

# Valor de la información de deuda en el mercado de créditos

Diego Beas · Carlos Pulgar · Sebastián Ramírez



Regulador y Supervisor Financiero de Chile

The Working Papers series is a publication of the Financial Market Commission (CMF), whose purpose is to disseminate preliminary research in the finance area for discussion and comments. These works are carried out by professionals of the institution or entrusted by it to third parties.

The objective of the series is to contribute to the discussion and analysis of relevant topics for financial stability and related regulations. Although the Working Papers have the editorial revision of the CMF, the analysis and conclusions contained therein are the sole responsibility of the authors.

La serie de Documentos de Trabajo es una publicación de la Comisión para el Mercado Financiero (CMF), cuyo objetivo es divulgar trabajos de investigación de carácter preliminar en el área financiera, para su discusión y comentarios. Estos trabajos son realizados por profesionales de esta institución o encargados por ella a terceros.

El objetivo de la serie es aportar a la discusión y análisis de temas relevantes para la estabilidad financiera y normativas relacionadas. Si bien los Documentos de Trabajo cuentan con la revisión editorial de la CMF, los análisis y conclusiones en ellos contenidos son de exclusiva responsabilidad de sus autores.

---

Documentos de Trabajo de la Comisión para el Mercado Financiero (CMF) Financial Market Commission (CMF) Av. Libertador Bernardo O'Higgins 1449, Santiago, Chile Teléfono: (56) 22617 4058

Copyright ©2021 CMF  
Todos los derechos reservados

## Valor de la información de deuda en el mercado de créditos\*

---

Diego Beas Lagos<sup>3</sup>, Carlos Pulgar Arata<sup>4</sup>, Sebastián Ramírez Venegas<sup>5</sup>

Mayo 2024

### RESUMEN

Para los oferentes de créditos, tener más y mejor información sobre los deudores les permitiría predicciones más precisas de su posible incumplimiento, aliviando posibles consecuencias de la asimetría de información derivados del riesgo moral y selección adversa, contribuyendo a una asignación más eficiente de recursos y al ofrecimiento de tasas de interés más competitivas. En línea con lo anterior, los resultados del estudio muestran que la información, tanto positiva como negativa de deuda, es clave para realizar una correcta cuantificación de los riesgos, permitiendo: 1) una mejor protección de los intereses de depositantes e inversionistas; 2) mantener el crecimiento de la cartera de créditos sin experimentar aumentos del incumplimiento que estén fuera del perfil de riesgo de la instituciones financieras; 3) facilitar el acceso a productos financieros principalmente a nuevos clientes; y 4) favorecer la competencia del mercado crediticio.

### ABSTRACT

For credit providers, having more and better information about debtors would allow for more precise predictions of their potential default, alleviating potential consequences of information asymmetry stemming from moral hazard and adverse selection. This contributes to a more efficient allocation of resources and the offering of more competitive interest rates. In line with the above, the study's results show that both positive and negative debt information are key to accurately quantifying risks, enabling: 1) better protection of depositor's and investor's interests; 2) maintaining credit portfolio growth without experiencing increases in defaults that fall outside financial institution's risk profile; 3) facilitating access to financial products, especially for new customers; and 4) promoting competition in the credit market.

---

\*/ Las opiniones emitidas en este trabajo, y sus errores y omisiones, son de exclusiva responsabilidad de los autores y no necesariamente reflejan la visión de la institución. Se agradecen los comentarios, consejos y sugerencias del referato interno, así como también los de otros participantes en seminarios internos.

<sup>3</sup>/ División Normativa de Regulación Prudencial, CMF, dbeas@cmfchile.cl

<sup>4</sup>/ División Normativa de Regulación Prudencial, CMF, cpulgar@cmfchile.cl

<sup>5</sup>/ División Normativa de Regulación Prudencial, CMF, sramirez@cmfchile.cl

## I. Introducción

Las asimetrías de información (Akerlof, 1970), corresponden a fallas de mercado presentes en una negociación o un contrato cuando uno de los participantes tiene mayor conocimiento que su contraparte. Las asimetrías de información dificultan alcanzar un acuerdo óptimo en términos sociales e, incluso, pueden llevar a impedir que se lleven a cabo transacciones.

En particular, en el mercado crediticio las asimetrías de información producen el fenómeno de selección adversa<sup>6</sup> entre prestamistas y prestatarios, debido a que los oferentes de crédito no son capaces de distinguir la calidad de los solicitantes de crédito ya que sólo observan señales ruidosas del prestatario y desconocen su capacidad de pago efectiva. Una de las herramientas que los oferentes de crédito utilizan para reducir esta asimetría, es mediante inferencias y evaluaciones basándose en la información parcial de los deudores de la cual disponen, exigiéndoles en algunos casos el envío de información adicional, por ejemplo, sobre comportamiento de pago pasado, nivel de deuda o carga financiera, entre otro tipo de información que puede ser observable. No obstante, existen características de los clientes que dificultan el cumplimiento de sus obligaciones financieras y que para las instituciones crediticias son difícilmente comprobables o predecibles, como lo son, por ejemplo, las dificultades en la capacidad de pago ante gastos no previstos por situaciones laborales o de salud.

Por otra parte, para los oferentes de crédito tampoco es observable la voluntad de pagar de sus clientes una vez otorgado el producto, siendo esta una asimetría de información que genera riesgo moral<sup>7</sup>. En este contexto, "*buenos pagadores*" son quienes se esfuerzan por devolver los créditos y cumplir las condiciones de los contratos firmados, en contraste con los "*malos pagadores*" o deudores negligentes, quienes ocultan información y dejan de pagar los créditos en las fechas estipuladas. Exigencias de garantías o registros históricos de comportamiento que sancionen a los "*malos pagadores*" con menor y más caro acceso a financiamiento en el futuro, son formas a través de las cuales se gestiona el riesgo moral.

En este marco, las asimetrías de información en el mercado crediticio conllevan a acuerdos sub-óptimos, debido a la selección adversa y el riesgo moral, generando: 1) menores niveles de crédito a buenos pagadores y a mayores tasas de interés, 2) mayores niveles de crédito a malos pagadores y a menores tasas de interés, y 3) racionamiento del crédito (Stiglitz y Weiss, 1981).

Dado lo anterior, la información de los deudores y de su comportamiento de pago tiene diferentes roles. Por un lado, la información de deuda completa e histórica, con una duración definida, actúa como una señal para los oferentes de crédito al momento de evaluar a sus clientes, reduciendo la selección adversa. Por otro lado, un registro histórico debería considerar un plazo razonable que permita balancear la protección a los deudores mediante el derecho al olvido y la sanción a los malos pagadores en futuros financiamientos, disminuyendo los efectos del riesgo moral. En un caso extremo de total ausencia de información, el oferente de crédito no tendría incentivos a otorgar

---

<sup>6</sup> La selección adversa en el contexto de otorgamiento de crédito se puede entender como una manifestación de la asimetría de información entre el oferente de deuda y el deudor. En particular, la selección adversa se relaciona con el fenómeno del desconocimiento de la calidad del deudor por parte de la institución financiera antes de que se curse la operación de crédito.

<sup>7</sup> El riesgo moral en el contexto de una operación de crédito se puede entender como una manifestación de la asimetría de información entre el oferente de deuda y el deudor. En particular, el riesgo moral se relaciona con el fenómeno del desconocimiento de la capacidad de pago del deudor por parte de la institución financiera una vez que ya se ha cursado la operación de crédito.

financiamiento por desconocer la capacidad de pago de sus potenciales clientes, lo que afectaría al buen funcionamiento del sistema financiero y, en consecuencia, de la economía. A su vez, y por criterios prudenciales o de estabilidad financiera, en estas condiciones no sería deseable que se otorguen créditos, toda vez que ellos se financian con recursos de depositantes e inversionistas como parte fundamental del fondeo de las instituciones crediticias.

Desde el punto de vista de los deudores, la ausencia de información de deuda aumentaría el riesgo de sobreindeudamiento, agravando un problema que ya es severo en algunos hogares de la economía local y debilitando los efectos de diversos esfuerzos públicos y privados en la materia. Además, eliminar la sanción propia de un registro histórico a los “malos pagadores”, debilitaría la cultura de pago y el comportamiento de los deudores.

Para paliar eventuales ausencias de información, los oferentes podrían tender a buscar otras tecnologías que reduzcan la asimetría de información, como son: 1) desarrollo de comités especializados que revisen los antecedentes de cada potencial cliente y analicen detalladamente la capacidad de pago del deudor, 2) modelos de negocio que, en base a préstamos pequeños a un amplio número de clientes, generen información de la capacidad de pago de dichos clientes. Si bien este tipo de tecnologías limita los efectos de la asimetría de información, también podrían aumentar el costo del crédito para todos los deudores y limitar la competencia si no existen mecanismos que permitan que la gente pueda “portar su historial crediticio”. Además, este mecanismo no eliminaría las consecuencias del riesgo moral, por lo que puede seguir existiendo el racionamiento crediticio.

En este contexto, es interesante profundizar sobre el valor que tiene la información de deuda dentro del mercado financiero chileno. Un primer estudio en esta línea es el de Forteza y Medina (2018), quienes muestran una correlación positiva entre las tasas de incumplimiento y la morosidad en el sistema. El propósito del presente artículo es aportar en la misma línea, ofreciendo nuevos antecedentes para la discusión en temas relacionados con el uso de información. Particularmente, esta investigación busca cuantificar en qué medida la disponibilidad de información contemporánea e histórica de deuda reduce las consecuencias de problemas de asimetría de información sobre la conducta del deudor y la capacidad de distinguir entre distintos tipos de deudores.

Los efectos de la información se miden en dos escenarios. El primero es en base a un modelo de comportamiento, para lo cual se usa como contexto la conducta de los deudores de consumo de los bancos. El segundo, analiza el impacto en un modelo de admisión, donde se estudia el valor de la información en la capacidad de distinguir entre los clientes nuevos de la institución financiera que entran en incumplimiento y los que no lo hacen. A partir de estos resultados, se analiza la sensibilidad de la tasa de interés a la menor asimetría que se logra con la información de los deudores, entre otros elementos.

Para lograr los objetivos del artículo, la sección II desarrolla una revisión de la literatura económica relacionada con el valor de la información en el mercado de crédito; la sección III presenta la metodología utilizada en este estudio para evaluar el impacto del uso de la información de deuda; la sección IV muestra los principales resultados; la sección V abre la discusión sobre la extensión de información histórica de la deuda; la sección VI presenta una discusión sobre las implicancias políticas de los resultados del estudio y, finalmente, la sección VII expone las principales conclusiones.

## II. Revisión de la literatura

La literatura económica muestra que la asimetría de información entre prestamistas y prestatarios afecta el desarrollo financiero, pues la existencia de selección adversa y de riesgo moral generan un racionamiento de crédito por parte de las instituciones financieras, reduciendo la asignación eficiente de crédito y generando equilibrios sociales subóptimos (Stiglitz y Weiss, 1981, Jappelli y Pagano, 2002).

Mishkin (1999) señala que el problema de la selección adversa incentiva a los bancos a crear los medios para tratar de identificar lo mejor posible a qué grupo pertenecen los eventuales clientes. No obstante, ante cambios en los términos y características del contrato de crédito, el tipo de prestatarios que atrae el banco puede variar. Por ejemplo, en periodos de inestabilidad, aumentos en la tasa de interés empeoraría la calidad de la cartera bancaria, ya que se atraerían a clientes más riesgosos debido a que los “buenos pagadores” renunciarían a pedir créditos antes de dejarlos de pagar, mientras que los negligentes no tendrían inconvenientes de pedir créditos, aunque luego no los devuelvan.

Stiglitz y Weiss (1992), y Mankiw (1992) señalan que el riesgo moral ocurre porque para los bancos es costoso monitorear la actividad del prestatario. Por tanto, cuando el banco atomiza la cantidad de prestatarios, la dificultad de monitoreo provoca un mercado incompleto y, en la mayoría de los casos, la decisión de prestar dependerá de la información brindada directamente por los prestatarios. En este sentido, esta falla de mercado causa un menor volumen en otorgamiento de crédito a nivel agregado y, por ende, un menor nivel de producción de la economía.

Jappelli y Pagano (2006) señalan que el intercambio de información entre instituciones financieras mediante los registros de créditos busca disminuir los problemas de asimetrías de información, generando importantes beneficios para el mercado crediticio. Primero, esta información mejora el conocimiento de los bancos sobre las características de los solicitantes y permite una mejor predicción sobre las probabilidades de pago y asignación de riesgo. Segundo, reduce las rentas informativas que los bancos podrían extraer de sus clientes. Tercero, puede funcionar como un dispositivo de disciplina del prestatario. Finalmente, elimina el incentivo de los deudores para sobre endeudarse mediante la obtención simultánea de créditos de varios bancos sin que ninguno de ellos se dé cuenta.

Un elemento clave en el diseño de un sistema de reporte crediticio es el tipo de datos incluidos en dicho reporte. En tal sentido, los mismos autores señalan que los diseños más simples contienen solamente información negativa, es decir, principalmente deudas en mora o castigos. Luego, sistemas intermedios, incluyen adicionalmente datos de montos totales de préstamos por deudor, de modo que los prestamistas puedan estimar más precisamente el grado de endeudamiento de los solicitantes. Finalmente, los sistemas más sofisticados incluyen información demográfica sobre los prestatarios, como, por ejemplo, edad, nivel educacional y nivel de ingreso; e información positiva como los pagos realizados oportunamente, entre otros. Cabe destacar que la información positiva genera un círculo virtuoso, ya que los deudores o potenciales deudores disponen de mejor información, ayudándolos a la toma de decisiones de endeudamiento responsable. Por su parte, las entidades del sector financiero podrán mejorar sus modelos de puntaje crediticio y tomar mejores decisiones de otorgamiento de crédito.

En el mundo existen dos tipos de registros de información de deudas: burós privados y públicos. Las actividades de los burós de crédito están reguladas casi en todas partes para evitar la violación de la

privacidad y las libertades civiles. Leyes de privacidad dan una amplia gama de garantías para el consumidor, como límites al acceso a los archivos por parte de usuarios potenciales; la eliminación obligatoria de información de los archivos después de un tiempo establecido; prohibición de recopilar ciertos tipos de información; y el derecho a acceder, verificar y corregir el propio archivo. Como se puede observar en la Tabla 1, a nivel latinoamericano; Chile está al debe en términos de registro de deuda, pues el buró privado tiene brechas en el tipo de información, mientras que el registro público tiene una limitación respecto al tipo de reportantes, esto, al contar solamente con la información de deudores de las instituciones financieras que son supervisadas por la CMF y que tienen la obligación de informar, de acuerdo con lo establecido en el artículo 14 de la Ley General de Bancos, incluyendo bancos, sociedades de apoyo al giro, cooperativas de ahorro y crédito, y empresas emisoras de tarjetas de crédito. En esta línea, la evaluación del Banco Mundial (2021) enfatiza los problemas de cobertura del sistema de información de deuda chileno.

**Tabla 1:** Registros de deudores a nivel latinoamericano

País	Buró Privado			Registro Público		
	Más que IF	Positiva y Negativa	Información Histórica*	Más que IF	Positiva y Negativa	Información histórica
Chile	Sí	No	No	No	Sí	No
Argentina	Sí	Sí	Sí	No	No	Sí
Bolivia	Sí	Sí	Sí	No	Sí	Sí
Brasil	Sí	No	No	No	Sí	Sí
Colombia	Sí	Sí	No	Sí	No	No
Costa Rica	Sí	Sí	Sí	No	Sí	Sí
Ecuador	Sí	Sí	Sí	No	No	No
Guatemala	Sí	No	Sí	No	Sí	Sí
Mexico	Sí	Sí	Sí	No	No	No
Nicaragua	Sí	Sí	Sí	No	No	No
Panamá	Sí	Sí	Sí	No	No	No
Paraguay	Sí	Sí	No	No	No	No
Perú	Sí	No	Sí	No	Sí	Sí
Uruguay	Sí	Sí	Sí	No	No	No
Venezuela	No	Sí	Sí	No	No	No

(\*) En Chile, el buró privado (Boletín Comercial) tiene brechas respecto el tipo de información y el registro público (LGB) tiene una limitación respecto al tipo de reportantes.

Fuente: "Doing Business in 2020", World Bank.

El número de años que un sistema de información crediticia "recuerda" el incumplimiento o los atrasos de un deudor, es un parámetro importante en el diseño de un sistema de información crediticia. En un extremo, un sistema con memoria infinita, donde los prestatarios no tienen posibilidad de salir de la "lista negra", incluso después de un pago tardío, puede crear un alto incentivo para pagar a tiempo, pero puede disuadir ex ante la decisión de contraer una deuda. El riesgo de aparecer eternamente en la lista negra en caso de incumplimiento puede ser tan grande

como para disuadir pedir prestado, incluso a personas con perspectivas financieras sólidas. En este sentido, una lista negra con una memoria muy larga ex post puede generar que los deudores regresen. Incluso si un deudor tiene el dinero para pagar un préstamo en mora, puede tener poco incentivo para hacerlo porque, en cualquier caso, su reputación se estropea permanentemente. En el otro extremo, un sistema en el que se mantienen registros durante muy poco tiempo y este es borrado inmediatamente después del pago tardío, ejercería muy poca disciplina sobre los prestatarios y, en consecuencia, proporcionan muy poca información sobre su historial para prestamistas.

El límite óptimo de tiempo en el que deben informarse las deudas en mora y/o estas prescriban de manera legal, difiere de país a país y depende de muchas características incluidas, por ejemplo, la persistencia de perturbaciones inductoras al incumplimiento. Cuando los derechos de los acreedores están menos protegidos, por ejemplo, debido a una deficiente aplicación judicial, la necesidad de disciplinar a los prestatarios puede ser más urgente que en otro lugar, y, por lo tanto, uno puede querer hacer que la memoria del sistema sea más larga y menos indulgente.

Distintos estudios empíricos encuentran que la información ayuda a las instituciones financieras a predecir mejor el cumplimiento de sus clientes, disminuyendo así los niveles de impago (Kallberg y Udell, 2003; Barron y Staten, 2003; Powell, et al., 2004). De esta forma, información de buena calidad y proveniente de diversas fuentes es clave para distinguir entre buenos y malos pagadores. En particular, sistemas de información que tienen una amplia cobertura en términos del universo de oferentes de crédito y del alcance de la información (considerando información positiva y negativa de los deudores), contribuyen en mayor medida a predecir el comportamiento de pago de los deudores.

Desde la perspectiva empírica, algunos de los beneficios del intercambio de información sobre deudores en el mercado de crédito comprenden: i) la mejora en la disponibilidad del crédito (Triki y Gajigo, 2014; Brown et al., 2009; Djankov et al., 2007; Galindo y Miller, 2001); ii) la reducción de los costos (Brown et al., 2009); y iii) la disminución de las tasas de incumplimiento (Jappelli y Pagano, 2002).

Desde la perspectiva de los acreedores, en mercados de capitales imperfectos debido a asimetrías de información, monitoreo costoso, problemas de incentivos o problemas de cumplimiento de contratos; más información genera: i) una mejor estructura de capital y solvencia, con menor riesgo de crédito (Galindo y Miller, 2001); ii) un manejo adecuado de riesgos (Acharya et al., 2011; Houston et al., 2010); y iii) una disminución en las quiebras de las instituciones financieras (Brockman y Unlu, 2009; Djankov et al., 2007; Claessens y Klapper, 2005).

A nivel nacional, y tras la promulgación de la Ley 20.575 en 2012, mejor conocida como "Ley del Borrónazo", los deudores morosos vivieron un alivio temporal al no hacerse públicos sus antecedentes. Sin embargo, de acuerdo con estudios posteriores<sup>8</sup>, se pudo establecer que como resultado de esta ley: i) existió una mayor propensión de incumplimiento en los deudores que salieron de los registros del buró de crédito. Así, a sólo seis meses de la restricción de la comunicación de obligaciones impagas, 700 mil deudores habían vuelto a los registros; y ii) las entidades crediticias establecieron condiciones más restrictivas en el otorgamiento de crédito, empeorando las condiciones de acceso al crédito a los individuos menos riesgosos. En el agregado hubo una reducción del crédito de la cartera de consumo, que se redujo un 3,5%, equivalente a \$20 mil millones, con

---

<sup>8</sup> Para más información sobre los resultados del estudio del "Borrónazo" del 2012, revisar el Anexo 1 de este documento.



mayores restricciones de acceso para las personas de bajos ingresos y saldos de préstamos más pequeños.

### III. Metodología para evaluar el impacto del uso de la información de deuda

Mediante distintos modelos y ejercicios cuantitativos, se busca evaluar el valor de la información de deuda, contemporánea e histórica, en los modelos desarrollados por los oferentes de crédito. Para el análisis se consideran modelos generales con información administrativa de los principales oferentes de crédito del mercado local, donde el valor de la información de deuda se puede analizar desde la perspectiva de tres dimensiones:

- 1) **Prudencial:** se busca cuantificar la mejora en el nivel de discriminación entre buenos y malos pagadores que logra el modelo gracias a la incorporación de la información de deuda. Lo anterior tiene un impacto directo en la correcta cuantificación de los riesgos de los deudores, lo que permite proteger a los acreedores: depositantes e inversionistas.
- 2) **Mercado:** se estima el efecto que podría tener la información de deuda en las tasas de interés ofrecidas por las instituciones y en la inclusión o exclusión financiera. En primera instancia, tras una restricción de información, los efectos pueden materializarse en una mayor tasa de interés, pues los oferentes de crédito internalizan como un costo el mayor riesgo de crédito que tendrían por la ausencia de información. En general, las personas que solicitan financiamiento podrían ser aprobadas, rechazadas, o podrían analizarse más detalladamente en un Comité de revisión cuando se requiera. En este caso, y por la falta de información, las tasas ofrecidas podrían ser más altas por los mayores gastos administrativos que implicaría el análisis de los deudores uno a uno en instancias de Comité.
- 3) **Competencia:** se analiza el valor de la información respecto a la promoción de la competencia, dependiendo del nivel de integración del deudor dentro del sistema financiero, es decir, número de entidades con las cuales tiene productos.

En este contexto, este trabajo considera distintas metodologías y set de información para cuantificar el valor de la información en las dimensiones presentadas con anterioridad:

- 1) **Modelo de admisión u otorgamiento:** Este tipo de modelos son requeridos para segmentar a los potenciales clientes mediante su capacidad de pago. Este trabajo se basa en un contexto de admisión de nuevos clientes en una institución financiera (IFI) y para todas las industrias. Estas son Bancos, Cooperativas de Ahorro y Crédito (CACs), Emisores de Tarjeta no Bancarios (ETNBs) y Sociedades de Apoyo al Giro Bancario (SAGs). Dado que el objetivo es realizar un modelo para clientes nuevos con la IFI, se usa como supuesto que estos corresponden a aquellos que tienen deuda en el mes, pero no tienen créditos con la respectiva institución en los 2 meses anteriores (en ningún producto, ni activa ni contingente).

Cabe tener presente que los clientes podrían tener algún tipo de producto pasivo con el banco, tales como un depósito a plazo, cuenta vista y/o cuenta corriente; y por ello no serían

sujetos desconocidos para la institución. Sin embargo, el acceso a dicha información es sensible, por lo que la búsqueda del cliente nuevo se hace sólo a través de sus colocaciones.

Para este caso, se usa como base de deudores a los que se informan en los archivos D10 "Información de deudores artículo 14 LGB" y D27 "Obligaciones de los arrendatarios en operaciones de leasing", y que luego se refunden a través de la nómina de deudores (archivos R04 y R05).

Luego, se determina si existe un incumplimiento prospectivo del deudor, es decir, si presenta morosidad mayor a 90 días en una ventana de 12 meses siguientes al periodo de observación. Esto se hace a nivel agregado, sin considerar una separación por producto. Siguiendo lo anterior, se modela la probabilidad de incumplimiento (PD, por sus siglas en inglés) con tres especificaciones:

- i. **Modelo Base:** considera sólo la razón deuda sobre ingreso (RDI). Se asume que, en el caso de clientes nuevos, la única información que conoce la IFI del deudor es su ratio de apalancamiento (RDI).
  - ii. **Información industria:** considera la RDI y se adicionan factores estimados con información de deuda de la industria a la cual pertenece la IFI (70 variables).
  - iii. **Información total:** considera la RDI y se adicionan factores estimados con información de deuda de todos los oferentes de crédito (70 variables).
- 2) **Modelo de comportamiento:** Este tipo de modelo se utiliza para medir los riesgos de crédito de la cartera de las instituciones y, de esa forma, cuantificar las provisiones y el capital requerido. Para efectos de este trabajo se utilizan los datos con los que se desarrolló la propuesta normativa del modelo estándar de provisiones de las colocaciones de consumo de las instituciones bancarias. En base a este modelo, se observa el desempeño al modelar la probabilidad de incumplimiento de la cartera considerando las siguientes tres especificaciones:
- i. **Modelo Base:** considera 23 variables construidas únicamente con información de deuda bancaria.
  - ii. **Información industria:** considera las variables del modelo anterior y se adicionan factores estimados con información del deudor en otras instituciones de la misma industria (83 variables).
  - iii. **Información total:** considera las variables del modelo base y se adicionan factores estimados con información del deudor en todo el sistema financiero (83 variables).

Para ambos modelos ("de admisión" y "de comportamiento"), la diferencia en desempeño entre las especificaciones ii) y i) permite cuantificar el efecto que tiene la incorporación de la información de deuda de la industria a la cual pertenece la IFI. Por su parte, al comparar iii) con ii), es posible determinar el impacto cuando la información de deuda considera la todos los oferentes de créditos vigentes en la nómina.

Las variables asociadas a las especificaciones ii) y iii) se construyen en base a los datos administrativos que existen en los archivos D10 y D27. Se debe tener presente que la información de deuda utilizada

correspondería sólo a las entidades que reportan estos archivos. Esto equivaldría a más de un 90% de la deuda formal de los hogares<sup>9</sup>.

Si bien los modelos no son necesariamente los que utilizan los oferentes de crédito, se considera que, al desarrollar una metodología robusta, podrían acercarse a los elaborados por los distintos participantes. Los resultados obtenidos en este documento son ejemplificadores y referenciales, de manera de lograr determinar el valor de la información de deuda en términos generales.

Una limitante de realizar este análisis con datos administrativos es que consideran solo la información de los solicitantes de crédito que fueron aceptados por las instituciones financieras; es decir, no se encuentra información de los rechazados. Por lo tanto, los modelos de *scoring* crediticio basados únicamente en solicitudes aceptadas pueden estar sesgados, teniendo un impacto estadístico en los resultados. Varios autores presentan alternativas para poder solucionar el sesgo de selección de estos modelos (Chen et al 2012; Li et al 2017; Mancenido, 2020). No obstante, el uso de esta técnica de inferencia de rechazo no justificaría su ganancia marginal en la calificación crediticia, por lo que, si los supuestos de los modelos son correctos, se puede ignorar la información de los rechazados en el proceso de estimación (Chen y Astebro, 2001; Feelders, 2000; Feelders, 2003).

Cabe destacar que, en gran medida los oferentes de crédito locales modelan la PD utilizando regresiones logísticas (discusión en Bishop, 2006). Sin embargo, en este análisis también se considerará el algoritmo *Gradient Boosting Machines* – GBM (Friedman, 1999; Ke et al 2017), el cual mejora el desempeño del modelo anterior, ya que permite considerar relaciones no lineales entre las variables.

En la metodología del estudio y para cada uno de los escenarios planteados (modelo base, modelo con información de la industria, y modelo con información completa), ambos algoritmos (RL y GBM) son utilizados para modelar la PD. Estos algoritmos retornan un valor puntaje o *score* que permite evaluar el desempeño del modelo en su capacidad de separar a buenos y malos deudores a través del AUROC. En cada algoritmo y escenario, se usan las 20 variables que más información aportan. Para la selección de estas variables se usa el set completo disponible en cada caso (equivalente a un total de 73 en admisión y 83 variables en comportamiento), y luego se aplica el algoritmo *Recursive Feature Elimination* (RFE). Particularmente, el algoritmo RFE actúa de manera iterativa, generando una perturbación aleatoria en cada característica, lo que permite medir su aporte en la bondad de ajuste del modelo. Luego de este proceso, el número final de variables consideradas se acota a 20, ya que luego de ello no se observan ganancias relevantes en el AUROC (detalles en anexo 2.1). Finalmente, cabe mencionar que este proceso de selección se realiza sólo en los escenarios de información industria y completa, ya que en el modelo base el número de variables ya está acotado a 20 aproximadamente.

## IV. Resultados

Para el análisis de los resultados, mostraremos las principales estimaciones de los modelos, discutiendo sobre el valor de la información de deuda en las tres dimensiones relevantes del estudio: 1) prudencial, 2) mercado y 3) competencia. El anexo 3.1 y 3.2 contiene una caracterización de los deudores y cartera de clientes del ejercicio de admisión y comportamiento, respectivamente. El

---

<sup>9</sup> Detalle en [https://www.cmfchile.cl/portal/prensa/615/articles-51456\\_doc\\_pdf.pdf](https://www.cmfchile.cl/portal/prensa/615/articles-51456_doc_pdf.pdf)

anexo 3.3 muestra una caracterización de cómo los deudores se distribuyen a través de las instituciones.

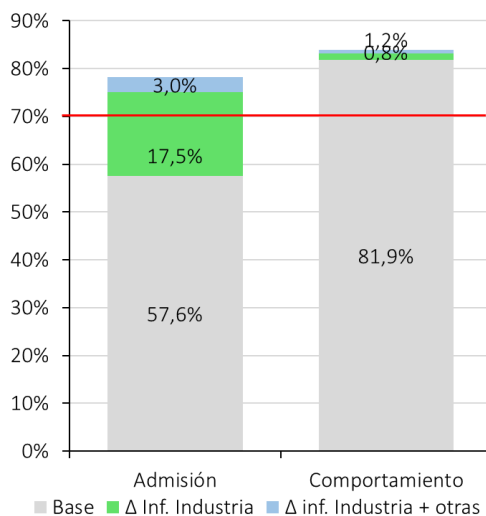
A continuación, se expondrán los resultados a nivel agregado, mientras que en el anexo (4.1, 4.2, 5.1 y 5.2) se tiene el detalle por industria. A nivel general, los resultados son robustos en todas las industrias que están presentes en los respectivos ejercicios, salvo para el caso de los ETNBs. La representación de este tipo de entidades es muy menor, ya que a diciembre de 2020 (último corte para la estimación de la PD), sólo una entidad era parte de la nómina de deudores y, por lo tanto, no se podrían extrapolar las conclusiones a estas instituciones. Luego de esta fecha más entidades ingresaron a la nómina, por lo que sería razonable verificar en el futuro los resultados particulares a estas instituciones.

### 1) Ámbito prudencial:

Para los ejercicios mencionados, se estudia el cambio en el poder de discriminación del modelo por la incorporación de información de deuda, haciendo uso de la medida AUROC<sup>10</sup> en una muestra de prueba (Hosmer y Lemeshow, 2000). Un buen nivel de discriminación implica una cuantificación correcta del riesgo de crédito, pues permite distinguir de mejor manera entre los buenos y los malos pagadores. Lo anterior permite que la institución pueda: i) realizar una adecuada selección de los clientes de acuerdo con el modelo de negocio y apetito por riesgo de la IFI, y ii) un adecuado cómputo de provisiones y capital que una IFI debe mantener, permitiendo resguardar de mejor forma el interés de los depositantes e inversionistas.

Los resultados en la figura 1 señalan que, en el modelo de admisión, la información de la RDI no es suficiente para lograr un desempeño aceptable (es decir, que el indicador AUROC sea mayor al 70%). Por ello, cobra relevancia el efecto de la información de deuda con otras entidades en permitir a las instituciones diferenciar entre buenos y malos pagadores.

**Figura 1: Cambios en el AUROC**



<sup>10</sup>El AUROC (“Area Under the Receiver Operating Characteristic curve”) es una métrica que permite medir el poder discriminatorio en un modelo y se puede interpretar como la probabilidad que el modelo clasifique correctamente una observación positiva aleatoriamente escogida por sobre una observación negativa aleatoriamente escogida.

Fuente: Elaboración propia en base a los resultados de los modelos generales. \*La línea roja representa el nivel mínimo de discriminación aceptable según la medida de AUROC.

Por el contrario, en el caso del modelo de comportamiento, la información del deudor con la propia IFI es suficiente para lograr un nivel de discriminación excelente (presentan un indicador AUROC promedio mayor del 80%). En este caso, el aporte de la información de otras entidades es acotado, sin perjuicio de las diferencias según integración del deudor que se detallan en la dimensión de competencia. Sin perjuicio de que este resultado podría indicar un bajo aporte de la información de deuda fuera de la institución, dicha conclusión es cuestionable por al menos dos razones: 1) la conclusión se basa en el resultado de un modelo de comportamiento realizado fundamentalmente con información bancaria y, por lo tanto, sesgada pues es una industria que concentra la gran mayoría del mercado de créditos. Luego, para una institución y/o industria diferente, la información de comportamiento en otras entidades podría tener un nivel de relevancia importante. Por este hecho, los resultados podrían ser diferentes según el tipo de entidad de la que se trate. Al menos, en el caso de los bancos, el comportamiento en la misma IFI es suficiente para discriminar entre los tipos de pagadores; 2) los modelos de comportamiento tienen implícito el sesgo de selección, en el sentido de que para su desarrollo se consideran deudores que ya fueron aprobados por el modelo de admisión.

Luego, los resultados del modelo de admisión son más generalizables y relevantes al llevarse a cabo sobre más de un tipo de entidad. También pueden resultar más interesantes por aplicarse en el escenario donde mayor incertidumbre existe respecto al tipo de cliente que se enfrenta, esto es, en la admisión. Por esto, los resultados siguientes consideran este escenario, salvo se señale lo contrario.

## 2) Ámbito de mercado:

En primera instancia, se estima el efecto de la información de deuda sobre las tasas de interés. En ausencia de información de deuda, y si se mantiene el número de deudores aceptados, el oferente de crédito asumirá mayor riesgo de crédito, el cual traspasaría a través de mayores tasas de interés. Para ello, para un nivel dado de aceptación de clientes que solicitan financiamiento con la institución, se estima la PD empírica que los oferentes de crédito tendrían tanto con el modelo base (modelo i) como con aquel que tiene información total de deuda (modelo iii)). Para este ejercicio, se asume una tasa de aceptación del 89%, que correspondería aproximadamente al valor que estarían considerando las IFIs de acuerdo con la información de la PD empírica en el último año, según se expondrá luego.

Para determinar el cambio en tasas de interés, es necesario determinar la variación en la prima por riesgo de crédito (PRC), la cuál es un componente de la tasa de interés que las instituciones cobran. Este valor se puede estimar a partir del valor PRC que iguala la siguiente ecuación (Pulgar et al, 2019):

$$(1 + GA + GF) = PD \cdot (1 - LGD) + (1 - PD) \cdot (1 + GA + GF + PRC)$$

Del caso anterior, GA corresponde a los gastos administrativos, GF al gasto financiero y LGD a la pérdida dado el incumplimiento de la cartera y/o producto en análisis.

A modo de ejemplo, se expondrá el potencial impacto en productos de consumo, considerando los resultados del modelo de admisión y los siguientes parámetros:

**Tabla 2:** Parámetros usados en la estimación del cambio en la PRC

Producto	LGD	GF	GA
Cuotas	36,2%	10,7%	5,1%
Tarjetas	51,0%	10,7%	5,1%

*Fuente:* Elaboración propia en base a información administrativa, considerando datos hasta agosto de 2022. Para GF y GA se considera información de las SAGs, cuyo negocio principal es la entrega de este tipo de producto.

Los resultados en la Tabla 4 muestran que la PD aumentaría en 347 pb si no se tuviera acceso a información de deuda, pasando desde un 6,8% a 10,3%. Luego, conocida la PD del modelo con y sin información, se estima el cambio en la PRC para tarjetas y créditos en cuotas de consumo, usando los parámetros de la Tabla 2. Los resultados en la Tabla 3 muestran que el cambio en PRC y, por lo tanto, en tasas de interés, sería de 283 pb y 521 pb en el caso de créditos en cuotas y tarjetas de crédito, respectivamente. Considerando que la tasa de interés en tarjetas fue de 31% a agosto de 2022, el impacto estimado implicaría un aumento de 17% en las tasas. En cuanto a créditos en cuotas, las tasas fueron de 20%, por lo que el impacto sería un aumento de 14%. El efecto en cuotas es más acotado, dado que tiene una LGD menor.

**Tabla 3:** Efecto en las tasas de interés de productos de consumo

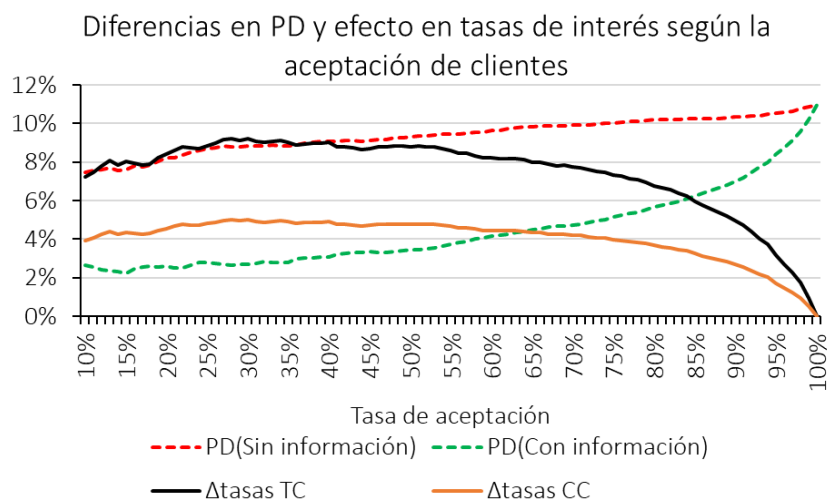
Ítem	Con información	Sin información	$\Delta$ pb
PD Consumo	6,8%	10,3%	347
PRC-Cuotas	5,6%	8,4%	283
PRC-Tarjetas	10,2%	15,5%	521

*Fuente:* Elaboración propia en base a resultados del modelo de admisión.

En términos prácticos, para un crédito en cuotas de \$5 millones a 60 meses plazo, se pagarían \$480 mil más en intereses por el incremento de la tasa de interés que se ha estimado. En el caso de una tarjeta, se pagarían \$65 mil más en intereses en una compra de \$1 millón a 24 meses plazo.

Finalmente, el cambio en PD podría ser más alto y, por lo tanto, el impacto más elevado en tasas, para aquellas entidades que tienen un perfil de riesgo más conservador, que son aquellas que aceptan menos clientes (figura 2). A medida que la tasa de aceptación baja, mayor es la brecha en PD que hay entre el modelo con y sin información de deuda, lo que se significa un efecto mayor en tasas de interés. En el caso más extremo, las tasas de interés de las tarjetas podrían subir en 9,24 puntos porcentuales (30% relativo); mientras que en cuotas el mayor efecto sería de 5,0 puntos porcentuales (25% en términos relativos).

**Figura 2:** Efecto en PD y tasas de consumo umbral de aceptación



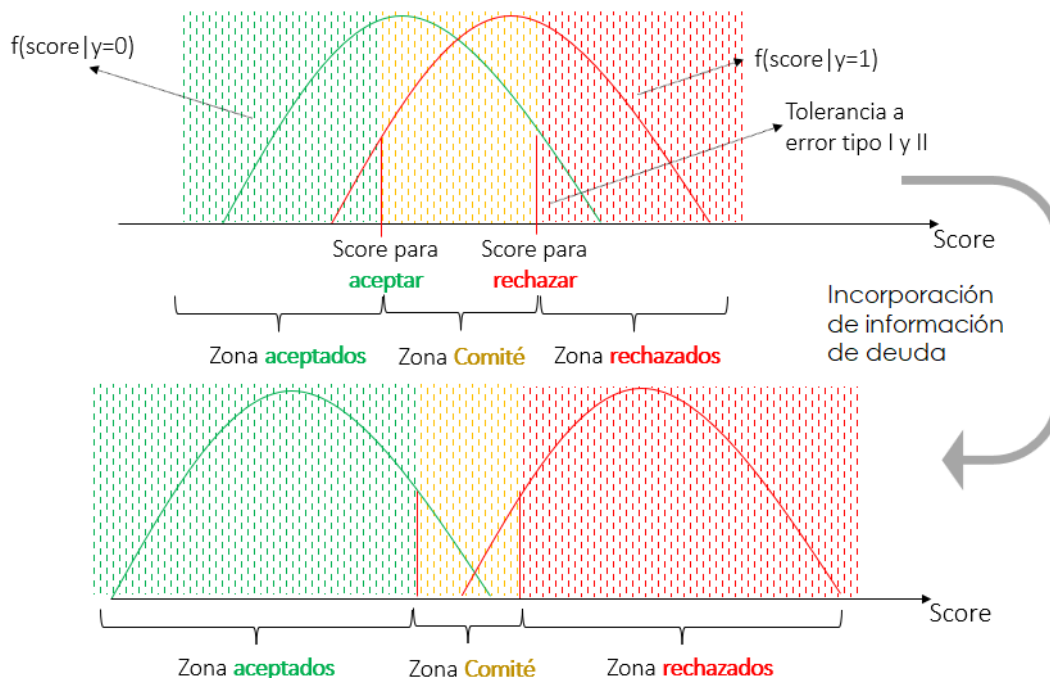
Fuente: Elaboración propia en base a resultados del modelo de admisión.

Ahora bien, el supuesto de que se mantenga el mismo nivel de aprobación sería poco realista en el largo plazo, debido a que los oferentes de crédito tendrían incentivos a actualizar sus modelos de otorgamiento. La menor bondad de ajuste que tendría el modelo sin información de deuda implicaría un mayor gasto administrativo, según se expone a continuación. Este es el segundo canal de transmisión del efecto de la ausencia de información sobre las tasas de interés.

Un modelo de PD entrega un puntaje para individuos que son buenos y malos pagadores, cuya distribución se denota por  $f(s|y = 0)$  y  $f(s|y = 1)$  respectivamente, tal como muestra el diagrama de la figura 3. Idealmente, se espera que estas distribuciones estén separadas de manera que, al establecer un umbral de aceptación de clientes (zona verde de la figura), se minimicen los aceptados que son malos pagadores (error tipo II) y que, al rechazar deudores (zona roja), se minimicen los rechazos de deudores que son en realidad buenos pagadores (error tipo I). En el intermedio, existe una zona de indeterminación (zona amarilla), que son los casos que una entidad analizaría en mayor profundidad, a través de un Comité.

La figura 3 da cuenta que incorporar más información de deuda mejora la discriminación del modelo, lo que se traduce en que, para un mismo nivel de tolerancia al riesgo: i) disminuyen los individuos que deben ir a un Comité de revisión, ii) más individuos catalogados como buenos pagadores son aceptados, y iii) una mayor cantidad de malos pagadores son detectados. Específicamente, el efecto i) y ii) se podría asociar a aumentos de la inclusión financiera, ya que permite un acceso al crédito de forma más fluida. Sin perjuicio de lo anterior, la presencia de un Comité de revisión significa costos relevantes para las entidades, por lo que, si tuvieran que examinar una gran cantidad de casos, es probable que no todos sean sujetos de evaluación, afectando el acceso a productos financieros. Por lo tanto, la existencia de información de deuda permitiría disminuir los costos en que incurrirían las IFIs para resolver la asimetría de información que existe en un contrato de crédito, ya que la instancia de Comité queda acotada a una menor cantidad de casos.

**Figura 3:** Diagrama sobre la mejora en la discriminación del modelo al incorporar información



Fuente: Elaboración propia.

Para definir el cambio de la proporción de clientes en cada categoría (aceptado, Comité, o rechazado) por la incorporación de información de deuda en el modelo, es necesario definir un nivel de tolerancia al error Tipo I (se identifica como “malo” a un deudor que no incumple sus obligaciones) y II (se identifica como “bueno” a un deudor que luego incumple sus obligaciones) que la IFI está dispuesta a tolerar. Esta decisión corresponde a un aspecto de la estrategia de negocios de cada oferente de crédito y, por lo mismo, sus niveles podrían ser distintos entre sí. Por lo anterior, la figura 4 muestra el cambio en la masa de deudores de cada categoría para diversos valores de tolerancia al error o apetito al riesgo<sup>11</sup>. Cabe tener presente que la figura 4 expone el resultado para el caso de admisión que es donde, de acuerdo con los resultados de la dimensión prudencial, la información de deuda tiene mayor relevancia. Sin perjuicio de ello, en el anexo 5 se detallan los resultados para el escenario de comportamiento y se complementan los resultados de admisión.

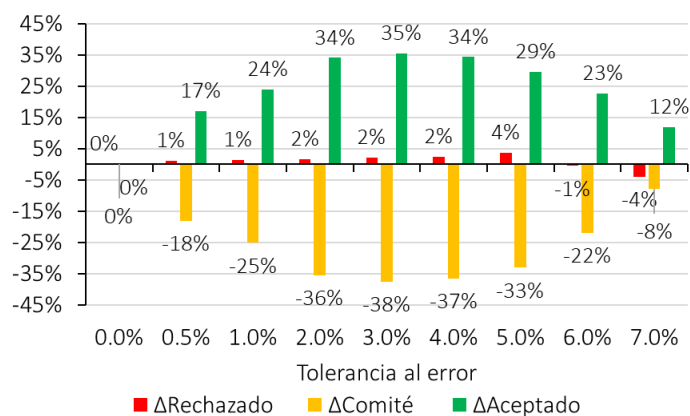
En un caso extremo, si la tolerancia al error tipo I y II es nula, todos los deudores deberían ser revisados por un Comité especializado. Dicho nivel de conservadurismo no permitiría extraer ganancias de la información de deuda, al no existir variaciones en la cantidad de clientes que son aceptados y rechazados por la incorporación de información de deuda. Ahora bien, es esperable que las IFIs toleren algo de error tipo I y II debido a que: i) empíricamente se observa incumplimiento en los deudores aceptados, y ii) el gasto de apoyo asociado a los Comités especializados podría ser elevado y, por lo tanto, podría ser razonable acotar su preponderancia. En el otro extremo, si existe un nivel elevado de tolerancia, la mayoría de los deudores serían aceptados sin necesidad de la existencia de un Comité de revisión.

<sup>11</sup> Los resultados se basan en el algoritmo GBM, el cual mejora el nivel de discriminación en comparación con la regresión logística.



En base a la discusión anterior, se observa que el efecto de la información es cóncavo al nivel de tolerancia al error. En la medida que este valor crece, la incorporación de la información permite aumentar los clientes que son aceptados automáticamente, disminuyendo la necesidad de pasar por un Comité de revisión. Es importante notar que, gracias a la incorporación de información de deuda se logra aminorar la cantidad de clientes que requieren ser evaluados por un Comité sin aumentar el nivel de error tipo I y II del modelo. Se observa que el valor de la información se maximiza en un nivel de tolerancia en torno al 3%, donde un 35% adicional de potenciales clientes serían automáticamente aceptados gracias al conocimiento de su comportamiento de pago a partir de la información de deuda.

**Figura 4:** Cambios en las tipologías de deudores según tolerancia al error, usando información de todos los oferentes de crédito



*Fuente:* Elaboración propia en base a los resultados del modelo de admisión.

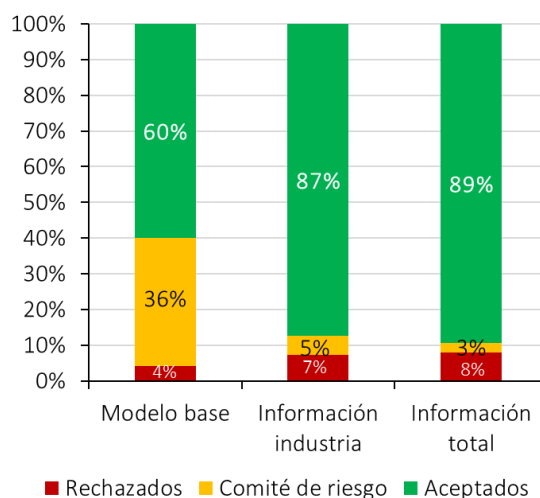
De acuerdo con la información administrativa, la PD empírica promedio de los clientes nuevos es de un 5% en el último año, lo cual, significa que, en el agregado, el sistema está dispuesto a tolerar un 5% aproximadamente de falsos negativos (error tipo II). Por lo cual, para el siguiente resultado se asumirá un nivel de tolerancia de 5%. En base a esto, los resultados de la figura 5 muestran que:

- a) Considerando la información de todos los oferentes, un potencial de 29% más de clientes podrían lograr un acceso más fluido a productos financieros en una nueva institución. Este valor se obtiene al observar que la cantidad de clientes aceptados cambia desde un 60%, cuando se conoce sólo la RDI, a un 89% cuando se conoce la información de deuda. Esto equivale a 2,3 millones de deudores beneficiados, considerando la cantidad de deudores vigentes en la nómina R04 a diciembre de 2020.
- b) La mayor información de las IFIs también afectaría el número de rechazados, pasando desde un 4% a un 8%, lo que representa aproximadamente 320.000 clientes. Ahora bien, dicho efecto sería marginal en comparación al total de nuevos clientes incluidos, pero prudencialmente deseable para así no generar incumplimientos elevados de acuerdo con el apetito por riesgo de la IFI.

Si una IFI no tiene ningún tipo de información, el 100% de los deudores pasarían por una instancia de Comité, ya que no existe un modelo para discriminar a los deudores cuantitativamente. Contar con información de la carga financiera logra que un 36% los casos vayan a Comité; mientras que, en el

caso de tener la información total del sistema financiero, los deudores en Comité disminuirían a sólo un 3% de acuerdo con los resultados anteriores (figura 5).

**Figura 5:** Ejercicio con todos los deudores y tolerancia al error de 5%



*Fuente:* Elaboración propia en base a los resultados del modelo de admisión y una tolerancia al error del 5%

Como se comentó anteriormente, la presencia de un Comité especializado generaría costos relevantes para una entidad, presionando eventualmente las tasas de interés al alza. Particularmente, el promedio del concepto “Gastos de apoyo” de los estados financieros de los bancos a diciembre de 2021, presenta un costo promedio de 1,54% de los activos totales.

**Tabla 4:** Importancia relativa de los gastos de apoyo promedio en relación con el (%) de activos

Item	% activo total
Margen de intereses	2,66%
Comisiones netas	0,62%
Gasto en provisiones totales	-0,76%
Gastos de apoyo	-1,54%
Otros	0,42%
<b>ROA (antes de impuestos)</b>	<b>1,40%</b>

*Fuente:* Elaboración propia en base al promedio de Estados financieros de las instituciones bancarias supervisadas por la Comisión.

Una menor necesidad de Comité permitiría reducir u optimizar el uso de los recursos asociados a gastos de apoyo, reduciendo los costos de resolver los problemas de asimetría de información. Para aproximar el impacto que genera una reducción de la presencia de un Comité especializado, se realiza una estimación de tasa de interés (TI). Tal como se mencionó en el ejercicio de tasas anterior, este valor se puede expresar como sigue:

$$TI = GA + GF + PRC$$

Con información de agosto de 2022, considerando la información del sistema bancario y de productos de consumo, los parámetros promedio de la ecuación anterior corresponden a:

$$TI = 1,9\% + 10,4\% + 10,2\% = 22,5\%$$

En el caso de que existiese ausencia de la información, si la IFI quisiera mantener una tasa de crecimiento similar de su cartera, el gasto de apoyo en Comités podría ser hasta 12 veces superior (36% versus 3%) de acuerdo con los resultados de la figura 5. Por lo anterior, la tasa de interés podría aumentar en 20,9% (11x1,9%) si es que existiera una relación completamente lineal entre los gastos administrativos y el tamaño del Comité. Sin embargo, es esperable encontrar economías de escala en este tipo de instancias, por lo que el efecto puede ser más acotado. Por otro lado, si es que la IFI no quisiera incurrir en más gastos de apoyo, eventualmente no evaluaría a un 33% de las solicitudes de crédito (36%-3%, según figura 5), lo que significaría en una reducción de la oferta de crédito a clientes nuevos en la misma proporción.

En base a la discusión anterior, es posible concluir que la información de deuda permite: i) que las instituciones ofrezcan tasas de interés más competitivas, gracias a que puede cuantificar de mejor manera la PRC y, además, realizar una asignación más eficiente de los recursos al no incurrir en mayores gastos de apoyo; y ii) en el agregado, mantener la oferta de crédito, gracias a que las instituciones incurren en menores costos para resolver los problemas de asimetría de información.

### 3) Ámbito de competencia:

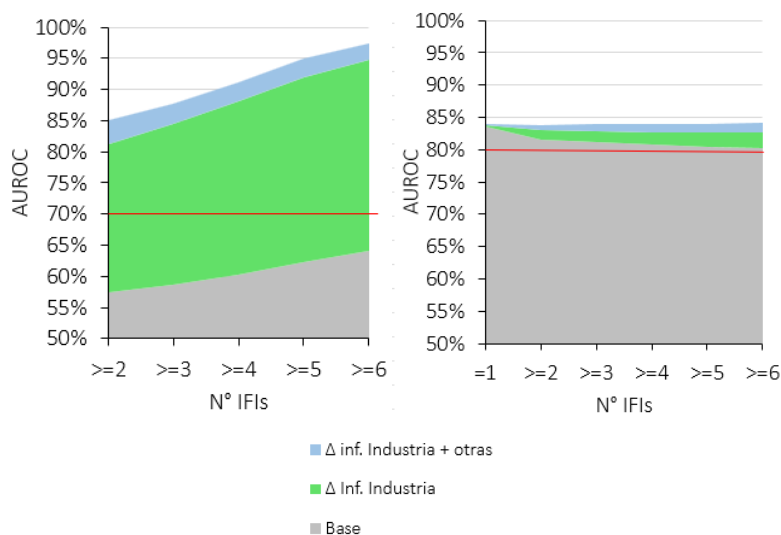
Respecto al factor competencia, la literatura indica que el intercambio de información crediticia reduce el monopolio de información que un prestamista tiene sobre sus prestatarios, favoreciendo la competencia (Padilla and Pagano, 1997). En ese sentido, la falta de acceso a la información del deudor actúa como una barrera de entrada para las otras IFIs.

La mayor integración de los clientes, que se refiere a tener productos en más instituciones del sistema financiero, genera un mayor poder de negociación del deudor pues puede acceder a los mismos servicios a través de diferentes IFIs. En el contexto de nuestras estimaciones, como muestra la figura 6, se observa que en la medida que el deudor es más integrado, mejora el nivel de discriminación que se produce gracias a la incorporación de la información de deuda, particularmente en el modelo de admisión.

En el ejercicio del modelo de comportamiento se observa que, mientras más integrado es el deudor, más importante es la información de deuda de la industria para mantener un nivel de discriminación por sobre el 80%. En este sentido, la información propia del oferente de crédito empieza a perder poder para lograr el nivel de desempeño que se obtiene en deudores integrados en otras IFIs; por lo que se vuelve necesaria la incorporación de información de deuda.

Considerando que, en ambos escenarios, el mayor efecto de la información ocurre en deudores más integrados en el sistema financiero, este tipo de información facilitaría que: i) individuos puedan acceder a otros productos de la misma institución o bien de otros oferentes de crédito y ii) los oferentes de crédito puedan identificar con mayor precisión potenciales nuevos clientes. Ambos efectos favorecen la competencia en el mercado de créditos, particularmente promoviendo la entrada de nuevos actores que podrían ofrecer sus productos con riesgos bien gestionados gracias a la información de deuda. Este hecho comprueba la relevancia que tiene la información para aquellos oferentes de menor tamaño, permitiendo que puedan entregar ofertas competitivas, elevando el nivel de competencia en el mercado de créditos.

**Figura 6:** AUROC según nº IFIs en la que tiene productos el deudor y escenario. A la izquierda, resultados para admisión y a la derecha de comportamiento



Fuente: Elaboración propia en base a los resultados de los modelos generales de admisión y comportamiento.

## V. Extensión histórica de la información

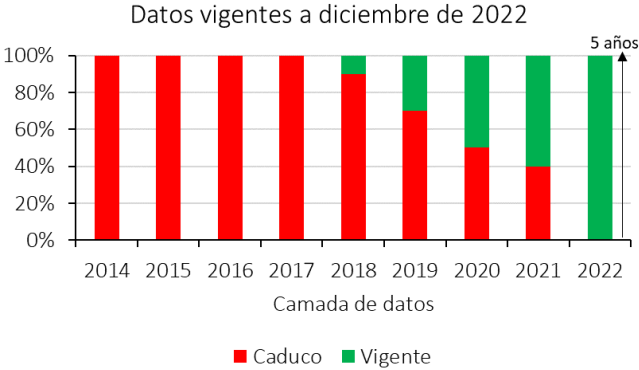
La información de deuda se utiliza fundamentalmente para monitorear el cumplimiento de las obligaciones. En ese sentido, la literatura académica divide a los morosos en dos grupos: los que “no pueden pagar” debido a que no tienen la capacidad de pagar y los que no tienen la voluntad de pagar, es decir, “no quieren pagar” (Bijak et al 2015). La literatura sugiere que para el primer grupo la pérdida de ingresos y/o el desempleo son los principales factores que los lleva a incumplir, mientras que aspectos que se orientan a la moral de pago del cliente, como por ejemplo si el deudor es responsable de pagar a tiempo sus deudas, se relacionan con las razones del segundo grupo para caer en morosidad. Esta diferencia es relevante a la hora de evaluar riesgo de crédito, ya que, si una persona cuenta con la capacidad de poder pagar su deuda, pero no tiene la voluntad de honrar dicho compromiso, simplemente no va a cumplir con la obligación de pagar. Dado que la voluntad de pago del segundo grupo es una característica directamente no observable, es que se hace relevante definir el grado óptimo de memoria y “perdón” de los registros de créditos considerando esta arista, y así, equilibrar la necesidad de disciplinar a los prestatarios con la necesidad de darles una “segunda oportunidad”.

Cuando los derechos de los acreedores están menos protegidos, la necesidad de disciplinar a los prestatarios puede ser más apremiante y, por lo tanto, es posible que sea deseable hacer que la memoria del sistema sea más larga y menos indulgente (Jappelli y Pagano, 2006). Luego, el tipo de información que se registra también cumple un rol al disciplinar la voluntad de pago, ya que, registros con información negativa prolongada generan mayor disciplina en el largo plazo aumentando los incentivos a pagar (Kovbasyuk y Spagnolo, 2021).

La Ley N°19.628 sobre protección de la vida privada, en su artículo 18, establece que: “En ningún caso pueden comunicarse los datos a que se refiere el artículo anterior, que se relacionen con una persona identificada o identificable, luego de transcurridos cinco años desde que la respectiva obligación se hizo exigible”. En otras palabras, en la calibración y aplicación de sus modelos internos las IFIs no pueden usar información personal que supere los 5 años de antigüedad. En este contexto, los datos que no se pueden utilizar reciben el nombre de “caducos”.

La restricción anterior implica que si, por ejemplo, se está aplicando o calibrando un modelo para datos personales con referencia a diciembre de 2022, no es posible incorporar información anterior a enero de 2018. Por otro lado, la mora histórica máxima posible de informar es menor para camadas más antiguas. Siguiendo con el mismo ejemplo, los datos de diciembre 2018 sólo podrán utilizar datos de mora histórica de máximo 1 año (no se podrán ver si hay morosidades superiores a ese plazo), ya que, de lo contrario, se estaría vulnerando el umbral de 5 años a diciembre de 2022. Por esto, la proporción de datos vigente es menor para camadas de datos antiguos (figura 7).

**Figura 7:** Datos vigentes y caducos a diciembre de 2022



Fuente: Elaboración propia.

El grado óptimo de años para registrar la información depende de muchas características, incluida, por ejemplo, la persistencia de *shocks* que inducen a la mora. De acuerdo con una revisión de otras jurisdicciones, el límite en años a informar difiere de país a país (tabla 5). Por ejemplo, en Dinamarca, las agencias de crédito tienen derecho a registrar y comunicar información de deuda con una historia máxima de 5 años; en Estados Unidos este máximo es de 7 años; y en Bélgica, en caso de las deudas vencidas, los datos se registran por un máximo de 10 años. En contraste, en Perú y El Salvador, la normativa solo permite mostrar el historial del cliente por dos y tres años, respectivamente. De acuerdo con la revisión de la experiencia internacional se puede estimar un promedio de 6 años, y una mediana de 5 años.

**Tabla 5:** Límite para informar deuda en registro consolidado de otros países

País	Límite en informar deuda (años)
Bélgica	10
Nicaragua	10
Estados Unidos	7
Francia	7
Reino unido	6
México	6
Australia	5
Bolivia	5
Brasil	5
Dinamarca	5
Colombia	4
Italia	3
El Salvador	3
Perú	2

*Fuente:* Elaboración propia en base a revisión de legislación internacional.

Finalmente, en base al modelo de admisión se puede estudiar la relevancia de la información histórica para estimar la capacidad de pago. Sin embargo, este resultado no permite sugerir un nivel de tiempo ideal para el reporte de deuda por los siguientes argumentos: 1) el ejercicio sólo considera la dimensión de la selección adversa (identificación de deudores al otorgamiento), sin tener presente si este plazo es razonable para aminorar los problemas de riesgo moral; 2) se basa en clientes que han sido aceptados y visibles en la nómina de deudores, por lo que tiene un sesgo que podría manifestarse en un umbral menor al que la industria considera efectivamente para rechazar a los deudores; y 3) el ciclo económico utilizado podría tener incidencia relevante en los resultados. Independientemente de este tipo de análisis, considerar 5 años de memoria en los registros como lo define actualmente nuestra legislación es un criterio consistente con la experiencia internacional revisada.

## VI. Implicancias de política pública

En Chile, la CMF consolida una nómina de deudores en base a la información sobre los créditos vigentes que reportan los bancos, las sociedades de apoyo al giro colocadoras de créditos, las cooperativas de ahorro y crédito supervisadas por la CMF, y las empresas emisoras de tarjetas de crédito no bancarias. No obstante, el Banco Mundial y el FMI (2004, 2011, 2021), en distintas evaluaciones del sistema financiero chileno, han destacado la importancia de ampliar el registro de deudores para evitar sobreendeudamiento, promover la competencia y fortalecer este registro. En esta línea, en el Congreso Nacional existen diversas iniciativas legales para potenciar la divulgación de la información de deuda, tales como el proyecto de ley que crea un Registro de Deuda Consolidada Boletín N°14.743-03 (REDEC) y otras para resguardar los derechos de privacidad de la información de deudores, como lo es el Boletín 11.144-07. A su vez, la recientemente aprobada Ley FINTEC, a través de la creación del Sistema de Finanzas Abiertas, obliga a las IFIs a compartir información de sus clientes con terceros que quieran ofrecerles sus productos financieros.

Específicamente, respecto al proyecto de Ley REDEC, se busca crear un registro de deuda administrado por la CMF que ampliaría la información de endeudamiento disponible al incluir nuevas entidades, tales como compañías de seguros y mutuarías, cooperativas no supervisadas por la CMF, cajas de compensación, securitizadoras y otras que la CMF establezca mediante norma de carácter general. Los aspectos relevantes del proyecto, hasta ahora, son: i) considera tanto personas naturales como personas jurídicas, ii) reportantes y mandatarios deben cumplir estrictos estándares de seguridad, lo que fortalece la protección de los datos de deuda, y iii) establece que el titular dé su consentimiento para el tratamiento de sus datos de deuda positiva y el derecho al olvido (ventana de 5 años de información). La CMF puede multar a los reportantes y prohibir el acceso al registro a entidades que infrinjan las disposiciones sobre seguridad de la información. En base a los resultados de este estudio se considera importante y urgente avanzar en un proyecto de deuda consolidada, así como de todas aquellas iniciativas que apunten a potenciar la divulgación de la información de deuda.

### **1) Importancia de un registro consolidado de deudores.**

Desde la perspectiva prudencial, los resultados muestran que es relevante la información de deuda para lograr una buena estimación de los parámetros de riesgo, particularmente en la admisión de nuevos clientes. Además, considerando el ámbito de mercado, no contar con acceso a la información de deuda reduce el nivel la bondad de ajuste de los modelos, lo que se podría traducir en un aumento de la PD en un 51% promedio en productos de consumo, implicando aumentos en tasas de interés y/o reducciones del crédito, con un impacto mayor para aquellas entidades que tienen un perfil de riesgo más conservador. Asimismo, la existencia de un registro consolidado de deudores permitiría disminuir los costos en que incurrirían las IFIs para resolver la asimetría de información que existe en un contrato de crédito, permitiendo que realicen una evaluación mediante Comités a una proporción menor de deudores y así optimizar el uso de los recursos asociados a gastos de apoyo.

Desde la arista de competencia de mercado, los resultados indican que más y mejor información tiene efectos mayores en los deudores más integrados al sistema financiero. Un registro de deuda facilitaría que: i) individuos puedan acceder a otros productos de la misma institución, o bien de otros oferentes de crédito y ii) los oferentes de crédito puedan evaluar con mayor precisión potenciales nuevos clientes. Ambos efectos favorecen la competencia en el mercado del crédito, particularmente promoviendo la entrada de nuevos actores que podrían ofrecer sus productos con riesgos bien gestionados gracias a la información de deuda completa.

El otorgamiento de créditos informales en entidades no reguladas (*shadow banking*), no permitiría explotar los beneficios de un registro de deuda, pues los oferentes de crédito del sistema formal no lograrían observar el comportamiento de pago de dichos deudores. Mostrar la capacidad de pago en la industria genera mejoras en las condiciones de acceso con otros oferentes de crédito. Este elemento es aún más relevante para que se puedan materializar ofertas competitivas en industrias y/o entidades emergentes.

### **2) Extensión histórica y seguridad de la información de deudores**

En términos generales, el límite en años a informar debe: i) balancear los elementos asociados al problema de riesgo moral asociado a la voluntad de pagar, y el derecho al olvido de los deudores; y ii) otorgar un piso suficiente de información para permitir a la institución aliviar el problema de selección adversa en relación con la capacidad de pago. En ese sentido, el umbral de 5 años definido por la legislación local es consistente con el nivel medio que se observa en la experiencia internacional.

Por otro lado, los estándares de Basilea e IFRS 9 exigen un ciclo económico completo de observación para la estimación de los parámetros de la PD y LGD. La literatura identifica que la duración promedio de los ciclos económicos varía entre países, donde los emergentes presentan una duración promedio de 10 años, pero con una mayor dispersión que los países desarrollados. En particular, se estima que para Chile la duración promedio de un ciclo económico es de 8 años (Rand y Tarp, 2002).

Una herramienta que ayuda a extender la ventana de datos aproximándose a la extensión del ciclo económico es la anonimización o disociación de los datos (Garrido et al, 2019). Este proceso permite almacenar y utilizar la información de deuda, pero respetando las legislaciones de protección de datos personales<sup>12</sup>.

### 3) “Borronazos” y eliminación de información de deudores.

En contraste a los esfuerzos legales para potenciar la divulgación de la información de deudores, actualmente en Chile se discuten proyectos de ley que buscan restringir y eliminar información de deudores del sistema. Estos corresponden a las iniciativas:

- i. “Chao Dicom”: iniciativa que busca establecer la prohibición de proporcionar información sobre deudas exigibles desde el 18 de octubre de 2019 al 31 de mayo de 2022, para deudores con un total adeudado inferior a los 2,5 millones de pesos.
- ii. “Chao Histórico”: iniciativa que busca consagrar el derecho al olvido financiero, mediante la eliminación de manera permanente de las bases de datos de todos los responsables y de los registros de deuda, de las deudas extinguidas, prescritas o saldadas hace cinco años. Además, obliga a las entidades financieras a entregar un informe cuando nieguen un crédito a las personas.

En cuanto estos proyectos de ley y los resultados de este estudio, se podría señalar que:

- Si los proyectos implican una eliminación parcial de registros de deuda, los modelos podrían subestimar la cuantificación del riesgo en el caso de aquellos individuos que se han beneficiado de la eliminación de sus deudas.
- Si la eliminación es de tipo permanente, los modelos se recalibrarían a la observación del comportamiento parcial de pago. Sin embargo, al no existir información “negativa”, no sería posible distinguir a los buenos pagadores. Como consecuencia, la información de deuda no lograría eliminar la incertidumbre en la cuantificación del riesgo, generando un aumento del riesgo en las carteras y, con ello, en tasas de interés.

En base a los proyectos que se mencionan como ejemplo, un “borronazo” del tipo “Chao Dicom”, en términos agregados, encarecería el crédito y existirían caídas en los niveles de préstamos producto de la mayor incertidumbre y los problemas de asimetría de información. En el segundo caso, proyectos del tipo “Chao Histórico”, pueden tener consecuencias negativas si el plazo de “olvido” es inferior a 5 años.

---

<sup>12</sup> Para más información sobre el proceso anonimización o disociación de datos personales revisar el Anexo 6 de este documento.



## VII. Conclusiones

La literatura económica y la evidencia empírica demuestran que más y mejor información sobre deudores permitiría una predicción más precisa de la probabilidad de incumplimiento por parte de las IFIs, mitigando los problemas de asimetría de información, disminuyendo con ello el riesgo moral y la selección adversa que se producen entre prestamistas y prestatarios. En concordancia, en base a los ejercicios para Chile presentados en este documento, se encuentra que la información de deuda, tanto positiva como negativa, es clave para realizar una correcta cuantificación de los riesgos, lo cual permite: 1) mejor protección de los intereses de depositantes e inversionistas; 2) mantener el crecimiento de la cartera de créditos sin experimentar aumentos del incumplimiento que estén fuera del perfil de riesgo de las IFIs; 3) facilitar el acceso a productos financieros, principalmente a nuevos clientes; y 4) favorecer la competencia en el mercado crediticio.

Contar con un registro de información consolidada de deudores que incluya a todo el sistema financiero, y presente información positiva y negativa, genera beneficios prudenciales, de desarrollo de mercado y de mayor competencia. Particularmente, más información: 1) permite mejorar los modelos de discriminación, 2) reducir los costos y generar una asignación más eficiente de los recursos para posibilitar que las IFIs ofrezcan tasas de interés más competitivas y una mayor oferta de crédito, y 3) permite a los clientes financieros acceder a otros productos de la misma institución, o bien de otros oferentes de crédito. Con todo, estos resultados respaldan la idea de avanzar con iniciativas que favorecen la divulgación de información de deuda, y de revisar en detalle aquellas que buscan restringir y/o eliminar información, pues estas pueden tener efectos negativos para los clientes financieros.

Sin perjuicio de la robustez de los resultados de este artículo, líneas de investigación futuras podrían expandir el estudio sobre entidades que recientemente se han incorporado en los registros de deuda; estudiar el efecto de la información de otros tipos de nóminas, tales como DICOM en la medida que dicha información sea posible de usar según la Ley; y segmentar los hallazgos según características de los deudores.

Finalmente, cabe destacar que, parecieran razonables los 5 años legales de plazo respecto del derecho al olvido, pues balancea la sanción al "mal pagador", reduciendo el riesgo moral y la protección al deudor; permitiendo que vuelvan a obtener endeudamiento. Sin perjuicio de ello, un modelo económico micro fundado que considere los costos que se originan debido a la presencia de asimetría de información en el mercado de crédito, podría permitir generar nuevos antecedentes para calibrar el umbral ideal que debería tener el registro. Por otro lado, se tiene que los estándares internacionales en materia prudencial exigen que se utilice información de un ciclo económico completo (8-10 años para Chile). En este sentido, la anonimización de datos es una herramienta que están desarrollando diferentes jurisdicciones para lograr que, al mismo tiempo que se cumpla con las normas de protección de datos personales, las entidades financieras puedan utilizar información histórica para modelar el riesgo de crédito de sus carteras.

## Referencias

- Acharya, V. V., Amihud, Y., y Litov, L. (2011). Creditor rights and corporate risk-taking. *Journal of Financial Economics*, 102(1), 150-166.
- Akerlof, G. (1970). The Market of Lemons: Qualitative Uncertainty and the Market Mechanism. *Quarterly Journal of Economics*, Vol 84(3), págs. 488-500.
- Banco Mundial (2021). Chile: Financial System Stability Assessment. <https://www.imf.org/en/Publications/CR/Issues/2021/12/09/Chile-Financial-Systems-Stability-Assessment-510866>
- Barron, John M., and Michael Staten. —The Value of Comprehensive Credit Reports: Lessons from the U.S. Experience.|| In *Credit Information Systems and the International Economy*, by Margaret J. Miller, 273-310. Cambridge, MA: The MIT Press, 2003.
- Bijak, K., Mues, C., So, M. C., & Thomas, L. (2015). Credit card market literature review: Affordability and repayment. In *Casualty Actuarial Society E-Forum*, Spring (Vol. 64, No. October, pp. 1-62).
- Bishop, Ch. (2006). “Linear Models for Classification”. Chapter 4 of *Pattern Recognition and Machine Learning*. New York :Springer, 2006.
- Brasher, E. A. (2018). Addressing the failure of anonymization: guidance from the European union's general data protection regulation. *Colum. Bus. L. Rev.*, 209.
- Brockman, P., y Unlu, E. (2009). Dividend policy, creditor rights, and the agency costs of debt. *Journal of Financial Economics*, 92(2), 276-299.
- Brown, M., Jappelli, T., y Pagano, M. (2009). Information sharing and credit: Firm-level evidence from transition countries. *Journal of Financial intermediation*, 18(2), 151-172.
- Bücker, M., van Kampen, M., & Krämer, W. (2013). Reject inference in consumer credit scoring with nonignorable missing data. *Journal of Banking & Finance*, 37(3), 1040-1045.
- Chen, G. G., & Åstebro, T. (2001). The economic value of reject inference in credit scoring. Department of Management Science, University of Waterloo.
- Chen, G. G., & Åstebro, T. (2012). Bound and collapse Bayesian reject inference for credit scoring. *Journal of the Operational Research Society*, 63(10), 1374-1387.
- Claessens, S., y Klapper, L. F. (2005). Bankruptcy around the world: Explanations of its relative use. *American Law and Economics Review*, 7(1), 253-283.
- Djankov, S., McLiesh, C., y Shleifer, A. (2007). Private credit in 129 countries. *Journal of Financial Economics*, 84(2), 299-329.
- Feelders, A. J. (2000). Credit scoring and reject inference with mixture models. *Intelligent Systems in Accounting, Finance & Management*, 9(1), 1-8.
- Feelders, A. J. (2003). An overview of model-based reject inference for credit scoring.
- FMI (2004). “Chile: Financial System Stability Assessment,” IMF Country Report No. 04/269, August.
- FMI (2011). “Chile: Financial System Stability Assessment,” IMF Country Report No. 11/261, August.
- FMI (2021). “Chile: Financial System Stability Assessment,” IMF Country Report No. 21/262, December.
- Forteza, J., y Medina, V. (2018). Historial de morosidad y riesgo de crédito: Implicancias de prohibir el uso de registros históricos de datos comerciales caducos.

- Forteza, J., Medina, V., & Pulgar, C. (2018). Marco general para el diseño de métodos estándar de provisiones por riesgo de crédito. Series de Estudios Normativos de la SBIF, (18/04).
- Friedman, J. H. (1999). "Greedy Function Approximation: A Gradient Boosting Machine".
- Galindo, A., y Miller, M. (2001). Can credit registries reduce credit constraints? Empirical evidence on the role of credit registries in firm investment decisions. In Annual Meetings of the Inter-American Development Bank, Santiago Chile.
- Garrido, J., Bergthaler, M. W., DeLong, M. C. M., Johnson, J., Rasekh, A., Rosha, A., & Stetsenko, N. (2019). The use of data in assessing and designing insolvency systems. International Monetary Fund.
- Houston, J. F., Lin, C., Lin, P., y Ma, Y. (2010). Creditor rights, information sharing, and bank risk taking. *Journal of Financial Economics*, 96(3), 485-512.
- Jappelli, T., & Pagano, M. (2002). Information sharing, lending and defaults: Cross-country evidence. *Journal of Banking & Finance*, 26(10), 2017-2045.
- Jappelli, Tullio & Pagano, Marco. (2006). Role and Effects of Credit Information Sharing. *Econ. Consum. Credit*
- Kallberg, J. G., & Udell, G. F. (2003). The value of private sector business credit information sharing: the US case. *Journal of Banking & Finance*, 27(3), 449-469.
- Ke, G., Meng, Q., Finley, T., Wang, T., Chen, W., Ma, W., et al. (2017). "Lightgbm: A highly efficient gradient boosting decision tree". *Advances in Neural Information Processing Systems*, 30, 3146–3154.
- Kovbasyuk, S., & Spagnolo, G. (2021). Memory and markets. Available at SSRN 2756540.
- Hosmer, D.W. and Lemeshow, S. (2000). "Assessing the Fit of the Model". Chapter 5 of *Applied Logistic Regression* (eds W.A. Shewhart, S.S. Wilks, D.W. Hosmer and S. Lemeshow).
- Li, Z., Tian, Y., Li, K., Zhou, F., & Yang, W. (2017). Reject inference in credit scoring using semi-supervised support vector machines. *Expert Systems with Applications*, 74, 105-114.
- Liberman, A., Neilson, C., Opazo, L., y Zimmerman, S. (2018). The equilibrium effects of information deletion: Evidence from consumer credit markets (No. w25097). National Bureau of Economic Research.
- Mancisidor, R. A., Kampffmeyer, M., Aas, K., & Jenssen, R. (2020). Deep generative models for reject inference in credit scoring. *Knowledge-Based Systems*, 196, 105758.
- Mankiw, G. (1992). "The Allocation of Credit and Financial Collapse." *New Keynesian Economics*. Third printing. MIT Press, Cambridge, Massachusetts. Vol. 2, pp. 277-291.
- Mishkin, F. S. (1999). International capital movements, financial volatility and financial instability.
- Ohm, P. (2009). Broken promises of privacy: Responding to the surprising failure of anonymization. *UCLA L. Rev.*, 57, 1701.
- Powell, Andrew, Nataliya Mylenko, Margaret Miller, and Giovanni Majnoni. *Improving Credit Information, Bank Regulation and Supervision: On the Role and Design of Public Credit Registries*. Washington, D.C.: World Bank Research Working Paper Series, 2004.
- Pulgar, C. & Rojas, C. "Estimación de prima por riesgo de crédito y otros costos implícitos en tasas de interés de créditos bancarios". Documento de Trabajo N°03/19 de la CMF. Julio 2019.

- Rand, J., & Tarp, F. (2002). Business cycles in developing countries: are they different? *World Development*, 30(12), 2071-2088.
- Stiglitz, J. E., & Weiss, A. (1981). Credit rationing in markets with imperfect information. *The American Economic Review*, 71(3), 393-410.
- Triki, T., y Gajigo, O. (2014). Credit bureaus and registries and access to finance: new evidence from 42 African countries. *Journal of African Development*, 16(2), 73-101.

## Anexos

### Anexo 1 – El caso del “borronazo” chileno de 2012

Mediante la información histórica de deuda los burós de crédito mejorarían el conocimiento que las entidades financieras tienen sobre la capacidad de pago de los solicitantes, permitiendo una predicción más precisa de la probabilidad de incumplimiento (PD, por sus siglas en inglés), facilitando a los prestamistas una mejor orientación de sus préstamos, y disminuyendo los problemas derivados de selección adversa entre las partes.

En 2012, los burós de crédito chilenos se vieron obligados a dejar de reportar incumplimientos de 2,8 millones de individuos, tras la aprobación de la Ley 20.575<sup>13</sup>, que establece el principio de finalidad en el tratamiento de datos personales. En particular, el artículo N°2 transitorio indica que las instituciones financieras no podrán comunicar los datos relativos obligaciones cuando se hayan hecho exigibles antes del 31 de diciembre de 2011 y se encuentren impagas, siempre que el total de obligaciones impagas del titular sea inferior a \$2.500.000 (excluyendo intereses y reajustes).

En este contexto, Liberman et al (2018) utilizando datos del mercado de crédito chileno, implementan técnicas de *machine learning* para generar predicciones sobre la PD bancaria de los deudores y, con ello, estudiar los efectos de esta política. Los principales resultados de los autores indican que el no reporte de los datos de deudores beneficiados con la política produjo sobreestimaciones sistemáticas de PD para los clientes sin mora, aumentando en 22%, y subestimaciones de la PD de un 29% en promedio para aquellos deudores con mora. Particularmente, el “borronazo” del 2012 generó pérdidas regresivas, ya que indujo un efecto mayor en los no morosos entre las entidades no bancarias que en las entidades bancarias. Lo anterior, debido a que los bancos continuaron observando incumplimientos bancarios después de la eliminación, pero la información de burós de crédito eliminada era la única información de incumplimiento disponible para los prestamistas no bancarios.

Además, con políticas de eliminación de información, los mayores perdedores son los individuos menos riesgosos que tienden a parecerse a los mayores ganadores, es decir, a los más riesgosos, en todas las características observables. Finalmente, en el agregado hubo una reducción del crédito, donde la cartera de consumo se redujo un 3,5%, equivalente a \$20 mil millones, con mayores restricciones de acceso para las personas de bajos ingresos y saldos de préstamos más pequeños.

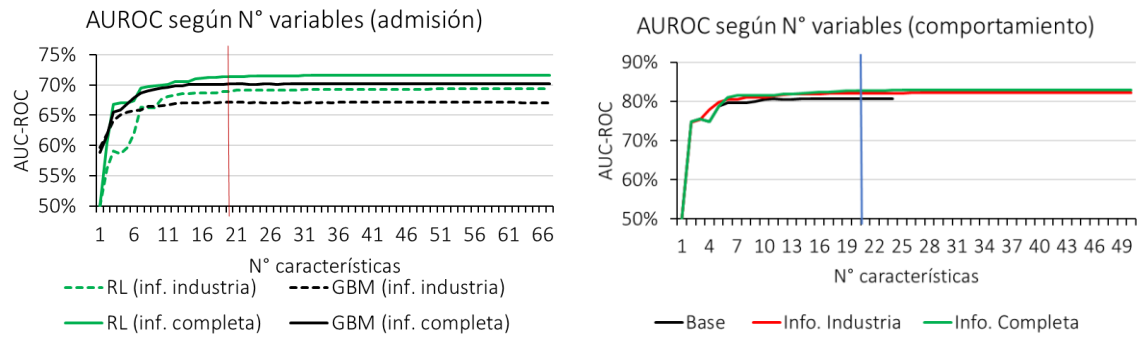
#### Anexo 2.1 – Resultados del algoritmo RFE

El resultado de la ejecución del algoritmo de selección (RFE) se observa en la figura A.1. Se observa que luego de 20 variables no existen efectos relevantes sobre el poder de discriminación del modelo. Este algoritmo, entonces, permite identificar las variables más relevantes para cada escenario.

---

<sup>13</sup> Fuente: <https://www.bcn.cl/leychile/navegar?idNorma=1037366>

**Figura A.1:** resultado de algoritmo RFE



\*RL: regresión logística, GBM: Gradient Boosting Machines.

Fuente: Elaboración propia en base a los resultados de los modelos generales de comportamiento y admisión.

## Anexo 2.2 – Variables seleccionadas en ejercicio de admisión

**Tabla A.1:** variables seleccionadas en las especificaciones del ejercicio de admisión

R	RL (industria)	GBM (industria)	RL (completa)	GBM (completa)
1	const	TEND_T0_T12_SIST	const	TEND_T0_T12_SIST
2	RDI	TEND_T3_T12_SIST	TEND_T3_T12_SIST	TEND_T3_T12_SIST
3	TEND_T0_T6_SIST	MAX_MORA_T12_SIST	MAX_MORA_T12_SIST	MAX_MORA_T12_SIST
4	TEND_T3_T6_SIST	RDI	TEND_T0_T12_SIST	TEND_T0_T6_SIST
5	TEND_T3_T12_SIST	TIPO_1	TEND_T0_T6_SIST	TIPO_1
6	TEND_T0_T12_SIST	TEND_T3_T6_SIST	TEND_T3_T6_SIST	RDI
7	MAX_MORA_T12_SIST	TEND_T0_T6_SIST	TIPO_1	N_IFI
8	BIN_DET60_SIST_T12	N_IFI	MAX_MORA_TO_SIST	TEND_T3_T6_SIST
9	BIN_DET30_SIST_T6	RATIO_30D_12M_SIST	TIPO_2	RATIO_30D_SIST
10	TIPO_1	TIPO_2	BIN_DET180_SIST_T36	MAX_MORA_TO_SIST
11	N_IFI	MAX_MORA_TO_SIST	RATIO_30D_12M_SIST	TIPO_2
12	BIN_DET180_SIST_T36	BIN_DET180_SIST_T36	BIN_DET1A_SIST	N_INDUSTRIAS
13	TIPO_2	BIN_DET30_SIST_T36	BIN_DET180_SIST	RATIO_30D_12M_SIST
14	TIPO_3	RATIO_30D_SIST	RATIO_60D_12M_SIST	BIN_DET180_SIST_T36
15	RATIO_60D_12M_SIST	BIN_DET90_SIST_T36	BIN_DET3A_SIST_T12	RATIO_60D_12M_SIST
16	MAX_MORA_TO_SIST	RATIO_60D_12M_SIST	BIN_DET90_SIST_T12	RATIO_180D_SIST
17	BIN_DET30_SIST_T24	BIN_DET60_SIST_T36	BIN_DET2A_SIST	BIN_DET30_SIST_T36
18	BIN_DET30_SIST_T36	BIN_DET30_SIST_T6	RDI	RATIO_60D_SIST
19	BIN_DET180_SIST	TIPO_3	N_INDUSTRIAS	BIN_DET30_SIST_T6
20	RATIO_30D_12M_SIST	RATIO_90D_SIST	BIN_DET60_SIST_T12	BIN_DET60_SIST_T24

### Anexo 2.3 – Variables seleccionadas en ejercicio de comportamiento

**Tabla A.2:** variables seleccionadas en las especificaciones del ejercicio de comportamiento

R	Base	Inf. Industria	Inf. Completa
1	const	const	const
2	MAX_MORA_T6_TO	MAX_MORA_T6_TO	MAX_MORA_T6_TO
3	RATIO_30D_12M_BCO	MAX_MORA_TO_TO	RATIO_30D_12M_BCO
4	MAX_MORA_TO_TO	TEND_T3_T12_SIST	MAX_MORA_TO_TO
5	TEND_T3_T12_BCO	BIN_DET30_SIST_T12	TEND_T3_T12_SIST
6	MAX_MORA_T12_TO	MAX_MORA_T12_TO	BIN_DET30_SIST_T6
7	MAX_MORA_T3_TO	TEND_T3_T6_SIST	MAX_MORA_T12_TO
8	TEND_T3_T6_BCO	TEND_T3_T12_BCO	TEND_TO_T6_SIST
9	BIN_GAR_CHIP	TEND_TO_T6_SIST	RATIO_60D_12M_SIST
10	CF	TEND_TO_T12_SIST	TEND_T3_T6_SIST
11	RDI	CF	TEND_TO_T12_SIST
12	TEND_TO_T12_BCO	RDI	TEND_T3_T12_BCO
13	TEND_TO_T6_BCO	MAX_MORA_T3_TO	RATIO_90D_12M_SIST
14	BIN_RENEGOCIADO_TO_TO	BIN_DET30_SIST	CF
15	RATIO_90D_BCO	RATIO_60D_12M_SIST	N_IFI
16	RATIO_180D_BCO	RATIO_90D_12M_SIST	BIN_DET30_SIST
17	RATIO_90D_12M_BCO	BIN_RENEGOCIADO_T12_TO	MAX_MORA_T3_TO
18	RATIO_60D_BCO	BIN_RENEGOCIADO_T6_TO	BIN_GAR_CHIP
19	RATIO_30D_BCO	BIN_RENEGOCIADO_T3_TO	RDI
20	BIN_RENEGOCIADO_T12_TO	BIN_RENEGOCIADO_TO_TO	BIN_DET30_SIST_T12
21	BIN_RENEGOCIADO_T6_TO	RATIO_30D_12M_SIST	BIN_RENEGOCIADO_T12_TO

### Anexo 2.4 – Diccionario de las variables

**Tabla A.3:** diccionario de las variables utilizadas en los modelos finales

Variable	Descripción
BIN_DETXX_SIST_TYY	Variable que toma valor 1 si el deudor tiene mora mayor o igual a XX en el sistema, observado hasta YY periodos atrás.
BIN_GAR_CHIP	Variable que toma valor 1 si la persona tiene crédito con garantía hipotecaria en el banco.
BIN_RENEGOCIADO_TY_TO	Variable que toma valor 1 si el deudor tiene una operación renegociada en los últimos Y meses.
CF	Carga financiera.
const	Constante (válido para regresión logística)
MAX_MORA_TY_SIST	Máximo nivel de mora observado en los últimos Y meses en el sistema.
MAX_MORA_TY_TO	Máximo nivel de mora observado en los últimos Y meses en el banco.

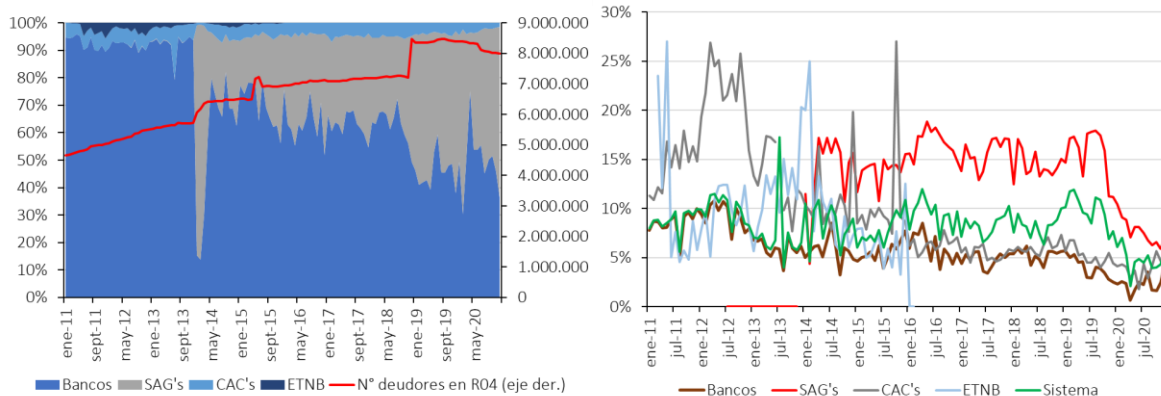
N_IFI	Número de instituciones en las cuáles el deudor presenta al menos un producto financiero.
N_INDUSTRIAS	Número de industrias en las cuáles el deudor tiene un producto financiero.
RATIO_XX_12M_YY	Promedio del ratio de deuda morosa a XX periodo respecto al total en los últimos 12 meses, medido a nivel de YY.
RATIO_XX_YY	Ratio de deuda morosa a XX periodo respecto al total, medido a nivel de YY.
RDI	Razón deuda a ingresos.
TEND_TX_TY_ZZ	Promedio de deuda en los últimos TX meses respecto al promedio de los últimos TY meses, medido a nivel ZZ.
TIPO_X	Binaria que marca el tipo de IFI: 1) bancos, 2) SAG, 3) CAC y 4) ETNB.

### Anexo 3.1: Caracterización de los nuevos clientes (ejercicio admisión)

Al momento de la admisión, un deudor resulta totalmente desconocido para la IFI, salvo información que se logre acreditar desde diversas fuentes. Una de ellas es la de comportamiento que se obtiene a partir del archivo R04 que consolida la información de deuda de los bancos, CACs y ETNB. En base a esto, resulta interesante determinar si se puede construir un modelo con discriminación aceptable con sólo esta información.

El número de deudores informados en el R04 llega a 8 millones de personas a diciembre de 2020 (este valor se elevó a 10 millones con la entrada de los emisores de tarjeta no bancarios en dic-21). Los clientes nuevos ocurren principalmente en Bancos y SAG's. La PD de este tipo de clientes es de un 8,3%, siendo más elevada en SAG's (detalle en figura A.3).

**Figura A.2:** a la izquierda: Distribución de los clientes nuevos entre los tipos de instituciones y número de deudores informados en el archivo R04. A la derecha: valor de la PD de clientes nuevos según el tipo de industria

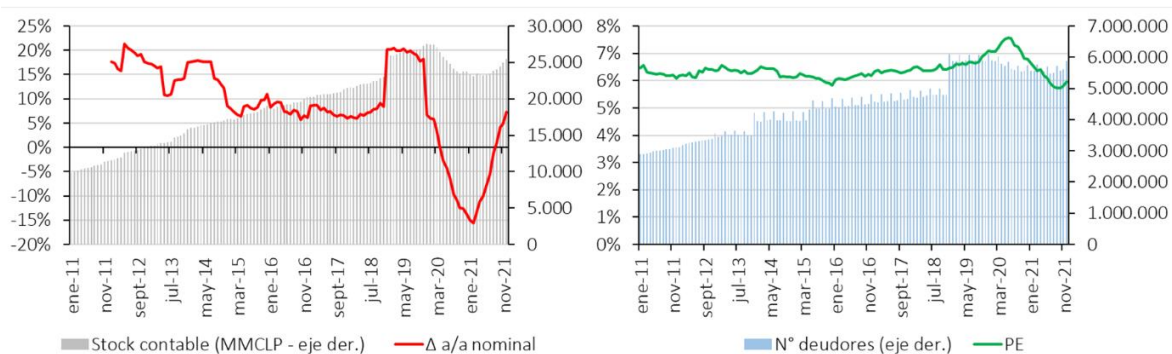


### Anexo 3.2: Caracterización de la cartera de consumo (ejercicio comportamiento)



A diciembre de 2021 la exposición activa de la cartera de consumo de los bancos totalizaba 25.000 millones de pesos. La variación nominal anual promedio era de un 8%, destacando la notable contracción durante el año 2020. En cuanto a los deudores, estos totalizaban alrededor de 5,6 millones de personas a la misma fecha. La pérdida esperada promedió un 6,4%.

**Figura A.3:** Stock contable de la cartera de consumo de los bancos y variación anual de su valor; y número de deudores de la cartera y la pérdida esperada (PE) promedio de la cartera



### Anexo 3.3: Distribución de los deudores entre industrias

Una gran proporción de clientes tiene productos en más de una entidad, particularmente en CAC's y SAG's (tabla A.4). Por otro lado, un 51,5% de los clientes tiene deuda en solo una industria (tabla A.5). A nivel de tipo de entidad, los deudores más integrados al sistema financiero son los de las CAC's, con un 81,3% de sus clientes con deuda en otra industria. Le siguen las SAG's con un 74,5%, bancos con un 64,6% y ETNB con un 62,3%.

El alto nivel de integración entre distintos tipos de industrias o entidades podría propiciar un escenario donde la información de deuda juegue un rol relevante en modelos de *scoring*, más aún en la evaluación de entrada.

**Tabla A.4:** Proporción de clientes con productos en más de una entidad por industria

IFI	Proporción
CAC's	81,8%
SAG's	80,4%
Banco	73,6%
ETNB	68,3%

**Tabla A.5:** Distribución de los deudores según las industrias en que participa (en color se marca la(s) industria(s) que hace(n) referencia)

Banco	SAG's	CAC's	ETNB	Proporción	Proporción acumulada
				22,8%	22,8%

				14,9%	37,7%
				0,9%	38,6%
				12,9%	51,5%
				25,1%	76,7%
				0,5%	77,2%
				3,7%	80,9%
				0,4%	81,3%
				0,5%	81,8%
				6,6%	88,3%
				0,4%	88,7%
				1,0%	89,7%
				9,2%	98,9%
				0,3%	99,2%
				0,8%	100,0%

#### Anexo 4.1 – Medidas de desempeño (ejercicio admisión)

**Tabla A.6:** Indicador Kolmogorov Smirnov (KS<sup>15</sup>) del escenario de admisión, en la muestra de validación y uso de regresión logística

Muestra	Modelo	Bancos	SAG's	CAC's	ETNB	Sistema
Total	Modelo base	8,3%	15,7%	5,4%	14,8%	9,7%
	Sólo industria	32,0%	27,2%	21,2%	9,4%	31,1%
	Info. completa	34,0%	32,6%	36,3%	19,8%	33,1%
> 1 IFI	Modelo base	4,7%	5,3%	5,4%	6,4%	5,3%
	Sólo industria	32,0%	27,2%	21,2%	9,4%	31,1%
	Info. completa	34,0%	32,6%	36,3%	19,8%	33,1%
> 1 Industria	Modelo base	4,3%	4,8%	5,4%	6,4%	4,2%
	Sólo industria	31,6%	23,0%	21,4%	9,4%	28,8%
	Info. completa	34,3%	30,4%	36,5%	19,8%	32,1%

**Tabla A.7:** AUROC<sup>14</sup> del escenario de admisión, en la muestra de validación y uso de regresión logística

Muestra	Modelo	Bancos	SAG's	CAC's	ETNB	Sistema
Total	Modelo base	55,2%	59,9%	52,8%	56,9%	57,6%
	Sólo industria	71,8%	67,5%	62,2%	50,1%	70,7%
	Info. completa	71,6%	70,6%	71,6%	61,0%	72,0%
> 1 IFI	Modelo base	52,2%	53,8%	52,9%	49,2%	53,3%
	Sólo industria	71,8%	67,5%	62,2%	50,1%	70,7%
	Info. completa	71,6%	70,6%	71,6%	61,0%	72,0%
	Modelo base	52,1%	53,0%	53,0%	49,2%	52,7%

> 1 Industria	Sólo industria	72,1%	65,3%	62,3%	50,1%	69,5%
	Info. completa	71,1%	69,7%	71,7%	61,0%	71,5%

**Tabla A.8:** Indicador Kolmogorov Smirnov ( $KS^{15}$ ) del escenario de admisión, en la muestra de validación y uso de algoritmo GBM

Muestra	Modelo	Bancos	SAG's	CAC's	ETNB	Sistema
Total	Modelo base	8,3%	15,7%	5,4%	14,8%	9,7%
	Sólo industria	38,8%	28,2%	26,8%	20,5%	35,1%
	Info. completa	42,3%	38,9%	46,4%	22,3%	42,0%
> 1 IFI	Modelo base	4,7%	5,3%	5,4%	6,4%	5,3%
	Sólo industria	38,8%	28,2%	26,8%	20,5%	35,1%
	Info. completa	42,3%	38,9%	46,4%	22,3%	42,0%
> 1 Industria	Modelo base	4,3%	4,8%	5,4%	6,4%	4,2%
	Sólo industria	37,9%	24,2%	27,0%	20,5%	31,4%
	Info. completa	44,0%	37,4%	46,7%	22,3%	42,2%

**Tabla A.9:** AUROC<sup>14</sup> del escenario de admisión, en la muestra de validación y uso de algoritmo GBM

Muestra	Modelo	Bancos	SAG's	CAC's	ETNB	Sistema
Total	Modelo base	55,2%	59,9%	52,8%	56,9%	57,6%
	Sólo industria	77,0%	70,2%	67,2%	61,2%	75,2%
	Info. completa	78,4%	75,7%	80,0%	62,9%	78,2%
> 1 IFI	Modelo base	52,2%	53,8%	52,9%	49,2%	53,3%
	Sólo industria	77,0%	70,2%	67,2%	61,2%	75,2%
	Info. completa	78,4%	75,7%	80,0%	62,9%	78,2%
> 1 Industria	Modelo base	52,1%	53,0%	53,0%	49,2%	52,7%
	Sólo industria	77,1%	68,0%	67,3%	61,2%	73,2%
	Info. completa	79,6%	75,1%	80,2%	62,9%	78,3%

#### Anexo 4.2 – Medidas de desempeño (ejercicio comportamiento)

**Tabla A.10:** Medidas de desempeño del escenario de comportamiento, en la muestra de validación

Modelo	Regresión logística		GBM	
	AUROC <sup>14</sup>	KS <sup>15</sup>	AUROC <sup>14</sup>	KS <sup>15</sup>

<sup>14</sup> El AUROC ("Area Under the Receiver Operating Characteristic curve") es una métrica que permite medir el poder discriminatorio en un modelo y se puede interpretar como la probabilidad que el modelo clasifique correctamente una observación positiva aleatoriamente escogida por sobre una observación negativa aleatoriamente escogida.

<sup>15</sup> Corresponde al valor del estadístico asociado a la prueba de Kolmogorov Smirnov, la cual evalúa si el puntaje predicho por el modelo, tanto para deudores buenos y malos, es significativamente distinto. La hipótesis nula es que ambas distribuciones del puntaje son generadas por la misma función y, por lo tanto, no difieren. El estadístico se calcula como  $\max [F_1(x) - F_2(x)]$ , donde  $F_1(x)$  es la densidad acumulada del puntaje de buenos pagadores y  $F_2(x)$  la de malos pagadores. Entre más cercano a 100% el indicador, más señal de que el algoritmo dividió de forma más distanciada al score de ambas clases.

Toda la muestra	Base	79,3%	44,1%	81,9%	48,2%
	Info. Industria	80,9%	47,4%	83,1%	50,7%
	Info. Completa	81,6%	48,2%	83,9%	51,8%
> 1 IFI	Base	78,8%	43,2%	81,5%	47,5%
	Info. Industria	80,9%	47,3%	83,0%	50,4%
	Info. Completa	81,6%	48,2%	83,9%	51,7%
> 1 Industria	Base	79,0%	43,1%	81,6%	47,4%
	Info. Industria	80,7%	46,9%	82,7%	49,9%
	Info. Completa	81,8%	48,3%	83,9%	51,7%

### Anexo 5.1 – Resultados en mercado (ejercicio admisión)

Tabla A.11: Resultados en la perspectiva de mercado en el ejercicio de admisión

Tipo entidad	Nivel de tolerancia	Modelo	Rechazados	Comité	Aceptados
Sistema completo	0,0%	Base	0,0%	100,0%	0,0%
		Inf. industria	0,0%	100,0%	0,0%
		Inf. completa	0,0%	100,0%	0,0%
	0,5%	Base	0,5%	92,3%	7,2%
		Inf. industria	1,5%	77,9%	20,6%
		Inf. completa	1,6%	74,2%	24,2%
	1,0%	Base	1,1%	85,4%	13,6%
		Inf. industria	2,3%	61,1%	36,6%
		Inf. completa	2,4%	60,2%	37,4%
	2,0%	Base	2,2%	71,3%	26,5%
		Inf. industria	3,6%	43,8%	52,6%
		Inf. completa	3,7%	35,7%	60,5%
	3,0%	Base	3,0%	59,8%	37,2%
		Inf. industria	5,0%	28,1%	67,0%
		Inf. completa	5,1%	22,2%	72,7%
	4,0%	Base	4,2%	47,0%	48,9%
		Inf. industria	6,4%	22,2%	71,4%
		Inf. completa	6,4%	10,4%	83,1%
	5,0%	Base	4,2%	35,9%	59,9%
		Inf. industria	7,3%	5,3%	87,4%
		Inf. completa	7,8%	2,8%	89,4%
	6,0%	Base	6,3%	21,9%	71,8%
		Inf. industria	6,4%	0,0%	93,6%
		Inf. completa	5,7%	0,0%	94,3%

		Base	6,3%	8,0%	85,7%
	7,0%	Inf. industria	2,3%	0,0%	97,7%
		Inf. completa	2,3%	0,0%	97,7%
<b>Bancos</b>	0,0%	Base	0,0%	100,0%	0,0%
		Inf. industria	0,0%	100,0%	0,0%
		Inf. completa	0,2%	99,8%	0,0%
	0,5%	Base	0,5%	91,7%	7,8%
		Inf. industria	1,5%	69,2%	29,3%
		Inf. completa	1,4%	73,9%	24,7%
	1,0%	Base	0,9%	82,6%	16,4%
		Inf. industria	2,2%	53,5%	44,2%
		Inf. completa	2,3%	49,6%	48,1%
	2,0%	Base	1,9%	70,5%	27,6%
		Inf. industria	3,5%	31,3%	65,3%
		Inf. completa	3,5%	29,8%	66,6%
	3,0%	Base	3,1%	53,2%	43,7%
		Inf. industria	4,5%	16,0%	79,5%
		Inf. completa	4,6%	13,9%	81,5%
	4,0%	Base	3,1%	38,7%	58,2%
		Inf. industria	6,0%	3,6%	90,5%
		Inf. completa	6,1%	3,0%	90,9%
	5,0%	Base	5,0%	21,6%	73,4%
		Inf. industria	3,8%	0,0%	96,2%
		Inf. completa	3,5%	0,0%	96,5%
	6,0%	Base	5,0%	6,1%	88,8%
		Inf. industria	0,7%	0,0%	99,3%
		Inf. completa	0,7%	0,0%	99,3%
7,0%	Base	0,0%	0,0%	100,0%	
	Inf. industria	0,0%	0,0%	100,0%	
	Inf. completa	0,0%	0,0%	100,0%	
<b>SAG's</b>	0,0%	Base	0,0%	100,0%	0,0%
		Inf. industria	0,0%	100,0%	0,0%
		Inf. completa	0,0%	100,0%	0,0%
	0,5%	Base	0,5%	93,7%	5,7%
		Inf. industria	1,6%	83,2%	15,1%
		Inf. completa	1,5%	84,5%	14,0%
	1,0%	Base	1,2%	86,2%	12,6%
Inf. industria		2,4%	72,6%	25,0%	
Inf. completa		2,4%	72,0%	25,6%	
2,0%	Base	2,1%	75,7%	22,2%	

		Inf. industria	3,6%	62,2%	34,2%	
		Inf. completa	4,0%	50,1%	46,0%	
	3,0%	Base	3,5%	64,4%	32,1%	
		Inf. industria	5,0%	55,1%	39,9%	
		Inf. completa	5,2%	33,8%	61,0%	
	4,0%	Base	4,6%	52,1%	43,3%	
		Inf. industria	6,5%	35,5%	58,0%	
		Inf. completa	6,7%	23,2%	70,1%	
	5,0%	Base	5,8%	43,2%	51,0%	
		Inf. industria	7,8%	34,2%	58,0%	
		Inf. completa	8,2%	14,7%	77,2%	
	6,0%	Base	6,3%	33,1%	60,6%	
		Inf. industria	8,9%	12,0%	79,0%	
		Inf. completa	9,4%	6,7%	83,9%	
	7,0%	Base	8,1%	23,3%	68,6%	
		Inf. industria	10,3%	3,4%	86,3%	
		Inf. completa	10,5%	1,0%	88,4%	
	CAC's	0,0%	Base	0,0%	100,0%	0,0%
			Inf. industria	0,2%	99,8%	0,0%
			Inf. completa	0,1%	99,2%	0,7%
		0,5%	Base	0,5%	95,9%	3,6%
			Inf. industria	1,7%	85,7%	12,6%
			Inf. completa	2,2%	79,8%	18,0%
		1,0%	Base	1,0%	92,0%	7,1%
Inf. industria			2,5%	77,3%	20,2%	
Inf. completa			3,5%	63,1%	33,4%	
2,0%		Base	1,7%	81,5%	16,8%	
		Inf. industria	4,5%	68,3%	27,2%	
		Inf. completa	5,1%	44,2%	50,6%	
3,0%		Base	2,9%	70,2%	26,9%	
		Inf. industria	6,1%	55,4%	38,5%	
		Inf. completa	6,7%	29,3%	64,0%	
4,0%		Base	3,4%	60,8%	35,7%	
		Inf. industria	7,4%	49,5%	43,1%	
		Inf. completa	8,0%	18,0%	73,9%	
5,0%		Base	3,4%	52,2%	44,3%	
		Inf. industria	8,4%	36,3%	55,3%	
		Inf. completa	9,3%	11,7%	79,0%	
6,0%		Base	5,8%	42,6%	51,6%	
		Inf. industria	9,6%	26,1%	64,3%	

		Inf. completa	10,8%	4,0%	85,2%
	7,0%	Base	7,4%	33,5%	59,1%
		Inf. industria	10,7%	7,8%	81,5%
		Inf. completa	10,8%	0,0%	89,2%
ETNB's	0,0%	Base	0,0%	98,5%	1,5%
		Inf. industria	0,2%	99,2%	0,6%
		Inf. completa	0,1%	99,9%	0,0%
	0,5%	Base	0,2%	93,1%	6,7%
		Inf. industria	0,7%	89,1%	10,3%
		Inf. completa	0,7%	84,5%	14,8%
	1,0%	Base	0,8%	84,7%	14,4%
		Inf. industria	0,7%	89,1%	10,3%
		Inf. completa	1,5%	75,9%	22,6%
	2,0%	Base	1,9%	76,7%	21,4%
		Inf. industria	2,3%	69,3%	28,4%
		Inf. completa	2,8%	64,9%	32,3%
	3,0%	Base	3,3%	61,5%	35,2%
		Inf. industria	3,2%	55,5%	41,3%
		Inf. completa	3,8%	44,0%	52,2%
	4,0%	Base	4,1%	50,6%	45,3%
		Inf. industria	3,2%	55,5%	41,3%
		Inf. completa	4,7%	31,5%	63,8%
	5,0%	Base	5,0%	38,5%	56,5%
		Inf. industria	3,2%	55,5%	41,3%
		Inf. completa	5,8%	23,9%	70,4%
	6,0%	Base	6,4%	20,5%	73,1%
		Inf. industria	6,4%	20,1%	73,4%
		Inf. completa	7,0%	11,5%	81,5%
	7,0%	Base	7,5%	13,3%	79,2%
		Inf. industria	6,4%	20,1%	73,4%
		Inf. completa	7,0%	0,0%	93,0%

### Anexo 5.2 – Resultados en mercado (ejercicio comportamiento)

El ejercicio se puede realizar considerando la información obtenida en el modelo de comportamiento. Particularmente, el acceso de información de deuda facilitaría el acceso a nuevos productos a los actuales clientes de las IFIs.

En este caso, si bien la mejora en discriminación es marginal, hasta un 7% de los deudores de consumo (392 mil personas) podrían lograr un acceso más fluido a nuevos productos en su institución. El efecto de la información es cóncavo al nivel de tolerancia, dado que, en bajos niveles,

no existen clientes que sean aceptados (todos en Comité); mientras que, en niveles más elevados, ya no existe la instancia intermedia.

**Tabla A.12:** Resultados en la perspectiva de mercado en el ejercicio comportamiento

Nivel de tolerancia	Modelo	Rechazados	Comité	Aceptados
0,0%	Base	0,0%	100,0%	0,0%
	Inf. industria	0,0%	100,0%	0,0%
	Inf. completa	0,0%	100,0%	0,0%
0,5%	Base	2,3%	73,8%	23,9%
	Inf. industria	2,3%	67,4%	30,3%
	Inf. completa	2,3%	71,7%	26,0%
1,0%	Base	3,4%	59,0%	37,6%
	Inf. industria	3,4%	52,8%	43,7%
	Inf. completa	3,5%	51,7%	44,9%
2,0%	Base	5,1%	34,9%	60,0%
	Inf. industria	5,2%	30,6%	64,2%
	Inf. completa	5,2%	28,5%	66,3%
3,0%	Base	6,6%	19,3%	74,1%
	Inf. industria	6,6%	16,8%	76,6%
	Inf. completa	6,7%	16,0%	77,3%
4,0%	Base	8,0%	8,7%	83,3%
	Inf. industria	7,9%	7,6%	84,5%
	Inf. completa	8,0%	7,2%	84,8%
5,0%	Base	9,3%	1,3%	89,4%
	Inf. industria	9,3%	1,1%	89,6%
	Inf. completa	9,3%	0,8%	89,9%

## Anexo 6 – Anonización de datos de deudores

La anonimización o disociación de datos de la identidad del deudor es el proceso de transformación de la información mediante la eliminación o el cifrado de esta, a fin de proteger la privacidad y confidencialidad del titular de estos datos. Estas técnicas permiten que los datos sean retenidos y utilizados por quienes manejan las bases de datos, pero rompiendo el vínculo entre un individuo y los datos almacenados. Debido a lo anterior, es la forma más estándar para almacenar o divulgar información personal respetando las leyes de protección de datos personales.

El proceso de anonimización requiere una adecuada comprensión del propósito final de la utilización de la información, así como de su nivel de utilidad, teniendo en cuenta que independientemente de las técnicas empleadas, una vez realizado el proceso de anonimización se reduce la información original del conjunto de datos. Sin embargo, en la práctica, la anonimización no lograría obstruir completamente las identidades de los interesados debido al potencial de desanonización o reidentificación (Ohm, 2009). Las legislaciones actuales no pueden simplemente ignorar el riesgo que conlleva la reidentificación. Durante las últimas décadas las leyes de protección de datos se han



basado en el supuesto de que la anonimización es la solución ideal, por lo que el desafío actual para los reguladores es responder rápida y contundentemente a los cambios tecnológico-disruptivos, como lo es la posibilidad de revertir datos innominados, y así restablecer el equilibrio de la ley, protegiendo a los titulares de información sensible de daños inminentes y significativos (Brasher, 2018; Ohm, 2009).

En línea con lo anterior, a nivel mundial existe una variedad de regulaciones que estipulan cómo se debe manejar la información confidencial del cliente financiero. Gran parte de estas políticas se relacionan con la utilización de métodos adecuados para proteger la información mediante la disociación de la información sensible. Particularmente, podemos destacar los siguientes casos:

- 1) *Estados Unidos*: las regulaciones que van desde *Gramm-Leach-Bliley Act* (GLBA) hasta *EU's General Data Protection Regulation* (GDPR) obligan a los bancos a auditar sus datos trimestralmente y eliminar los datos que se mantienen sin el consentimiento de sus clientes. Sin embargo, los bancos no necesitan borrar toda esa información. Ellos pueden anonimizar los datos de forma de poder seguir utilizándolos con otros fines, abordando al mismo tiempo el problema de la privacidad.
- 2) *La Unión Europea*: la anonimización de datos es una técnica fundamental para cumplir con el Reglamento general de protección de datos, regulación que entró en vigor en 2018 e impone fuertes multas a quienes violen sus estándares de privacidad y seguridad. Particularmente, considerando que la anonimización de datos personales se ha vuelto realidad, sus normativas consideran garantías adicionales para proteger la privacidad del consumidor. En este sentido, para el cumplimiento de las leyes de protección de datos personales, la normativa se enfoca en la gestión de riesgos en lugar de una anonimización perfecta, y en políticas internas sólidas de divulgación y tratamiento de datos por parte de las entidades.

En Chile, en julio de 2022 la Comisión para el Mercado Financiero (CMF) publicó una norma que exige a las instituciones financieras una Política Interna de Seguridad y Manejo de la Información de Deudores. Esta norma autoriza a las instituciones a realizar procedimientos de disociación de datos para que puedan utilizar información no vigente y restringida por ley para fines estadísticos o para el desarrollo de modelos de riesgo.

Respecto a las técnicas de anonimización, estas dependen tanto de los atributos y características de la información, como del tratamiento que se le pretende dar a los datos anonimizados. La elección de las técnicas de anonimización seleccionadas deben ser estudiadas y realizadas por un equipo de trabajo, previa evaluación y aprobación del equipo evaluador de riesgos, así como tener pruebas permanentes que permitan verificar su efectividad. En particular, las técnicas de anonimización de datos se pueden agrupar según diferentes características del proceso (por ejemplo: técnicas con o sin perturbación) o por enfoques (aleatorización, generalización, seudoanonimización).

Las 6 técnicas comunes más utilizadas a nivel internacional en la anonimización de datos personales corresponden a:

- 1) *Generalización*: Consolida y agrega la información, creando una categorización amplia para brindar una vista más generalizada del contenido de los datos.
- 2) *Seudonimización*: reemplaza información directamente identificable con un identificador artificial, denominado "seudónimo".

- 3) *Datos sintéticos*: generación de datos que se parecen mucho a los datos confidenciales. A menudo se usa para entornos de prueba y para validar o entrenar modelos para matemáticas o aprendizaje automático. Puede utilizar desviaciones estándar, medianas, regresión lineal u otras técnicas estadísticas para generar los datos sintéticos.
- 4) *Intercambio/permutación o mezcla de datos*: reposiciona datos de un conjunto para que los valores no coincidan con los datos originales.
- 5) *Privacidad diferencial o perturbación*: método que ocasiona que los datos sean más ambiguos al aleatorizar elementos, modificándolos ligeramente mediante la aplicación de técnicas que redondean números y/o agregar ruido aleatorio. Cuando esta técnica se realiza de manera deliberada y controlada, sirve para que los registros individuales sean menos sensibles mientras tienen efectos predecibles y que se pueden revertir para así conocer el dato original, si fuera necesario. Sin embargo, después de agregar ruido, el resultado del análisis se convierte en una aproximación, no en el resultado exacto que se habría obtenido solo si se hubiera realizado sobre el conjunto de datos reales.
- 6) *Reducción o anulación de data*: Los datos confidenciales simplemente se eliminan, convirtiendo esta información en valores nulos.



Regulador y Supervisor Financiero de Chile

[www.cmfchile.cl](http://www.cmfchile.cl)

