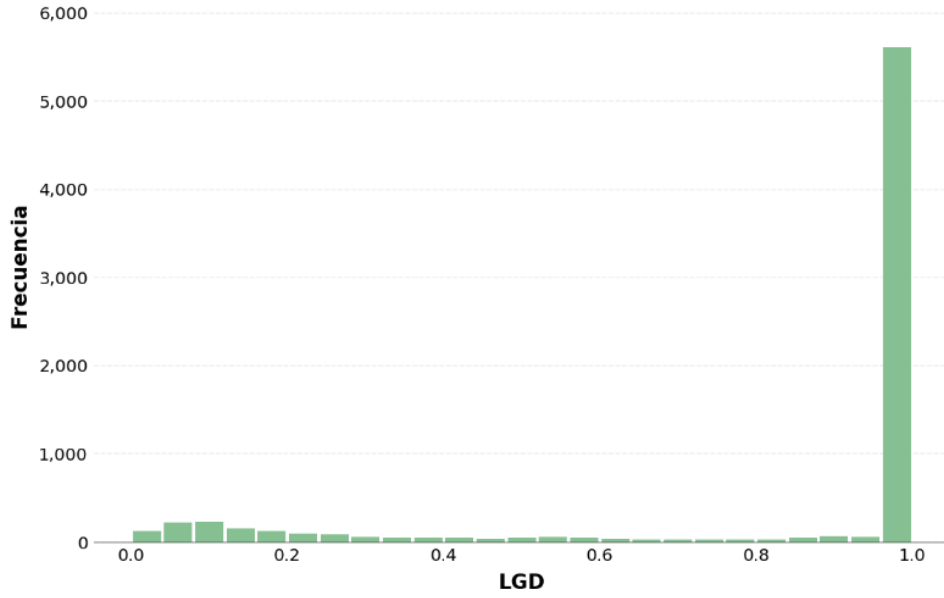


Figura 7. Histograma de LGD, valores mayores a cero.

En este caso nuevamente se busca modelar una probabilidad, por lo que la regresión logística usada en la sección 4.2.1 es un modelo adecuado. La variable dependiente que separa a ambos grupos es:

$$LGD\_2_i = \begin{cases} 0, & \text{si } LGD_i \in (0,1) \\ 1, & \text{si } LGD_i = 1 \end{cases}$$

La distribución de esta variable en los datos se muestra en la figura 8. Un 25,2% de los datos tiene una LGD entre cero y uno y un 74,8% tiene LGD=1.

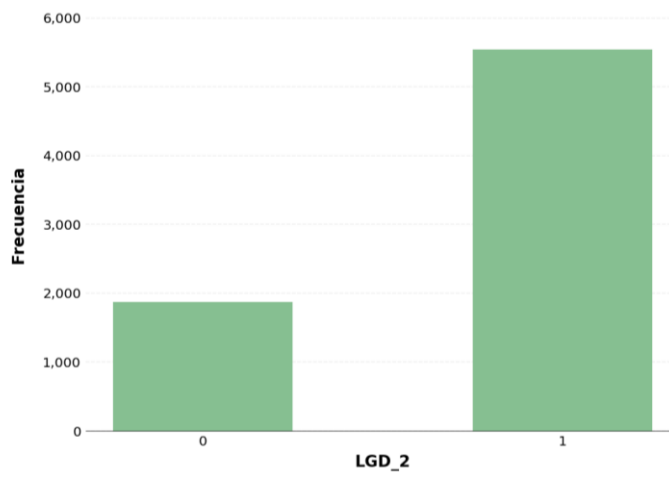


Figura 8. Histograma variable dependiente de regresión logística para segunda probabilidad.

Los créditos de ambos grupos fueron castigados, por lo que adicionalmente a las variables usadas en la sección 2.1 se pueden utilizar las de castigo de la tabla 3. Estas variables son:

- Renta: Renta anual del deudor en pesos.
- Colocación: Colocaciones totales de consumo en dicho banco.
- Hipotecario: 1 si tiene crédito hipotecario, 0 si no.
- Castigo: Monto castigado en UF.
- Antigüedad: Meses entre el inicio del crédito y su castigo.

Donde se toma logaritmo natural de las variables Renta, Colocación y Castigo. La calibración del modelo se muestra en la tabla 6.

Variable	Coficiente	P-value
Constante	10,58	$1,8 \cdot 10^{-72}$
Renta	-0,65	$1,7 \cdot 10^{-63}$
Colocación	-0,05	0,039
Hipotecario	-1,07	$5,5 \cdot 10^{-13}$
Castigo	0,26	$2,6 \cdot 10^{-23}$
Antigüedad	-0,04	$8,1 \cdot 10^{-9}$

Tabla 6. Resultados regresión logística de la segunda probabilidad.

Todas las variables y la constante del modelo son significativas al 5%.

Análogamente a la parte anterior, la probabilidad  $p_i$  de que un deudor  $i$  pertenezca al grupo LGD=1 es:

$$p_i = \frac{e^{(\beta_0 + \beta_1 \cdot Renta_i + \beta_2 \cdot Colocación_i + \beta_3 \cdot Hipotecario_i + \beta_4 \cdot Castigo_i + \beta_5 \cdot Antigüedad_i)}}{1 + e^{(\beta_0 + \beta_1 \cdot Renta_i + \beta_2 \cdot Colocación_i + \beta_3 \cdot Hipotecario_i + \beta_4 \cdot Castigo_i + \beta_5 \cdot Antigüedad_i)}}$$

Mientras que la probabilidad  $(1 - p_i)$  de que un deudor  $i$  pertenezca al grupo LGD entre 0 y 1 es:

$$(1 - p_i) = \frac{1}{1 + e^{(\beta_0 + \beta_1 \cdot Renta_i + \beta_2 \cdot Colocación_i + \beta_3 \cdot Hipotecario_i + \beta_4 \cdot Castigo_i + \beta_5 \cdot Antigüedad_i)}}$$

Interpretando los coeficientes del modelo, la probabilidad  $p_i$  de que el deudor pertenezca al grupo LGD = 1 aumenta si:

- La Renta disminuye. Lo que dice este resultado es que deudores con ingresos menores tendrán una mayor probabilidad de que, si es que existe una pérdida para el banco, ésta sea total. Este resultado concuerda con la intuición de que un deudor que tenga un ingreso menor tendrá una peor capacidad de pago.



- La colocación disminuye. Esto dice que los créditos que de menor monto tienen una mayor probabilidad de que, en caso de que exista pérdida para el banco, ésta sea del 100%. Este resultado difiere de las intuiciones de esta misma variable en la regresión logística de la sección 4.2.1, donde un menor monto estaba relacionado con una recuperación total más probable. Una hipótesis para esta diferencia es que los bancos tienen más incentivos para buscar y realizar acciones de cobranza para créditos más grandes, por lo que, dado que el crédito fue castigado, montos más pequeños podrían no ser atractivos para el banco y su pérdida sería total. Esto comprueba lo descrito al inicio de la sección, de que al modelar tres procesos distintos la relación entre las variables y la LGD no tiene por qué ser necesariamente la misma.
- La variable Hipotecario disminuye. Con esto, deudores que no posean un crédito hipotecario tendrán una mayor probabilidad de que, en caso de existir una pérdida para el banco, ésta sea total. Una hipótesis para explicar este comportamiento es que cuando no existe un crédito hipotecario, el banco no posee ninguna garantía de dicho deudor, por lo que cuando éste deja de pagar un banco no tendría un colateral que exigir.
- La variable Castigo aumenta. Recordando que el castigo es cuando un banco saca de su contabilidad a un crédito al considerar que sus recuperaciones son poco probables, lo que el coeficiente dice es que cuando el banco castiga un monto mayor del crédito, es más probable que no se recupere nada de dicho monto. Una posible explicación para este comportamiento es que un mayor monto castigado refleja que el banco considera a dicho deudor como más riesgoso, por lo que se explicaría que éste tenga una peor capacidad de pago en caso de un incumplimiento. Otra posible explicación es respecto al tamaño absoluto del monto castigado, donde deudas más grandes serían más difíciles de pagar para un deudor, por lo que sería más probable que éste no pague nada después de ser castigado.
- La variable Antigüedad disminuye. Lo que esto dice es que entre menos meses han pasado entre el inicio del crédito y cuando fue castigado, más probable será que de existir una pérdida ésta sea total. Una hipótesis para explicar esto es que créditos con mayor antigüedad pueden representar a deudores que han tenido un comportamiento de pago relativamente estable hasta que al final cayeron en default. Dicho lo anterior, de forma implícita se podría suponer que créditos más antiguos son menos riesgoso y aquellos créditos más nuevos son más riesgosos.

Al igual que en la sección 4.2.1, se calibra una versión alternativa de este modelo, donde se utilizan dummies para segmentar las variables Renta, Colocación y Castigo, pero no se observa monotonía en sus coeficientes por lo que este modelo no es usado.

### 4.2.3. LGD entre 0 y 1

Al dejar de considerar los casos de LGD igual a uno o cero, la distribución de los datos restantes se muestra en la figura 9.

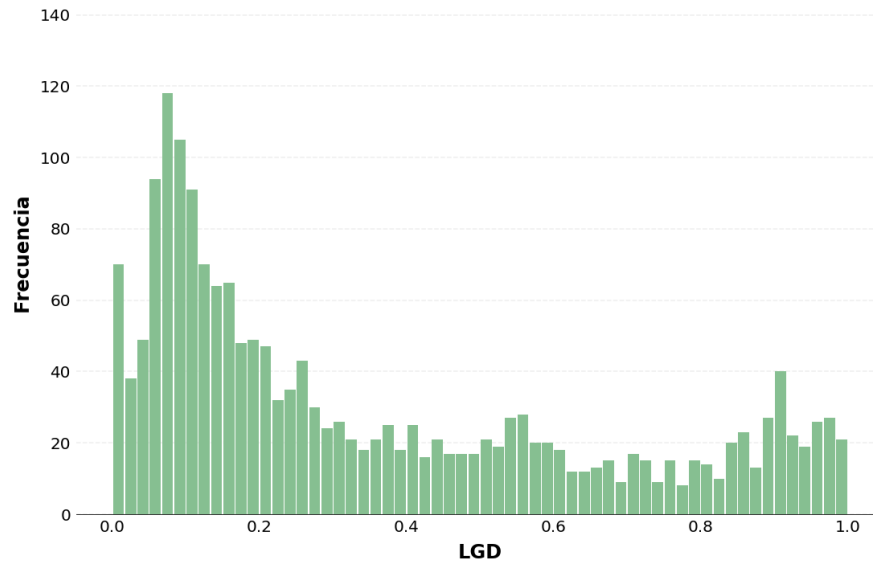


Figura 9. Histograma de LGD, valores distintos de cero y uno.

Al igual que la probabilidad de las secciones 4.2.1 y 4.2.2, se busca modelar una variable que solo toma valores entre cero y uno, pero la diferencia es que en ese caso la variable dependiente toma dos valores según cada grupo, mientras que en este caso se busca modelar una variable continua que pueda tomar cualquier valor entre cero y uno.

En esta sección se proponen tres modelos para la estimación de LGD. Al igual que para la sección 4.2.2, estos créditos fueron castigados por lo que se pueden usar todas las variables:

- Renta: Renta anual del deudor en pesos.
- Colocación: Colocaciones totales de consumo en dicho banco.
- Hipotecario: 1 si tiene crédito hipotecario, 0 si no.
- Castigo: Monto castigado en UF.
- Antigüedad: Meses entre el inicio del crédito y su castigo.

Cabe destacar que al no considerar los casos de LGD igual a 0 o 1, el número de créditos se reduce a 1.869. Un 15,6% del total de créditos.

#### 4.2.3.1. Regresión lineal

Una opción es realizar una regresión lineal directamente sobre los datos. La variable dependiente en este caso es la variable LGD, sin ninguna transformación.

Los resultados de calibrar este modelo se muestran en la tabla 7.

Variable	Coefficiente	P-value
Constante	1,31	$7 \cdot 10^{-24}$
Renta	-0,09	$1,4 \cdot 10^{-25}$
Hipotecario	-0,08	0,03
Castigo	0,06	$1,2 \cdot 10^{-42}$
Antigüedad	0,01	$2,8 \cdot 10^{-17}$

Tabla 7. Resultados regresión lineal para LGD entre 0 y 1.

Donde todas las variables menos Colocación son significativas al 5%. Analizando el signo de los coeficientes, el valor esperado de la LGD aumenta si:

- La renta del deudor es menor, resultado en línea con los modelos anteriores de que un deudor con menos ingresos pudiese tener menor capacidad de pagar su deuda y luego generar una pérdida mayor para el banco.
- No posee un crédito hipotecario. En este caso el banco nuevamente no tendría una vivienda como colateral para recuperar pérdidas.
- Su monto castigado fue mayor. Una hipótesis es que, si el monto castigado a pagar es alto, éste sería más difícil de pagar para el deudor, por lo que la pérdida sería mayor.
- Es un crédito más antiguo. Intuitivamente este resultado difiere de del obtenido en el modelo de la sección 4.2.2. En ese caso una mayor antigüedad está asociada con una pérdida menor en caso de default, mientras que en este caso la pérdida será mayor.

Los resultados descritos deben ser tomados con precaución, considerando los pocos datos que fueron usados para la calibración del modelo. La desventaja de usar este método es que al calcular la LGD esperada de un individuo, ésta pueda tomar valores negativos o mayores a uno. En particular, al evaluar los datos en el modelo calibrado se obtienen 4 casos con LGD negativa, lo que no es el comportamiento buscado.

#### 4.2.3.2. Regresión logística

Pese a que una regresión logística usa como variable dependiente una variable binaria y en este caso los valores de LGD son continuos entre cero y uno, podemos redefinir la LGD como:

$$LGD_{-}m_i = \begin{cases} 1, & \text{si } LGD_i \geq LGD_{promedio} \\ 0, & \text{si } LGD_i < LGD_{promedio} \end{cases}$$

Que divide la variable en valores que están sobre y bajo la media. La LGD promedio de los valores distintos de cero y uno es 34,97%. La distribución de esta variable en los datos se muestra en la figura 10.

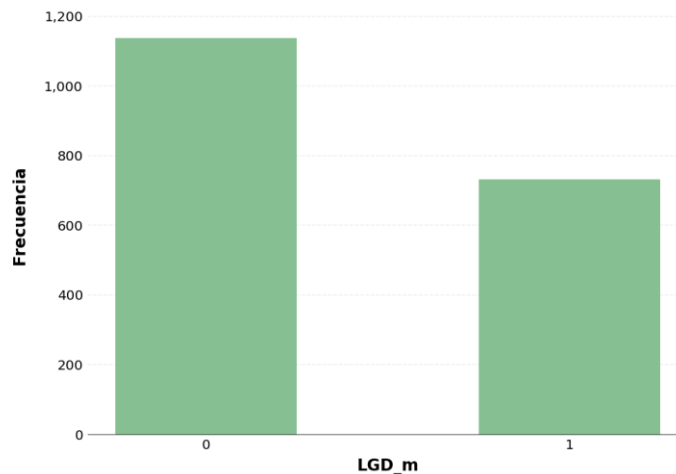


Figura 10. Histograma variable dependiente de regresión logística.

Donde un 39,2% de los deudores tienen una LGD sobre la media y un 60,8% bajo la media. Al calibrar la regresión logística por primera vez, tanto Colocación como Hipotecario no fueron significativas. Se extrae del modelo la variable Colocación, que tiene el mayor p-valor, y se calibra el modelo por segunda vez. Nuevamente la variable Hipotecario no es significativa al 5%. Finalmente, se calibra por tercera vez el modelo sin la variable Hipotecario, donde todas las variables restantes son significativas. Los resultados se muestran en la tabla 8.

Variable	Coeficiente	P-value
Constante	8,76	$4 \cdot 10^{-15}$
Renta	-0,78	$2,7 \cdot 10^{-25}$
Castigo	0,43	$5,5 \cdot 10^{-29}$
Antigüedad	0,07	$3,1 \cdot 10^{-10}$

Tabla 8. Resultados regresión logística para LGD entre 0 y 1.

La interpretación de estos tres coeficientes es la misma que la explicada para la regresión lineal en la sección 4.2.3.1.

Una desventaja de usar este modelo es que transformar todos los valores de LGD a dos valores genera una pérdida de información: un deudor con una LGD de 40% y otro con una LGD de 90% en este modelo tienen una LGD\_m igual a 1.

Adicionalmente, la elección de dividir la muestra por la LGD promedio es una división arbitraria. Por ejemplo, como alternativa podría usarse la mediana de los datos, y a priori tampoco se sabe si ésta es la mejor opción.

### 4.2.3.3. Regresión beta

Al querer modelar la LGD, se busca modelar una variable que está entre 0 y 1 y que adicionalmente es continua, en el sentido de que puede tomar cualquier valor en dicho intervalo.

Los dos modelos anteriores no permiten capturar este comportamiento de forma directa. La regresión lineal genera estimados que podían ser mayores a 1 o menores a 0, y la regresión logística genera una pérdida de información de los datos al reducir todos los valores de LGD a solo dos valores: 1 o 0.

Un modelo que sí se adapta a ese comportamiento es una regresión beta<sup>7</sup>, que asume que la variable dependiente distribuye beta, distribución que adicionalmente es lo suficientemente flexible para adaptarse a distintos datos.

Esta distribución depende de dos parámetros, los cuales determinan todas las formas que ésta puede tomar, como se observa en la figura 11. El primer parámetro es su esperanza  $\mu$  y el segundo es  $\phi$ , que corresponde a un parámetro de precisión: dado un  $\mu$  fijo, si  $\phi$  es mayor, menor será la varianza de los datos.

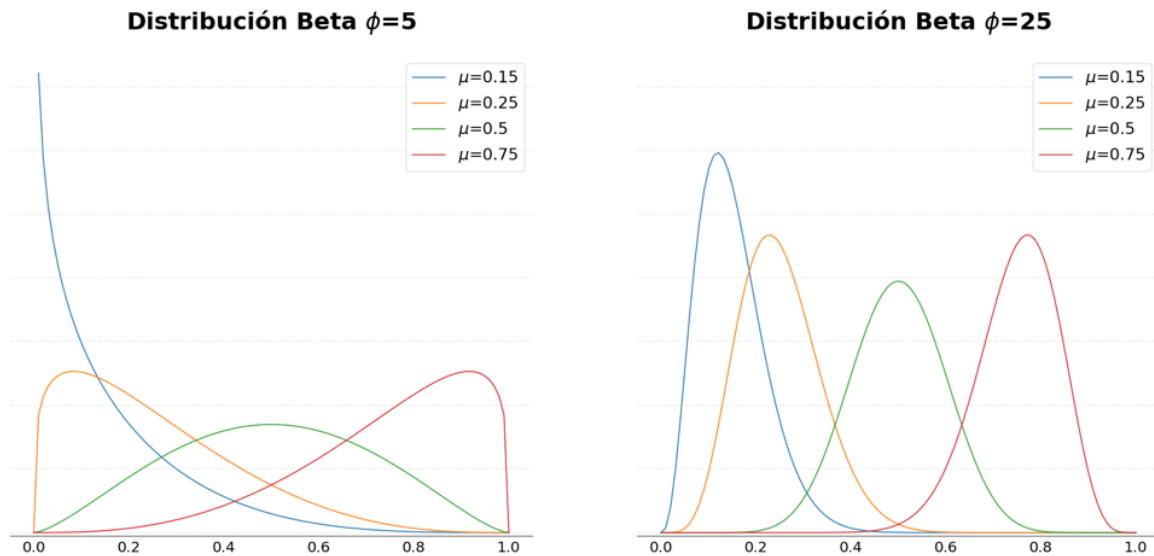


Figura 11. Densidad de probabilidad de una distribución beta para distintos parámetros.

Como se observa, una de las ventajas de la distribución beta es su flexibilidad, que permite que se adapte a distintas distribuciones de los datos, incluyendo, por ejemplo, aquellas que no son simétricas.

El modelo a estimar tiene una estructura similar a una regresión normal, y es de la forma:

$$\text{Logit}(\mu_i) = \beta_0 + \beta_1 \cdot \text{Renta}_i + \beta_2 \cdot \text{Colocacion}_i + \beta_3 \cdot \text{Hipotecario}_i + \beta_4 \cdot \text{Castigo}_i + \beta_5 \cdot \text{Antiguedad}_i$$

Donde se busca estimar los seis coeficientes de esta regresión y el parámetro  $\phi$  de toda la muestra por Máxima Verosimilitud.

La función Logit( $\cdot$ ) de dicha ecuación es de la forma

$$\text{Logit}(\mu_i) = \ln\left(\frac{\mu_i}{1 - \mu_i}\right)$$

Que busca transformar la variable entre 0 y 1 en cualquier número real, pero otras funciones que cumplan dicho propósito pueden ser usadas, como probit, o log-log. Este modelo también permite agregar una segunda regresión sobre el parámetro  $\phi$ , lo que no es utilizado en este caso.

Los resultados de la calibración de este modelo se muestran en la tabla 9.

Variable	Coefficiente	P-value
Constante	3,09	$2,2 \cdot 10^{-9}$
Renta	-0,3	$1,4 \cdot 10^{-18}$
Colocación	-0,05	0,02
Castigo	0,23	$4,5 \cdot 10^{-22}$
Antigüedad	0,06	$2,8 \cdot 10^{-22}$
$\phi$	2,13	$5,9 \cdot 10^{-264}$

Tabla 9. Resultados regresión beta para LGD entre 0 y 1.

La variable Hipotecario no fue significativa. La interpretación para los coeficientes de Renta, Castigo y Antigüedad es la misma que la explicada para la regresión lineal en la sección 4.2.3.1. El coeficiente de la variable Colocación indica que el valor esperado de la LGD será mayor para montos de créditos menores, lo que estaría en línea con la hipótesis que estos créditos no serán atractivos para que el banco realice más acciones de cobranza.

La calibración del modelo nos dice que cada crédito  $i$  tendrá una distribución beta con parámetros  $\mu_i$  y  $\phi$ . En la figura 12 se muestra una comparación entre la distribución de los datos con una distribución beta. El primer parámetro será el promedio de todos los  $\mu_i$ ,  $\mu = 0,38$  y el segundo será el de la muestra completa,  $\phi = 2,13$ .

La distribución beta efectivamente entrega una mayor flexibilidad funcional respecto a otras distribuciones comunes como la normal, por lo que este modelo parecería ser adecuado al tratar con datos que no son simétricos.

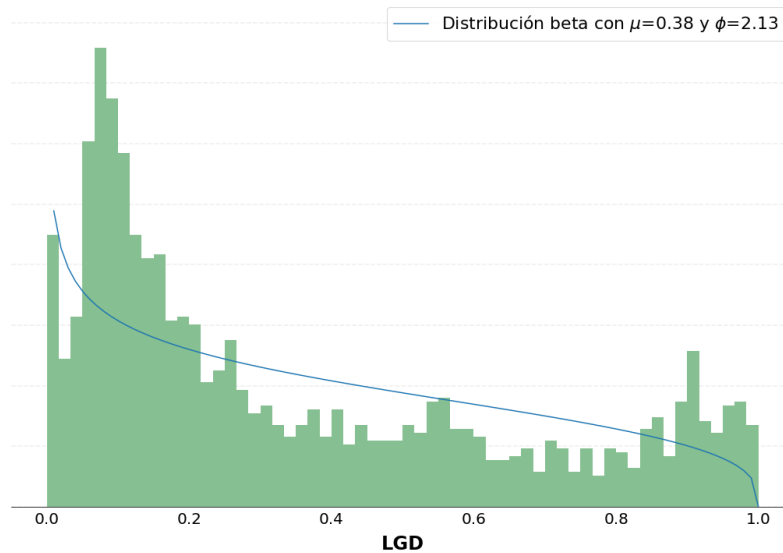


Figura 12. Comparación de la distribución de los datos y el ajuste de una distribución beta.

Respecto a la distribución de la LGD, la regresión beta cumple con modelar una variable continua entre 0 y 1, pero la inferencia que podemos hacer no fue distinta a la que se puede hacer con solo una regresión lineal. La sofisticación del modelamiento debe elegirse según sea necesario en cada problema.

#### 4.2.4. Comentarios generales

Los modelos expuestos en esta sección logran explicar la relación entre LGD y algunas variables de cada crédito, pero pueden complejizarse si se dispone de más variables que puedan describir de mejor manera el riesgo del deudor.

Las variables de los modelos presentados están vinculadas al deudor y su capacidad de pago, pero las recuperaciones no solo dependen del deudor. Otro factor que afecta las recuperaciones son las estrategias de cobranza de cada institución financiera.

Algunas de las posibilidades de cobranza son la cobranza interna, el contratar una agencia externa o vender la deuda. Como se dijo anteriormente cada uno de estos procesos limita los valores que pueda tomar la LGD sin importar el comportamiento de pago del deudor.

La cobranza hecha por un agente externo implica que la LGD es al menos el porcentaje de comisión a pagar por ese servicio. La venta de la deuda implica que la LGD es fija e igual al precio cobrado por dicha deuda.

Adicionalmente cada banco cuenta con herramientas específicas para realizar cobranza. Por ejemplo, ante un default el primer paso a tomar podría ser realizar una llamada al deudor, acción que puede resultar en un pago o un compromiso de pago a futuro. Ejemplo

de otras posibilidades pueden ser renegociaciones, envío de correos o procesos judiciales.

Cada banco tiene procesos de cobranza específicos, por lo que el tratar de realizar un modelo general para todas las instituciones puede no ser factible. Cada banco decide en qué momento es adecuado llamar al deudor, si es cuando muestra atraso en algún pago o cuando lleva meses sin pagar. A qué deudores contactar también depende de cada institución. Por ejemplo, un banco podría decidir no tomar acciones de cobranza para un crédito de menor monto y focalizar dichos esfuerzos en otros créditos de mayor monto.

Con lo dicho anteriormente, un deudor tiene ciertas características que determinarán su capacidad de pago, pero esto podría no traducirse de forma directa en un valor de LGD. Por ejemplo, para un deudor cuyas variables explicativas muestran que tiene una buena capacidad de pago, se esperaría que ante un default su LGD sea menor. Si el proceso de cobranza es eficiente este podría ser el caso, pero si no lo es, o si el banco elige no tomar acciones para dicho crédito, éste tendrá menos recuperaciones y la LGD será mayor.

Estos modelos tienen como objetivo describir de mejor manera la LGD y su relación con las variables expuestas, pero son fácilmente modificables para hacer predicciones. Para esto, las variables utilizadas deben ser conocidas al momento del incumplimiento. Pese a que en este caso se utilizan dos variables conocidas después del default, cuando estaban en castigo, se pueden buscar equivalentes en variables previas al incumplimiento. Por ejemplo, se esperaría que el comportamiento de la variable antigüedad sea el mismo si se define como los meses entre el inicio del crédito y el castigo o los meses entre el inicio del crédito y el default.

Por otro lado, la LGD entre 0 y 1 corresponde a solo al 15,6% del total de datos, por lo que toda sofisticación que pueda usarse para modelar esta porción perderá valor con un mal modelamiento de las probabilidades del resto de los datos.

Finalmente, los datos utilizados corresponden a datos agregados de varias instituciones financieras, lo que permitió obtener resultados significativos para las variables. Este no será necesariamente el caso para cada institución financiera por separado, dado que solo contarán con sus datos internos. Esto podría presentar problemas de poca cantidad de datos que suelen aparecer al modelar la LGD.

#### **4.2.5. Comparación con otros trabajos de LGD**

A diferencia de otros tipos de colocaciones, en los Acuerdos de Basilea no se muestran valores sugeridos de LGD para créditos de consumo sin garantías. Se sugiere generar segmentaciones de créditos homogéneas y con suficientes datos para proponer estimaciones dentro de cada segmento.

Thomas et al<sup>2</sup> plantea un modelo de árbol de decisión usando datos créditos a personas sin garantía en Reino Unido. A diferencia de los datos usados en el presente trabajo,



menos de un 10% de los créditos tienen una pérdida total, por lo que no se distingue de forma clara que esos créditos fuesen un grupo diferente.

Para separar entre los grupos  $LGD = 0$  o  $LGD > 0$  se utilizan otras variables con las que no se dispone para este trabajo. El autor concluye que, si el monto inicial del crédito es más alto, menor es la probabilidad de tener una LGD igual a cero. Este resultado es el mismo obtenido en la sección 4.2.1 pero con la variable Colocación, donde si el monto que queda por pagar al momento del default es más alto, menor será la probabilidad de tener una LGD igual a cero.

Otras relaciones encontradas son que la probabilidad de tener una LGD igual a cero es mayor para deudores que llevan más tiempo viviendo en su residencia actual o que han tenido más meses con mora durante la duración del crédito.

Para el modelo de LGD entre 0 y 1, sin describir su relación con la LGD, se identifica como buen predictor el tiempo del crédito hasta el momento del default, análogo a la variable Antigüedad usada en los modelos. También se presenta como buen predictor el monto al inicio del crédito, análogo a la variable Colocación, aunque en este caso solo es significativa al 5% de confianza para la regresión beta.

El trabajo de Bellotti et al<sup>8</sup>, que usa una base de tarjetas de crédito de Reino Unido, prueba distintos modelos de LGD y determina que una regresión lineal ajusta mejor a los datos que el resto de los modelos, incluido un árbol de decisión.

Dentro de las variables usadas, a mayor el tiempo del cliente en el banco o del cliente con la tarjeta previo al default, menor será la LGD esperada. Esto es consistente con la variable Antigüedad de los tres modelos para LGD entre 0 y 1 de la sección 4.2.3.

Adicionalmente, se muestra que el monto y el número de tarjetas de crédito tienen una relación positiva con la LGD, dado que montos mayores serán más difíciles de pagar. Esta intuición coincide con la de la variable Colocación de la sección 4.2.1, pero difiere en los otros modelos. Esta explicación también coincide con el coeficiente de la variable Castigo en todos los modelos.

Por otro lado, a menor ingreso mayor será la LGD esperada, que es el mismo resultado de la variable Renta en todos los modelos.

Finalmente, LaCour-Little<sup>9</sup>, usando datos de créditos hipotecarios, encuentra que la LGD disminuye con el tamaño del crédito. Este resultado difiere de los otros trabajos discutidos en esta sección. Una explicación dada es que créditos hipotecarios de menor monto tienen mayor severidad de pérdida por los altos costos asociados a una venta de la garantía.

### 4.3. Otros modelos

En esta sección se enumeran otros modelos que pueden describir de forma directa el valor de LGD, sin necesidad de utilizar las separaciones del árbol de decisión descrito en la sección 4.2.

El trabajo de Li et al. (2017)<sup>10</sup> compara modelos comúnmente usados en la literatura para la LGD:

- Zero-One-Inflated Beta (ZOIB): Modelo que toma una distribución beta en  $(0,1)$  y agrega una distribución Bernoulli que puede tomar los valores 0 y 1. Lo descrito tiene una forma similar al histograma mostrado en la figura 4.
- Regresión lineal, la cual no permite restringir los valores calculados en el rango  $[0,1]$  ni describir una distribución bimodal.
- La regresión Gaussiana inversa (IG smearing) y el modelo FRR, que permiten valores en el rango  $[0,1]$  pero no describir una distribución bimodal.
- La regresión Gamma (CG), el modelo TGG, Beta Inflada (IB) y regresión Beta (BR), los cuales son sofisticados y específicamente diseñados para distribuciones bimodales.

Los resultados de su trabajo muestran que al usar simulaciones con datos generados aleatoriamente según una distribución ZOIB, todos los modelos tienen un mal ajuste y tienen una precisión similar al realizar predicciones.

Sin embargo, al comparar la capacidad de predecir la distribución de la LGD, y no solo un valor esperado, los modelos CG, TGG, IB y BR tienen un desempeño similar y mejor que la regresión lineal y el modelo IG. El modelo FRR no será apropiado para este propósito por la incertidumbre que tiene al generar predicciones de distribución.

Lo anterior es un ejemplo de que la utilización de modelos más sofisticados no necesariamente entregará mejores resultados que modelos más simples como una regresión lineal. Se debe evaluar la pertinencia de cada modelo según sea el propósito de las estimaciones de la LGD, como describir su comportamiento, entregar predicciones o estimar su distribución.

## **5. FACTORES QUE AFECTAN LA ESTIMACIÓN DE LGD**

La estimación de LGD, y en particular la consolidación de la base de datos de la sección 3, tuvo supuestos implícitos y decisiones que se tomaron para poder lograr su cálculo. El objetivo de esta sección es testear en qué magnitud cambiarían los valores de la LGD al cambiar estos supuestos.

El valor de referencia será la LGD promedio calculada en la sección 4.1, que es un 51,702% y que aumenta a 57,009% al incorporar el gasto promedio.

Las implicancias de un aumento de 1% en la estimación de LGD se traducen de manera directa en un aumento de un 1% en la Pérdida Esperada o provisiones. Cuando hablamos de carteras con muchos créditos, ese 1% puede significar sumas de dinero considerables para el banco.

Adicionalmente, la valorización de la cartera de créditos de un banco no es la suma de todo el capital que se le debe, ya que se le deben descontar las provisiones para reflejar lo que se espera se pierda con deudores que no pagarán. Siguiendo con el ejemplo anterior, si el monto provisionado aumenta en un 1%, esta diferencia se restaría del activo del banco.

Además, la nueva Ley General de Bancos aprobada en el año 2018 que busca adoptar estándares regulatorios de Basilea III, incluye el cálculo de activos ponderados por riesgo usando, entre otras variables, la LGD. Luego, otra implicancia en una diferencia en la estimación de la LGD es una diferencia en el valor de los activos ponderados por riesgo, que a su vez implica diferencias en los requerimientos de capital del banco.

Cuando se habla de un cambio de la LGD se habla de una diferencia en los valores de la variable LGD en los datos de la sección 3, previo a calibrar algún modelo con dicha base. Por lo tanto, otro potencial problema es que el nivel de sofisticación al modelar la LGD pierda valor si, por ejemplo, al cambiar la tasa de descuento utilizada, su valor cambie de manera considerable.

### **5.1. Definición de default**

El default o incumplimiento corresponde a un punto en que se considera que las recuperaciones de un deudor y sus créditos serán remotas, dado que muestra una deteriorada o nula capacidad de pago. Existen muchas condiciones que tanto un banco de forma interna como la entidad reguladora pueden usar para considerar a un deudor en dicha categoría.

La base de datos usada en los modelos fue constituida usando la definición de incumplimiento dictada por la CMF, descrita en la sección 2.1. Alternativamente, para poder observar cambios en el valor de LGD según cambios en la definición de default, se propone una definición alternativa. En este caso, un crédito estará en default si tiene un

atraso igual o superior a 90 días en el pago de intereses o capital, es decir, solo la primera condición de la definición normativa.

La comparación entre los histogramas de LGD de ambos modelos se muestra en la figura 13. Se observa que ambas distribuciones son similares, salvo que en el caso de la definición alternativa se cuenta con menos créditos. Efectivamente la definición de default alternativa es un subconjunto de la normativa, dado que la primera contiene todos los créditos con más de 90 días de mora, pero la segunda adicionalmente también considera créditos de deudores que tengan algún otro crédito con más de 90 días de mora, independiente de si el crédito mismo tiene días de mora o no. Adicionalmente a estos créditos, la definición normativa incluye también deudores y sus créditos que cumplan la segunda y tercera condición.

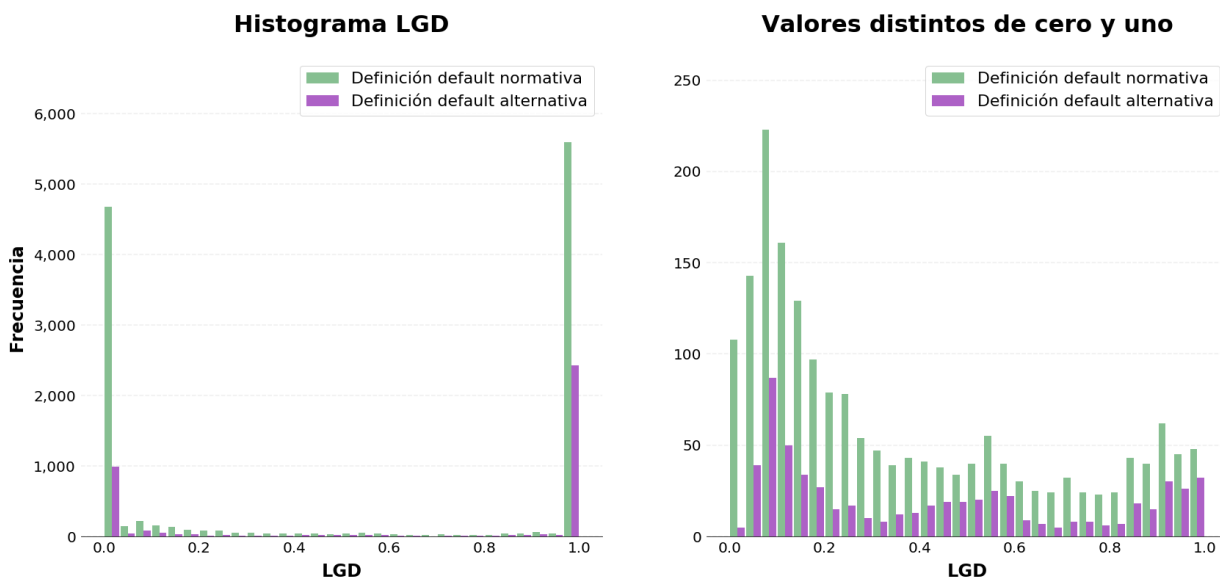


Figura 13. Comparación del histograma de LGD usando las dos definiciones de default.

En el caso de la definición de default normativa, la LGD promedio tal como se calculó en la sección 4.1, es de un 51,702% y bajo la definición de default alternativa es un 66,543%. Esto muestra un aumento gracias al cambio de definición, pero esto no es el caso general. Al cambiar este supuesto solo cambiarán la cantidad de créditos que se consideran para los modelos, lo que implica un posible cambio en el valor de LGD promedio y no si éste necesariamente aumenta o disminuye.

Bajo la definición de default normativa, de los 11.995 créditos considerados:

- Un 38,17% tiene una LGD igual a cero.
- Un 15,58% tiene una LGD entre cero y uno.
- Un 46,25% tiene una LGD igual a 1.

Bajo la definición de default alternativa, son considerados 3.996 créditos, 7.999 menos que en el caso anterior. Del total de créditos:

- Un 24,625% tiene una LGD igual a cero.
- Un 15,265% tiene una LGD entre cero y uno.
- Un 60,11% tiene una LGD igual a 1.

Donde se observa que el aumento del promedio de LGD para la definición alternativa está explicado porque ésta considera más créditos con LGD de 100%.

Pese a que en la actualidad la definición normativa es fija, ésta también ha cambiado durante los años. La definición actual fue introducida en el año 2015.

Lo anterior muestra que la forma en que una institución financiera categoriza como remota las recuperaciones de un deudor es variable, dado que, por ejemplo, regulatoriamente podría fijarse una definición más conservadora. Esto muestra la importancia de entender cuáles serán los efectos de potenciales cambios en los valores de LGD utilizados.

## 5.2. Definición de cura

La definición que se considera en este trabajo para elegir los casos con LGD igual a cero es, del total de créditos no castigados, cuáles no están en incumplimiento cuatro meses después. Esta definición es utilizada en esta sección para demostrar la sensibilidad que tiene el cálculo de LGD al número de casos con LGD igual a cero que son considerados.

Lo anterior describe el proceso de cura, según el cual un deudor deja de estar en default sin necesidad que el banco castigue a sus créditos. En el Compendio de Normas Contables<sup>1</sup> se define que para que un deudor deje de estar en incumplimiento, deben cumplirse las siguientes condiciones:

- Ninguna obligación del deudor presenta un atraso en su pago superior a 30 días corridos.
- No se le han otorgado nuevos refinanciamientos para pagar sus obligaciones.
- Al menos uno de los pagos efectuados incluye amortización de capital.
- Si el deudor tuviere algún crédito con pagos parciales en periodos inferiores a seis meses, ya ha efectuado dos pagos.
- Si el deudor debe pagar cuotas mensuales por uno o más créditos, ha pagado cuatro cuotas consecutivas.
- El deudor no aparece con deudas directas impagas en la información que refunde la CMF, salvo por montos insignificantes.

Según la definición utilizada en la sección 3, de todos los créditos que cayeron en default, se consideran con LGD=0 a los que no fueron castigados y que ya no estuvieran en dicho estado cuatro meses después. Los créditos que no están en incumplimiento cuatro meses después representan un 38,1% del total de créditos que no fueron castigados.

Esta definición puede ser razonable al considerar que un crédito dejó de estar en default si en 4 meses más ya no lo está, pero a priori no se puede decir que sea la más adecuada, ni menos una definición correcta.

Se podría argumentar que, si un crédito está en default y no fue castigado de manera directa, todos estos casos se debiesen considerar con LGD igual a cero. En la figura 14 se compara el histograma de LGD de la definición de cura original usada en la sección 2 con la definición alternativa aquí propuesta.

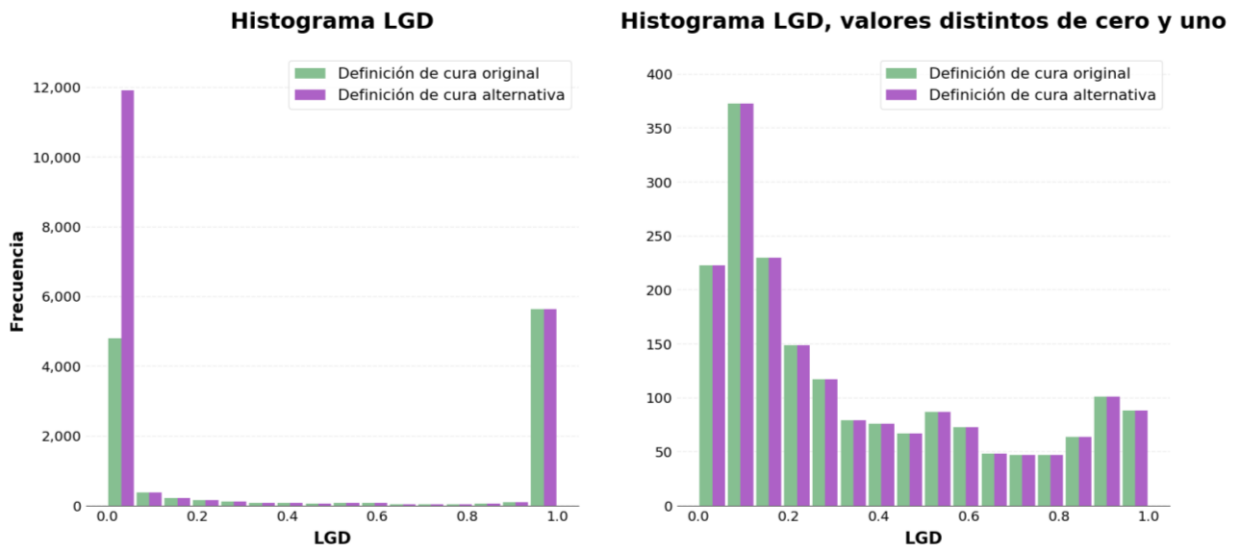


Figura 14. Histogramas de LGD para ambas definiciones de cura.

Se observa que la única diferencia entre las distribuciones de LGD en ambos casos es que se están considerando más casos con LGD igual a cero. La definición original considera 4.578 créditos con LGD igual a cero, mientras que con la definición alternativa estos aumentan a 11.704.

El efecto de aumentar los casos de  $LGD=0$  hará que la LGD promedio disminuya. En la figura 15 se muestra el cambio en el valor de la LGD promedio según el número de casos considerados. Si son cero casos, se dice que solo los créditos castigados debiesen ser considerados para el cálculo de LGD y el considerar el total es la definición alternativa propuesta.

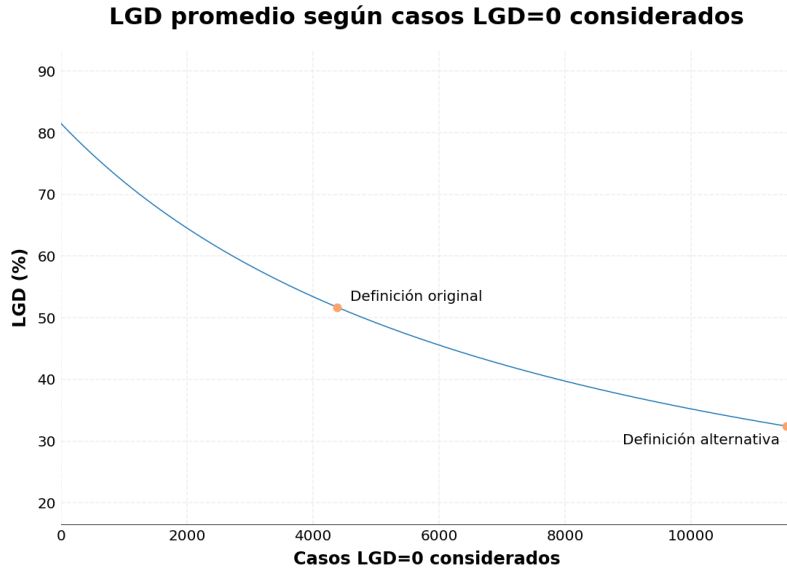


Figura 15. Cambios en el valor de la LGD promedio según casos LGD igual a cero considerados.

Al pasar de la definición de cura a la definición alternativa, la LGD promedio disminuye de un 51,702% a un 32,433%, una diferencia de un 19,3%. En particular, si no se considera ningún caso, la LGD promedio es superior al 80%. Esto deja en evidencia que el valor de LGD es altamente sensible a cuántos casos de LGD igual a cero se elijan considerar.

La definición de cura solo afecta el modelamiento de la probabilidad para separar los grupos LGD=0 y LGD>0 de la sección 4.2.1. El resto del modelamiento es exactamente el mismo, porque la distribución de LGD distinta de cero sigue siendo la misma.

En la tabla 10 se muestra la comparación de los valores de LGD al combinar las dos definiciones de cura con las dos definiciones de default de la sección 5.1.

	Definición cura original	Definición cura alternativa
Definición default normativa	51,702%	32,43%
Definición default alternativa	66,54%	45,12%

Tabla 10. Comparación de los valores de LGD promedio al combinar las dos definiciones de cura y las dos de default.

Al usar la definición de default alternativa la LGD tiene una variación de un 21,42% entre ambas definiciones de cura. Todo lo anterior muestra la importancia de la existencia de una definición regulatoria de tanto default como de las condiciones necesarias para salir

de dicho estado. De no existir, las diferencias en la estimación de la LGD pueden ser considerables.

### 5.3. Tasa de descuento

Cuando un crédito cae en default, y un tiempo después el banco logra recuperar la mitad de ese monto, no se puede decir que la LGD será 50%. El monto recuperado corresponde a dicho monto en la fecha de recuperación, por lo que no es comparable con montos en la fecha de default.

Dicho lo anterior, todo monto recuperado o gastado debe descontarse a la fecha de default para que así sea comparable, y es con esa consideración que la fórmula para el cálculo de LGD está construida.

En la figura 16 se muestra la LGD de un crédito ficticio que cayó en default con una exposición de 100, del que el banco logró recuperar 50 un año después, usando tasas de descuento entre 0% y 20%.

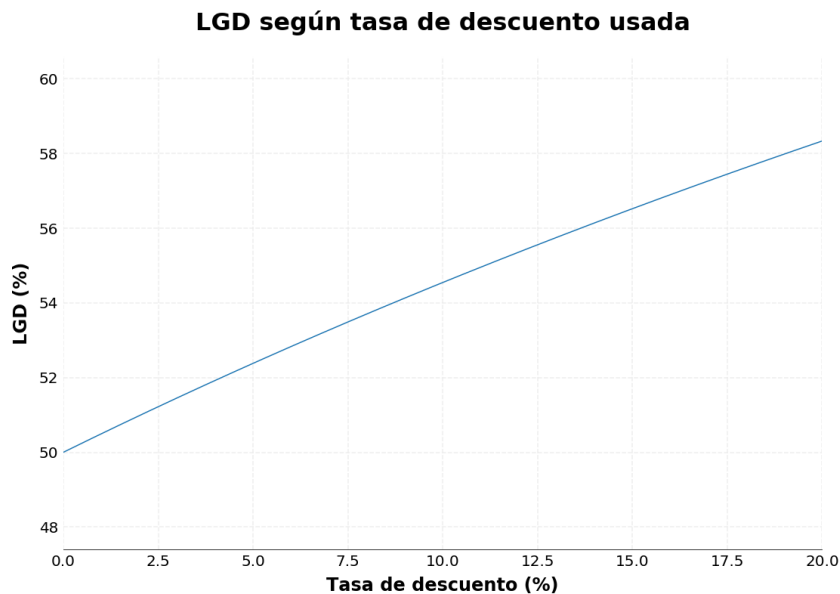


Figura 16. Cambios en la LGD de un crédito con EAD de 100 y una recuperación de 50 un año después para distintas tasas de descuento.

El caso en que la tasa es igual a 0% es cuando efectivamente la LGD sería un 50%, dado que el valor de las recuperaciones sería el mismo en la fecha de default y en la fecha de recuperación. A partir de ese punto, se observa que cuando la tasa de descuento es mayor, la LGD calculada será mayor. Esta observación se cumple en este caso, donde solo se están considerando las recuperaciones y no los gastos asociados.



Para el presente trabajo se utiliza la tasa de créditos de consumo promedio del banco. La elección del uso de esta tasa no es a priori adecuada ni menos una elección correcta.

Para comparar los efectos de la tasa elegida en los datos, se compara el valor de LGD al descontar las recuperaciones usando la tasa de consumo, la tasa promedio de activos bancarios y la tasa promedio de pasivos bancarios.

En los datos la tasa de pasivos bancarios es siempre menor a las otras dos tasas. Esto es razonable dado que el pasivo bancario está compuesto, por ejemplo, por depósitos a plazo, cuyas tasas suelen ser menores a las de los activos del banco, que incluyen créditos y en particular créditos de consumo. Adicionalmente, en los datos la tasa de consumo es mayor que la tasa de activos en la mayoría de los casos. La figura 17 muestra el histograma de la LGD en estos 3 casos.

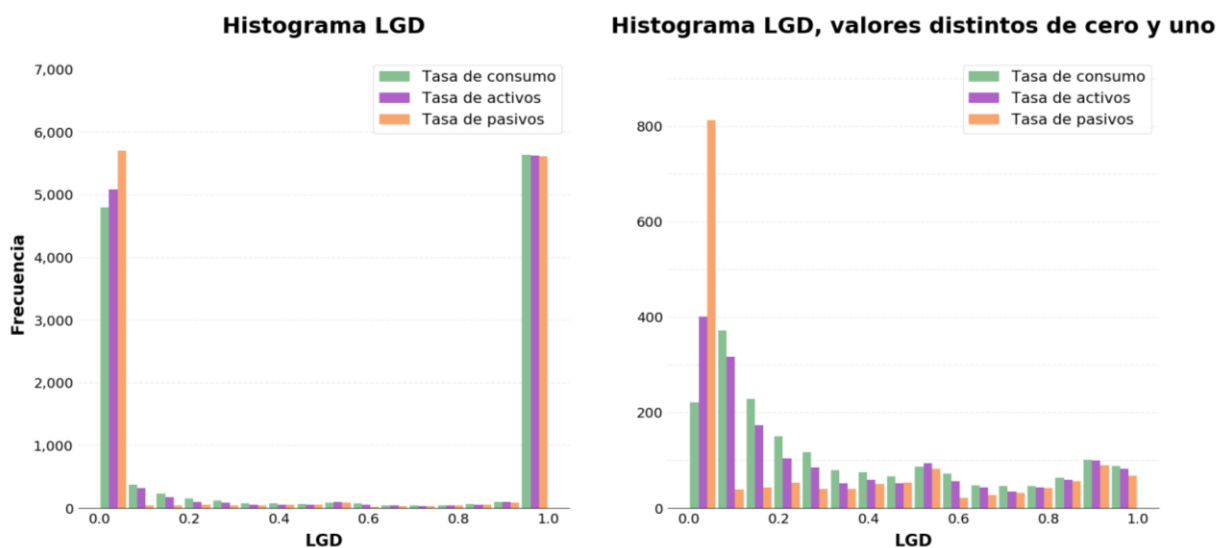


Figura 17. Histograma de LGD al descontar las recuperaciones usando la tasa de consumo, de activos y de pasivos del banco.

Como era de esperarse por la relación entre las tasas, la distribución de LGD con la tasa de activos se ve desplazada a la izquierda respecto a la LGD con la tasa de consumo. La LGD con tasa de pasivos se ve desplazada a la izquierda respecto a las otras dos tasas, con una mayor concentración de valores de LGD cercanos a cero. Cabe destacar que, si un crédito no tiene ninguna recuperación, la pérdida será total y sin importar la tasa de descuento su  $LGD=100\%$ . Esto se puede observar en el histograma de la izquierda, donde la barra de  $LGD=1$  es igual en los 3 casos.

La LGD promedio al descontar usando la tasa de consumo es 51,702%, y disminuye a 51,058% al usar la tasa de activos y al 49,82% al usar la tasa de pasivos. Al cambiar la tasa solo cambian los valores de LGD, pero el número de datos sigue siendo exactamente el mismo.

No existe una recomendación regulatoria como es el caso de la definición de default de la sección 5.1, por lo que como se dijo la elección de la tasa de descuento a utilizar queda a criterio de cada institución.

Siguiendo la lógica de evaluación de proyectos, la tasa de descuento a utilizar en dicho caso sería el WACC, o Costo Medio Ponderado del Capital. A nivel aproximado se estaría hablando de una tasa con la que el banco logra financiarse, por lo que en este caso sería razonable utilizar la tasa promedio de pasivos bancarios.

Schuermann (2004)<sup>3</sup> por otro lado argumenta que, en el caso de la LGD, no sería adecuado utilizar la tasa libre de riesgo, sino la tasa de un activo de similar riesgo. Cabe destacar que la tasa del costo de capital puede ser similar a la tasa libre de riesgo en el caso particular de un banco.

La recomendación de Schuermann sería, por ejemplo, utilizar la Tasa de Rentabilidad Exigida por un banco. En dicho caso, dado que las recuperaciones de un crédito en default tienen asociado un riesgo, la elección debiese ser la tasa libre de riesgo más una prima por riesgo. Para esto podría ser más adecuado usar la tasa de activos promedio, o la de consumo.

A nivel general, la recomendación sería definir a nivel regulatorio indicaciones sobre la tasa de descuento a utilizar, lo que permitiría evitar diferencias en el cálculo de LGD entre bancos.

#### **5.4. Costos**

Toda recuperación que un banco logre no es realmente recuperada en su totalidad si se incurrieron en gastos para conseguirla. Es por eso que la fórmula para el cálculo de LGD incluye no solo las recuperaciones descontadas a la fecha de default, sino también los costos.

Gastos típicos en los que se incurre luego de un default son acciones de cobranza de, por un ejemplo, una persona llamando o enviando correos al deudor por su deuda no pagada. La cobranza también puede ser externalizada a otra entidad que realice dicha acción, la cual podría cobrar una comisión del monto recuperado.

Otro caso es cuando la deuda se vende a otra entidad, donde el banco recupera una fracción determinada y el comprador de la deuda tendrá la responsabilidad de cobrarla obteniendo las recuperaciones para sí mismo. En ese caso dicha fracción será la LGD.

También el banco podría incurrir en acciones judiciales contra el deudor para conseguir alguna recuperación, las cuales tienen un costo asociado. Adicionalmente, pese a que no es el caso de los créditos de consumo usados en este informe, otros créditos como el hipotecario tienen una garantía, por lo que el banco podría incurrir en gastos para la venta de dicho colateral.

Para el presente trabajo no se dispone de este tipo de detalles en los costos, ni tampoco de un gasto total para cada crédito. La información con la que se dispone es el monto total que gastó y recuperó el banco para cada trimestre, por lo que la tasa de recuperación  $Rf_i$  utilizada para el cálculo de la LGD de cada crédito solo incluía las recuperaciones.

Para resolver este problema, se introduce la tasa de recuperación efectiva de la forma:

$$h_{it} = \frac{Rec_{it}(r_{it}) - Gas_{it}(r_{it})}{Rec_{it}(r_{it})}$$

Es decir, el porcentaje recuperado sin gastos. Con esto, la LGD promedio real, que incluye una tasa de recuperación efectiva  $h$  será:

$$LGD = (1 - R) = (1 - h \cdot Rf)$$

Los detalles de estos cálculos se muestran en la sección 3.2. La pregunta a realizar es qué tasa de recuperación efectiva  $h$  se debería usar. Los valores trimestrales de  $h_{it}$  para 8 de las instituciones financieras se muestran en la figura 18.

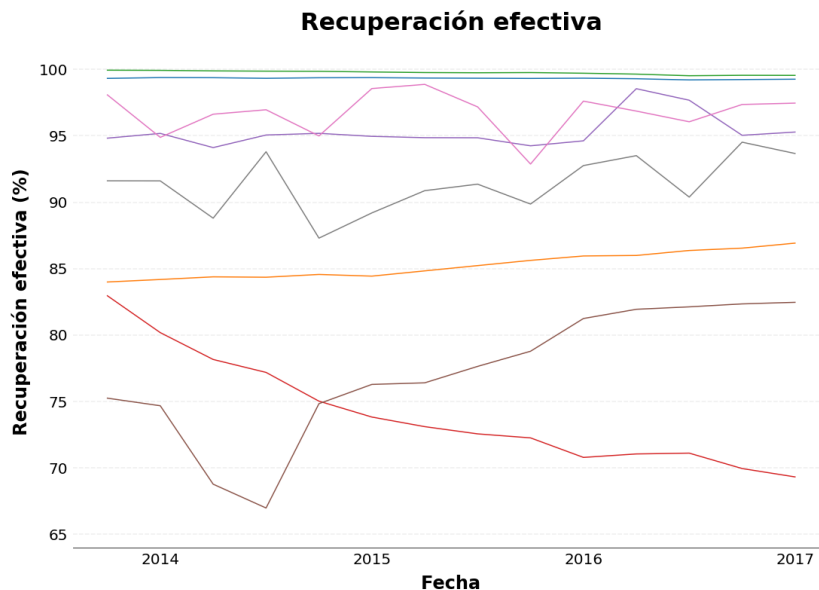


Figura 18. Tasa de recuperación efectiva trimestral para 8 instituciones financieras.

Donde se observan diferencias importantes entre instituciones con tasas menores a 85%, y otras cuya tasa es mayor al 95%. La tasa promedio de los 12 bancos durante todo este periodo se muestra en la tabla 11.

Tasa de recuperación efectiva promedio	
99,32%	95,32%
85,24%	99,87%
97,33%	75,57%
76,38%	77,13%
99,75%	96,74%
74,12%	91,38%

Tabla 11. Tasa de recuperación efectiva promedio durante todo el periodo para las 12 instituciones financieras.

Se observa que se confirma la alta variabilidad en la tasa de recuperación efectiva observada en la figura 18, con valores que oscilan entre un 74,12% y un 99,87%. La pregunta es qué tasa se debería usar para que el valor de LGD promedio final incluya los gastos.

Durante el presente informe se utiliza la tasa promedio de todas las instituciones, que es un 89,011%. En la figura 19 se muestra el valor de la LGD corregida por costos para distintas tasas de recuperación efectiva, incluyendo la tasa mínima, máxima y promedio de los 12 bancos.

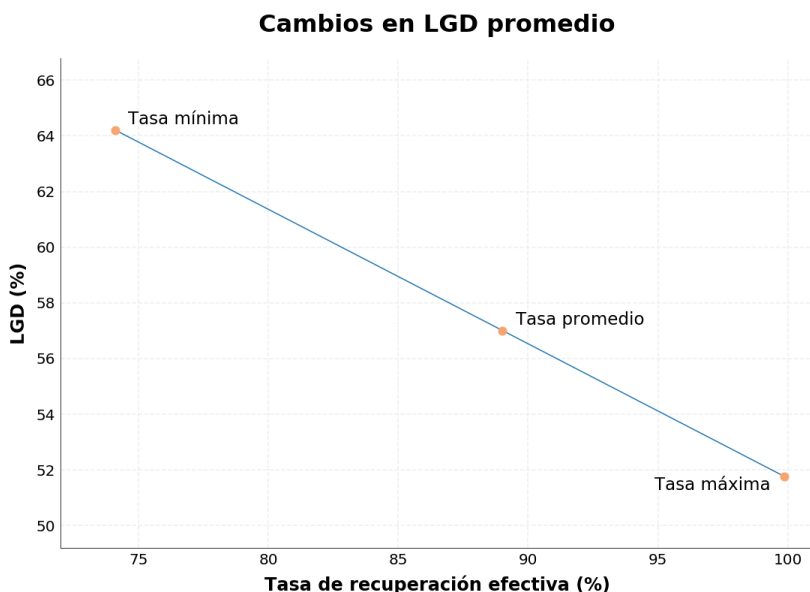


Figura 19. Cambios en LGD para distintas tasas de recuperación efectivas.

Se observa que nuevamente existe una alta variabilidad en los valores corregidos de la LGD, donde al usar la tasa mínima la LGD es 64,21%, al usar la tasa promedio es 57,01% y al usar la tasa máxima un 51,77%.

Pese a que pueda ser razonable usar la tasa promedio para la muestra completa, el banco que reportó la tasa mínima efectivamente corregiría en dicha magnitud sus estimaciones de LGD, dado que esa es la información con la que dispone, y lo mismo ocurrirá con las otras instituciones financieras.

Luego, pese a estimaciones sofisticadas y precisas de LGD, sus valores son muy sensibles a la corrección que se le haga para incluir costos. Una hipótesis propuesta para explicar la diferencia en los valores reportados es que efectivamente existan diferencias significativas en la eficiencia de sus procesos de cobranza. Un banco puede tener recuperaciones de las cuales un porcentaje cercano a cero fue destinado a las distintas acciones de cobranza, mientras que para otro casi un 25% fue destinado a dicho propósito.

Otra posible hipótesis propuesta es que las mediciones y reportes trimestrales no sean homogéneos o comparables entre todas las instituciones financieras. Los bancos podrían tener formas distintas de contabilizar sus gastos o algunos considerar cosas que otros no.

Al seguir de forma precisa la fórmula para el cálculo de LGD, de forma ideal cada gasto en el que se incurrió para algún tipo de cobranza debe ser descontado y contabilizado para cada crédito. En la práctica esto podría ser difícil y también estar sujeto a nuevos supuestos. Por ejemplo, ¿cuál sería el costo de una llamada de cobranza para un crédito? ¿Una fracción del sueldo de la persona encargada?

Aun realizando reportes trimestrales, la recomendación propuesta es buscar formas de estandarizar la información que entregan los bancos, por ejemplo, pidiendo un desglose de las distintas categorías de cobranza en las que incurrió. Esto permitiría identificar en qué categorías en específico se producen las diferencias entre cada banco, en lugar de solo ser un monto trimestral.

## **5.5. Horizonte de recuperación**

Dado que un crédito está en default, teóricamente el banco podría seguir obteniendo recuperaciones varios años después, las que deberían ser descontadas a la fecha de incumplimiento según cada plazo.

Esto haría que el valor de la LGD nunca sea final, dado que con cada nueva recuperación su valor podría cambiar. Es por esto que se fija un horizonte de recuperación, donde luego de un default solo se consideran las recuperaciones dentro de dicho plazo, con las cuales se calcula la LGD, y no consideran las que podrían recibirse luego de eso.

Dicho lo anterior, el valor de LGD es sensible a cambios en el horizonte de recuperación, donde para horizontes más largos se contabilizan más recuperaciones, por lo que la LGD es menor, y en horizontes más cortos se contabilizan menos recuperaciones y la LGD es mayor.

En el caso de este trabajo, las tasas de recuperaciones usadas son construidas con recuperaciones recibidas dentro de los 24 meses luego del castigo. Todas las recuperaciones recibidas luego de ese plazo, de existir, no son consideradas.

El supuesto implícito de elegir un plazo es que las recuperaciones recibidas luego de eso serán lo suficientemente pocas para que no haya cambios mayores en los valores de LGD. Un método propuesto<sup>6</sup> para estandarizar y homogeneizar el plazo de recuperación, consiste en elegir un plazo tal que minimice la omisión de variables de recuperación.

El horizonte de recuperación también afecta si los datos más recientes pueden o no ser utilizados, donde al considerar como plazo 24 meses luego del castigo, los créditos que fueron castigados hace menos de 24 meses no tienen un valor de LGD final, y por lo tanto no podrían ser utilizados para los modelos.

## **5.6. Recesiones**

Los valores de LGD utilizados en el presente informe corresponden a créditos de consumo que cayeron en default entre julio del 2009 hasta diciembre del 2011. Durante ese tiempo las condiciones económicas fueron favorables, si las comparamos con la crisis subprime en los años 2007 y 2008.

La LGD promedio durante todo ese periodo fue de un 51,702%, lo que se interpreta como una pérdida esperada de aproximadamente la mitad del capital del crédito en caso de un default. Nada permite afirmar que dicho valor se mantendrá ante condiciones económicas distintas.

Existe mucha evidencia de que las recuperaciones serán más bajas durante una recesión<sup>10</sup>, por lo que en dichas condiciones sería esperable que la LGD tuviera valores mayores a los estimados.

Asumiendo que los datos utilizados sean de horizontes más largos, de forma ideal se podría contar con créditos en default en varios años, incluyendo también periodos de recesión. En dicho caso, al estimar la distribución de LGD se podría no utilizar el valor esperado, si no algún percentil pesimista con un valor de LGD mayor.

El Comité de Basilea propone la aplicación de una función<sup>10</sup> que tome como variables la LGD calculada por el banco, junto a variables que representen el estado de la economía, tipo de crédito, y tipo y valor de la garantía. Esta función entregaría como resultado un estimado pesimista de la LGD que se tendría en caso de una recesión, llamada Downturn LGD.

Otra explicación para la necesidad de una estimación pesimista, es que existe correlación entre la LGD y la Probabilidad de Default<sup>11</sup>, por lo que se necesita aumentar la estimación de LGD para compensar por la no existencia de esta correlación en las fórmulas de requerimientos de capital de los Acuerdos de Basilea.

Dicho lo anterior, las estimaciones y modelos propuestos en el presente informe deben ser tomados con precaución, dado que sus resultados pueden no ser extrapolables a situación de estrés económico.

## 6. CONCLUSIÓN

El presente trabajo tiene como objetivo acercar la literatura internacional sobre LGD a la industria financiera chilena. Se logran presentar algunos de los modelos utilizados para su estimación, explicando de qué forma estos pueden ser o no adecuados, según las ventajas y desventajas que tendría su aplicación.

Los modelos son calibrados usando datos reales de créditos de consumo chilenos, lo que permite identificar algunos factores de riesgo que logran explicar los valores de la LGD. La renta es consistentemente significativa en los distintos modelos, permitiendo establecer la relación de que deudores con una menor renta tendrán una mayor LGD esperada.

La variable castigo también es consistentemente significativa, mostrando que montos castigados más altos tendrán relación con valores de LGD más altos. Para alguno de los modelos se logra establecer que deudores sin un crédito hipotecario tendrán una mayor LGD esperada, mientras que la relación entre LGD y las variables colocación y antigüedad varía dependiendo del modelo. Esto último pone en evidencia que los factores de riesgo no tienen por qué ser necesariamente los mismos para los tres procesos en los que se puede dividir la LGD.

Durante todo este trabajo se utilizan cinco factores de riesgo, pero cada banco contará con más variables internas que le permitirán describir de mejor manera el riesgo de cada deudor y de su cartera. Este trabajo también deja la puerta abierta a incluir no solo el comportamiento de pago de los deudores, sino también a modelar las acciones de cobranza del banco, que también tienen efectos en el cálculo de la LGD.

Estos modelos tienen como objetivo describir de mejor manera la LGD y su relación con las variables expuestas, pero son fácilmente modificables para hacer predicciones. Para esto, las variables utilizadas deben ser conocidas al momento del incumplimiento.

Un hallazgo de este trabajo es que la porción de datos de LGD entre 0 y 1 corresponde solo al 15,6% de los datos, por lo que el mayor peso del modelamiento recae en las regresiones logísticas para estimar las probabilidades de pertenecer a los distintos grupos de LGD.

Todo esto confirma que la LGD es una variable desafiante de modelar, cuya distribución es inusual respecto a modelos clásicos: continua, con valores entre 0 y 1, bimodal y con una porción importante de datos iguales a cero y a uno.

Los datos utilizados corresponden a datos agregados de varias instituciones financieras. Este no será necesariamente el caso para cada institución por separado, dado que solo contarán con sus datos internos. Esto podría presentar problemas de poca cantidad de datos que suelen aparecer al modelar la LGD, lo que podría remediarse si los bancos tienen acceso a información común sobre clientes o modelos usados.



Este trabajo además permite entender en mayor detalle qué factores pueden cambiar los valores de LGD, donde se observan diferencias porcentuales importantes al cambiar los supuestos de su construcción.

La definición de default y de cura es fija por la normativa, pero todas las otras definiciones podían cambiar según el criterio de cada banco. El valor de LGD disminuye en un 1,9% al cambiar la tasa de descuento utilizada.

Al considerar los costos para el cálculo, la LGD corregida oscila entre un 51,77% y un 64,21%, dependiendo de las mediciones de gastos de la institución financiera que se elija, lo que pone en evidencia una alta variabilidad en los valores reportados por cada banco.

Estas diferencias tienen un impacto directo en provisiones, donde el cambio porcentual en la estimación de LGD será exactamente el cambio porcentual que tendrá la estimación de las provisiones. El cambio de la LGD también tiene incidencia en los cálculos de capital.

La recomendación para evitar esto es estandarizar todas estas definiciones para todos los bancos, definiendo de la misma forma que para el default de un deudor, qué tasa de descuento se debería usar y la forma en que se deberían reportar los gastos de cobranza.

Finalmente, este trabajo presenta valores de LGD en condiciones económicas normales, donde en caso de una recesión se esperaría que su valor aumentara. El desafío para trabajos futuros es buscar formas de vincular los valores de LGD con otras variables que permitan modelar cómo se comportaría ante un escenario económico estresado.

## 7. BIBLIOGRAFÍA

1. CMF. Compendio de Normas Contables.
2. Lyn Thomas, Christophe Mues, Anna Matuszyk (2010). Modellig LGD for unsecured personal loans: Decision tree approach.
3. Til Schuermann (2004). What do we know about Loss Given Default.
4. David Pacheco, Carlos Pulgar, Eduardo Valdebenito (2014). Modelo estándar de provisiones para la cartera hipotecaria local.
5. Jaime Forteza, Víctor Medina, Carlos Pulgar (2018). Modelo estándar de provisiones para la cartera comercial grupal.
6. Jaime Forteza, Víctor Medina, Carlos Pulgar (2018). Marco general para el diseño de métodos estándar de provisiones por riesgo de crédito.
7. Silvia Ferrari, Francisco Cribari-Neto (2004). Beta regression for modeling rates and proportions.
8. Tony Bellotti, Jonathan Crook (2009). Loss Given Default models for UK retail credit cards.
9. Michael LaCour-Little, Yanan Zhang (2014). Default Probability and Loss Given Default for Home Equity Loans.
10. Phillip Li, Xiaofei Zhang, Xinlei Zhao (2018). Modeling Loss Given Default.
11. Peter Miu, Bogie Ozdemir (2005). Basel Requirement of Downturn LGD: modeling and estimating PD & LGD correlations.
12. BIS (2005). An explanatory note on the Basel II IRB Risk Weight Functions.