



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

EVALUACIÓN DEL RIESGO REPUTACIONAL EN EL NEGOCIO BANCARIO
MEMORIA PARA OPTAR AL TÍTULO DE INGENIERO CIVIL INDUSTRIAL

MICHAEL THOMAS CLEMANS MONJES

PROFESOR GUÍA:
CARLOS PULGAR ARATA

MIEMBROS DE LA COMISIÓN:
CARLOS REYES RUBIO
FELIPE VILDOSO CASTILLO

SANTIAGO DE CHILE

2023

**RESUMEN DE LA MEMORIA
PARA OPTAR AL TÍTULO DE:**

Ingeniero Civil Industrial

POR: Michael Thomas Clemans Monjes

PROF. GUÍA: Carlos Pulgar A.

FECHA: 2023

EVALUACION DEL RIESGO REPUTACIONAL EN EL NEGOCIO BANCARIO

En el mundo del negocio bancario el riesgo reputacional se refiere a cualquier tipo de amenaza o peligro que pueda afectar negativamente la reputación de la empresa. El riesgo reputacional es particularmente difícil de estudiar, puesto que su estudio usualmente involucra data altamente subjetiva y difícil de cuantificar como son las opiniones.

La Comisión del Mercado Financiero (CMF) está en proceso de implementar los estándares de Basilea III en Chile, lo que incluye la posible introducción de requisitos de capital adicionales bajo el Pilar II. En este marco, se requiere que los bancos regulados desarrollen metodologías para evaluar riesgos materiales, incluyendo el riesgo reputacional. Sin embargo, hasta la fecha, no existe una metodología pública en Chile para calcular el riesgo reputacional en el ámbito bancario.

La presente memoria, en este contexto, tuvo por objetivo evaluar riesgo reputacional en el negocio bancario en el contexto nacional a través del desarrollo e implementación de dos metodologías alternativas que buscaron medir riesgo reputacional para instituciones bancarias en el contexto local y estimar requerimiento de capital directo. Estas metodologías nombradas como Método T y Método R se enfocaron en analizar el riesgo reputacional a través de fuentes como Twitter y la página web Reclamos.cl, respectivamente. Estos métodos establecieron una conexión entre el riesgo reputacional y el sector bancario al elaborar índices de reputación que se contrastaron con las variables financieras de Ganancia, ROA y ROE.

Los resultados obtenidos de la investigación sugieren la viabilidad de calcular el riesgo reputacional en el sector bancario chileno mediante el Método T, aunque no mediante el Método R. El Método T logró evidenciar correlación para entre los índices diseñados y las variables de ROA y ROE, pudiendo calcularse exitosamente requisitos de capital. Estos valores fueron coherentes con los capitales regulatorios T1 y T2 exigidos por la LGB (Ley General de Bancos), estimando requerimiento de capital desde 0,11% APR hasta 1,68% APR a lo largo de 8 bancos utilizados como muestra.

Para trabajos futuros se plantea volver a efectuar el procedimiento del Método T pero mejorándose la calidad de la data. Se propone aumentar el número de variables financieras asociadas a desempeño bancario utilizadas en el Método T y desarrollar nuevos Métodos alternativos que utilicen fuentes de información alternativa (por ejemplo, noticias o redes sociales diferentes a Twitter).

Agradecimientos

Agradezco a todos los que me apoyaron y ayudaron con la realización de esta memoria, desde los primeros días de universidad, hasta este último período universitario.

A mis tres profesores que me acompañaron a lo largo del curso asociado de Trabajo de Título; Carlos Pulgar, Carlos Reyes y Hugo Sánchez, quienes siempre con muy buena disposición me brindaron consejos y guiaron en todo momento, siendo a veces duros con las evaluaciones, pero siempre velando por que desarrollase el mejor trabajo posible.

A Diego Beas por haber sido quien en primer lugar conceptualizo el trabajo de tesis, y quien me brindó numeroso material de apoyo para efectuarla. Quién, sin ninguna obligación estricta, me dio su tiempo en forma de reuniones regulares para resolver dudas y sugerir nuevas ideas.

A Jonás Peñailillo, Alfonso Ordóñez y Sebastián Urbina por la retroalimentación e inspiración para continuar trabajando. A Franco Martini y Carlos Lazo por su ayuda con los aspectos financieros de la tesis.

A mis padres y hermanos que siempre estuvieron presentes en este período universitario, apoyándome de forma incondicional.

Tabla de Contenido

1	Introducción	1
1.1	Motivación	1
1.2	Objetivo General	3
1.3	Objetivos Específicos	3
1.4	Metodología	4
1.5	Limitaciones	8
2	Marco teórico	10
2.1	Riesgo reputacional.....	10
2.1.1	Fuentes de origen de riesgo reputacional	11
2.1.2	Efectos del riesgo reputacional.....	11
2.2	Valoración de empresa	13
2.2.1	Variables financieras.....	13
2.2.2	Características.....	15
2.3	Índice	16
2.3.1	Índice, riesgo reputacional y valoración de empresa.....	16
2.3.2	Métodos	16
2.3.3	Diseño.....	18
2.4	Método R	19
2.4.1	Reclamos.cl	19
2.4.2	Scraping	19
2.4.3	Selenium	19
2.4.4	Chromium.....	20
2.4.5	Índice por reclamo	20
2.5	Método T.....	22
2.5.1	Twitter.....	22
2.5.2	Snsrape.....	24
2.5.3	NLP y Análisis de Sentimiento.....	24
2.5.4	Transformadores	26
2.5.5	Pysentimiento	27
2.5.1	Índice por Twitter	28
2.6	Correlación.....	30
2.6.1	Regresiones lineales	30

2.6.2	Modelo de datos de panel	30
2.7	Requisito de Capital	31
2.7.1	Capital regulatorio	31
2.7.2	Activos Ponderados por Riesgo.....	32
2.7.3	Evaluación de capital.....	32
2.7.4	Calculo.....	33
3	Implementación y análisis de resultados	35
3.1.1	Extracción de data	35
3.1.2	Procesamiento de data	38
3.1.3	Análisis Exploratorio.....	41
3.1.4	Minería de datos	43
3.1.5	Regresiones lineales	44
3.1.6	Regresiones lineales por datos de panel	48
3.1.7	Requisito de capital sugerido.....	51
4	Conclusiones	57
5	Glosario	60
6	Bibliografía.....	61
	Anexos	63
	Anexo A: Índice Base en el tiempo	63
	Anexo B: Resultados por datos de panel	69
	Anexo C: APR, Patrimonio y Activos	73

Índice de Tablas

Tabla 1: Tests Macro F1 para RoBERTuito. (Pérez, 2022)	28
Tabla 2: Ejemplo de base de datos. (Método R).....	36
Tabla 3: Ejemplo de base de datos. (Método T)	37
Tabla 4: Ejemplo de Índice Base. (Método R)	39
Tabla 5: Ejemplo de Índice Base. (Método T).....	40
Tabla 6: Estadísticas Índice Base. (Método R).....	41
Tabla 7: Estadísticas Índice Base. (Método T)	42
Tabla 8: Correlaciones individuales según Ganancia. (Método R)	44
Tabla 9: Correlaciones individuales según ROA. (Método R).....	45
Tabla 10: Correlaciones individuales según ROE. (Método R)	45
Tabla 11: Correlaciones individuales según Ganancia. (Método T).....	46
Tabla 12: Correlaciones individuales según ROA. (Método T)	47
Tabla 13: Correlaciones individuales según ROE. (Método T).....	47
Tabla 14: Correlaciones por panel según ROE. (Método R).....	48
Tabla 15: Correlaciones por panel según ROA. (Método T).....	49
Tabla 16: Correlaciones por panel según ROE. (Método T)	50
Tabla 17: β ROE por Índice. (Método R)	51
Tabla 18: Parámetros por Banco. (Método R).....	52
Tabla 19: Capital Requisito ROE. (Método R).....	52
Tabla 20: β ROE y β ROA por Índice. (Método T)	54
Tabla 21: Parámetros por Banco. (Método T)	54
Tabla 22: Capital Requisito ROE. (Método T).....	55
Tabla 23: Capital Requisito ROA (Método T).....	56

Índice de Ilustraciones

Ilustración 1: Resumen de la audiencia disponibles para anuncios. (Branch, 2022).....	22
Ilustración 2: Arquitectura del Transformador. (Ashish Vaswani, 2017)	27
Ilustración 3: Evaluación de solvencia en la nueva LGB. (Carlos Pulgar, 2020)	33

Índice de Ecuaciones

Ecuación 1: Índice Método R. (Elaboración propia)	21
Ecuación 2: Índice Método T. (Elaboración propia)	29
Ecuación 3: Índice Método R. (Elaboración propia)	29
Ecuación 4: Regresión lineal	30
Ecuación 5: Requisito de capital	33
Ecuación 6: Requisito de capital (Efecto β)	34
Ecuación 7: Requisito de capital (Efecto β y δ)	34
Ecuación 8: Requisito de capital (ROA)	34
Ecuación 9: Requisito de capital (ROE).....	34

1 Introducción

1.1 Motivación

Los Acuerdos de Basilea son una serie de normativas y estándares internacionales diseñados para fortalecer y mejorar la regulación y supervisión del sistema financiero a nivel global. Estos acuerdos son elaborados por el Comité de Supervisión Bancaria de Basilea (también conocido como el Comité de Basilea), un organismo compuesto por representantes de bancos centrales y autoridades de supervisión bancaria de diversos países. Los Acuerdos de Basilea tienen como objetivo principal garantizar la estabilidad del sistema financiero y proteger a los bancos y a los depositantes frente a los riesgos asociados con las actividades bancarias. Los acuerdos también buscan promover la transparencia, la integridad y la eficiencia del sector bancario a nivel mundial.

A la fecha ha habido 3 versiones del acuerdo de Basilea, habiéndose publicado la versión final de Basilea III a finales de 2017. Basilea III fue resultado de una revisión al marco de adecuación de capital, en contexto de prevenir nuevamente situaciones como crisis subprime de 2008, que puso en relieve los riesgos que afectan a la industria bancaria y el alto costo asociado para la economía en caso de incorrecta su gestión y mitigación. (Gabriela Aguilera, 2020)

Chile se encuentra actualmente en proceso de implementar Basilea III, habiendo iniciado la Comisión para el Mercado Financiero (CMF) el proceso normativo asociado el enero de 2019 con la publicación Ley N° 21.130, la cual tiene por objetivo modernizar la Ley General de Bancos (LGB). Se espera que esta modernización alinee la LGB con los estándares y mejores prácticas internacionales, fortaleciendo la solvencia de la banca, mejorando la competitividad de los bancos chilenos, tanto a nivel local como internacional, se facilite acceso a nuevas fuentes de financiamiento, más diversificadas, de menor costo o mayor plazo, construyendo así a una mayor estabilidad del sistema. También se espera armonizar requerimientos entre filiales de bancos extranjeros y bancos locales para eliminar asimetrías regulatorias en la jurisdicción chilena. (Gabriela Aguilera, 2020)

En contexto de que la Ley N° 21.130 ha generado un marco legal más flexible para la rápida adopción por la vía normativa de los estándares de Basilea, se facultó a la CMF previo acuerdo favorable del Consejo del Banco Central de Chile (BCCCh), para definir las nuevas metodologías de cálculo de los activos ponderados por riesgo de crédito, mercado y operacional; las condiciones de emisión de instrumentos híbridos AT1, y la determinación y cargos de capital para bancos de importancia sistémica local. Asimismo, introdujo los colchones de conservación y contra cíclicos y amplió las facultades de la CMF para hacer descuentos prudenciales al capital regulatorio y exigir medidas adicionales, entre ellas mayor capital, a bancos que presentan deficiencias en el proceso de evaluación supervisora. La implementación de Basilea III permite focalizar la gestión de riesgos hacia una visión

más comprensiva de ellos, con foco en la suficiencia de capital (Carlos Pulgar, 2020) (Gabriela Aguilera, 2020) (CMF, 2020).

Desde la promulgación de la Ley N° 21.130, la CMF inició el proceso normativo para la implementación los 3 pilares de Basilea III, los cuales se refieren:

Pilar 1: Requisitos mínimos de capital que los bancos deben mantener para cubrir los riesgos crediticios, de mercado y operativos. Establece normativas para calcular las ratios de capital, como el Ratio de Capital Principal y el Ratio de Capital Total, con el fin de asegurar la solidez financiera de las instituciones bancarias.

Pilar 2: Evaluación de la capacidad de los bancos para gestionar riesgos y mantener un colchón de capital adecuado en relación con sus perfiles de riesgo específicos. Implica pruebas de estrés, evaluaciones internas y supervisión de los reguladores para garantizar que los bancos puedan absorber pérdidas inesperadas y mantener operaciones estables.

Pilar 3: Mejora de la transparencia y divulgación de información por parte de los bancos. Esto implica proporcionar detalles sobre la estructura de capital, los riesgos asumidos, los activos ponderados por riesgo y otros aspectos relevantes. Esta divulgación permite a inversores, depositantes y reguladores tomar decisiones informadas y supervisar la salud financiera de las instituciones bancarias.

Por parte de las exigencias del llamado Pilar 2, la CMF incorporó a la RAN (Recopilación Actualizada de Normas para bancos) el Capítulo 21-13 que estipula que los bancos regulados deberán contar con metodologías para medir todos sus riesgos materiales, las cuales debe documentar en el informe de autoevaluación de patrimonio efectivo (IAPE). Además, la CMF podrá exigirles requerimiento de patrimonio efectivo respecto de su índice de solvencia (o índice de adecuación de capital) si estima que existen riesgos materiales no suficientemente cubiertos.

Dentro de los posibles riesgos materiales se encuentra el llamado riesgo reputacional. Este a la fecha es un riesgo que se considera difícil de medir, del cual la literatura es incipiente a nivel mundial y del cual no existe una metodología pública en Chile para medirlo.

El presente trabajo tendrá el propósito de brindar valor al desarrollar una metodología replicable que permita medir riesgo reputacional para instituciones bancarias en Chile, para posteriormente calcular los requerimientos de capital adicionales asociados. Estos requerimientos de capitales adicionales servirán para fortalecer la solvencia de la banca nacional, según lo ya mencionado en el Capítulo 21-13 de la RAN, particularmente ayudando con el proceso de autoevaluación de capital o que los propios bancos puedan determinar patrimonio efectivo para cubrir sus riesgos en un horizonte de al menos tres años y con la evaluación de la CMF sobre la suficiencia de patrimonio efectivos de los bancos como respaldo de perfil de riesgo. A un nivel más macro, todo lo anterior ayudará con la finalidad última de la implementación de Basilea III en Chile, y potencialmente el desarrollo de una metodología nueva que sirva para el estudio de riesgo reputacional en el negocio bancario chileno.

1.2 Objetivo General

Desarrollar una metodología replicable que permita medir requerimiento de capital por una materialización del riesgo reputacional para bancos chilenos.

1.3 Objetivos Específicos

Los objetivos específicos son los siguientes:

1. Crear e implementar índices que determinen y representen bien el sentimiento promedio de los clientes hacia la muestra de bancos chilenos a lo largo de un horizonte de tiempo bien definido, sea por frecuencia de reclamos o por polaridad de tweets.
2. Estudiar el riesgo reputacional de la muestra de los bancos chilenos, explorando potencial correlación entre los índices y una selección de variables financieras asociadas al desempeño bancario.
3. Cálculo del capital requisito sugerido para la muestra de bancos chilenos a partir de las correlaciones obtenidas.
4. Análisis de resultados. Determinar si los capitales requisitos calculados son coherentes y útiles para la evaluación de capital a partir de riesgo reputacional en el Capítulo 21-13 de la RAN. Explicar por qué los resultados obtenidos fueron de tal forma, si hubo fallas en la metodología y aristas potenciales para explorar en trabajos futuros.

1.4 Metodología

El estudio de riesgo reputacional se efectuó con información obtenida de dos fuentes distintas, Reclamos.cl y Twitter. Dado que estas dos fuentes presentan naturalezas distintas, se divide la metodología en dos enfoques paralelos, los cuales son denominados "Método R" para la información procesada de Reclamos.cl y "Método T" para Twitter.

El Método R empieza por guardar la información asociada a los reclamos en bases de datos de formato .csv para cada uno de los bancos a lo largo de 8 años de tiempo, con punto de partida 1 de abril del año 2015 y punto de fin 31 de marzo del año 2023. En total se crean 8 bases de datos con las columnas Fecha y Contenido, representando la primera la fecha y la segunda el contenido del reclamo mismo. Se prescinde el título del reclamo, el nombre de usuario y demás atributos de Reclamos.cl, como solo se busca contar frecuencia.

Las 8 bases de datos luego son procesadas con un programa creado en Python para obtener los índices asociados por mes. El programa empieza por separar la información de la columna Fecha, creando las columnas Año y Mes. El programa ordena la base de datos según mes, teniendo por resultado 96 filas con el conteo mensual de los reclamos a lo largo del período de 8 años. El programa luego importa otras 8 bases de datos paralelas que contienen el número de clientes asociado por cada banco según mes y las fusiona con la base de datos anterior, para posteriormente calcular el índice como una columna de nombre Índice que contiene la división del número de reclamos en el número de clientes. La columna Índice para facilitar interpretación se multiplica por 1.000.000, equivalente a dejar el coeficiente por cada millón de personas. Se exportan las 8 bases de datos resultantes en formato .csv para efectuar los análisis

El Método T comienza guardando tweets con sus principales atributos en bases de datos tipo .csv por cada banco, con el mismo intervalo de tiempo del método R. En total se crean 8 bases de datos con las columnas Date, User, Tweet, replyCount, likeCount y retweetCount. La columna Date representa fecha, User el usuario, Tweet el contenido del tweet y replyCount, likeCount, retweetsCount el número de respuestas, me gustas y retweets respectivamente.

Al igual que en el método R, las 8 bases de datos se procesan con un programa creado en Python que calcula los índices asociados por mes. Este difiere en que calcula el índice a partir de polaridad o sentimiento mes a mes, basado en el promedio mensual resultante de sumar el sentimiento obtenido de cada tweet efectuado dentro del mes, a partir del modelo de Procesamiento de Lenguaje Natural de nombre RoBERTuito. En concreto el programa añade la columna de nombre Sentimiento a la base de datos, que contiene el valor de polaridad determinado para cada uno de los tweets extraídos (positivo, negativo o neutro). El programa organiza la columna de fechas (Date) basándose en el valor promedio de polaridad. Este valor promedio se calcula teniendo en cuenta la proporción de tweets según sus polaridades, y se pondera adicionalmente por la importancia de cada tweet. La importancia se determina a través de un factor que considera la cantidad de respuestas, "me gusta" y retweets asociados a cada tweet. Se exportan las 8 bases de datos resultantes en formato .csv para efectuar los análisis.

Tanto las bases de datos con los valores de índice calculado para el método R como T se importan a Rstudio para continuar el procesamiento. Se importan esta vez por separado las 8 bases de datos respectivas, trabajándose en 16 archivos .rmd distintos. En cada archivo se importan en paralelo las bases de datos correspondientes a los valores mensuales de ROE, ROA y Ganancia (en caso de que aplique) y se fusionan con la base de datos de índice. Dentro de Rstudio se formatean y limpian correctamente los datos según necesidad, modificándose el formato de fecha para ser trabajable en el lenguaje R, y parcheando con filas de índice = 0 para aquellos meses sin reclamos o valor de polaridad. A cada una de las bases de datos combinadas se le agregan 4 columnas nuevas con variaciones del índice; Indice_Rezagado_3,Indice_Rezagado_6,Indice_Acumulado_3, Indice_Acumulado_6. Estas representan el valor del índice considerando un rezago del sentimiento del público en 3 y 6 meses y el índice como el sentimiento promedio de los últimos 3 y 6 meses. A cada una de las bases de datos combinadas también se le agrega una columna de nombre Delta, que corresponde al parámetro δ de estrés del índice original para el cálculo de pérdida. Este se calcula como la diferencia entre el máximo y el mínimo valor del índice original a lo largo de un período móvil de 1 año. Finalmente se agregan a las bases de datos combinadas una columna de nombre Banco, con el nombre del banco correspondiente de cada una, para permitir regresiones por panel.

La base de datos combinada resultante de todo el proceso se exporta en formato .csv. Estas quedan con las columnas Año, Mes, Ganancia, Indice, ROA, ROE, Banco, Indice_Rezagado_3,Indice_Rezagado_6,Indice_Acumulado_3,Indice_Acumulado_6, Delta.

En los mismos archivos .Rmd, se aplican regresiones lineales para las bases de datos combinadas, alternándose entre Ganancia (de ser aplicable),ROA y ROE como variables dependientes e Indice, Indice_Rezagado_3,Indice_Rezagado_6,Indice_Acumulado_3,Indice_Acumulado_6 como variables independientes. De las correlaciones se consiguen los valores β y el nivel de significancia para cada uno de los bancos muestra de forma individual.

A forma de estudiar la evolución del índice base a lo largo del tiempo se importan los .csv resultantes del método R y T a la herramienta de visualización de data de nombre PowerBI. Se obtiene gráficos que señalan cuando el índice base alcanza su valor máximo y mínimo a lo largo del período de estudio 2015-2023.

Para medir correlación conjunta se utiliza la técnica de datos de panel. Esta técnica requiere tener todos los bancos en una sola base de datos, por lo que se importan las bases de datos combinadas nuevamente a archivos .rmd en Rstudio. Aquí en ambos métodos se fusionan las 8 bases de datos combinadas de cada banco en una super base de datos del sistema entero, ordenada según la columna de Banco. Esta super base de datos es exportada en .csv y luego se ingresa a un programa en Python. En Python se importa el método PanelOLS de la librería linearmodels.panels y se efectúan las regresiones por panel para las distintas combinaciones posibles de ROA y ROE como variables dependientes e Indice_Rezagado_3,Indice_Rezagado_6,Indice_Acumulado_3,Indice_Acumulado_6 como variables independientes. De las correlaciones por panel nuevamente se consiguen los valores β y el nivel de significancia, pero esta vez de forma grupal.

Con el β obtenidos de la regresión por panel y los valores más extremos de δ calculados con anterioridad, se calcula requerimiento de capital en los 2 métodos para los 8 bancos respecto a los distintos índices y las variables financieras. Esto se hace mediante la herramienta Excel, prefiriéndose por mayor simpleza el registrarse los resultados de forma manual en vez de automatizada. El requerimiento de capital por ROA se calcula multiplicando el β ROA por índice con el δ y los activos por banco dividido en los APR por banco. El requerimiento de capital por ROE se calcula con la misma lógica, pero cambiando el BROA por BROE y los Activos por Patrimonio. Los requerimientos de capital para dejarlos porcentualmente se multiplican por 100. Como resultado final se expresan los resultados en tablas con una columna para el tipo de índice y dos columnas de nombre Cap Req. ROE y Cap Req. ROA, con los resultados de los requerimientos de capital por ROE y ROA expresados en % de APR.

Dadas las necesidades, y en consecuencia expectativas, que se desprenden de la necesidad de la CMF por medir requerimientos de capital, se plantean los siguientes test:

Método R

Test R.GAN

H0 – Índice por Reclamos.cl no correlacionará negativamente con ganancia

H1 – Índice por Reclamos.cl correlacionará negativamente con ganancia

Test R.ROA

H0 – Índice por Reclamos.cl no correlacionará negativamente con la eficiencia (ROA)

H1 – índice por Reclamos.cl correlacionará negativamente con la eficiencia (ROA)

Test R.ROE

H0 – Índice por Reclamos.cl no correlacionará negativamente con la rentabilidad (ROE)

H1 – índice por Reclamos.cl correlacionará negativamente con la rentabilidad (ROE)

Método T

Test T.GAN

H0 – Índice por Twitter no correlacionará positivamente con ganancia

H1 – Índice por Twitter correlacionará positivamente con ganancia

Test T.ROA

H0 – Índice por Twitter no correlacionará positivamente con la eficiencia (ROA)

H1 – índice por Twitter correlacionará positivamente con la eficiencia (ROA)

Test T.ROE

H0 – Índice por Twitter no correlacionará positivamente con la rentabilidad (ROE)

H1 – índice por Twitter correlacionará positivamente con la rentabilidad (ROE)

La lógica subyacente siendo que un mayor nivel de reclamos (Método R) resulta en un peor desempeño del banco reflejado en las variables financieras y una mejor opinión del público (Método T) produce lo opuesto.

1.5 Limitaciones

Debido al vertiginoso avance en el desarrollo de modelos para el procesamiento de lenguaje natural, es importante destacar que parte de la metodología empleada en este estudio podría volverse obsoleta en un corto período de tiempo. Particularmente en el Método T, se emplea un modelo denominado RoBERTuito, el cual se fundamenta en RoBERTa. Este último, a su vez, encuentra sus bases en el modelo original BERT, lanzado por Google en 2018. RoBERTuito compite en el campo del procesamiento de lenguaje natural con otros modelos destacados, incluida la serie GPT de OpenAI. La versión más reciente de esta serie, GPT-4, fue lanzada el 14 de marzo de 2023 y ha demostrado logros sobresalientes en tareas de procesamiento de lenguaje. Sin embargo, hay una diferencia crucial entre GPT-4 y RoBERTuito: mientras que RoBERTuito es un modelo de código abierto, GPT-4 opera a través de una API que requiere solicitudes a los servidores de OpenAI, con un costo asociado por cada solicitud. Este enfoque puede resultar costoso al procesar grandes volúmenes de información, lo que el tesista estima requerir para el trabajo completo un gasto de más de 100 dólares a la fecha y varios días de procesamiento, siendo esta una de las razones por la que se prefirió RoBERTuito.

Recientemente, tras la adquisición de Twitter por parte del magnate Elon Musk, se han implementado modificaciones significativas en la plataforma que podrían hacer imposible replicar el Método T. A partir de febrero de 2023, Twitter discontinuó la opción de extraer tweets de manera gratuita a través de su API. Además, en marzo del mismo año, se incrementaron los precios de las membresías disponibles y se eliminaron funcionalidades clave, como la capacidad de extraer tweets específicos mediante nombres de cuentas y fechas precisas. Un cambio importante es que Twitter también restringió el acceso público a su endpoint, lo que ha inutilizado la aplicación de herramientas como SNScrape para la extracción de datos. Aunque no está claro si estos cambios serán permanentes, todavía existe una posibilidad de encontrar nuevas formas o metodologías alternativas para realizar el scraping de tweets, considerando la popularidad y relevancia de la plataforma.

Es importante reconocer que la investigación puede estar influenciada por sesgos inherentes al tipo de usuario y a los objetivos que persiguen al utilizar plataformas como Twitter o Reclamos.cl. Twitter, en particular, ha adquirido una reputación por albergar discursos negativos y por ser un espacio donde los usuarios a menudo expresan quejas o insatisfacciones en diversos aspectos de sus vidas. Esta dinámica puede ser exacerbada por el grado de anonimato que algunos usuarios aprovechan en la plataforma. En contraste, Reclamos.cl presenta una perspectiva más específica, ya que se centra en recopilar únicamente quejas y reclamos. Esta singularidad podría evitar algunos de los problemas asociados a emplear análisis de sentimientos, ya que en esta plataforma todos los contenidos son intrínsecamente negativos en naturaleza, aunque con variaciones en la intensidad, como han explorado trabajos previos (Yerko Moena, 2019).

El uso de técnicas de análisis de sentimientos podría presentar limitaciones considerables, dado que, en el caso de Twitter, existe la preocupación por la presencia de cuentas "trolls" que pueden distorsionar artificialmente la polaridad de los sentimientos expresados en un período determinado, ya sea positiva o negativamente. En contraposición, Reclamos.cl aborda en cierta medida esta problemática al requerir la validación de la identidad de

quienes presentan reclamos, lo que podría conferir mayor seriedad y legitimidad a los contenidos almacenados en la plataforma.

Dicho lo anterior, es crucial mencionar que, si bien Reclamos.cl proporciona una calidad de datos potencialmente más sólida debido a sus medidas de validación y autenticidad, su volumen de datos es notablemente menor en comparación con la vasta cantidad de información disponible en Twitter, que figura entre las redes sociales más prominentes a nivel global. Este contraste podría implicar que la cantidad de datos extraída de Reclamos.cl podría resultar insuficiente para un análisis exhaustivo, planteando la posibilidad de que los resultados estén afectados por el tamaño reducido de la muestra.

Con relación a la construcción del índice, es importante destacar que se exploró la posibilidad de crear diferentes versiones de este, utilizando como criterio la inclusión de tweets y reclamos correspondientes a empresas filiales de las 8 instituciones bancarias seleccionadas. Lamentablemente, una exploración inicial reveló que ninguna de estas filiales contaba con una muestra suficiente de comentarios para justificar la creación de versiones distintas del índice. De manera similar, se analizó la opción de agregar un factor a las polaridades, especialmente en el caso del análisis de sentimiento para los tweets, con el fin de otorgar importancia a los tweets realizados por personalidades financieras en el contexto chileno. Sin embargo, una vez más, se encontraron problemas relacionados con el tamaño de la muestra disponible.

2 Marco teórico

En el siguiente capítulo se presentarán una serie de conceptos financieros y de ciencias de datos asociados a la metodología por desarrollarse. El capítulo justificará la construcción de los índices que permitan calcular requisito de capital.

2.1 Riesgo reputacional

El riesgo reputacional es un concepto de gran importancia en el mundo financiero, pero también es difícil de medir debido a la falta de uniformidad y a la falta de investigación exhaustiva en torno a este tema (Ferreira, 2015). La Real Academia Española (RAE) define el riesgo reputacional como el "riesgo de que una entidad de crédito vea dañada su reputación en el mercado" (Rae, 2023). Según el Deutsche Bank, el riesgo reputacional se refiere al posible daño a la marca y reputación de una entidad, así como al riesgo asociado a las ganancias, el capital o la liquidez que surgen de asociaciones, acciones o inacciones que puedan ser percibidas por los accionistas como inapropiadas, no éticas o inconsistentes con los valores y creencias del banco (Deutsche Bank, 2017). Peter Mitic, por su parte, indica que la reputación es "La percepción de los stakeholders hacia una organización que puede afectar, ya sea de manera positiva o negativa, la relación comercial entre el stakeholder y la organización" (Mitic, 2018).

Cuantitativamente, Fontnouvelle y Perry indican que el riesgo reputacional puede ser evaluado mediante la reacción del precio de las acciones de una empresa ante el anuncio de una pérdida operativa significativa. Si el valor de mercado de la empresa disminuye más de lo esperado en relación con las pérdidas operativas anunciadas, se interpreta como una señal de pérdida de reputación. Los autores justifican esta definición al explicar que, experimentalmente, se ha observado una correlación de 1 a 1 entre la caída del valor de mercado y los anuncios de pérdidas causadas por factores externos, mientras que la correlación es de 2 a 1 en casos de pérdidas causadas por fraudes internos (Jason Perry, 2005).

Sin embargo, el riesgo reputacional y sus formas de medición, especialmente en el ámbito financiero, han recibido poca atención en general. Históricamente, la industria financiera se ha centrado más en el estudio del riesgo operacional y sus consecuencias en la reputación, en lugar de investigar directamente el riesgo reputacional (Ferreira, 2015). Los estudios que se han realizado específicamente sobre la reputación de instituciones financieras, como los bancos, se han centrado principalmente en países desarrollados, con un enfoque particular en Estados Unidos y Europa. Estos estudios son escasos en otras regiones del mundo, lo que puede deberse a un desinterés o falta de recursos para llevar a cabo investigaciones en esos lugares (Ferreira, 2015) (David Adeabah, 2021).

Descrito lo anterior y teniendo por contexto que este trabajo se enfoca en el sector bancario de Chile, será necesario desarrollar un método nuevo para la representación del riesgo reputacional, que logre captar las características propias de la realidad nacional.

2.1.1 Fuentes de origen de riesgo reputacional

Las fuentes de riesgo reputacional son diversas y no se limitan únicamente a las pérdidas en valoración causadas por riesgos operacionales. Requieren que las instituciones bancarias sean altamente flexibles. Según Carter y Power, el valor de la franquicia y la postura reputacional de las empresas están positivamente relacionados (Richard B Carter, 2012). Walter identifica fuentes de riesgo reputacional como la asimetría entre ganancias y riesgos, la remuneración basada en bonificaciones, la responsabilidad social, los fallos operativos, la productividad y la disponibilidad de recursos (Walter, 2013). Michaelowa identifica otras fuentes de riesgo reputacional, como el descuido o retraso en contribuir a la construcción de un futuro mejor, el comportamiento gerencial irresponsable, el sistema de gobierno ineficiente, la estrategia defectuosa, la mala gestión y liderazgo, la supervisión inadecuada, conflicto de intereses, cultura corporativa problemática y una interpretación laxa de la integridad ambiental (Axel Michaelowa, 2020). Del mismo modo, Zaby y Pohl identifican fuentes de riesgo reputacional, como los requisitos sociales, la satisfacción del cliente, la calidad de los procesos internos, las crisis en otros bancos, la orientación hacia el mercado de capitales y los requisitos legislativos y regulatorios (Simon Zaby, 2019). Otras fuentes de riesgo reputacional incluyen la escala y rentabilidad, así como el financiamiento de proyectos controvertidos (David Adeabah, 2021).

2.1.2 Efectos del riesgo reputacional

Los cambios en la reputación de una empresa pueden generar diversas acciones y reacciones por parte de los grupos de interés, como clientes, inversores y empleados. Estas respuestas a menudo resultan en cambios en los costos y el volumen de negocio de la empresa, pero son difíciles de cuantificar de antemano.

Un buen manejo del riesgo reputacional o una sólida reputación empresarial genera un círculo virtuoso de beneficios en variables financieras clave. Esto incluye un mayor atractivo para atraer talento en los puestos de trabajo, una mayor confianza y disponibilidad de pago por parte de los inversionistas incluso cuando el nivel de riesgo se mantiene constante, y expectativas de beneficios positivas. Además, se observa una rentabilidad superior al promedio del sector y un éxito financiero generalizado, como se refleja en indicadores como el retorno de la inversión. Un buen manejo del riesgo reputacional también conduce a un aumento en la tasa de nuevas operaciones, una reducción de los costos de capital y una mayor facilidad para obtener financiamiento a través de recursos propios. Además, se fortalece la motivación de los empleados y se genera una imagen pública más favorable del banco en los medios de comunicación (Klaus-Peter Wiedmann, 2001) (Simon Zaby, 2019).

Por el contrario, un mal manejo de riesgo reputacional o mala reputación puede generar un círculo vicioso negativo que afecta negativamente la salud financiera de la empresa. Esto se manifiesta en diversas variables financieras, como una disminución en la tasa de adquisición de nuevos negocios y la pérdida de clientes actuales y potenciales. Una mala reputación también puede ocasionar la pérdida de empleados y gerentes dentro de la

organización, lo que conlleva costos adicionales de contratación y períodos de inactividad del personal. Además, la empresa enfrenta mayores costos de financiamiento en términos de crédito o emisión en el mercado de valores. Los prestamistas externos exigen una prima de riesgo más alta, lo que se traduce en mayores costos de capital para el banco. La regulación gubernamental puede generar incrementos en los costos para las empresas, como multas u otras sanciones. A su vez, los accionistas pueden considerar razonable retirar su capital social o vender sus acciones en el mercado secundario, lo que resulta en una disminución en el precio de las acciones (especialmente para los bancos cotizados en bolsa). Además, una mala reputación disminuye la motivación de los empleados y su identificación con el banco (Simon Zaby, 2019) (Jason Perry, 2005). Es importante destacar que una pérdida de reputación causada por el anuncio de una pérdida operativa también puede generar retornos negativos anormales significativos. Además, puede surgir un riesgo sistemático que afecte a bancos que no hayan experimentado una pérdida de reputación (Ferreira, 2015).

A corto plazo, los eventos que afectan la reputación pueden tener un impacto significativo en las ventas y ganancias, aunque en términos absolutos su magnitud suele ser pequeña en comparación con las ventas totales. Sin embargo, es importante tener en cuenta que una reputación negativa tiene un impacto más fuerte que una reputación positiva. Por lo tanto, es crucial gestionar de manera efectiva los eventos que podrían potencialmente dañar la reputación, ya que su prevención es más relevante que aquellos que generan una percepción positiva. Es importante considerar que los eventos de reputación negativa afectan más fuertemente a las empresas de mayor tamaño (Mitic, 2018).

Por último, los clientes tienden a pasar por alto los eventos que afectan la reputación a menos que les impacten directamente. En cambio, su principal preocupación se centra en la ganancia monetaria y la calidad del servicio recibido, en lugar de los problemas asociados con malas conductas o eventos negativos de reputación (Mitic, 2018).

2.2 Valoración de empresa

La valoración de una empresa se refiere al proceso de determinar el valor intrínseco o justo de una entidad comercial en un momento dado. Esta valoración es fundamental en diversos contextos, como fusiones y adquisiciones, inversiones, valoración de activos, planificación financiera y determinación de precios de acciones en los mercados financieros. La valoración de una empresa se basa en una serie de factores y métodos que pueden influir en sus resultados.

Recordando que uno de los objetivos específicos es estudiar el impacto del riesgo reputacional en el desempeño bancario, será necesario también una forma de cuantificar la valoración de empresa.

2.2.1 Variables financieras

La valoración de una empresa se puede llevar a cabo a partir de diversos indicadores financieros, entre los que se incluyen el nivel de liquidez, endeudamiento, rentabilidad, rotación de activos y gestión de inventario.

Por algunas ventajas que se explicarán más adelante en este informe, se elige representar valoración de empresa a partir de ratios de rentabilidad, específicamente la diferencia de precios entre las acciones de los bancos (Ganancia), Retorno sobre los Activos (ROA) y Retorno sobre el Patrimonio Neto (ROE). Estas ratios se calcularán **mes a mes**.

Las definiciones son las siguientes:

Ganancia:

La "ganancia" es la diferencia positiva entre el precio al que una persona compró una acción y el precio al que la vendió. Representa el beneficio que obtiene cuando el precio de venta supera el precio de compra. La ganancia es una medida de éxito en la inversión y significa que la persona ha ganado dinero al vender sus acciones a un precio mayor al que las adquirió. En contraste, si el precio de venta es menor que el precio de compra se produce una "pérdida", lo que significa que la persona ha perdido dinero en la transacción. La ganancia o pérdida resultante es un componente clave en el cálculo del rendimiento o retorno de la inversión en acciones.

En concreto se tratará la ganancia de los bancos muestra, para un período particular, como la diferencia entre el precio de su acción del período correspondiente con el precio de su acción del período anterior. El precio de las acciones se obtendrá de información dispuesta públicamente por la CMF.

ROE (Return on Equity):

El ROE, o Retorno sobre el Patrimonio Neto, es una métrica financiera que evalúa la eficiencia con la que una empresa utiliza su capital propio (patrimonio neto) para generar ganancias. Se calcula dividiendo la utilidad neta de la empresa entre el patrimonio neto promedio de la misma durante un período específico. La fórmula para calcular el ROE es la siguiente:

$$\text{ROE} = \frac{\text{PatrimonioNetoPromedio}}{\text{UtilidadNeta}}$$

- Utilidad Neta: Es el beneficio neto que obtiene la empresa después de deducir todos los gastos, impuestos e intereses.
- Patrimonio Neto Promedio: Es el promedio de los valores del patrimonio neto al comienzo y al final del período.

El ROE es importante porque mide la rentabilidad de los fondos propios de la empresa. Un ROE alto generalmente indica que la empresa es eficiente en la utilización de su capital propio para generar ganancias. La información del ROE mensual se obtendrá de información pública dispuesta por la CMF.

ROA (Return on Assets):

El ROA, o Retorno sobre los Activos, es una métrica que evalúa la eficiencia de una empresa en la utilización de sus activos para generar ganancias. Mide la rentabilidad de los activos totales de la empresa, independientemente de cómo está financiada. La fórmula para calcular el ROA es la siguiente:

$$\text{ROA} = \frac{\text{ActivosTotales}}{\text{UtilidadNeta}}$$

- Utilidad Neta: Igual que en el caso del ROE.
- Activos Totales: Es el valor total de los activos que posee la empresa.

El ROA es importante porque indica cuánto beneficio se genera por cada unidad de activos utilizados. Un ROA alto sugiere que la empresa es eficiente en la gestión de sus activos para generar ganancias. Al igual que en los casos anteriores, los valores de ROA se obtendrán de información pública dispuesta por la CMF.

2.2.2 Características

De las ventajas asociadas al uso de ratios de rentabilidad para representar desempeño y valoración de empresas están:

- **Medición de la eficiencia financiera:** Las ratios de rentabilidad proporcionan una visión directa de la eficiencia con la que una empresa utiliza sus recursos para generar beneficios. Esto es fundamental para evaluar su capacidad para generar ganancias en función de sus activos y capital.
- **Perspectiva a largo plazo:** Las ratios de rentabilidad, en particular ROA y ROE, permiten evaluar el desempeño financiero de una empresa a lo largo del tiempo. Esto es esencial para comprender la consistencia de su rentabilidad y su capacidad para mantenerla en el tiempo.
- **Enfoque en los intereses de los accionistas:** ROE, en particular, se centra en la rentabilidad en relación con el capital de los accionistas, lo que es esencial para los inversionistas, ya que les muestra cuánto retorno están obteniendo sobre su inversión en la empresa.
- **Comprender la eficiencia operativa:** Las ratios de rentabilidad también ofrecen una idea de la eficiencia operativa de la empresa, lo que es valioso para los inversores y gerentes que desean identificar áreas de mejora en la gestión de costos y la generación de ingresos.
- **Acceso público:** Las ratios de rentabilidad, particularmente precio de acción, ROA y ROE, están dispuestos en Chile como información pública por la CMF, lo que facilita su obtención y posiblemente evita otros problemas del ámbito legal.

2.3 Índice

2.3.1 Índice, riesgo reputacional y valoración de empresa

Una estrategia posible para representar y medir riesgo reputacional es a través de un índice derivado de un proceso de minería de datos que se centre en las opiniones de los clientes. (Mitic, 2018). Esto se puede hacer asociando opiniones negativas con presencia de riesgo reputacional y viceversa.

Recordando el capítulo 2.1.2, el riesgo reputacional puede afectar la valoración de una empresa, observándose mayores valores de franquicia, rentabilidades superiores promedio y mejor éxito financiero generalizado a mejor reputación y, por el contrario, menor precio de acciones y ganancias a peor reputación.

Sabiendo que se puede utilizar un índice para representar riesgo reputacional, entonces se vuelve factible construir un índice de riesgo reputacional para los bancos muestra, que este correlacionado en algún grado con la valoración de los bancos y consecuentemente con las variables financieras de Ganancia, ROA y ROE.

2.3.2 Métodos

Hoy en día existen múltiples medios por los cuales la gente expresan sus opiniones sobre empresas, como lo son las encuestas, foros, noticias, redes sociales, blogs, etc. Para efectos de esta memoria se decide utilizar las pagina web de reclamos Reclamos.cl y la red social Twitter, dados los siguientes motivos:

Reclamo

Reclamo.cl se elige al ser una fuente directa de satisfacción, siendo uno de los foros web nacionales más importantes en materia de transparencia y relaciones de consumo, a la fecha con más de 10 años de funcionamiento y habiendo alcanzado en algún momento más de 40.000 visitantes promedios diarios. (Yerko Moena, 2019) La particularidad de Reclamo.cl radica en que cada comentario de los consumidores en la plataforma se trata de una queja dirigida a la empresa, lo que permite establecer una conexión directa entre la frecuencia de reclamos en un período determinado y el nivel de riesgo reputacional que enfrenta el banco. De esa forma períodos de tiempo con mayor frecuencia de reclamos relativa se pueden usar para denotar mayor riesgo reputacional y viceversa. Todo lo anterior considerando la necesidad de normalizar según la cantidad de clientes de cada banco, considerando que mientras más grande sea el banco, es de esperar un mayor volumen de quejas.

Twitter

Twitter tiene una serie de ventajas a la hora de extraer opiniones. Al ser una red social esta es una alternativa costo-efectiva para que las empresas comuniquen y creen relaciones fuertes con sus clientes, ajenas a las barreras geográficas. Twitter siendo una de las redes sociales más grandes del mundo, acapara un público masivo, presentando volúmenes gigantescos de información utilizables como retroalimentación. Twitter ha mostrado tener influencias en varias áreas, tales como resultados de elecciones y debates políticos, comunicaciones académicas, reputación académica, volatilidad de acciones, retornos y volúmenes. (Thomas Fujiwara, 2023). Se sabe que la opinión pública en Twitter puede influenciar significativamente en el mercado financiero, impactando en las decisiones de inversionistas y la aparición de retornos anormales, entre otros (Francisco Guijarro, 2019) (Gabriele Ranco, 2015). Cada uno de los tweets hecho es guardado de forma pública, lo que hace una fuente de información legal y fácil de obtener. Las opiniones a cada uno de los tweets pueden expresar un rango de emociones complejas difíciles de clasificar, pero utilizando el acercamiento adecuado, se puede calcular riesgo reputacional a lo largo del tiempo clasificando tweets como positivos, negativos o neutros.

2.3.3 Diseño

Considerando que en Chile existen numerosos bancos de diversos tamaños, es lógico suponer que la cantidad de información disponible para cada banco en forma de opiniones recopiladas de Reclamos.cl y Twitter variará considerablemente. Esto tiene por resultado que el volumen de información disponible para algunos bancos podría resultar insuficiente para realizar un análisis de riesgo reputacional representativo. Por lo tanto, en lugar de analizar todo el panorama de la industria bancaria en Chile, se llevará a cabo un muestreo de bancos que el investigador considere, según su propio criterio, que posean un tamaño considerable (un gran número de clientes) y una notable presencia en las redes sociales (un alto número de reclamos registrados en Reclamos.cl y seguidores en Twitter).

Debido a consideraciones de privacidad, no se revelarán directamente los nombres de los bancos en el estudio. En su lugar, se utilizarán nombres genéricos como BancoA, BancoB, BancoC, BancoD, BancoE, BancoF, BancoG y BancoH para referirse a ellos de ahora en adelante. El horizonte de tiempo seleccionado para este estudio será de 8 años, basado en la consideración de un ciclo económico en el país y la necesidad de contar con una muestra de datos amplia para obtener resultados significativos. Se utilizará como punto de partida el 1 de abril de 2015 y punto de termino el 31 de marzo de 2023.

Por último, a forma de estudiar riesgo reputacional en distintos escenarios posibles de acuerdo con la naturaleza de la pérdida o ganancia de riesgo reputacional en el tiempo, se plantearán 5 versiones del índice para ambos métodos. Se tendrá un caso base (sin alteraciones) que asumirá que los efectos de pérdida o ganancia reputacional se ve reflejada de forma inmediata. Se tendrá el promedio del índice durante los últimos 3 y 6 meses, lo que permitirá evaluar efecto acumulado del perdida o ganancia reputacional. Se generará una versión del índice que refleje su estado retrasado en 3 y 6 meses, permitiendo evaluar retraso en el efecto de pérdida o ganancia de perdida reputacional. El caso que reporte los “mejores” resultados para el cálculo de requisito de capital debería en teoría reflejar cual es el comportamiento real del índice.

2.4 Método R

2.4.1 Reclamos.cl

Reclamos.cl es una plataforma web que ofrece a los usuarios la posibilidad de presentar quejas y reclamos relacionados con productos o servicios. Su objetivo principal es facilitar la comunicación entre los consumidores y las empresas, con el fin de resolver problemas y mejorar la calidad de atención al cliente. Este sitio web actúa como intermediario en el proceso de reclamación, brindando información sobre los derechos de los consumidores en Chile.

Reclamos.cl pone al alcance del público información abierta sobre reclamos realizados por los usuarios. Esta característica permite acceder de manera conveniente y rápida a una amplia compilación de quejas, las cuales son indicativas de la satisfacción de los clientes en distintas situaciones. Esta información se convierte en una valiosa fuente de referencia para evaluar el desempeño de las empresas y comprender las necesidades y preocupaciones de los consumidores de una manera significativa.

2.4.2 Scraping

El "scraping", o web scraping, se refiere a la práctica de recopilar información de sitios web de forma automatizada. Utilizando programas y herramientas especializadas, esta técnica permite extraer datos estructurados o no estructurados desde páginas web de manera sistemática y eficiente. El proceso implica enviar solicitudes a una página web, analizar el código fuente HTML resultante y seleccionar y extraer la información deseada, como texto, imágenes, enlaces o cualquier otro contenido relevante. El scraping encuentra aplicaciones en diversas áreas, como la recopilación de datos para análisis, investigación, seguimiento de precios, comparación de productos y mucho más. En el caso de esta memoria, el uso de scraping será imprescindible para construir las bases de datos del método R y método T.

2.4.3 Selenium

Selenium es una librería de Python ampliamente utilizada para la automatización de pruebas y la interacción con navegadores web. Su principal objetivo es proporcionar una API sencilla para controlar y manipular navegadores web de manera programática.

Con Selenium, es posible escribir programas en Python que automatizan diversas tareas, como cargar páginas web, completar formularios, hacer clic en elementos y extraer información. Además, estas funciones brindan la posibilidad de extraer información usando técnicas de scraping.

2.4.4 Chromium

Chromium es un proyecto de código abierto que proporciona la base para el desarrollo de navegadores web populares como Google Chrome y otros navegadores similares. En el contexto de scraping de datos, Chromium puede integrarse con Selenium mediante el uso de un controlador específico del navegador llamado WebDriver.

Al utilizar Chromium con Selenium, es posible simular la interacción con el sitio web de Reclamos.cl de manera similar a como lo haría un usuario. Esto implica navegar por las diferentes subpáginas que contienen los reclamos y extraer la información relevante de forma automatizada.

2.4.5 Índice por reclamo

Para el diseño del índice, en este caso para los reclamos de Reclamos.cl, se debe considerar la naturaleza de la plataforma web y el contenido extraído mediante scraping. Cada uno de los reclamos de Reclamos.cl contiene las iniciales del usuario, lugar de residencia, un título, fecha, el reclamo en cuestión y el objetivo del reclamo o la petición del usuario dirigida a la empresa. La mayoría de estos parámetros, sin embargo, no aportan valor para la elaboración del índice, puesto que la naturaleza de un reclamo es naturalmente negativa, tal y como se ha explorado en trabajos previos. (Yerko Moena, 2019). En resumidas cuentas, al declarar el reclamo como negativo el contenido de este mismo se puede prescindir.

Entonces, partiendo de la premisa de que cada uno de los reclamos opera como un comentario con polaridad negativa, se propone la creación de un índice que evalúe la frecuencia de estos reclamos en un periodo de un mes. Esta elección de intervalo de tiempo se justifica por dos razones principales: primero, facilita comparaciones directas con el índice obtenido mediante el Método T; segundo, permite la acumulación de una muestra más sólida de reclamos, ya que un periodo menor podría no generar resultados estadísticamente significativos.

Además, sería beneficioso aplicar una normalización basada en la cantidad de clientes, dado que es lógico suponer que bancos con más clientes generarían un mayor número de reclamos, y viceversa. Por lo tanto, el índice propuesto dividirá las frecuencias de reclamos por la cantidad de clientes registrados en cada banco, mensualmente. Para facilitar la interpretación del índice, se multiplicará por 1,000,000, lo que permitirá concebirlo como la cantidad de reclamos por cada millón de clientes de los respectivos bancos.

Este enfoque operará bajo la suposición de que los clientes de estos bancos no han presentado más de un reclamo por banco durante el período de estudio de 8 años. Aunque esto no debería impactar en los cálculos de manera significativa, es importante mencionar que puede existir cierto grado de superposición de clientes entre las distintas instituciones financieras.

El índice diseñado es el siguiente:

$$IndiceReclamos_{[i,t]} = \frac{FrecReclamos_{it} * 1.000.000}{TotalClientes_{it}}$$

$i \in MuestraBancos, t = tiempo(mes)$

Ecuación 1: Índice Método R. (Elaboración propia)

Donde de forma auto explicativa *IndiceReclamos* representa el índice, *FrecReclamos* la cantidad de reclamos y *TotalClientes* el total de clientes, para alguno de los 8 bancos de muestra estudiados (i), en cada mes (t).

2.5 Método T

2.5.1 Twitter

Twitter es una plataforma de redes sociales y microblogging que permite a los usuarios publicar y leer mensajes cortos, conocidos como "tweets". Estos tweets tienen una limitación de caracteres, de 280 caracteres por mensaje para cuentas normales y 4.000 para cuentas pagadas.

Los usuarios de Twitter pueden seguir a otros usuarios y recibir actualizaciones de sus tweets en su línea de tiempo. También pueden interactuar con los tweets de otros usuarios a través de respuestas, retweets (compartir un tweet en tu propia línea de tiempo) y "me gusta" (indicar que te gusta un tweet).

Twitter se utiliza para una variedad de propósitos, como compartir noticias, opiniones, pensamientos, enlaces, imágenes y videos. También se utiliza como una plataforma para seguir a personas influyentes, celebridades, organizaciones, medios de comunicación y participar en conversaciones sobre temas de interés.

La plataforma se ha convertido en una fuente importante de información en tiempo real, con muchas empresas de noticias, organizaciones y figuras públicas utilizando Twitter para comunicarse directamente con su audiencia y mantenerse actualizados sobre eventos actuales.

Según el informe de Hootsuite y WeAreSocial 2022, Twitter tiene una audiencia disponible para hacer anuncios de 2,9 millones de usuarios para Chile o un 15% de la población y el 18% de los usuarios de Internet del país. (Branch, 2022)



Ilustración 1: Resumen de la audiencia disponibles para anuncios. (Branch, 2022)

Datareportal por otro lado en su reporte “Digital” más reciente, señala que Twitter tiene cercano a 3,65 millones usuarios chilenos a principios 2023, equivalentes a 18,6 por ciento de la población. (DataReportal, 2023)

Aunque la información previa sugiere que Twitter no ha sido adoptado por la mayoría de la población chilena, los datos presentados indican que la cantidad de usuarios chilenos en esta plataforma es sustancial. Estas cifras respaldan la factibilidad de obtener una muestra significativa de tweets a lo largo del período de 8 años propuesto.

2.5.2 Snsrape

Para extraer la información de los tweets de los clientes de los bancos, Twitter ofrece una interfaz de programación de aplicaciones (API, por sus siglas en inglés). Sin embargo, debido a recientes cambios en las políticas empresariales tras la adquisición de la empresa por parte de Elon Musk, el acceso a esta herramienta se ha visto restringido. Ahora se requiere una cuenta de nivel empresarial para realizar tareas de recolección de datos o *scraping*, que anteriormente estaban disponibles incluso para usuarios no suscritos a planes de pago. Ante esta limitación, se decidió utilizar la librería *snsrape*, escrita en Python y enfocada en el *scraping* de diversas redes sociales. Esta herramienta, similar a la API de Twitter, permite realizar búsquedas en un volumen considerable y de manera automatizada. Además, al estar programada en Python, ofrece aún más opciones de personalización en las búsquedas. Una ventaja adicional es que no hay restricciones en la cantidad de búsquedas que se pueden realizar en Twitter, lo cual es crucial para poder estudiar los bancos chilenos seleccionados a lo largo del horizonte de tiempo de 8 años propuesto.

2.5.3 NLP y Análisis de Sentimiento

El procesamiento de lenguaje natural (NLP, por sus siglas en inglés) es un campo de estudio que se centra en la interacción entre las computadoras y el lenguaje humano. Consiste en desarrollar modelos y algoritmos que permitan a las máquinas comprender, interpretar y generar lenguaje natural de manera efectiva.

El análisis de sentimiento se refiere a la determinación automática de la subjetividad, polaridad y fuerza de un texto, sin importar si está escrito con intencionalidad objetiva o subjetiva (Carlos Perales González, 2017). Esta es una subárea del NLP que ha experimentado un crecimiento significativo en el último tiempo, impulsado en gran medida por el reciente interés del público en la inteligencia artificial, a raíz del éxito de herramientas como ChatGPT de OPENAI.

El análisis de sentimiento se ha utilizado ampliamente en diversos contextos, como las redes sociales, el mercado de acciones, las elecciones y la gestión de desastres (Terblanche, 2020) .

Existen varios desafíos que plantean dificultades en el análisis de sentimiento:

- **Prevalencia del inglés:** Los modelos de análisis de sentimiento se han desarrollado principalmente en inglés debido a la disponibilidad de información y el interés en este idioma, lo que dificulta su aplicabilidad en otros idiomas, como el español.
- **Ambigüedad lingüística:** La presencia de sarcasmo, expresiones idiomáticas y el contexto cultural pueden llevar a interpretaciones erróneas por parte de los programas de análisis de sentimiento.

- **Variabilidad en las expresiones de sentimiento:** Las personas pueden expresar sentimientos de diversas formas que son difíciles de capturar de manera precisa en texto escrito, como tonos y matices emocionales.
- **Disponibilidad de datos:** La mayoría de los datos disponibles en Internet carecen de etiquetas y se encuentran desorganizados, lo que dificulta el entrenamiento de modelos de aprendizaje y la obtención de representaciones adecuadas, ya que se requiere una gran cantidad de datos.
- **Limitaciones de dominio:** Los modelos de análisis de sentimiento suelen estar entrenados en un dominio o contexto específico, lo que limita su capacidad para generalizarse a otros dominios o contextos.

A la fecha existen múltiples estrategias para efectuar análisis de sentimiento, pero las principales son por lexicón o aprendizaje de máquina:

Lexicón o diccionario

Método diseñado para determinar la polaridad o la carga emocional de un texto. Este enfoque se basa en el uso de lexicones, que son listas de palabras predefinidas con anotaciones de sentimiento asociadas a ellas. Depende en gran medida de la calidad y cobertura del lexicón utilizado. No tiene en cuenta el contexto o la estructura sintáctica de las frases, lo que puede llevar a resultados menos precisos en ciertos casos.

Esta metodología es relativamente simple y rápida, pero depende en gran medida de la calidad y cobertura del lexicón utilizado.

Aprendizaje de máquina

Esta metodología utiliza algoritmos de aprendizaje automático para entrenar modelos que puedan identificar y clasificar el sentimiento en un texto. Se necesita un conjunto de datos etiquetados, es decir, un conjunto de textos previamente clasificados con su sentimiento correspondiente (positivo, negativo, neutro, etc). A partir de estos datos etiquetados, se extraen características relevantes del texto (como palabras clave, frecuencia de palabras, estructura gramatical, etc.) y se utilizan como entradas para entrenar un modelo de aprendizaje automático

Esta metodología es más flexible y puede capturar el contexto y las relaciones complejas entre las palabras, pero requiere un conjunto de datos etiquetados y más recursos computacionales para el entrenamiento del modelo.

Como el Método T procesará sentimiento desde Twitter, necesitando así una estrategia que logre comprender contextos complejos, sutilezas de lenguajes y detectar elementos como ironía, se empleará **aprendizaje de máquina** para análisis de sentimiento.

2.5.4 Transformadores

En el contexto del aprendizaje de máquina, los transformadores son una arquitectura de modelo que se utiliza para tareas de procesamiento de lenguaje natural y otras tareas de aprendizaje automático. El concepto de los transformadores fue introducido en el popular paper "Attention Is All You Need" (Ashish Vaswani, 2017).

En este *paper*, los autores proponen una arquitectura de red neuronal basada completamente en mecanismos de atención, eliminando las capas recurrentes o convolucionales utilizadas en enfoques tradicionales de NLP. Los transformadores se diseñaron para superar las limitaciones de las redes recurrentes en términos de paralelismo y eficiencia computacional.

La idea clave detrás de los transformadores es el mecanismo de atención, que permite a la red aprender a enfocarse en partes relevantes de la entrada durante la tarea de procesamiento. En lugar de procesar la secuencia de entrada en orden secuencial, los transformadores calculan las representaciones de todas las posiciones de entrada simultáneamente mediante operaciones de atención.

En los transformadores, la entrada se divide en vectores de palabras (embeddings) y se someten a múltiples capas de atención. La atención se calcula mediante productos escalares entre los embeddings de consulta, clave y valor. Esto permite que el modelo se centre en partes específicas de la secuencia de entrada, identificando relaciones y dependencias importantes.

El enfoque basado en transformadores ha demostrado ser altamente efectivo en varias tareas de NLP, como traducción automática, generación de texto, reconocimiento de voz y más. Los transformadores han logrado avances significativos en términos de precisión y velocidad de entrenamiento en comparación con las arquitecturas anteriores, haciéndolos ideales para efectos de esta investigación.

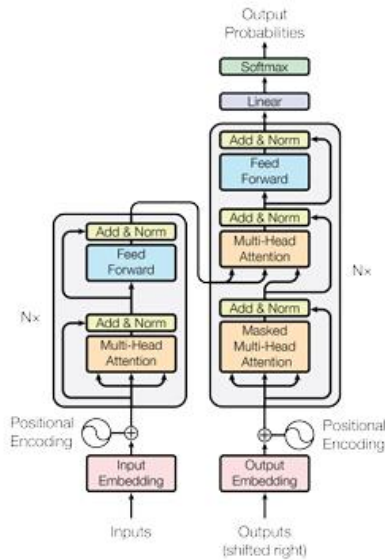


Ilustración 2: Arquitectura del Transformador. (Ashish Vaswani, 2017)

2.5.5 Pysentimiento

dialectos de español, que usa como base un modelo de nombre RoBERTuito, que a su vez se basa en el modelo transformador de NLP “RoBERTa” pero entrenado en 500 millones tweets (Juan Manuel Pérez, 2021) (Pérez, 2022) (García-Vega, 2020)

RoBERTa (Robustly Optimized BERT approach) es un modelo de lenguaje basado en la arquitectura de los transformadores y se basa en la popular arquitectura de BERT (Bidirectional Encoder Representations from Transformers). RoBERTa fue presentado en el paper "RoBERTa: A Robustly Optimized BERT Pretraining Approach" (Yinhan Liu, 2019).

Test de robustez

Comparado con otros modelos de procesamiento de lenguaje natural en español, RoBERTuito se destaca en áreas como detección de discurso de odio, análisis de sentimientos, emoción e identificación de ironía. En pruebas que abarcan estas tareas, RoBERTuito ha demostrado un rendimiento superior a modelos como RoBERTa, BERTin y BETO. Esto se puede ver en la siguiente tabla donde se efectuaron pruebas por puntuación Macro F1:

Model	Emoción	Hate speech	Ironía	sentimiento
robertuito	0,560 ± 0,010	0,759 ± 0,007	0,739 ± 0,005	0,705 ± 0,003
roberta	0,527 ± 0,015	0,741 ± 0,012	0,721 ± 0,008	0,670 ± 0,006
bertin	0,524 ± 0,007	0,738 ± 0,007	0,713 ± 0,012	0,666 ± 0,005
beto_uncased	0,532 ± 0,012	0,727 ± 0,016	0,701 ± 0,007	0,651 ± 0,006
beto_cased	0,516 ± 0,012	0,724 ± 0,012	0,705 ± 0,009	0,662 ± 0,005
mbert_uncased	0,493 ± 0,010	0,718 ± 0,011	0,681 ± 0,010	0,617 ± 0,003
biGRU	0,264 ± 0,007	0,592 ± 0,018	0,631 ± 0,011	0,585 ± 0,011

Tabla 1: Tests Macro F1 para RoBERTuito. (Pérez, 2022)

Donde el puntaje relevante a análisis de sentimiento en español indica un valor de 0,705 +- 0,003, sugiriendo ser el mejor modelo de este *benchmark*.

2.5.1 Índice por Twitter

Cuando se aborda el diseño de un índice destinado a emplearse en este método, resulta fundamental tener en consideración la naturaleza propia de Twitter, así como la información de relevancia que puede ser extraída de cada publicación. En cada tweet, se engloban elementos tales como la fecha de su creación, el usuario emisor, el contenido textual en sí, el recuento de "me gusta", la cantidad de retweets y el número de respuestas generadas. El proceso de análisis de sentimiento se llevará a cabo naturalmente sobre el contenido textual de los mensajes. Cada uno de estos mensajes será procesado mediante el modelo de RoBERTuito, lo que posibilitará obtener una evaluación con respecto a si el comentario en cuestión presenta una polaridad negativa, neutra o positiva.

En el entorno de Twitter, es innegable que las figuras prominentes, ya sean del ámbito financiero u otros sectores, pueden ejercer un impacto considerable en la percepción pública debido a la extensa cantidad de seguidores que atraen. Este impacto se traduce en atributos clave de los tweets, como la cantidad de "me gusta", retweets y respuestas que logran generar. Para reflejar adecuadamente esta influencia en el índice, se introducirá un factor adicional que otorgará una ponderación progresiva a los comentarios, proporcional a la magnitud de estas propiedades. En otras palabras, a medida que aumenta el nivel de estas métricas (es decir, "me gusta", retweets y respuestas), el valor de ponderación asignado al comentario también se incrementará en proporciones similares. Este enfoque permitirá capturar y representar de manera precisa el grado de influencia ejercido por las personalidades notables en la opinión pública, incluso en comentarios de tamaños comparables.

Además de lo expuesto previamente, se propone realizar mediciones periódicas del índice en contraposición a mantener un índice estático. Esta estrategia permitiría la creación de una serie temporal que, en teoría, facilitaría la identificación de eventos particulares, como pérdidas operativas, que pudieran impactar la reputación de las entidades bancarias. (Mitic, 2018) (Carlos Perales González, 2017). El cálculo del índice se llevará a cabo en intervalos mensuales, promediando los sentimientos registrados diariamente a lo largo de dicho período. Esta elección se basa en la premisa de que intervalos de tiempo más cortos podrían arriesgar la representatividad de la muestra, mientras que intervalos más largos proporcionarían muy pocos datos a lo largo de la serie temporal.

Con la capacidad de obtener polaridades positivas, neutras y negativas a partir de Pysentimiento, se propone el índice:

$$Indice_{[i,t,t_f]} = \frac{1}{t} \sum_t^{t_f} \frac{pos_{it} + (-1) * neg_{it}}{pos_{it} + neu_{it} + neg_{it}}$$

$i \in MuestraBancos, t = tiempo(dias)$

Ecuación 2: Índice Método T. (Elaboración propia)

Donde pos, neg y neu representan los valores de polaridad positiva, neutra y negativa respectivamente, asociados a la empresa i en un día específico t . Estos valores se calculan considerando la frecuencia de comentarios etiquetados por RoBERTuito como positivos, negativos o neutrales. Posteriormente, se les suma un valor constante de 1, junto con el factor de popularidad del usuario que realizó cada comentario, para obtener los valores finales.

Esto se ve de la siguiente forma:

$$pos = \sum_1^f 1 + factor * (nrespuestas + nmegustas + nretweets)$$

$$neg = \sum_1^f 1 + factor * (nrespuestas + nmegustas + nretweets)$$

$$neu = \sum_1^f 1 + factor * (nrespuestas + nmegustas + nretweets)$$

Ecuación 3: Índice Método R. (Elaboración propia)

Donde f representa la frecuencia del tweet que fluctúa dependiendo de polaridad del tweet.

2.6 Correlación

Para estudiar potencial correlación entre los índices y las variables financieras en el Método R y Método T, se requiere el procesamiento de la data obtenida a partir de técnicas estadísticas. Se estudiará la correlación en de la muestra individualmente y como sistema.

2.6.1 Regresiones lineales

Una de las técnicas más comunes para evaluar la correlación en análisis de datos es la regresión lineal.

La regresión lineal es un método estadístico utilizado para modelar la relación entre una variable dependiente (también llamada variable de respuesta) y una o más variables independientes (también llamadas variables predictoras o explicativas), asumiendo que esta relación puede aproximarse mediante una función lineal. En otras palabras, la regresión lineal busca encontrar la mejor línea recta que se ajuste a los datos observados y pueda utilizarse para predecir valores de la variable dependiente en función de los valores de las variables independientes.

La ecuación de la regresión lineal es:

$$Y = \beta_0 + \beta_1 X + \epsilon$$

Ecuación 4: Regresión lineal

Donde Y corresponde a la variable dependiente (Ganancia, ROA o ROE), β_0 al intercepto, β_1 pendiente, X a la variable independiente (índice base, índice promedio acumulado o índice rezagado) y ϵ el error o variación no explicada por la relación lineal.

2.6.2 Modelo de datos de panel

Para expandir el estudio de correlación a nivel grupal, se puede valer de la técnica estadística conocida como modelo de datos de panel. Esta permite estudiar comportamiento conjunto, reflejado en un parámetro β útil para el cálculo de requisito de capital.

Más en detalle, un modelo de datos de panel se utiliza para analizar conjuntos de datos que incluyen observaciones repetidas a lo largo del tiempo para múltiples entidades o individuos. El modelo de datos de panel esencialmente combina elementos de análisis de series temporales y análisis transversales para obtener una comprensión más profunda de los patrones y las relaciones en datos que cambian tanto con el tiempo como entre diferentes unidades o sujetos.

Un modelo de datos de panel puede implementar regresiones lineales en la estructura del análisis de datos de panel, permitiendo la estimación de los coeficientes relacionados a las variables independientes para cada entidad individual y en cada período temporal. Esta aproximación permitirá en teoría capturar de manera efectiva los efectos de la variable independiente en la variable dependiente a lo largo del período estudiado, para todas las entidades.

2.7 Requisito de Capital

El requisito de capital se entiende como el capital básico (patrimonio) y otros tipos de instrumentos (bonos, acciones preferentes) que los bancos deben conformar para asegurar su solvencia. Estos requisitos se miden como un porcentaje de los activos de los bancos, que a su vez son ponderados por el riesgo (APR), requiriéndose así conformar un capital regulatorio mayor para un activo más riesgoso. (Verbanaz, 2019). El requisito de capital es entonces es la cifra específica que un banco debe mantener para cumplir con reglas general y estándares que los bancos deben seguir en relación con su capital.

2.7.1 Capital regulatorio

El capital regulatorio es la cantidad mínima de fondos propios que un banco o una institución financiera debe mantener para respaldar sus operaciones y proteger a los depositantes y acreedores en caso de situaciones adversas. El capital regulatorio se establece por las autoridades financieras, como los bancos centrales o las agencias de supervisión bancaria, con el objetivo de garantizar la estabilidad y solidez del sistema financiero. Adicionalmente el capital regulatorio debe ser igual o superior al requisito de capital exigido a los bancos.

El capital regulatorio se utiliza para proteger a los depositantes y otros acreedores del banco, ya que estos fondos son una especie de colchón de seguridad en caso de pérdidas financieras. Si un banco enfrenta problemas o pérdidas en sus activos, el capital regulatorio actúa como una reserva para cubrir esas pérdidas y mantener la continuidad de las operaciones.

Los acuerdos de capital del Comité de Basilea distinguen dos niveles de capital, según su capacidad de cobertura de pérdidas. El capital de nivel 1 (T1) constituye el capital de funcionamiento (going-concern), conformado a su vez por el capital básico (CET1) y el capital adicional de nivel 1 (AT1). El capital de nivel 2 es aquel a utilizar en la liquidación de un banco (gone-concern), constituido principalmente por bonos subordinados y provisiones voluntarias. El capital regulatorio es igual a la suma del capital de nivel 1 y el capital de nivel 2. Un elemento esencial en el marco de capital en Basilea III, es la mayor importancia relativa que adquiere el capital de mayor calidad o capital básico pues, además de elevar su piso, suma una serie de buffers o colchones adicionales constituidos

exclusivamente con este tipo de capital, corrigiendo una de las principales deficiencias del marco de capital de Basilea II, que abrió un mayor espacio al capital híbrido (Gabriela Aguilera, 2020).

2.7.2 Activos Ponderados por Riesgo

Los Activos Ponderados por Riesgo (APR) se utilizan para determinar la cantidad de capital que un banco o entidad financiera debe mantener en reserva para cubrir sus riesgos crediticios, de mercado y operativos. Básicamente, los APR reflejan la cantidad de riesgo asociada con los activos de un banco y son una medida para asegurarse de que la entidad tenga suficiente capital para protegerse contra posibles pérdidas. (Gabriela Aguilera, 2020)

El capital regulatorio se expresa como un porcentaje de los APR del banco. Los activos ponderados por riesgo reflejan el riesgo asociado con los diferentes tipos de activos que posee un banco. Cuanto mayor sea el riesgo percibido de un activo, mayor será su ponderación y, por lo tanto, requerirá más capital.

La LGB impone un CET1 a con un piso de 4,5% de los APR. CET1 + AT1 de un 6% y un 8% para el capital regulatorio (T1 + T2).

2.7.3 Evaluación de capital

Uno de los propósitos del Pilar 2 de los Acuerdos de Basilea, es garantizar un equilibrio apropiado entre el perfil de riesgo y el capital real que las instituciones financieras mantienen. Esto se aplica tanto en términos absolutos como en la composición del capital (CET1, T1 y T2), considerando también cómo se distribuye entre diferentes entidades independientes en un grupo consolidado de instituciones financieras. Este ajuste entre el capital real y los riesgos inherentes ayuda a mejorar la gestión de riesgos. Para lograr este objetivo, las instituciones deben llevar a cabo una autoevaluación de capital (PAC). En este proceso, identificarán, cuantificarán y sumarán sus riesgos, calcularán el capital necesario para cubrirlos, incorporarán una planificación de capital a mediano plazo y establecerán un objetivo interno de capital real. Esto permitirá mantener constantemente un margen adecuado por encima de los requerimientos legales establecidos en el Pilar 1. (Carlos Pulgar, 2020).

La nueva LGB busca cerrar las brechas existentes en los principios del Pilar 2, que ya están presentes en la regulación actual sobre clasificación de gestión y solvencia (capítulo 1-13 de la RAN). Específicamente, la ley otorga al supervisor la capacidad de exigir capital regulatorio adicional a instituciones con evaluaciones deficientes. Según el artículo 66 quinquies de la LGB, si la supervisión revela riesgos insuficientemente cubiertos por los requisitos patrimoniales generales, el supervisor puede requerir que los bancos mantengan más capital efectivo (ya sea de nivel 1 o 2, según lo autorice la Comisión). Esta decisión debe basarse en una resolución fundamentada, siguiendo los criterios generales del nuevo Capítulo 21-13 de la RAN. Además, al menos cuatro de los cinco Comisionados deben

votar a favor de esta medida, que no puede exceder el 4% de los Activos Ponderados por Riesgo (APR) del banco. (Carlos Pulgar, 2020)

Según la nueva LGB la clasificación de solvencia de un banco entonces está distribuido como se señala en la siguiente imagen:

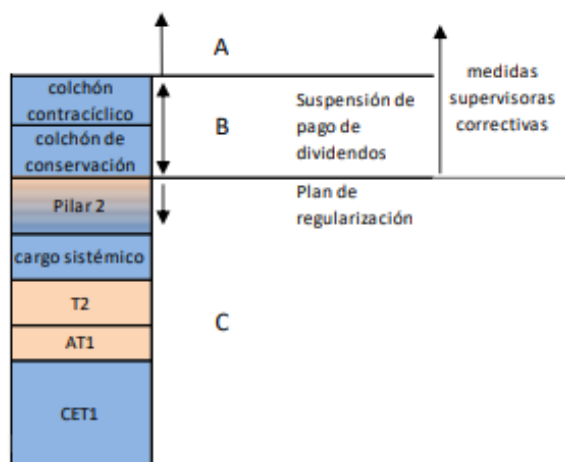


Ilustración 3: Evaluación de solvencia en la nueva LGB. (Carlos Pulgar, 2020)

Por tanto, esta memoria de calcular exitosamente requisito de capital por riesgo reputacional ayudaría con la implementación de los requerimientos del Pilar 2, permitiendo llegar a una solvencia B.

2.7.4 Calculo

La expresión básica para determinar el requisito de capital (o capital regulatorio correspondiente) se define como:

$$Req. capital = \frac{Fondos}{APR}$$

Ecuación 5: Requisito de capital

Se quiere ponderar el efecto del riesgo reputacional, por lo que recordando que las fluctuaciones en los índices de riesgo reputacional (tanto para Método R como Método T) pueden correlacionar con la ganancia de los bancos, este efecto también debería poderse expresarse en términos monetarios (Mitic, 2018).

El requisito de capital entonces se multiplica por el coeficiente β del sistema de la muestra de bancos, calculado a través de regresiones lineales por panel. Este coeficiente indica la proporción de cambio de la variable dependiente en función de la independiente, por lo que

al multiplicarse por el requerimiento de capital se consigue la proporción del requerimiento de capital que corresponde a reputación.

$$Req. capital = \frac{\beta * Fondos}{APR}$$

Ecuación 6: Requisito de capital (Efecto β)

Además, considerando que el objetivo de fijar un requerimiento de capital es de evitar problemas de solvencia, se debe plantear la ecuación pensando en el caso pesimista. Esto se puede lograr representar agregando el efecto del índice estresado por un valor extremo o δ .

$$Req. capital = \frac{\beta * \delta * Fondos}{APR}$$

Ecuación 7: Requisito de capital (Efecto β y δ)

El valor extremo se calculará como la diferencia máxima entre los valores de los índices para 12 meses o 1 año consecutivo, bajo la lógica de identificar los cambios más pronunciados en la serie de tiempo.

El requerimiento de capital cambiara dependiendo como se seleccione el β . Los planteamientos según βROA y βROE son los siguientes:

$$Req. capital. ROA_{[i,j]} = \frac{\beta ROA_j * \delta_i * Activos_i * 100}{APR_i}$$

Ecuación 8: Requisito de capital (ROA)

$$Req. Capital. ROE_{[i,j]} = \frac{\beta ROE_j * \delta_i * Patrimonio_i * 100}{APR_i}$$

Ecuación 9: Requisito de capital (ROE)

Donde i y j corresponden en ambos casos a los 8 bancos de muestra y las 5 versiones del índice. Se multiplica por 100 para expresar los valores porcentualmente a forma de facilitar la lectura.

3 Implementación y análisis de resultados

3.1.1 Extracción de data

Método R

Se desarrolló un programa en Python que utilizó la librería Selenium en combinación con Chromium para llevar a cabo un proceso de recolección de datos. El programa se encargó de acceder a las páginas de reclamos de ocho bancos diferentes. Para ello, se proporcionó al controlador (driver) la dirección web de la subpágina de Reclamo.cl asociada a cada banco.

La metodología empleada implicó la utilización de funciones específicas proporcionadas por el controlador de Chromium, que permitieron interpretar y navegar por elementos de la página utilizando selectores CSS. De esta manera, el programa accedió a cada uno de los reclamos presentes en una tabla y, posteriormente, almacenó la fecha y el contenido de cada reclamo en una lista vacía.

El programa fue diseñado para abarcar un período de tiempo específico, comprendido entre el 1 de abril de 2015 y el 31 de marzo de 2023. Durante este intervalo, el proceso de búsqueda y extracción de datos se realizó de manera automatizada y ordenada.

La información recopilada durante este proceso se gestionó y almacenó en una base de datos utilizando la librería Pandas, una herramienta popular para la manipulación y el análisis de datos en Python. Una vez que los datos se organizaron y estructuraron adecuadamente, el programa exportó la información en un archivo en formato CSV. Se optó por la codificación UTF-8 para garantizar la compatibilidad con caracteres especiales y asegurar la integridad de los datos almacenados.

	Fecha	Contenido
0	17-05-23	Estimados cuando se cumplió el año del seguro se llamo al número entregado por F
1	15-05-23	El día 06-04-23 se compro con tarjeta de credito F en la tienda Adidas un par de zapatillas,
2	08-05-23	En el mes de marzo refinancie la deuda que tenía con F o banco F y mi fecha de pago para la primera cuota esta para el 05/05 del presente año
3	05-05-23	Pésima la atención en el banco F del mall ARAUCO Maipú, tienen poco personal y lo poco que tienen es muy ineficiente, se demoran demasiado en atender, hay 10 módulos
4	27-04-23	contrate un hipotecario banco F en el año 2008, repentinamente cerraron todas las sucursales, quedando solamente con atención en Santiago centro, no tienen un sistema
5	22-04-23	Estoy muy molesta porque me engañaron compre unas almohadas estaban en oferta, inclusive di un abono de 7.000 quiero pagar las cuotas son 2 de 6.490. Cual fue mi
6	20-04-23	El día 20 de abril de 2023 me dirijo a banco F sucursal Moneda para hacer el depósito de un dinero, ingreso al banco me saluda el guardia y le pregunto donde puedo
7	19-04-23	Estimados Buenas noches, Junto con saludar quiero expresar mi descontento, situación que paso a detallar cuando quise utilizar mi tarjeta F, el ejecutivo me señalo que para
8	19-04-23	Buenas Tardes, Realice el cambio de tarjeta ya que tenía la antigua pero NADIE aviso que la mantención subía a 7.000 mil pesos. cuando realice el cambio la persona me dijo
9	16-04-23	Buenos días el banco F me tiene bloqueada la tarjeta diciendo que la razón es por NO USO. Pero al preguntar al servicio cliente dicen que la puedo usar y que no está

10	02-04-23	Pelota de marca Puma me duró 60 minutos de juego , está se pincho de una no se supone ke son para jugar fútbol ?? No alcance ni a pagar una cuota y se pincho pésima
11	30-03-23	Me llegó un promoción dirigida, que consiste en comprar en algún comercio que no sea F y me devuelven el cobro de mantención (usurero de sobre 8500), activo la promoción,
12	27-03-23	Buenas noches,Mi nombre es ... A continuación explico mi situación:Después de muchos intentos mediante correos y diversos llamados al servicio al cliente y Cobranza
13	23-03-23	Page el total de mi deuda todo al día, asta page de mas como un abono. No registro deuda, pero al sacar, informe de deuda salgo como deudor algo que sorprende, asta ellos indicaron eso.
14	22-03-23	Sres. Banco F:Para pagar cuota de crédito de consumo, se solicita información vía telefónica, ya que monto de la cuota, supera el monto máximo de su pagina web.
15	19-03-23	Estoy muy molesta, soy ... cliente de F y hace una semana hice una compra de una lavadora porque se me había echado a perder la mía , lo cual tuvo arreglo y anulé la
16	17-03-23	Llaman a mi trabajo y en ningún momento ingresé el teléfono de mi trabajo ni de mi casa, como es posible que sepan esa información si yo no la dí, con que trafico de información trabajan esta empresa si obtienen datos particulares.
17	15-03-23	Hice una compra el 29 de diciembre y me pidieron validar un cupón de descuento con la tarjeta de la tienda. La compra la pagué en el momento en efectivo, salí de la tienda con la
18	10-03-23	el mes pasado pague la cuota mensual adelantada, por lo que este mes debio salir cero, pero para mi sorpresa tengo un costo de mantencion sobre 8 mil pesos, reviso meses
19	08-03-23	Después de devolver unos productos con nota de crédito , más una compra que no llego y en la cual también se realiza nota de crédito , la empresa me cobra a la fecha \$40.000.
20	08-03-23	Buenas tarde, necesito que quesaquen del seguro dental, ya que B llamando desde enero y aun no me dan una solución fui al banco F de valparaiso y nada muy mala la
21	08-03-23	Me tiene cansada que las empresas y bancos, se den el lujo de hacer uso del teléfono personal de uno, para llamar libremente y ofrecer porquerías que NO ME INTERESAN!
22	01-03-23	Buenas tardes, escribo por este medio para informar que el día 25/02 hice un pago en banco F via servipag y el pago figura doble en mi cuenta y en banco F. Espere las 72

Tabla 2: Ejemplo de base de datos. (Método R)

Método T

Se desarrolló un programa en Python para extraer información y tweets utilizando la librería snsrape. Este programa utilizó un criterio de búsqueda o query para indicar a snsrape que recorriera todos los tweets de los usuarios que mencionaron las cuentas oficiales de los bancos seleccionados. Se aseguró de excluir los tweets realizados por los propios bancos hacia sí mismos. El período de búsqueda abarcó desde el 1 de abril de 2015 hasta el 31 de marzo de 2023, es decir, los últimos 8 años.

La información obtenida de cada tweet, realizada por los usuarios considerados como clientes de los bancos respectivos, se almacenó en una lista vacía. Para cada tweet, se guardó información relevante como la fecha, el nombre de usuario, el contenido del tweet, el número de respuestas, el número de "me gusta" y el número de retweets.

Una vez completado este proceso, los datos recopilados se convirtieron en una base de datos utilizando la librería pandas, y se exportaron en formato .csv con codificación UTF-8. Debido a un problema propio del programa snsrape, resultó imposible realizar la extracción completa de tweets para los 8 años en una sola iteración, lo que requirió la extracción de la información en "fragmentos". Es posible que, debido a esta situación, se haya perdido cierta cantidad de información a lo largo de este proceso.

Date	Tweet	replyCount	likeCount	retweetCount
2016-05-26 23:49:26+00:00	@BancoC no he podido realizar transferencias	3	0	0
2016-05-26 23:46:49+00:00	@Cooperativa y desde el lunes protestaremos	0	8	12
2016-05-26 23:43:40+00:00	@LealFleal1986 @BancoC	0	4	6
2016-05-26 23:43:07+00:00	@nomadiko1 @BancoC #huelgaC	0	3	6
2016-05-26 23:41:38+00:00	@sindicatoC @BancoC comparte este #yarur	1	3	3
2016-05-26 23:39:10+00:00	@sindicatoC @BancoC #huelgaC Yarur es el	0	17	23
2016-05-26 23:37:19+00:00	Su manual de "beneficios" es una práctica	3	13	25
2016-05-26 23:35:46+00:00	@BancoC @CNovaOficial puras falacias apoyamos	0	0	0
2016-05-26 23:33:39+00:00	Siempre hemos estado abiertos a conversar....	2	8	16
2016-05-26 23:28:43+00:00	@GabyTorrealbaM @sindicatoC	0	3	3
2016-05-26 23:18:58+00:00	No entiendo como #BancoC kiere ensuciar	1	9	10
2016-05-26 23:00:06+00:00	@MichelPesquera @PalabrasTiernas difunde	0	0	0
2016-05-26 22:57:21+00:00	@barretojacks @sindicatoC @EvitaMe5	0	5	6
2016-05-26 22:55:54+00:00	@barretojacks @sindicatoC @EvitaMe5 @lissetrop	1	2	3
2016-05-26 22:55:16+00:00	@barretojacks @sindicatoC @EvitaMe5 @lissetrop	1	1	3
2016-05-26 22:53:54+00:00	@barretojacks @sindicatoC @EvitaMe5 @lissetrop	0	2	2
2016-05-26 22:45:44+00:00	Zarko Luksic subsecretario del trabajo y diputado DC ,	0	5	10
2016-05-26 22:43:44+00:00	@BancoC #somosindiferentes a sus	0	0	0
2016-05-26 22:31:55+00:00	@BancoC vende 3 valores:respeto integridad y	0	0	0
2016-05-26 22:30:17+00:00	@Ursula_Riquelme @sindicatoC @BancoC	0	8	12
2016-05-26 22:18:11+00:00	@BancoC @CNovaOficial @TBanc #yarur Queriai	0	0	0
2016-05-26 22:14:21+00:00	@BancoC estimados; buenas tardes. Cuanto más	5	0	0
2016-05-26 22:10:21+00:00	No más a la Pulpería de @BancoC @sindicatoC	0	5	11
2016-05-26 21:58:01+00:00	@tannganini @EvitaMe5 @lissetrop @sindicatoC	1	0	0
2016-05-26 21:57:01+00:00	@tannganini @EvitaMe5 @lissetrop @sindicatoC	0	1	0
2016-05-26 21:56:01+00:00	@tannganini @EvitaMe5 @lissetrop @sindicatoC	0	0	0

Tabla 3: Ejemplo de base de datos. (Método T)

3.1.2 Procesamiento de data

Método R

Se crea un programa en Python que carga la base de datos con pandas para reclamos en formato .csv. En este caso se importa tanto las bases de datos resultantes del scraping, como un set de bases de datos con la información del número de clientes totales por cada una de las 8 instituciones bancarias estudiadas.

El algoritmo en este caso es de la siguiente forma:

1. **Agrupación de Reclamos por Año y Mes:** Se agrupan los reclamos según Año y Mes.
2. **Creación de una Nueva Base de Datos:** Se crea una nueva base de datos que agrupa frecuencia de reclamos según Año y Mes.
3. **Combinación de Bases de Datos:** Se cruza por Año y Mes la base de datos anterior de frecuencia agrupada con la base de datos de número de clientes mensuales correspondiente. Para ello también se fuerza la segunda base de datos a tener el mismo formato de fecha.
4. **Cálculo del Índice:** Se añade una columna de nombre 'Índice' a la base de datos combinada, que resultante de aplicar el índice de reclamo a la frecuencia, lo que es a su vez equivalente a multiplicar la frecuencia por 1,000,000 y dividir el resultado según el número de clientes correspondiente.
5. **Exportación de Resultados:** Se retorna la base de datos resultante en formato .csv.

Año	Mes	Índice
2015	4	10.02
2015	5	8.53
2015	6	8.46
2015	7	8.42
2015	8	12.59
2015	9	12.55
2015	10	12.48
2015	11	11.06
2015	12	20.61
2016	1	5.56

2016	2	10.97
2016	3	15.05
2016	4	6.81
2016	5	13.62
2016	6	10.82
2016	7	6.74
2016	8	14.77
2016	9	14.75
2016	10	4.00
2016	11	20.02
2016	12	14.62
2017	1	15.86
2017	2	10.55
2017	3	11.85
2017	4	7.87

Tabla 4: Ejemplo de Índice Base. (Método R)

Método T

Se crea un nuevo programa en Python que carga con pandas la base de datos de tweets en formato .csv extraída con anterioridad y se importa ahora la librería pysentimiento.

El algoritmo en este caso es de la siguiente forma:

1. **Preprocesamiento de Tweets:** Llama al método preprocess_tweet de pysentimiento. Este transforma elementos de tweets difíciles de interpretar a etiquetas fácilmente legibles por RoBERTuito. Se simplifican direcciones url, caracteres repetidos, risas, hashtags y emojis.
2. **Predicción de Sentimiento:** Llama al método predict y devuelve el sentimiento identificado con la etiqueta POS, NEU o NEG para un texto positivo, neutro o negativo respectivamente. Este se guarda en una lista vacía creada con anterioridad y al final del bucle se añade la base de datos original como una nueva columna de nombre "polaridad".
3. **Separación en Base de Datos por Fecha:** Se separa la nueva base de datos con las polaridades en múltiples bases de datos más pequeñas según el formato de fecha día-mes-año.

4. **Aplicación de Ecuación de Índice:** Se crear y se llama un método auxiliar que aplica a cada sub base de datos la ecuación del índice, resultando en un valor único de polaridad promedio diario por cada base de datos.
5. **Reagrupación de la Base de Datos:** Se reagrupa la base de datos.
6. **Exportación de Resultados:** Se retorna la base de datos resultante en formato .csv.

El proceso mencionado anteriormente llevó alrededor de 24 horas, principalmente debido al tiempo requerido para obtener el sentimiento asociado con cada tweet. Para acelerar este proceso, se ejecutó localmente en la computadora del tesista y también se aprovecharon los servicios de Google Colaboratory.

Fecha	Indice	replyCount	likeCount	retweetCount
25-04-2017	-0.72	1	2	0
26-04-2017	-0.35	0	1	0
27-04-2017	-0.56	1	0	0
28-04-2017	-0.48	4	1	0
29-04-2017	-0.31	3	4	0
30-04-2017	-0.48	0	1	0
01-05-2017	-0.22	4	7	12
02-05-2017	-0.36	5	2	0
03-05-2017	-0.04	0	7	1
04-05-2017	-0.13	3	3	0
05-05-2017	-0.62	11	8	1
06-05-2017	-0.50	2	3	2
07-05-2017	0.00	0	0	0
08-05-2017	-0.44	3	1	2
09-05-2017	-0.23	0	1	0
10-05-2017	-0.28	2	2	1
11-05-2017	-0.35	3	10	71
12-05-2017	-0.58	4	5	29
13-05-2017	-0.36	3	3	1
14-05-2017	-0.48	2	5	10
15-05-2017	-0.48	0	3	3
16-05-2017	-0.51	3	1	1
17-05-2017	-0.44	2	1	0
18-05-2017	-0.32	0	1	0
19-05-2017	-0.63	4	1	0
20-05-2017	-0.56	0	0	0
21-05-2017	-0.76	1	0	1

Tabla 5: Ejemplo de Índice Base. (Método T)

3.1.3 Análisis Exploratorio

A forma de tener una comprensión inicial de la naturaleza de los índices calculados se efectúa un análisis exploratorio de los datos. Este dará una vista general del nivel de opinión de los clientes bancarios en ambos métodos al visualizar el valor promedio, máximo y mínimo de Índice Base. Se estudia solo Índice Base como representante de todas las variedades de Índice, anticipándose que el comportamiento de los valores de promedio, máximo y mínimo serán idénticos en los 5 casos. Esta información permitirá potencialmente detectar anomalías y/o servir de apoyo para análisis posteriores.

Los resultados se obtienen a partir del programa enfocado en visualizaciones de data por excelencia de Microsoft, PowerBI. Se adjunta en Anexo A las visualizaciones correspondientes por banco.

Método R

Índice Base			
Banco	Promedio	Máximo	Mínimo
A	8,81	21,43	2,27
B	59,92	245,25	13,18
C	11,03	25,88	3,14
D	21,96	61,88	3,63
E	16,26	56,33	0
F	2,86	14,72	0
G	17,80	57,09	5,64
H	12,21	56,50	1,02

Tabla 6: Estadísticas Índice Base. (Método R)

Lo primero que se observa es como la magnitud del Índice Base a lo largo de los 8 bancos oscila fuertemente. En términos de promedio este puede ir de un valor relativamente pequeño de 2,86 en Banco F hasta un 59,92 en Banco B, siendo este segundo valor 20 veces más que el primero. Como máximo índice Base puede ir de 14,72 en Banco F hasta 245,25 hasta Banco B o 16 veces más. En mínimo de 0 en Banco E y F hasta 13,18 en Banco B. El cálculo de los índices por Método R penaliza el valor por el número de clientes de banco, pero pareciera que esta no logra normalizar el nivel de reclamos según el tamaño de los bancos. En caso de que el método de normalización haya sido el correcto los resultados insinuarían que por algún motivo existe una diferencia considerable e incluso extrema entre la opinión de los clientes hacia los bancos muestra.

Los cambios de magnitud dentro de los mismos bancos también oscilan fuertemente. Ejemplos como el Banco E yendo de 0 a 56,33 o incluso Banco B de 13,18 a 245,25, podrían evidenciar la existencia de eventos de gran pérdida operacional que hayan afectado

temporalmente la opinión de los bancos, aunque también podría indicar una tendencia al empeoramiento de la opinión o aumento de riesgo reputacional a lo largo del tiempo.

Método T

Índice Base			
Banco	Promedio	Máximo	Mínimo
A	-0,33	0	-0,59
B	-0,49	-0,3	-0,66
C	-0,36	0,01	-0,61
D	-0,45	-0,14	-0,71
E	-0,42	0,14	-0,76
F	-0,30	0,03	-0,71
G	-0,46	-0,1	-0,65
H	-0,13	0	-0,53

Tabla 7: Estadísticas Índice Base. (Método T)

En este caso se evidencia de forma inmediata una oscilación menos pronunciada entre banco y banco. Siendo en parte esto efecto de cotas inferiores y superiores mejor demarcados (los índices por Método van de -1 a 1), Índice Base se mueve en promedio similares de entre 0,3 y 0,5 para la mayoría de los bancos, con excepción de Banco H. Esto insinúa que, a diferencia de Método R, la diferencia de reputación banco a banco no es tan extrema.

Sin embargo, los cambios de magnitud dentro de los mismos bancos son considerables. Banco F por ejemplo parte un máximo de 0,03 y llega a un mínimo de -0,71. Como se puede apreciar en Anexo A, la tendencia universal del valor del índice de la muestra a lo largo del tiempo fue a una disminución constante, sin evidencia que sugiera una recuperación futura. Dado que la disminución es constante no se puede atribuir todo a un evento de gran pérdida operacional, pero quizá recientes contingencias que han afectado a Chile en los últimos años (estallido social, COVID-19), podrían explicar un sentimiento general de malestar o de negatividad expresado precisamente a través de los tweets analizados.

3.1.4 Minería de datos

Método R

La base de datos en .csv resultante del procesamiento de datos se convierte a formato .xlsx y se importa a Rstudio. En paralelo se importan además las bases de datos correspondientes a los valores de roa, roe y las ganancias mensuales reportados por los bancos. Las ganancias se calculan como la diferencia del precio de cierre, y solo es pertinente a 5 de ellos que están abiertos a la bolsa.

Como primer paso se importan las librerías openxlsx, lubridate, readxl, ggplot2, tidyr y dplyr. Estos serán necesarias para llamar los métodos necesarios para las regresiones.

Se formatean correctamente y limpian las 3 bases de datos importadas. Se hacen conversiones al formato de fecha en los 3 casos. Cuando corresponda, se rellenan los valores faltantes de ganancia con el promedio entre la fecha anterior y posterior. Se llenan los valores faltantes de índice con un valor de 0 para denotar polaridad neutra en días donde no se hayan emitidos tweets para determinado banco. Roe y Roa no tienen datos faltantes así que no se modifican.

Las bases de datos parcheadas ahora se unen por fecha mediante el método merge de R. Esta base de datos fusionada contiene las columnas Año, Mes, Ganancia, índice, Roa y Roe.

Con la base de datos fusionada, se calculan y se añaden nuevas columnas con los valores del índice promedio y retardado para 3 y 6 meses. Se calcula y añade también la columna de nombre δ para la diferencia entre el máximo y mínimo móvil para 12 meses o 1 año. Se añade una columna de nombre Banco que tiene el nombre del banco asociado a cada tupla para usarse posteriormente en las regresiones de panel.

Se exporta la base de datos final en .xlsx para usarse nuevamente en Rstudio en otro proyecto.

Método T

El procedimiento es igual al Método R, con la única diferencia es que se importa la base de datos correspondiente al índice por tweets en vez del índice por reclamos.

3.1.5 Regresiones lineales

Utilizando Rstudio, en un proyecto .rmd nuevo, se importa el archivo .xlsx resultante del proceso de minería de datos para cada banco. Utilizando el método nativo lm de Rstudio, se efectúan regresiones utilizando alternadamente el índice base, índice promedio acumulado 3 meses, índice promedio acumulado 6 meses, índice rezagado 3 meses, índice rezagado 3 meses como variable dependiente y Ganancia (en casos aplicables), ROA y ROE como variable independiente. Los coeficientes de regresión estimados son guardados en un archivo Excel en la forma de una tabla con todas las distintas combinaciones posibles.

A continuación se presentarán los resultados, donde los coeficientes en verde representan magnitudes positivas, rojo magnitudes negativas y *** denota significancia al nivel de 0.001, ** al nivel de 0.01, * a nivel del 0.05 y . a nivel del 0.1. Para efecto de este trabajo se considerará un nivel de significancia del 0.05.

Para evitar análisis redundantes en las próximas secciones se aprovechará de descartar o conservar las combinaciones de índice con variables financieras caso a caso de acuerdo con los resultados obtenidos.

Método R

Correlaciones

Banco	Ganancia				
	base	prom-3	prom-6	rez-3	rez-6
A	-3,77x10 ⁻⁰⁵	-7,45x10 ⁻⁰⁵	-6,83x10 ⁻⁰⁵	-6,72x10 ⁻⁰⁵	3,35x10 ⁻⁰⁶
C	1,22x10 ⁻⁰⁴	9,34x10 ⁻⁰⁵	1,39x10 ⁻⁰⁴	6,26x10 ⁻⁰⁵	9,05x10 ⁻⁰⁵
E	2,33x10 ⁻⁰⁵	3,98x10 ⁻⁰⁶	-1,35x10 ⁻⁰⁶	-6,86x10 ⁻⁰⁶	-1,28x10 ⁻⁰⁵
G	5,02x10 ⁻⁰⁵	-4,07x10 ⁻⁰⁵	-5,83x10 ⁻⁰⁵	-3,14x10 ⁻⁰⁵	-4,88x10 ⁻⁰⁵
H	-1,22x10 ⁻⁰⁴	-7,64x10 ⁻⁰⁵	1,45x10 ⁻⁰⁴ *	-6,33x10 ⁻⁰⁵	1,11x10 ⁻⁰⁴ *

Tabla 8: Correlaciones individuales según Ganancia. (Método R)

Ninguna de las 5 distintas versiones del índice señalan efecto significativo de los índices en la variable Ganancia, con excepción del banco H en el caso prom-6 y lag-6. Los coeficientes además son extremadamente pequeños, no superando la orden de magnitud de la milésima, lo que da a entender que de existir algún efecto bien podría ser despreciable.

Para ningún banco en ningún caso existe evidencia para rechazar H0 para Test R.GAN. Se **descarta** Ganancia para el Método R.

ROA					
Banco	base	prom-3	prom-6	rez-3	rez-6
A	$-2,37 \times 10^{-04}$ *	$-2,97 \times 10^{-04}$ **	$-2,79 \times 10^{-04}$ *	$-1,98 \times 10^{-04}$ *	$-8,12 \times 10^{-05}$
B	$-1,29 \times 10^{-05}$.	$-1,14 \times 10^{-05}$	$-8,28 \times 10^{-06}$	$-7,67 \times 10^{-06}$	$1,23 \times 10^{-06}$
C	$1,22 \times 10^{-05}$	$6,37 \times 10^{-05}$	$1,44 \times 10^{-04}$	$6,41 \times 10^{-05}$	$9,78 \times 10^{-05}$.
D	$-9,37 \times 10^{-06}$	-	$3,31 \times 10^{-05}$	$7,15 \times 10^{-06}$	$8,14 \times 10^{-05}$.
E	$-1,02 \times 10^{-04}$	$-8,49 \times 10^{-05}$	$-1,20 \times 10^{-04}$	$-6,81 \times 10^{-05}$	$-1,17 \times 10^{-04}$
F	-	-	-	-	-
	$2,04 \times 10^{-03}$ ***	$2,89 \times 10^{-03}$ ***	$3,06 \times 10^{-03}$ ***	$2,40 \times 10^{-03}$ ***	$2,36 \times 10^{-03}$ ***
G	$-7,99 \times 10^{-05}$ **	-	$-1,06 \times 10^{-04}$ **	$-7,68 \times 10^{-05}$ **	$-4,67 \times 10^{-05}$.
H	$-1,83 \times 10^{-05}$	$-3,45 \times 10^{-05}$	-	$-3,24 \times 10^{-05}$	-
		$8,23 \times 10^{-05}$ ***		$5,91 \times 10^{-05}$ ***	

Tabla 9: Correlaciones individuales según ROA. (Método R)

Los resultados muestran efecto significativo principalmente en los bancos A, F, G. Los coeficientes en su mayoría son negativos. Banco B, C y D toman valores positivos, pero ninguno logra evidenciar correlación.

Si bien la mayoría de los resultados rechazan H0 para Test R.ROA, las magnitudes de los coeficientes son nuevamente extremadamente bajas, por lo que se decide **descartar** ROA para el Método R.

ROE					
Banco	base	prom-3	prom-6	rez-3	rez-6
A	$-1,65 \times 10^{-03}$	$-1,82 \times 10^{-03}$	$-1,36 \times 10^{-03}$	$-1,10 \times 10^{-03}$	$1,57 \times 10^{-04}$
B	$-2,08 \times 10^{-04}$	$-1,45 \times 10^{-04}$	$-5,01 \times 10^{-05}$	$-8,65 \times 10^{-05}$	$1,37 \times 10^{-04}$
C	$-1,17 \times 10^{-03}$.	$-1,94 \times 10^{-03}$ *	$-1,89 \times 10^{-03}$.	$-6,19 \times 10^{-04}$	$-1,47 \times 10^{-04}$
D	$2,32 \times 10^{-03}$ ***	$2,81 \times 10^{-03}$ ***	$3,33 \times 10^{-03}$ ***	$2,13 \times 10^{-03}$ ***	$2,28 \times 10^{-03}$ ***
E	$-1,67 \times 10^{-03}$	$-1,42 \times 10^{-03}$	$-1,96 \times 10^{-03}$	$-1,13 \times 10^{-03}$	$-1,84 \times 10^{-03}$
F	-	-	-	-	-
	$7,37 \times 10^{-03}$ ***	$1,09 \times 10^{-02}$ ***	$1,11 \times 10^{-02}$ ***	$9,34 \times 10^{-03}$ ***	$-7,93 \times 10^{-03}$ **
G	$2,43 \times 10^{-04}$	$6,63 \times 10^{-04}$	$1,30 \times 10^{-03}$ **	$6,57 \times 10^{-04}$.	$1,28 \times 10^{-03}$ ***
H	$2,51 \times 10^{-05}$	$-1,72 \times 10^{-04}$	$-8,60 \times 10^{-04}$ **	$-2,49 \times 10^{-04}$	$-7,32 \times 10^{-04}$ **

Tabla 10: Correlaciones individuales según ROE. (Método R)

Los coeficientes no muestran ningún patrón reconocible. Banco A, B, C, E y F tienen coeficientes principalmente negativos, pero solo F rechaza H0 para el Test R.ROE en las 5 versiones del índice. De forma inusual todas las versiones del índice presentan efecto significativo en banco D, pero con coeficiente positivo.

Dicho lo anterior, el valor de los coeficientes no es tan bajo como en ROA, por lo que se **conserva** ROE para el Método R.

Método T

Correlaciones

Banco	Ganancia				
	base	prom-3	prom-6	rez-3	rez-6
A	-2,87x10⁻⁰⁴	2,14x10 ⁻⁰³	2,83x10 ⁻⁰³	1,64x10 ⁻⁰³	3,02x10 ⁰³
C	1,05x10 ⁻⁰³	4,27x10 ⁻⁰⁴	7,72x10 ⁻⁰⁴	4,83x10 ⁻⁰⁴	2,28x10 ⁰³
E	-4,44x10⁻⁰⁴	-3,05x10⁻⁰³	-1,705x10⁻⁰³	-2,30x10⁻⁰³	-7,23x10⁰⁵
G	2,68x10 ⁻⁰³	3,77x10 ⁻⁰³	3,42x10 ⁻⁰³	2,80x10 ⁻⁰³	4,24x10 ⁰⁴
H	6,95x10 ^{-03*}	6,80x10 ⁻⁰³	4,72x10 ⁻⁰³	1,01x10 ^{02**}	1,20x10 ⁰³

Tabla 11: Correlaciones individuales según Ganancia. (Método T)

Los resultados obtenidos sugieren (salvo el banco H) que ninguna de las 5 distintas versiones del índice presentan efecto significativo en la variable Ganancia. Los coeficientes son en su mayoría positivos para A,C,G y H y completamente negativos para E, pero por falta de evidencia es difícil afirmar si es que esta relación es real y no debido al azar.

Los resultados a lo mucho permite a Banco H rechazar H0 para el Test T.GAN, y solo para algunas versiones del índice, por lo que se decide **descartar** Ganancia como variable de estudio para riesgo reputacional para el Método T.

ROA					
Banco	base	prom-3	prom-6	rez-3	rez-6
A	-4,30x10⁻⁰³	-7,86 x10⁻⁰³*	-9,27x10⁰³*	-7,17x10⁻⁰³*	-4,97x10⁻⁰³
B	-8,98x10⁻⁰³*	-1,19x10⁰²**	-1,41x10⁰²**	-1,08x10⁰²**	1,42x10⁰²***
C	7,85x10⁰³***	8,19x10⁰³***	8,32x10⁰³**	5,73x10⁰³**	5,90x10⁻⁰³**
D	1,89x10⁰²***	1,83x10⁰²***	1,67x10⁰²***	1,59x10⁰²***	1,25x10⁻⁰²**
E	8,30x10⁻⁰³	7,76 x10⁻⁰³	5,77x10⁻⁰³	4,50 x10⁻⁰³	3,87x10⁻⁰³
F	2,36x10⁰²***	2,37x10⁰²***	2,32x10⁰²***	2,15x10⁰²***	1,98x10⁰²***
G	8,37x10⁰³***	9,89x10⁰³***	1,14x10⁰²***	8,33x10⁰³***	7,57x10⁰³***
H	4,81 x10⁻⁰⁴	7,47x10⁻⁰⁴	3,23x10⁻⁰³	4,83x10⁻⁰⁴	1,57 x10⁻⁰³

Tabla 12: Correlaciones individuales según ROA. (Método T)

Contrastando fuertemente con la situación anterior, los resultados para ROA sugieren que en la gran mayoría de los casos las 5 distintas versiones del índice presentan efecto significativo en el ROA a distintos niveles de significancia. Los bancos C,D,F y G rechazan H0 para el Test T.ROA, pero los bancos A,B, E y H no.

Para A y B el coeficiente es en su mayoría estadísticamente significativo, pero siempre negativo. Esto sugiere que un aumento del índice o una mejor opinión de los clientes a los bancos A y B estaría acompañada de una reducción del ROA, lo que resulta contraintuitivo.

Si bien la mitad de los casos no rechazan H0, se decide **conservar** ROA para el Método T, considerando que algunos bancos si rechazan H0 y que la magnitud de los coeficientes es relativamente alta.

ROE					
Banco	base	prom-3	prom-6	rez-3	rez-6
A	-8,09x10⁰²*	-1,37x10⁰¹**	1,61x10⁰¹***	-1,22x10⁰¹**	-9,61x10⁻⁰²*
B	-2,30x10⁰¹**	-2,95x10⁰¹**	3,55x10⁰¹***	-2,70x10⁰¹**	3,55x10⁰¹***
C	4,63x10⁻⁰²	3,96x10⁻⁰²	2,47x10⁻⁰²	2,22x10⁻⁰²	1,26x10⁻⁰²
D	3,55x10⁰¹***	3,58x10⁰¹***	3,47x10⁰¹***	3,30x10⁰¹***	2,92x10⁰¹***
E	1,06x10⁻⁰¹	1,13x10⁻⁰¹	9,39x10⁻⁰²	6,90x10⁻⁰²	6,33x10⁰²
F	5,96x10⁻⁰²*	6,11x10⁻⁰²*	5,78x10⁻⁰²*	5,20x10⁻⁰²*	4,52x10⁻⁰².
G	-6,12x10⁻⁰²	-7,36x10⁻⁰²*	-9,29x10⁻⁰²*	-6,80 x10⁻⁰²*	-7,03 x10⁻⁰²*
H	1,05 x10⁻⁰²	2,92 x10⁻⁰³	1,39 x10⁻⁰²	-1,02 x10⁻⁰²	-3,01 x10⁻⁰²

Tabla 13: Correlaciones individuales según ROE. (Método T)

Los resultados de las correlaciones indican correlación en la mayoría de los casos, pero esta vez solo banco D y F logran rechazar H0 para el Test T.ROE correspondiente. Banco G sugiere que un aumento del índice por Twitter resultaría en una disminución de ROE.

Aunque no muchos resultados rechazan H0, La magnitud de las correlaciones son relativamente altas, por lo que decide una vez más **conservar**.

3.1.6 Regresiones lineales por datos de panel

Para obtener las regresiones lineales por datos de panel se aplicaron los siguientes pasos tanto para al Método R como el Método T.

Primero se creó un nuevo proyecto .rmd, importándose las 8 base de datos correspondientes a los 8 bancos. Se combinaron estas 8 bases de datos en una super tabla, aprovechándose de calcular e imprimir el δ asociado para uso posterior. Se exportaron las bases de datos a .csv para procesarse con Python. En Python se importó la función PanelOLS de la librería linearmodels.panel y se aplicó este método a la super tabla previa. Como parámetros se le indicó al programa la fecha de inicio de abril 2015 y fin de marzo 2023. Se creó un índice múltiple consistente de la columna Banco y Fecha. Se definió como variable independiente el índice y se fue alternando entre ROE y ROA para la variable dependiente. Se ajustó el modelo de efectos fijos para las variables usando el método PanelOLS. Se obtuvieron los resultados (ver Anexo B).

Los β de los sistemas son los siguientes:

Método R

ROE	Coficiente	t	p
Indicebase	$2,00 \times 10^{-04}$	1,27	0,200
Indicepro3	$4,00 \times 10^{-04}$	2,50	0,010
Indicepro6	$5,00 \times 10^{-04}$	3,18	0,002
Indicerez3	$4,00 \times 10^{-04}$	2,73	0,010
Indicerez6	$5,00 \times 10^{-04}$	3,76	0,002

Tabla 14: Correlaciones por panel según ROE. (Método R)

Los resultados obtenidos sugieren que 4 de los 5 índices; Indicepro3, Indicepro6, Indicerez3 e Indicerez6, tienen impacto significativo en la variable financiera (ROE). Los datos proporcionan evidencian de que la mayoría de los índices influyen de manera significativa en el ROE.

Los valores de los coeficientes en cada caso son relativamente pequeños. Los valores de t se mueven más allá de 1 desviación estándar, alcanzado hasta 3 de ellas en Indicepro6, indicando que el coeficiente estimado se aleja significativamente de la media muestral

Los valores de p son bajos, excepto por el Indicebase. Estos valores p se mueven por debajo de un valor de significancia de 0,05, sugiriendo que existe en su mayoría significancia estadística.

Como la evidencia sugiere significancia, no se puede atribuir la correlación a simple azar, pero al mismo tiempo como las correlaciones son positivas, tampoco se puede rechazar H0 para Test T.ROE. Esto hace pensar que un aumento en el nivel de reclamos de alguna forma produce un aumento del ROE al sistema, lo que absoluto contraintuitivo y contradictorio al objetivo final de este trabajo. Es posible que esto también haya sido por una falla en el diseño índice por parte del Método R, quizá habiéndose fallado en representar la opinión real de los clientes de los bancos.

Método T

ROA	Coeficiente	t	p
Indicebase	0,0118	5,55	0,00
Indicepro3	0,0129	5,10	0,00
Indicepro6	0,0121	4,28	0,00
Indicerez3	0,0083	3,63	0,00
Indicerez6	0,0071	2,99	0,00

Tabla 15: Correlaciones por panel según ROA. (Método T)

Los resultados obtenidos sugieren que cada uno de los índices tiene impacto significativo en el ROA. Los datos disponibles presentan evidencia de que todos de los índices tienen influencia en el ROA.

El valor de los coeficientes para cada uno de los casos es de magnitud apreciable, superando el coeficiente estimado en Indicebase e Indicepro3 hasta 5 desviaciones estándares por sobre la media muestral

Los valores p son funcionalmente nulos para todos los casos, indicando que existe claramente evidencia de significancia estadística.

ROE	Coeficiente	t	p
Indicebase	0,0896	3,97	0,00
Indicepro3	0,1047	3,94	0,00
Indicepro6	0,0943	3,20	0,00
Indicerez3	0,0663	2,79	0,00

Indicerez6	0,0533	2,19	0,00
------------	--------	------	------

Tabla 16: Correlaciones por panel según ROE. (Método T)

Los resultados obtenidos para ROE son muy similares al caso ROA. Estos nuevamente sugieren un impacto significativo en la variable financiera estudiada, de que todos los índices tienen influencia en el ROE.

El valor de los coeficientes es apreciable, aunque en promedio los valores de t son menores que para el ROA. Índicebase alcanza por ejemplo hasta unas 4 desviaciones estándar sobre la media aproximadamente.

Los valores p nuevamente son funcionalmente nulos, indicando significancia estadística.

En resumen, los resultados indican que existe una relación positiva y estadísticamente significativa entre las diferentes versiones del índice y las métricas financieras ROA y ROE. Cada incremento en el valor de los índices se asocia con aumentos correspondientes en los valores de ROA y ROE.

Esto contrasta fuertemente con los resultados obtenidos en el Método R, puesto que no solo se encontró evidencia de significancia para los indicadores con las dos variables financieras, sino que también el sentido de la relación es coherente en todos los casos con las hipótesis asociadas de los Test (Test T.ROA y Test T.ROE) y la intensidad del efecto es mucho mayor.

3.1.7 Requisito de capital sugerido

El cálculo de requisito de capital (y la data relevante) se expresará mediante Excel, al considerar la necesidad de entrecruzar información en formatos distintos

Recordando la Ecuación 8 y la Ecuación 9, para poder calcular requisito de capital se necesitan las variables de β , δ , APR y Activos o Patrimonio, dependiendo si se calcula requisito de capital respecto a ROA o ROE. Respectivamente, los β fueron calculados a través de la función `linearmodels.panel` de `PanelOLS` en Python, δ directamente de `Rstudio`. Patrimonio, Activos y APR de información pública obtenida dispuesta por la CMF (Anexo C), siendo estos cuantificados en **millones de pesos**. La información o parámetros anteriores se expresarán en dos tablas distintas, una con los β ROE según los 5 casos de Índice y otra con los Delta, APR, Activos y/o Patrimonio.

En cuanto a los resultados mismos, se descarta en ambos métodos calcular requisito de capital por Ganancia, según los valores no significativos de β Ganancia calculados en las regresiones lineales individuales. Similarmente se descarta calcular requisito de capital por ROA específicamente en el Método R según lo justificado en el capítulo 3.1.5. Se efectúa, si bien con valores de β ROE que no rechazan el H_0 para Test R.ROE, los cálculos de requisito de capital para ROE en el método R por haber evidencia de correlación significativa grupal en 4 de 5 casos.

Considerando las diferencias en la orientación entre los índices por método R y método T (los índices por método R denotan nivel de reclamos por lo que son iguales o mayores a 0, mientras los índices por método T denotan polaridad por lo que van de -1 a 1), el requisito de capital calculado por la Ecuación 8 o Ecuación 9 para el método R será **multiplicado por -1** a forma de hacer coherente la interpretación. De no hacer este ajuste, por ejemplo, método R sugerirá reducir requisito de capital en caso de que la evidencia indique que un mayor nivel de reclamos perjudica el desempeño bancario (BetaROE negativo), lo que sería ilógico.

Para ambos métodos son 5 tablas de requisito de capital las que se construyen, una para el Índice Base, Índice acumulado promedio de 3 y 6 meses e Índice retardado de 3 y 6 meses. Las filas corresponden a cada uno de los 8 bancos y los encabezados al requisito de capital por ROE o ROA cuando es aplicable.

Método R

	BetaROE
Indice_base	0,0002
Indice_pro_3	0,0003
Indice_pro_6	0,0004
Indice_rez_3	0,0003
Indice_rez_6	0,0004

Tabla 17: β ROE por Índice. (Método R)

Con un valor aproximado de β ROE de 0,0002 para el Índice Base y β ROE de 0,0003 para Índice_pro_3 e Índice_rez_3, el método R evidencia que el efecto de la opinión del público al desempeño bancario es de mayor intensidad como un efecto acumulativo (pro) o con cierto rezago. Adicionalmente con β ROE de 0,0004 para Índice_pro_6 y Índice_rez_6 el método R sugiere que el efecto de la opinión en el desempeño bancario también aumenta a un lapso mayor de tiempo

	Delta	Patrimonio (M CLP)	APR (M CLP)
A	16,48	4.792.573	36.643.744
B	223,34	3.060.494	32.348.716
C	19,37	4.873.608	48.942.325
D	53,67	984.465	5.086.894
E	45,58	3.416.717	27.253.695
F	10,99	208.575	1.276.719
G	48,11	4.128.339	38.564.165
H	52,92	3.257.026	31.299.006

Tabla 18: Parámetros por Banco. (Método R)

Como se puede observar los valores de δ varían fuertemente entre banco, lo que es coherente con lo observado anteriormente en la sección de Análisis Exploratorio, respecto a la disparidad en las subidas, bajadas y la magnitud misma del Índice Base en el tiempo para los distintos bancos. La proporción Patrimonio vs APR también resalta, con diferencias sustanciales al considerar que el Banco D es el máximo con un valor de 0,1935 (984.465/5.086.894) vs Banco B el mínimo con un valor de 0,0946 (3.060.494/32.348.716).

Cap Req. ROE (%APR)	Indice_base	Indice_pro_3	Indice_pro_6	Indice_rez_3	Indice_rez_6
A	-0,04%	-0,06%	-0,09%	-0,06%	-0,09%
B	-0,42%	-0,63%	-0,85%	-0,63%	-0,85%
C	-0,04%	-0,06%	-0,08%	-0,06%	-0,08%
D	-0,21%	-0,31%	-0,42%	-0,31%	-0,42%
E	-0,11%	-0,17%	-0,23%	-0,17%	-0,23%
F	-0,04%	-0,05%	-0,07%	-0,05%	-0,07%
G	-0,10%	-0,15%	-0,21%	-0,15%	-0,21%
H	-0,11%	-0,17%	-0,22%	-0,17%	-0,22%

Tabla 19: Capital Requisito ROE. (Método R)

Lo primero que salta a la vista en los %APR obtenidos es que son negativos, lo que indica que, según el Método R, se debería requerir un %APR menor como requisito de capital para cada una de las muestras de bancos, en cada uno de los casos posibles de índices posibles. Estos resultados nuevamente son contraintuitivos y por ello difíciles de interpretar, recordando que los β ROE calculados por las regresiones de panel indicaron que como sistema la muestra de bancos tiene un mejor desempeño a un nivel más alto de reclamos, lo que no rechaza H_0 Test R.ROE y que consecuentemente pierde su utilidad para efectos de sugerir un aumento de requisito de capital a los bancos para mitigar riesgo.

Dicho lo anterior, podría haber valor en términos cuantitativos al contrastarse con el Método T. En primer lugar, se nota una gran variabilidad en los resultados de requisito de capital. Este alcanza sus menores niveles en los bancos A,C y F, mientras alcanza sus mayores niveles en el Banco B. Mientras en los primeros tres llega a un mínimo de -0,04%, este también llega a un máximo de -0,85% Banco B. Coherente con lo visto en los β ROE, Índice_pro_6 e Índice_rez_6 tienen asociados los APR de mayor magnitud.

Método T

Se replica la metodología, siendo ahora relevante el β ROA.

	BetaROE	BetaROA
Indice_base	0,0896	0,0118
Indice_pro_3	0,1045	0,0129
Indice_pro_6	0,0943	0,0121
Indice_rez_3	0,0668	0,0083
Indice_rez_6	0,0533	0,0071

Tabla 20: β ROE y β ROA por Índice. (Método T)

β ROE

Se observa gran variabilidad a para los 5 casos. Indice_pro_3 reporta el mayor β ROE, alcanzando un valor de 0,1045 considerando 3 meses, seguido por Indice_pro_6 con 0,0943 para 6 meses. Esto sugiere que el efecto de la opinión de los clientes bancarios en Twitter es de mayor intensidad como un efecto acumulado al relativamente corto plazo (3 meses vs 6 meses). Esto contrasta fuertemente con lo calculado con Método R, que indicaba que el β ROE tenía mayor correlación a los 6 meses y que no había diferencia entre el efecto acumulado y rezagado.

β ROA

Nuevamente el índice por efecto acumulado a los 3 meses reporta resultados de mayor magnitud, con una variabilidad similar (aquí el máximo es solo 1,66 (0,0118/0,0071) veces superior al mínimo, comparado a los 1,96 (0,1045/0,0533) del caso β ROE). El mayor β ROA se encuentra en el Indice_pro3 con 0,0129, seguido cercanamente por el Indice_pro_6 con 0,0121 e Indice_base con 0,0118.

	Delta	Patrimonio (M CLP)	Activos (M CLP)	APR (M CLP)
A	0,48	4.792.573	55.384.781	36.643.744
B	0,22	3.060.494	57.463.222	32.348.716
C	0,51	4.873.608	78.504.757	48.942.325
D	0,35	984.465	7.048.685	5.086.894
E	0,87	3.416.717	40.784.929	27.253.695
F	0,48	208.575	1.265.474	1.276.719
G	0,39	4.128.339	69.486.857	38.564.165
H	0,48	3.257.026	45.548.014	31.299.006

Tabla 21: Parámetros por Banco. (Método T)

Como se puede observar los valores de δ varían sustancialmente entre banco, pero proporcionalmente esta variabilidad es mucho menos extrema que la obtenida para el Método R, con una proporción entre el máximo y mínimo de 3,95 (0,87/0,22) vs los más de 20,3 (223,34/10,99) del caso anterior. Estos resultados son nuevamente coherentes con lo observado en el Análisis exploratorio. La proporción Activos vs APR alcanza un mínimo de 0,9912 (1.265.474/1.276.719) para el Banco F y un máximo en Banco G con 1,800 (69.486.857/38.564.165).

Cap Req. ROE (%APR)	Indice_base	Indice_pro_3	Indice_pro_6	Indice_rez_3	Indice_rez_6
A	0,56%	0,65%	0,59%	0,42%	0,33%
B	0,19%	0,22%	0,20%	0,14%	0,11%
C	0,45%	0,53%	0,48%	0,34%	0,27%
D	0,61%	0,72%	0,65%	0,46%	0,36%
E	0,98%	1,14%	1,03%	0,73%	0,58%
F	0,70%	0,82%	0,74%	0,52%	0,42%
G	0,37%	0,44%	0,39%	0,28%	0,22%
H	0,45%	0,52%	0,47%	0,33%	0,26%

Tabla 22: Capital Requisito ROE. (Método T)

β ROE

A diferencia del requisito de capital calculado por β ROE en el Método R, los %APR por β ROE en Método T si toman valores positivos, significando que se debería requerir un %APR mayor como requisito de capital para la muestra de bancos, lo que se ajusta bien al objetivo de este trabajo. En este caso los %APR de capital requisito son asimismo en promedio de mayor magnitud y fluctúan de forma menos extrema. Por ejemplo, en Método R el %APR de Indice_base aumenta más de 10 veces de mínimo a máximo (-0,42%/-0,04%), mientras en Método T el %APR aumenta unas 5 veces (0,98%/0,19%).

Cuantitativamente los mayores niveles de %APR corresponden al Banco E, mientras los menores están asociados a Banco B. No se repiten valores entre bancos, y los %APR calculados son mayores en Indice_pro_3 y menores en Indice_rez_6, lo que condice con lo hablado de β ROE.

Cap Req. ROA (%APR)	Indice_base	Indice_pro_3	Indice_pro_6	Indice_rez_3	Indice_rez_6
A	0,85%	0,93%	0,87%	0,60%	0,51%
B	0,47%	0,51%	0,48%	0,33%	0,28%
C	0,96%	1,05%	0,99%	0,68%	0,58%
D	0,58%	0,63%	0,59%	0,41%	0,35%
E	1,53%	1,68%	1,57%	1,08%	0,92%
F	0,56%	0,61%	0,58%	0,39%	0,34%
G	0,83%	0,90%	0,85%	0,58%	0,50%
H	0,82%	0,90%	0,84%	0,58%	0,49%

Tabla 23: Capital Requisito ROA (Método T)

β ROA

Los %APR de requisito de capital por β ROA toman valores positivos, sin embargo, los valores son muy distintos a los calculados por β ROE tanto para Método R como Método T. Los %APR calculados son en promedio mayores a los de β ROE, con un mínimo de 0,47% en Banco B y un máximo de 1,53% en Banco E vs 0,19% y 0,98% en β ROE.

Resalta como tanto los %APR calculados por β ROE y β ROA en el Método T encuentran su magnitud mínima (absoluto) en Banco B, pero Método R encuentro su máximo (absoluto) en este mismo banco. Aunque es complicado identificar la razón detrás de esta diferencia en comportamiento, esto subraya las discrepancias en las opiniones expresadas por los clientes a través de Reclamos.cl y Twitter. Quizá la página web preferida en que los clientes de los bancos expresan insatisfacción varía caso a caso, resultando en que por ejemplo si los clientes de Banco B prefieran utilizar Reclamos.cl, esta página termine sobrerrepresentando negatividad y Twitter positividad. Alternativamente la gran diferencia en los %APR es una simple diferencia en factores como el patrimonio/activos o el delta de estrés.

Para efectos de la evaluación de capital los resultados parecieran ser coherentes para Método T, no así Método R. Como ya se mencionó, Método R indica que se debería disminuir la cantidad %APR exigido, lo que contradice la lógica de que exista riesgo reputacional en primer lugar y que por lo mismo da a entender que Método R no es útil para la evaluación de capital. Método T indica que si se debiese aumentar la cantidad de %APR asociado a los riesgos que están expuestos los bancos muestra, con resultados de magnitud coherentes que van desde un mínimo de 0,11% (β ROE, Banco B, rez6) y un máximo de 1,68% APR (β ROA, Banco E, pro3). Con lo anterior el capital regulatorio exigido por la LGB a los bancos de la muestra correspondería al 8% base del T1 y T2, más el requerimiento de pilar 2 sobre riesgos, yendo desde un **8,11% APR + %APR** de otros riesgos en Banco B hasta un **9,68% APR + %APR** de otros riesgos en Banco E.

4 Conclusiones

A forma de implementar Basilea III en Chile, la CMF inicio el proceso normativo relevante en enero de 2019 con la publicación de la Ley N° 21.130, que tiene por objetivo alinear la LGB con los estándares internacionales. Como parte de la Ley N° 21.130 la CMF inició el proceso normativo para implementar los llamados 3 pilares de Basilea III, siendo las exigencias del segundo de estos el motivo por el cual se incorporó el Capítulo 21-13 a la RAN. El Capítulo 21-13 indica que los bancos deben chilenos debe contar con metodologías para medir riesgos materiales que deben informar en el llamado IAPE, y que la CMF podrá exigir requerimiento de patrimonio efectivos adicionales en caso de estimar riesgos materiales no suficientemente cubiertos. Uno de los riesgos considerados para medir es riesgo reputacional, sin embargo, este a la fecha no cuenta con una metodología pública estandarizada para su medición en Chile. En este contexto la presente memoria diseño una metodología para la medición de riesgo reputacional, conforme la data disponible en redes sociales (Twitter) y la página más prominente de reclamos en Chile (Reclamos.cl), para luego calcular los requisitos de capital correspondientes.

En el análisis exploratorio los resultados de Método R sugieren que el nivel de opinión de los clientes varia de forma extrema banco a banco, mientras Método T sugiere que varían modestamente. Si bien los índices son fundamentalmente distintos y no es posible una comparación perfecta, en Método R se observó como el nivel de reclamos promedio podía ser hasta 20 veces mayor (Banco B vs Banco F), mientras en Método T la opinión fluctuó de forma conservadora a lo largo de un nivel de -0,3 y -0,4 (con excepción de Banco H). Este fenómeno se podría deber a que el índice de Método R no fue correctamente normalizado según la “importancia” de los bancos. Este fue resultado de la división del numero de reclamos por período en cantidad de clientes, pudiendo haberse utilizado otra métrica en vez de cantidad de clientes para medir la “importancia” del banco.

Los resultados del análisis exploratorio para Método R y T también sugieren una caída constante y fuerte del nivel de opinión del público hacía los bancos, sobre todo en los últimos 5 años (detallado en el Anexo A). Esto quizá podría explicarse dadas las recientes contingencias a nivel nacional e internacional, como el estallido social y COVID-19, que han provocado un sentimiento generalizado de malestar y de desconfianza del público a instituciones como los bancos.

Referente a las regresiones lineales, tanto el Método R como el Método T no lograron rechazar la hipótesis nula (H_0) en R.GAN y T.GAN para casi todas las combinaciones de índices y bancos correspondientes. Esto hace suponer que no existe relación alguna entre el precio de cierre de los bancos y su nivel de riesgo.

Las regresiones de Método R y Método T mostraron en su mayoría significancia estadística para las variables ROE y ROA, sin embargo, fueron inconsistentes en términos magnitudes de correlación muy baja (β menores a una milésima) y de sentido (algunos bancos indicaban correlación directa y otra inversa). Por un lado, los bajos niveles de correlación podrían deberse a que la valoración de una empresa depende de muchos factores, y el riesgo reputacional representa solo una pequeña parte de ellos. El problema de sentido es más difícil de diagnosticar, ya no hubo un patrón claro respecto a que combinación de

Banco o índice resulta en una correlación positiva o negativa, pero se podría suponer que hay alguna falla en el diseño de los índices.

Por el lado de las regresiones lineales por panel, el Método R señaló existir una pequeña correlación positiva entre el ROE y los índices acumulados y rezagados. Esto no permite rechazar H_0 para Test R.ROE, implicando que de alguna forma un aumento en el nivel de reclamos del sistema resulta en una mejora de la valoración de los bancos. Esto nuevamente es difícil de explicar, dada que como se observó en las regresiones lineales por banco no hubo un patrón o comportamiento reconocible, por lo que se supondrá nuevamente que hubo una falla en el diseño de Método R. En contraste Método T si valida los Test T.ROA y Test.ROE, lo que hace pensar que este método podría ser de utilidad para efectos de calcular requisitos de capital.

Los %APR calculados para Método R son todos negativos (como efecto del sentido de los β) y tienen valores muy dispares banco a banco. Los %APR calculados para Método T son positivos y la diferencia es menos marcado. Método R insinúa que los reguladores deberían reducir el capital regulatorio exigido en función del riesgo reputacional, mientras Método T insinúa que este debería aumentar.

En Método T los %APR calculados son mayores en β ROA que β ROE, pero el contraste es aún mayor al comparar %APR por métodos; los %APR alcanzan su máximos y mínimos absolutos en bancos distintos al punto de contradicción. Esto hace pensar que quizá en promedio los clientes de cada uno de los bancos tienen preferencia por usar Twitter o Reclamos.cl para expresar insatisfacción.

Considerando los resultados favorables de Método T se propone seleccionar los %APR calculados por esta metodología, considerando activos (β ROA) o patrimonio (β ROE). De acuerdo a los lineamientos del pilar 2 de Basilea III tomados por la LGB, para cubrir el riesgo reputacional de la muestra de bancos se necesitará de entre un mínimo de capital regulatorio de 0,11% APR hasta un máximo de 1,68% APR. Esto se añadirá a los capitales T1, T2 y cualquier %APR adicional basado en otros riesgos materiales.

Posibles fallas

Como mencionado el Método R fue incapaz de dar resultados coherentes o útiles para los objetivos planteados. Esto podría ser atribuible a diversos factores, como podrían haber sido mala calidad de data, mal procesamiento de datos o la ecuación planteada. Método T dio comparativamente buenos resultados, que permitieron en última instancia medir requisito de capital, pero dio resultados inconsistentes en las regresiones lineales, fallando en algunos casos con los tests planteados. Esto hace pensar que la data suministrada pudo no haber sido la mejor, que se necesitó un proceso más profundo de limpieza de datos o que simplemente elegir otra fuente de datos pudo haber sido aún mejor.

Es posible también que la selección de variables financieras no haya sido la correcta para ambos métodos. Se trabajó riesgo reputacional asumiendo que la opinión de los clientes de los bancos es quienes determinan la valoración de la empresa, cuando se pudo haber considerado de alguna forma otro segmento, como los mismos accionistas. Adicionalmente puede ser que el efecto de una mala reputación se sienta más fuerte en variables como los costos de rotación, causado por ejemplo por los empleados que renuncian de un banco por la mala imagen que este tiene.

Trabajo futuro

Se podría volver a explorar el Método R implementando cambios en función de los factores descritos como los potenciales culpables de la falta de correlación significativa encontrada. Así mismo se podría mejorar la calidad de la data del Método T, haciendo hincapié en los resultados inusuales de las regresiones lineales, suponiendo que hay espacio de mejora.

Considerando que el modelo 1 trata con información solo de Twitter y el modelo 2 con solo información de Reclamos.cl, se podrían explorar fuentes de información alternativas para calcular riesgo reputacional. Queda abierta la posibilidad de desarrollar modelos adicionales que calculen riesgo reputacional a partir de artículos y/o noticias que cubran eventos relacionados a los bancos muestra. Alternativamente se podrían reimplementar o modificar los modelos preexistentes para probarlos en otras redes sociales como Facebook o Instagram.

Según lo visto en las correlaciones individuales y por panel, no pareciera existir uniformidad en el comportamiento de los bancos, lo que hace suponer que de haber tomado otra muestra o combinación de bancos pudo haber cambiado sustancialmente los coeficientes obtenidos para calcular requisito de capital. Por lo anterior se sugiere que en caso de querer implementar alguno de los métodos para efectos de regulación bancaria, esto sea considerando el universo completo de bancos.

5 Glosario

API - Application Programming Interface

APR – Activos Ponderados por Riesgo

AT1 – Capital Adicional Nivel 1

BERT - Bidirectional Encoder Representations from Transformers

CET1 – Capital Ordinario de Nivel 1

CMF – Comisión para el Mercado Financiero

LGB – Ley General de Bancos

NLP – Natural Language Processing

PAC - Proceso de autoevaluación de capital

RAN - Recopilación Actualizada de Normas

ROA - Return on Assets

ROE - Return on Equity

T1 – Capital Nivel 1

T2 – Capital Nivel 2

6 Bibliografía

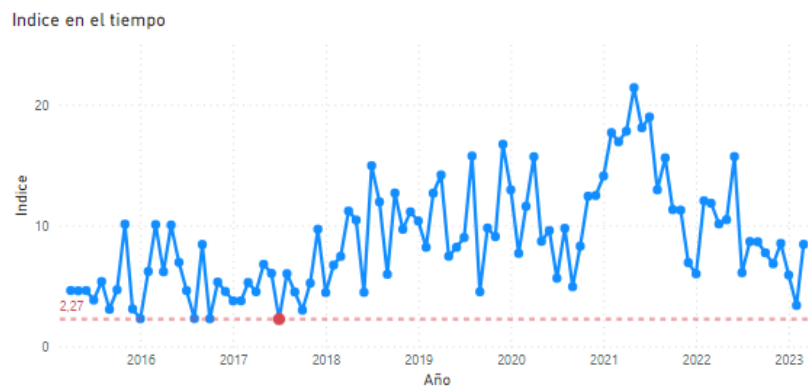
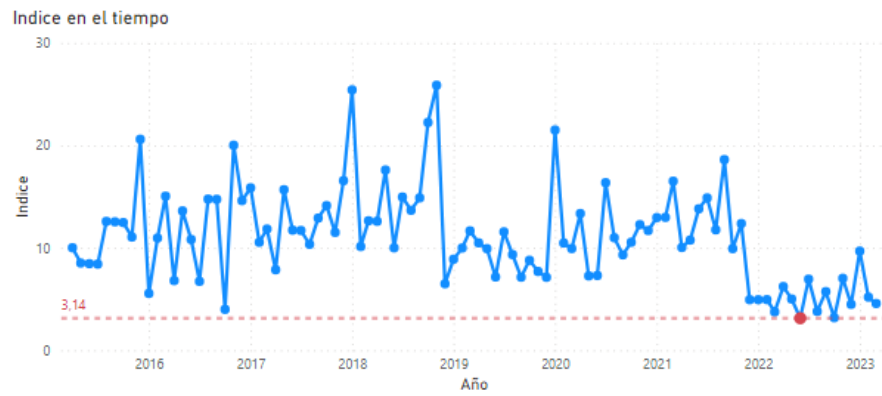
- Ashish Vaswani, N. S. (2017). *Attention Is All You Need*.
- Axel Michaelowa, K. M. (2020). *Catalysing private and public action for climate change mitigation: the World Bank's role in international carbon markets*.
- Branch. (2022). *Estadísticas de la situación digital de Chile en el 2021-2022*.
- Carlos Perales González, H. M.-O. (2017). *Sentiment analysis as reputational risk indicator*.
- Carlos Pulgar, N. S. (2020). *Criterios y directrices Generales para determinar Requerimientos patrimoniales Adicionales como resultado del Proceso de supervisión (pilar 2)*.
- CMF. (2020). *EVALUACION DE LA SUFICIENCIA DE PATRIMONIO EFECTIVO DE LOS BANCOS*. Obtenido de https://www.cmfchile.cl/portal/principal/613/articulos-29808_doc_pdf.pdf
- DataReportal. (2023). *Digital 2023: Chile*.
- David Adeabah, C. A. (14 de mayo de 2021). Obtenido de SSRN: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3844453
- Deutsche Bank. (2017). *Reputational risk management*. Obtenido de <https://www.db.com/cr/en/concrete-management-of-reputational-risks.htm>
- Ferreira, S. (noviembre de 2015). *Measuring reputational risk in the South African banking sector*.
- Francisco Guijarro, I. M.-C. (2019). *Liquidity Risk and Investors' Mood: Linking the Financial Market Liquidity to Sentiment Analysis through Twitter in the S&P500 Index*.
- Gabriela Aguilera, C. A. (Diciembre de 2020). *Implementación de Basilea III en Chile: evaluación de cierre*.
- Gabriele Ranco, D. A. (2015). *The Effects of Twitter Sentiment on Stock*.
- García-Vega, M. D.-G.-C.-R.-Z. (2020). *roceedings of the Iberian Languages Evaluation Forum (IberLEF 2020) Co-Located with 36th Conference of the Spanish Society for Natural Language Processing*. Malaga.
- Jason Perry, P. d. (30 de octubre de 2005). Obtenido de SSRN: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=861364
- Juan Manuel Pérez, J. C. (2021). *A Python Toolkit for Sentiment Analysis and SocialNLP task*.
- Klaus-Peter Wiedmann, C. J. (2001). Obtenido de https://link.springer.com/chapter/10.1007/978-3-8349-9164-5_16.
- Lucas Montesinos García, M. O. (agosto de 2014). *ANÁLISIS DE SENTIMIENTOS Y PREDICCIÓN DE EVENTOS EN TWITTER*.

- Mitic, P. (2018). REPUTATION RISK: MEASURED. *WITPress*, 171-180.
- Pérez, J. M. (2022). *RoBERTuito: A pre-trained language model for social media text in Spanish*. Marsella: European Language Resources Association.
- Rae. (23 de mayo de 2023). <https://dpej.rae.es/lema/riesgo-en-la-reputaci%C3%B3n>.
Obtenido de panhispánico: <https://dpej.rae.es/lema/riesgo-en-la-reputaci%C3%B3n>
- Richard B Carter, M. L. (2012). *Reputational Signals and Capital acquisition When insurance Companies Go Public*.
- Roland Gillet, G. H. (enero de 2010). Obtenido de ScienceDirect: <https://www.sciencedirect.com/science/article/abs/pii/S0378426609001873?via%3Dihub>
- Simon Zaby, M. P. (4 de julio de 2019). Obtenido de Safe journals: <https://journals.sagepub.com/doi/10.1177/2158244019861479>
- Terblanche, M. (diciembre de 2020). Financial Sentiment Analysis: an NLP approach towards reputation management.
- Thomas Fujiwara, K. M. (2023). *The Effect of Social Media on Elections: Evidence from the United States**.
- Verbanaz, S. A. (marzo de 2019). Requisitos de capital de la Ley que Moderniza la Legislación Bancaria .
- Walter, I. (2013). *The value of reputational capital and risk in Banking and finance*.
- Yerko Moena, C. V. (11 de marzo de 2019). Aplicación de un proceso de Análisis de sentimientos basado en aspectos para opiniones del portal reclamos.cl.
- Yinhan Liu, M. O. (2019). *RoBERTa: A Robustly Optimized BERT Pretraining Approach*.

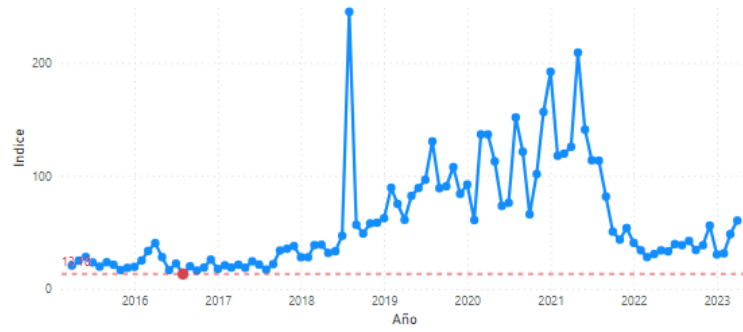
Anexos

Anexo A: Índice Base en el tiempo

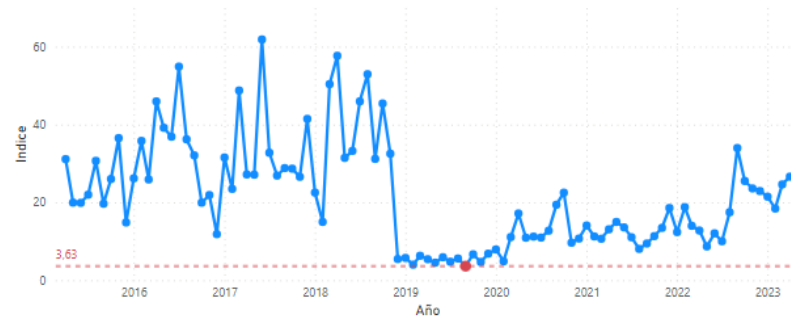
Método R



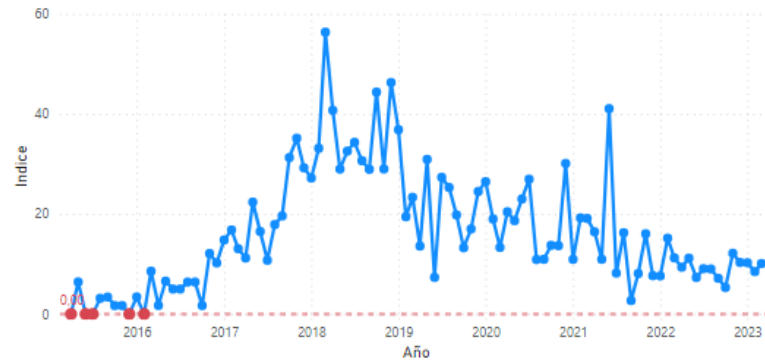
Indice en el tiempo



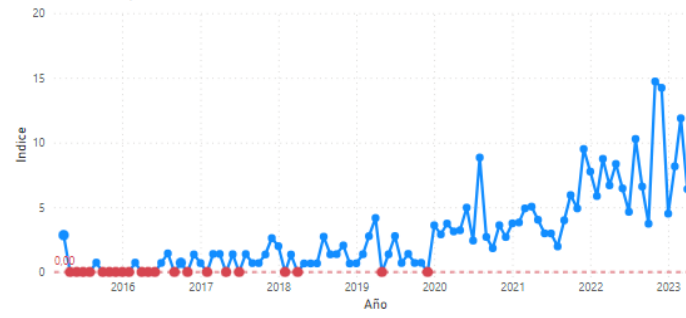
Indice en el tiempo



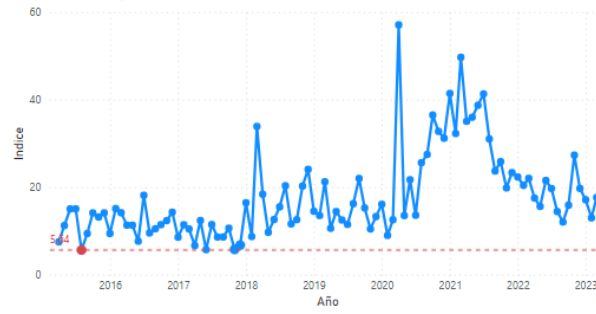
Indice en el tiempo



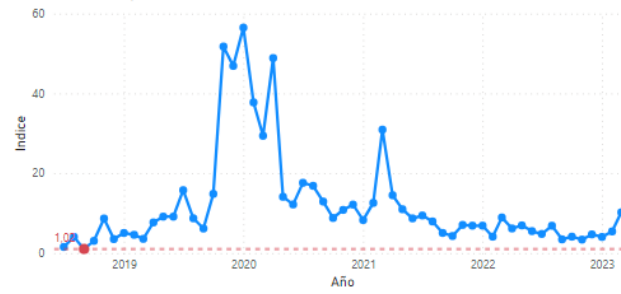
Indice en el tiempo



Indice en el tiempo



Indice en el tiempo



Método T

Índice en el tiempo BancoA



Índice en el tiempo BancoB



Índice en el tiempo BancoC



Indice en el tiempo BancoD



Indice en el tiempo BancoE



Indice en el tiempo BancoF



Indice en el tiempo BancoG



Indice en el tiempo BancoH



Anexo B: Resultados por datos de panel

Método R: Indice Base

PanelOLS Estimation Summary						PanelOLS Estimation Summary							
Dep. Variable:	ROA	R-squared:	0.0807	Dep. Variable:	ROE	R-squared:	0.0026						
Estimator:	PanelOLS	R-squared (Between):	-0.0125	Estimator:	PanelOLS	R-squared (Between):	0.0377						
No. Observations:	729	R-squared (Within):	0.0834	No. Observations:	729	R-squared (Within):	-0.0053						
Date:	Mon, Aug 07 2023	R-squared (Overall):	-0.0102	Date:	Mon, Aug 07 2023	R-squared (Overall):	0.0303						
Time:	13:40:34	Log-likelihood	2742.8	Time:	13:42:17	Log-likelihood	1029.6						
Cov. Estimator:	Unadjusted	F-statistic:	0.4560	Cov. Estimator:	Unadjusted	F-statistic:	1.6184						
Entities:	8	P-value	0.4998	Entities:	8	P-value	0.2038						
Avg Obs:	91.125	Distribution:	F(1,625)	Avg Obs:	91.125	Distribution:	F(1,625)						
Min Obs:	57.000	F-statistic (robust):	0.4560	Min Obs:	57.000	F-statistic (robust):	1.6184						
Max Obs:	96.000	P-value	0.4998	Max Obs:	96.000	P-value	0.2038						
Time periods:	96	Distribution:	F(1,625)	Time periods:	96	Distribution:	F(1,625)						
Avg Obs:	7.5938			Avg Obs:	7.5938								
Min Obs:	7.0000			Min Obs:	7.0000								
Max Obs:	8.0000			Max Obs:	8.0000								
Parameter Estimates						Parameter Estimates							
Parameter	Std. Err.	T-stat	P-value	Lower CI	Upper CI	Parameter	Std. Err.	T-stat	P-value	Lower CI	Upper CI		
Indice	-8.778e-06	1.3e-05	-0.6753	0.4998	-3.43e-05	1.675e-05	Indice	0.0002	0.0001	1.2722	0.2038	-9.427e-05	0.0004
F-test for Poolability: 16.251						F-test for Poolability: 7.2293							
P-value: 0.0000						P-value: 0.0000							
Distribution: F(102,625)						Distribution: F(102,625)							

Método R: Indice Pro3

PanelOLS Estimation Summary						PanelOLS Estimation Summary							
Dep. Variable:	ROA	R-squared:	8.646e-05	Dep. Variable:	ROE	R-squared:	0.0100						
Estimator:	PanelOLS	R-squared (Between):	-0.0048	Estimator:	PanelOLS	R-squared (Between):	0.0819						
No. Observations:	726	R-squared (Within):	0.0012	No. Observations:	726	R-squared (Within):	-0.0103						
Date:	Mon, Aug 07 2023	R-squared (Overall):	-0.0039	Date:	Mon, Aug 07 2023	R-squared (Overall):	0.0658						
Time:	14:36:24	Log-likelihood	2730.7	Time:	14:36:57	Log-likelihood	1027.4						
Cov. Estimator:	Unadjusted	F-statistic:	0.0538	Cov. Estimator:	Unadjusted	F-statistic:	6.2582						
Entities:	8	P-value	0.8167	Entities:	8	P-value	0.0126						
Avg Obs:	90.750	Distribution:	F(1,622)	Avg Obs:	90.750	Distribution:	F(1,622)						
Min Obs:	54.000	F-statistic (robust):	0.0538	Min Obs:	54.000	F-statistic (robust):	6.2582						
Max Obs:	96.000	P-value	0.8167	Max Obs:	96.000	P-value	0.0126						
Time periods:	96	Distribution:	F(1,622)	Time periods:	96	Distribution:	F(1,622)						
Avg Obs:	7.5625			Avg Obs:	7.5625								
Min Obs:	7.0000			Min Obs:	7.0000								
Max Obs:	8.0000			Max Obs:	8.0000								
Parameter Estimates						Parameter Estimates							
Parameter	Std. Err.	T-stat	P-value	Lower CI	Upper CI	Parameter	Std. Err.	T-stat	P-value	Lower CI	Upper CI		
Indice_pro_3	-3.436e-06	1.481e-05	-0.2319	0.8167	-3.253e-05	2.566e-05	Indice_pro_3	0.0004	0.0002	2.5016	0.0126	8.323e-05	0.0007
F-test for Poolability: 15.917						F-test for Poolability: 7.3310							
P-value: 0.0000						P-value: 0.0000							
Distribution: F(102,622)						Distribution: F(102,622)							

Método R: Indice Pro6

PanelOLS Estimation Summary						PanelOLS Estimation Summary							
Dep. Variable:	ROA	R-squared:	7.02e-05	Dep. Variable:	ROE	R-squared:	0.0156						
Estimator:	PanelOLS	R-squared (Between):	0.0045	Estimator:	PanelOLS	R-squared (Between):	0.1059						
No. Observations:	723	R-squared (Within):	-0.0011	No. Observations:	723	R-squared (Within):	-0.0114						
Date:	Mon, Aug 07 2023	R-squared (Overall):	0.0037	Date:	Mon, Aug 07 2023	R-squared (Overall):	0.0856						
Time:	14:34:51	Log-likelihood	2719.8	Time:	14:35:32	Log-likelihood	1025.8						
Cov. Estimator:	Unadjusted	F-statistic:	0.0435	Cov. Estimator:	Unadjusted	F-statistic:	9.8079						
Entities:	8	P-value	0.8349	Entities:	8	P-value	0.0018						
Avg Obs:	90.375	Distribution:	F(1,619)	Avg Obs:	90.375	Distribution:	F(1,619)						
Min Obs:	51.000	F-statistic (robust):	0.0435	Min Obs:	51.000	F-statistic (robust):	9.8079						
Max Obs:	96.000	P-value	0.8349	Max Obs:	96.000	P-value	0.0018						
Time periods:	96	Distribution:	F(1,619)	Time periods:	96	Distribution:	F(1,619)						
Avg Obs:	7.5312			Avg Obs:	7.5312								
Min Obs:	7.0000			Min Obs:	7.0000								
Max Obs:	8.0000			Max Obs:	8.0000								
Parameter Estimates						Parameter Estimates							
Parameter	Std. Err.	T-stat	P-value	Lower CI	Upper CI	Parameter	Std. Err.	T-stat	P-value	Lower CI	Upper CI		
Indice_pro_6	3.255e-06	1.561e-05	0.2085	0.8349	-2.741e-05	3.392e-05	Indice_pro_6	0.0005	0.0002	3.1318	0.0018	0.0002	0.0008
F-test for Poolability: 15.817						F-test for Poolability: 7.4188							
P-value: 0.0000						P-value: 0.0000							
Distribution: F(102,619)						Distribution: F(102,619)							

Método R: Indice Retard3

PanelOLS Estimation Summary						PanelOLS Estimation Summary							
Dep. Variable:	ROA	R-squared:	0.0002			Dep. Variable:	ROE	R-squared:	0.0118				
Estimator:	PanelOLS	R-squared (Between):	0.0059			Estimator:	PanelOLS	R-squared (Between):	0.0781				
No. Observations:	726	R-squared (Within):	-0.0016			No. Observations:	726	R-squared (Within):	-0.0102				
Date:	Mon, Aug 07 2023	R-squared (Overall):	0.0048			Date:	Mon, Aug 07 2023	R-squared (Overall):	0.0627				
Time:	14:00:41	Log-likelihood	2730.7			Time:	14:01:53	Log-likelihood	1028.0				
Cov. Estimator:	Unadjusted	F-statistic:	0.1066			Cov. Estimator:	Unadjusted	F-statistic:	7.4319				
Entities:	8	P-value	0.7442			Entities:	8	P-value	0.0066				
Avg Obs:	90.750	Distribution:	F(1,622)			Avg Obs:	90.750	Distribution:	F(1,622)				
Min Obs:	54.000					Min Obs:	54.000						
Max Obs:	96.000	F-statistic (robust):	0.1066			Max Obs:	96.000	F-statistic (robust):	7.4319				
		P-value	0.7442					P-value	0.0066				
Time periods:	96	Distribution:	F(1,622)			Time periods:	96	Distribution:	F(1,622)				
Avg Obs:	7.5625					Avg Obs:	7.5625						
Min Obs:	7.0000					Min Obs:	7.0000						
Max Obs:	8.0000					Max Obs:	8.0000						
Parameter Estimates						Parameter Estimates							
Parameter	Std. Err.	T-stat	P-value	Lower CI	Upper CI	Parameter	Std. Err.	T-stat	P-value	Lower CI	Upper CI		
Indice_retardo_3	4.227e-06	1.295e-05	0.3265	0.7442	-2.12e-05	2.965e-05	Indice_retardo_3	0.0004	0.0001	2.7262	0.0066	0.0001	0.0006
F-test for Poolability: 16.293						F-test for Poolability: 7.3752							
P-value: 0.0000						P-value: 0.0000							
Distribution: F(102,622)						Distribution: F(102,622)							

Método R: Indice Retard6

PanelOLS Estimation Summary						PanelOLS Estimation Summary							
Dep. Variable:	ROA	R-squared:	0.0020			Dep. Variable:	ROE	R-squared:	0.0224				
Estimator:	PanelOLS	R-squared (Between):	0.0197			Estimator:	PanelOLS	R-squared (Between):	0.1043				
No. Observations:	723	R-squared (Within):	-0.0036			No. Observations:	723	R-squared (Within):	-0.0056				
Date:	Mon, Aug 07 2023	R-squared (Overall):	0.0162			Date:	Mon, Aug 07 2023	R-squared (Overall):	0.0853				
Time:	14:02:30	Log-likelihood	2720.5			Time:	14:01:29	Log-likelihood	1028.3				
Cov. Estimator:	Unadjusted	F-statistic:	1.2683			Cov. Estimator:	Unadjusted	F-statistic:	14.152				
Entities:	8	P-value	0.2605			Entities:	8	P-value	0.0002				
Avg Obs:	90.375	Distribution:	F(1,619)			Avg Obs:	90.375	Distribution:	F(1,619)				
Min Obs:	51.000					Min Obs:	51.000						
Max Obs:	96.000	F-statistic (robust):	1.2683			Max Obs:	96.000	F-statistic (robust):	14.152				
		P-value	0.2605					P-value	0.0002				
Time periods:	96	Distribution:	F(1,619)			Time periods:	96	Distribution:	F(1,619)				
Avg Obs:	7.5312					Avg Obs:	7.5312						
Min Obs:	7.0000					Min Obs:	7.0000						
Max Obs:	8.0000					Max Obs:	8.0000						
Parameter Estimates						Parameter Estimates							
Parameter	Std. Err.	T-stat	P-value	Lower CI	Upper CI	Parameter	Std. Err.	T-stat	P-value	Lower CI	Upper CI		
Indice_retardo_6	1.446e-05	1.284e-05	1.1262	0.2605	-1.076e-05	3.960e-05	Indice_retardo_6	0.0005	0.0001	3.7619	0.0002	0.0002	0.0008
F-test for Poolability: 16.529						F-test for Poolability: 7.5558							
P-value: 0.0000						P-value: 0.0000							
Distribution: F(102,619)						Distribution: F(102,619)							

Método T: Indice Base

PanelOLS Estimation Summary						PanelOLS Estimation Summary							
Dep. Variable:	ROA	R-squared:	0.0485			Dep. Variable:	ROE	R-squared:	0.0254				
Estimator:	PanelOLS	R-squared (Between):	-0.5038			Estimator:	PanelOLS	R-squared (Between):	-0.4557				
No. Observations:	709	R-squared (Within):	0.0058			No. Observations:	709	R-squared (Within):	0.0198				
Date:	Wed, Aug 09 2023	R-squared (Overall):	-0.4261			Date:	Wed, Aug 09 2023	R-squared (Overall):	-0.3856				
Time:	19:13:51	Log-likelihood	2678.2			Time:	19:34:43	Log-likelihood	1003.0				
Cov. Estimator:	Unadjusted	F-statistic:	30.064			Cov. Estimator:	Unadjusted	F-statistic:	15.784				
Entities:	8	P-value	0.0000			Entities:	8	P-value	0.0001				
Avg Obs:	88.625	Distribution:	F(1,605)			Avg Obs:	88.625	Distribution:	F(1,605)				
Min Obs:	37.000					Min Obs:	37.000						
Max Obs:	96.000	F-statistic (robust):	30.064			Max Obs:	96.000	F-statistic (robust):	15.784				
		P-value	0.0000					P-value	0.0001				
Time periods:	96	Distribution:	F(1,605)			Time periods:	96	Distribution:	F(1,605)				
Avg Obs:	7.3854					Avg Obs:	7.3854						
Min Obs:	7.0000					Min Obs:	7.0000						
Max Obs:	8.0000					Max Obs:	8.0000						
Parameter Estimates						Parameter Estimates							
Parameter	Std. Err.	T-stat	P-value	Lower CI	Upper CI	Parameter	Std. Err.	T-stat	P-value	Lower CI	Upper CI		
Indice	0.0118	0.0021	5.5555	0.0000	0.0076	0.0160	Indice	0.0096	0.0226	3.9729	0.0001	0.0453	0.1340
F-test for Poolability: 17.074						F-test for Poolability: 17.074							
P-value: 0.0000						P-value: 0.0000							
Distribution: F(102,605)						Distribution: F(102,605)							

Método T: Indice Pro3

PanelOLS Estimation Summary							PanelOLS Estimation Summary															
Dep. Variable:	ROA	R-squared:	0.0427	Dep. Variable:	ROE	R-squared:	0.0259	Dep. Variable:	ROE	R-squared:	0.0259	Dep. Variable:	ROE	R-squared:	0.0259							
Estimator:	PanelOLS	R-squared (Between):	-0.5551	Estimator:	PanelOLS	R-squared (Between):	-0.5380	Estimator:	PanelOLS	R-squared (Between):	-0.5380	Estimator:	PanelOLS	R-squared (Between):	-0.5380							
No. Observations:	685	R-squared (Within):	0.0771	No. Observations:	685	R-squared (Within):	0.0185	No. Observations:	685	R-squared (Within):	0.0185	No. Observations:	685	R-squared (Within):	0.0185							
Date:	Wed, Aug 09 2023	R-squared (Overall):	-0.4712	Date:	Wed, Aug 09 2023	R-squared (Overall):	-0.4587	Date:	Wed, Aug 09 2023	R-squared (Overall):	-0.4587	Date:	Wed, Aug 09 2023	R-squared (Overall):	-0.4587							
Time:	19:39:24	Log-likelihood	2576.3	Time:	19:39:53	Log-likelihood	965.73	Time:	19:39:53	Log-likelihood	965.73	Time:	19:39:53	Log-likelihood	965.73							
Cov. Estimator:	Unadjusted	F-statistic:	26.057	Cov. Estimator:	Unadjusted	F-statistic:	15.514	Cov. Estimator:	Unadjusted	F-statistic:	15.514	Cov. Estimator:	Unadjusted	F-statistic:	15.514							
Entities:	8	P-value	0.0000	Entities:	8	P-value	0.0001	Entities:	8	P-value	0.0001	Entities:	8	P-value	0.0001							
Avg Obs:	85.625	Distribution:	F(1,584)	Avg Obs:	85.625	Distribution:	F(1,584)	Avg Obs:	85.625	Distribution:	F(1,584)	Avg Obs:	85.625	Distribution:	F(1,584)							
Min Obs:	34.000	F-statistic (robust):	26.057	Min Obs:	34.000	F-statistic (robust):	15.514	Min Obs:	34.000	F-statistic (robust):	15.514	Min Obs:	34.000	F-statistic (robust):	15.514							
Max Obs:	93.000	P-value	0.0000	Max Obs:	93.000	P-value	0.0001	Max Obs:	93.000	P-value	0.0001	Max Obs:	93.000	P-value	0.0001							
Time periods:	93	Distribution:	F(1,584)	Time periods:	93	Distribution:	F(1,584)	Time periods:	93	Distribution:	F(1,584)	Time periods:	93	Distribution:	F(1,584)							
Avg Obs:	7.3656	Parameter Estimates		Avg Obs:	7.3656	Parameter Estimates		Avg Obs:	7.3656	Parameter Estimates		Avg Obs:	7.3656	Parameter Estimates								
Min Obs:	7.0000	Parameter	Std. Err.	T-stat	P-value	Lower CI	Upper CI	Parameter	Std. Err.	T-stat	P-value	Lower CI	Upper CI	Parameter	Std. Err.	T-stat	P-value	Lower CI	Upper CI			
Max Obs:	8.0000	Index_pro_3	0.0129	0.0025	5.1046	0.0000	0.0080	0.0179	Index_pro_3	0.1047	0.0266	3.9387	0.0001	0.0525	0.1569	Index_pro_3	0.1047	0.0266	3.9387	0.0001	0.0525	0.1569
F-test for Poolability: 16.369							F-test for Poolability: 7.2068							F-test for Poolability: 7.2068								
P-value: 0.0000							P-value: 0.0000							P-value: 0.0000								
Distribution: F(99,584)							Distribution: F(99,584)							Distribution: F(99,584)								

Método T: Indice Pro6

PanelOLS Estimation Summary							PanelOLS Estimation Summary															
Dep. Variable:	ROA	R-squared:	0.0315	Dep. Variable:	ROE	R-squared:	0.0179	Dep. Variable:	ROE	R-squared:	0.0179	Dep. Variable:	ROE	R-squared:	0.0179							
Estimator:	PanelOLS	R-squared (Between):	-0.5175	Estimator:	PanelOLS	R-squared (Between):	-0.4779	Estimator:	PanelOLS	R-squared (Between):	-0.4779	Estimator:	PanelOLS	R-squared (Between):	-0.4779							
No. Observations:	661	R-squared (Within):	0.0647	No. Observations:	661	R-squared (Within):	0.0043	No. Observations:	661	R-squared (Within):	0.0043	No. Observations:	661	R-squared (Within):	0.0043							
Date:	Wed, Aug 09 2023	R-squared (Overall):	-0.4394	Date:	Wed, Aug 09 2023	R-squared (Overall):	-0.4094	Date:	Wed, Aug 09 2023	R-squared (Overall):	-0.4094	Date:	Wed, Aug 09 2023	R-squared (Overall):	-0.4094							
Time:	19:44:56	Log-likelihood	2476.8	Time:	19:45:25	Log-likelihood	930.77	Time:	19:45:25	Log-likelihood	930.77	Time:	19:45:25	Log-likelihood	930.77							
Cov. Estimator:	Unadjusted	F-statistic:	18.329	Cov. Estimator:	Unadjusted	F-statistic:	10.275	Cov. Estimator:	Unadjusted	F-statistic:	10.275	Cov. Estimator:	Unadjusted	F-statistic:	10.275							
Entities:	8	P-value	0.0000	Entities:	8	P-value	0.0014	Entities:	8	P-value	0.0014	Entities:	8	P-value	0.0014							
Avg Obs:	82.625	Distribution:	F(1,563)	Avg Obs:	82.625	Distribution:	F(1,563)	Avg Obs:	82.625	Distribution:	F(1,563)	Avg Obs:	82.625	Distribution:	F(1,563)							
Min Obs:	31.000	F-statistic (robust):	18.329	Min Obs:	31.000	F-statistic (robust):	10.275	Min Obs:	31.000	F-statistic (robust):	10.275	Min Obs:	31.000	F-statistic (robust):	10.275							
Max Obs:	90.000	P-value	0.0000	Max Obs:	90.000	P-value	0.0014	Max Obs:	90.000	P-value	0.0014	Max Obs:	90.000	P-value	0.0014							
Time periods:	90	Distribution:	F(1,563)	Time periods:	90	Distribution:	F(1,563)	Time periods:	90	Distribution:	F(1,563)	Time periods:	90	Distribution:	F(1,563)							
Avg Obs:	7.3444	Parameter Estimates		Avg Obs:	7.3444	Parameter Estimates		Avg Obs:	7.3444	Parameter Estimates		Avg Obs:	7.3444	Parameter Estimates								
Min Obs:	7.0000	Parameter	Std. Err.	T-stat	P-value	Lower CI	Upper CI	Parameter	Std. Err.	T-stat	P-value	Lower CI	Upper CI	Parameter	Std. Err.	T-stat	P-value	Lower CI	Upper CI			
Max Obs:	8.0000	Index_pro_6	0.0121	0.0028	4.2812	0.0000	0.0066	0.0177	Index_pro_6	0.0943	0.0294	3.2054	0.0014	0.0365	0.1521	Index_pro_6	0.0943	0.0294	3.2054	0.0014	0.0365	0.1521
F-test for Poolability: 15.849							F-test for Poolability: 7.3286							F-test for Poolability: 7.3286								
P-value: 0.0000							P-value: 0.0000							P-value: 0.0000								
Distribution: F(96,563)							Distribution: F(96,563)							Distribution: F(96,563)								

Método T: Indice Ret3

PanelOLS Estimation Summary							PanelOLS Estimation Summary															
Dep. Variable:	ROA	R-squared:	0.0220	Dep. Variable:	ROE	R-squared:	0.0132	Dep. Variable:	ROE	R-squared:	0.0132	Dep. Variable:	ROE	R-squared:	0.0132							
Estimator:	PanelOLS	R-squared (Between):	-0.3309	Estimator:	PanelOLS	R-squared (Between):	-0.3229	Estimator:	PanelOLS	R-squared (Between):	-0.3229	Estimator:	PanelOLS	R-squared (Between):	-0.3229							
No. Observations:	685	R-squared (Within):	0.0525	No. Observations:	685	R-squared (Within):	0.0078	No. Observations:	685	R-squared (Within):	0.0078	No. Observations:	685	R-squared (Within):	0.0078							
Date:	Wed, Aug 09 2023	R-squared (Overall):	-0.2804	Date:	Wed, Aug 09 2023	R-squared (Overall):	-0.2756	Date:	Wed, Aug 09 2023	R-squared (Overall):	-0.2756	Date:	Wed, Aug 09 2023	R-squared (Overall):	-0.2756							
Time:	19:50:14	Log-likelihood	2569.0	Time:	19:51:16	Log-likelihood	961.30	Time:	19:51:16	Log-likelihood	961.30	Time:	19:51:16	Log-likelihood	961.30							
Cov. Estimator:	Unadjusted	F-statistic:	13.165	Cov. Estimator:	Unadjusted	F-statistic:	7.8095	Cov. Estimator:	Unadjusted	F-statistic:	7.8095	Cov. Estimator:	Unadjusted	F-statistic:	7.8095							
Entities:	8	P-value	0.0003	Entities:	8	P-value	0.0054	Entities:	8	P-value	0.0054	Entities:	8	P-value	0.0054							
Avg Obs:	85.625	Distribution:	F(1,584)	Avg Obs:	85.625	Distribution:	F(1,584)	Avg Obs:	85.625	Distribution:	F(1,584)	Avg Obs:	85.625	Distribution:	F(1,584)							
Min Obs:	34.000	F-statistic (robust):	13.165	Min Obs:	34.000	F-statistic (robust):	7.8095	Min Obs:	34.000	F-statistic (robust):	7.8095	Min Obs:	34.000	F-statistic (robust):	7.8095							
Max Obs:	93.000	P-value	0.0003	Max Obs:	93.000	P-value	0.0054	Max Obs:	93.000	P-value	0.0054	Max Obs:	93.000	P-value	0.0054							
Time periods:	93	Distribution:	F(1,584)	Time periods:	93	Distribution:	F(1,584)	Time periods:	93	Distribution:	F(1,584)	Time periods:	93	Distribution:	F(1,584)							
Avg Obs:	7.3656	Parameter Estimates		Avg Obs:	7.3656	Parameter Estimates		Avg Obs:	7.3656	Parameter Estimates		Avg Obs:	7.3656	Parameter Estimates								
Min Obs:	7.0000	Parameter	Std. Err.	T-stat	P-value	Lower CI	Upper CI	Parameter	Std. Err.	T-stat	P-value	Lower CI	Upper CI	Parameter	Std. Err.	T-stat	P-value	Lower CI	Upper CI			
Max Obs:	8.0000	Index_retardo_3	0.0083	0.0023	3.6204	0.0003	0.0038	0.0128	Index_retardo_3	0.0668	0.0239	2.7945	0.0054	0.0198	0.1137	Index_retardo_3	0.0668	0.0239	2.7945	0.0054	0.0198	0.1137
F-test for Poolability: 16.230							F-test for Poolability: 7.1161							F-test for Poolability: 7.1161								
P-value: 0.0000							P-value: 0.0000							P-value: 0.0000								
Distribution: F(99,584)							Distribution: F(99,584)							Distribution: F(99,584)								

Método T: Indice Ret6

PanelOLS Estimation Summary							PanelOLS Estimation Summary							
Dep. Variable:	ROA	R-squared:	0.0156				Dep. Variable:	ROE	R-squared:	0.0084				
Estimator:	PanelOLS	R-squared (Between):	-0.2715				Estimator:	PanelOLS	R-squared (Between):	-0.2481				
No. Observations:	661	R-squared (Within):	0.0403				No. Observations:	661	R-squared (Within):	0.0039				
Date:	Wed, Aug 09 2023	R-squared (Overall):	-0.2304				Date:	Wed, Aug 09 2023	R-squared (Overall):	-0.2134				
Time:	20:07:44	Log-likelihood	2471.4				Time:	20:08:13	Log-likelihood	927.59				
Cov. Estimator:	Unadjusted	F-statistic:	8.9197				Cov. Estimator:	Unadjusted	F-statistic:	4.7746				
Entities:	8	P-value	0.0029				Entities:	8	P-value	0.0293				
Avg Obs:	82.625	Distribution:	F(1,563)				Avg Obs:	82.625	Distribution:	F(1,563)				
Min Obs:	31.000	F-statistic (robust):	8.9197				Min Obs:	31.000	F-statistic (robust):	4.7746				
Max Obs:	90.000	P-value	0.0029				Max Obs:	90.000	P-value	0.0293				
Time periods:	90	Distribution:	F(1,563)				Time periods:	90	Distribution:	F(1,563)				
Avg Obs:	7.3444						Avg Obs:	7.3444						
Min Obs:	7.0000						Min Obs:	7.0000						
Max Obs:	8.0000						Max Obs:	8.0000						
Parameter Estimates							Parameter Estimates							
	Parameter	Std. Err.	T-stat	P-value	Lower CI	Upper CI		Parameter	Std. Err.	T-stat	P-value	Lower CI	Upper CI	
	Indice_retardo_6	0.0071	0.0024	2.9856	0.0029	0.0024	0.0117	Indice_retardo_6	0.0533	0.0244	2.1851	0.0293	0.0054	0.1013
F-test for Poolability: 15.899							F-test for Poolability: 7.2772							
P-value: 0.0000							P-value: 0.0000							
Distribution: F(96,563)							Distribution: F(96,563)							

Anexo C: APR, Patrimonio y Activos

ADECUACIÓN DE CAPITAL CONSOLIDADA GLOBAL DEL SISTEMA BANCARIO CHILENO AL MES DE ABRIL DE 2023 (BASILEA III)

(Cifras en millones de pesos)

Activos Ponderados por Riesgo (APR)

7,341,772
2,640,510
5,042,248
36,643,744
48,342,325
32,348,716
5,086,834
2,864,263
27,253,635
1,276,719
38,564,165
8,157,158
23,501
254,428
584,852
1,160,071
31,239,006

250,084,065

PRINCIPALES ACTIVOS, PASIVOS Y PATRIMONIO DEL ESTADO DE SITUACIÓN FINANCIERA CONSOLIDADO POR INSTITUCIONES AL MES DE ABRIL DE 2023

(Cifras en millones de pesos)

Patrimonio	Total activos
------------	---------------

827,519	11,363,992
501,055	3,211,960
640,857	8,259,097
4,792,573	55,384,781
4,873,608	78,504,757
3,060,494	57,463,222
984,465	7,048,685
282,377	4,342,464
3,416,717	40,784,929
208,575	1,265,474
4,128,339	69,486,857
793,869	10,582,163
81,457	34,362
130,264	282,703
121,142	1,528,500
407,012	2,409,205
3,257,026	45,548,014

28,507,351

397,561,166