

ESCUELA POLITÉCNICA NACIONAL

FACULTAD DE CIENCIAS

DESAFÍOS DE LA INDUSTRIA BANCARIA EN LA ESTIMACIÓN DE LA SEVERIDAD DE PÉRDIDA (LGD - LOSS GIVEN DEFAULT)

TRABAJO DE TITULACIÓN PREVIO A LA OBTENCIÓN DEL TÍTULO DE
INGENIERA MATEMÁTICA

PROYECTO DE INVESTIGACIÓN

MARÍA BELÉN ROSERO SORIA
maria.rosero@epn.edu.ec

Directora: ADRIANA UQUILLAS ANDRADE, PHD
adriana.uquillas@epn.edu.ec

QUITO, AGOSTO 2023

DECLARACIÓN

Yo MARÍA BELÉN ROSERO SORIA , declaro bajo juramento que el trabajo aquí escrito es de mi autoría; que no ha sido previamente presentado para ningún grado o calificación profesional; y que he consultado las referencias bibliográficas que se incluyen en este documento.

A través de la presente declaración cedo mis derechos de propiedad intelectual, correspondientes a este trabajo, a la Escuela Politécnica Nacional, según lo establecido por la Ley de Propiedad Intelectual, por su reglamento y por la normatividad institucional vigente.

MARÍA BELÉN ROSERO SORIA

CERTIFICACIÓN

Certifico que el presente trabajo fue desarrollado por MARÍA BELÉN ROSERO SORIA , bajo mi supervisión.

ADRIANA UQUILLAS ANDRADE, PhD
Director del Proyecto

AGRADECIMIENTOS

A mis padres Nory y William, por todo el amor y su apoyo, por nunca dejarme sola.

A mi hermana Xiomy, por siempre ayudarme y estar incondicionalmente para mí y mi hijo.

A Sebastián, por ser quién toma de mi mano y me acompaña.

A Leslie y Evita, porque a pesar de que no están, guían mi vida.

A mi directora, Dra. Adriana Uquillas, por su paciencia y apoyo, en especial por ser un ejemplo para las mujeres en la ciencia.

A mis amigas Karen, Tefa, Eve y en especial a Liz por siempre estar para mí.

DEDICATORIA

A Benjamín, luz de mi vida.

Índice general

Índice de figuras	VIII
Índice de cuadros	X
Resumen	XII
Abstract	XIII
1. Introducción	1
1.1. Antecedentes	1
1.2. Justificación	3
1.2.1. Justificación Teórica	3
1.2.2. Justificación Metodológica	4
1.2.3. Justificación Práctica	5
1.3. Objetivos	6
1.3.1. Objetivo General	6
1.3.2. Objetivos Específicos	6
2. Marco Teórico	7
2.1. Conceptos financieros	7
2.1.1. Severidad de Pérdida	8
2.1.2. Factores Microeconómicos	9
2.1.3. Factores Macroeconómicos	12
2.1.4. Riesgo País	16
2.2. Estructuras de representación de datos	18

2.2.1.	Datos de Series de tiempo	18
2.2.2.	Sección cruzada o datos de corte transversal	19
2.2.3.	Datos en Panel	19
2.3.	Metodología para la estimación del modelo de regresión de datos en panel	19
2.3.1.	Modelo de Efectos Comunes	20
2.3.2.	Modelo de Efectos Fijos	20
2.3.3.	Modelo de Efectos Aleatorios	21
2.4.	Método de selección del modelo de regresión de Datos en Panel	22
2.4.1.	Test de Chow	22
2.4.2.	Test de Hausman	23
2.4.3.	Test de Multiplicadores de Lagrange	23
2.5.	Supuestos de prueba e idoneidad del modelo	24
2.5.1.	Prueba de Multicolinealidad	25
2.5.2.	Prueba de normalidad	25
2.5.3.	Prueba de heteroscedasticidad	26
2.5.4.	Prueba de autocorrelación	26
2.6.	Prueba de significación de los parámetros de regresión	27
2.6.1.	Prueba simultánea	27
2.6.2.	Prueba parcial	28
2.7.	Backtesting en datos en panel	28
2.7.1.	Backtesting de ajuste del modelo	28
3.	Implementación	30
3.1.	Descripción de las variables	30
3.1.1.	Descripción Estadística de los Datos	31
3.1.2.	Análisis de Estacionariedad	33
3.2.	Construcción del modelo de datos en panel	36
3.2.1.	Estructura	36
3.2.2.	Modelo de efectos comunes	37

3.2.3.	Modelo de efectos aleatorios	43
3.2.4.	Modelo de efectos fijos	45
3.3.	Método de selección	52
3.3.1.	Test de Hausman	52
3.3.2.	Test de Chow	53
3.4.	Idoneidad del modelo	53
3.5.	Pruebas de significación	55
3.5.1.	Prueba simultánea	55
3.5.2.	Prueba Parcial	55
3.6.	Predicción	57
4.	Conclusiones y Recomendaciones	61
4.1.	Conclusiones	61
4.2.	Recomendaciones	63
	Bibliografía	64
	Anexos	69
A.		70
A.1.	Funciones de R	70
A.1.1.	Funciones de datos de panel	71
A.1.2.	Pruebas	71

Índice de figuras

2.1. Ejemplo de la historia de un crédito. Fuente: (Cuevas, 2019)	8
2.2. Tasa de desempleo en el Ecuador 2014-2021. Fuente: https://es.theglobaleconomy.com/ecuador/unemployment_rate/	12
2.3. IPC en los meses de febrero en el Ecuador 2013-2022. Fuente: Ecuador en cifras. Boletín Técnico N.º 09-2022-IPC	13
2.4. Índice de actividad económica coyuntural Fuente: Banco Central del Ecuador 2012-2020. https://www.bce.fin.ec/index.php/component/k2/item/313-indice-de-actividad-econ%C3%B3mica-coyuntural-ideac	15
2.5. Evolución del Precio petróleo West Texas Intermediate Fuente: https://datosmacro.expansion.com/materias-primas/petroleo-wti	16
2.6. Riesgo País en América latina 2020. Fuente: Revista Panorámica 2020.	17
2.7. PIB 2017-2020 Fuente: (Banco Central del Ecuador, 2020)	18
2.8. Elección de estimación para la regresión de datos en panel. Fuente: (Zulfikar, 2018)	24
3.1. Estructura de la base de datos. Fuente: Propia	31
3.2. Serie de tiempo Estacionaria. Fuente: Propia	34
3.3. Serie de tiempo No Estacionaria. Fuente: Propia	35
3.4. Comportamiento de la Severidad de Pérdida. Fuente: Propia	36
3.5. Base de datos Balanceada Fuente: Propia	37
3.6. Test de Durbin Watson Fuente: propia	54
3.7. Test de Jarque Bera Fuente: propia	54
3.8. Estimación de la LGD, Bancos Grandes Fuente: Propia	58
3.9. Estimación de la LGD, Bancos Medianos Fuente: Propia	58
3.10. Estimación de la LGD, Bancos Pequeños Fuente: Propia	59

3.11. Predicción de la LGD, Bancos Grandes Fuente: Propia	59
3.12. Predicción de la LGD, Bancos Medianos Fuente: Propia	60
3.13. Predicción de la LGD, Bancos Pequeños Fuente: Propia	60

Índice de cuadros

2.1. Descripción de variables Microeconómicas	10
2.2. Descripción de variables Microeconómicas	11
2.3. Valor crítico de autocorrelación por Durbin Watson	27
3.1. Variables microeconómicas	32
3.2. Variables macroeconómicas	33
3.3. Modelo MCO	38
3.4. Modelo efectos aleatorios	43
3.5. Modelo de efectos fijos. Parte 1	46
3.6. Modelo de efectos fijos. Parte 2	47
3.7. Resultados de la prueba simultánea	55
3.8. Resultados del t-valor de las variables predictoras	56

Resumen

La industria bancaria en el Ecuador se rige a estándares regulatorios de la Superintendencia de Bancos, por lo que, las entidades deben establecer esquemas eficientes para el control del riesgo mediante la PD, EAD y LGD. El presente estudio se enfoca en el estudio de la LGD, o Loss Given Default, que corresponde una fracción de un crédito que no es recuperable en caso de que esté en incumplimiento o de factores externos a la entidad (factores macroeconómicos). Existen diferentes técnicas de estimación mediante diferentes variables, en este caso factores microeconómicos proporcionados por la Superintendencia de Bancos y Banco Central del Ecuador (BCE) en un período de tiempo trimestral de 2017 al 2020, estos indicadores se utilizarán para la construcción del conjunto de datos y así usar técnicas estadísticas como el análisis de regresión para estimar la LGD. Actualmente, en el Ecuador no existen estudios públicos para estimar la LGD con modelos estadísticos, por lo que, este proyecto es un aporte eficiente a la introducción de más técnicas estadísticas al sistema financiero.

Se utilizó el método de datos en panel, una técnica que controla los efectos a nivel de individuos y de tiempo. El modelo de efectos fijos nos ayudó a estimar la LGD y poder concluir que los factores macroeconómicos como el IPC, ICC y el PIB real influyen para que las entidades financieras tengan un LGD más elevado, lo que provoca que las tasas de recuperación sean más bajas, por ende más pérdidas. El incluir estas variables nos permite ajustar un modelo con un backtesting acertado con predicciones a corto plazo de la variable dependiente.

Palabras clave: Severidad de pérdida, datos en panel, modelos de efectos comunes, efectos fijos, efectos aleatorios.

Abstract

The banking industry in Ecuador is governed by the regulatory standards of the Superintendency of Banks, so entities must establish efficient schemes for risk control through PD, EAD and LGD. This study focuses on the study of the LGD, or Loss Given Default, which corresponds to a fraction of a loan that is not recoverable in case it is in default or due to factors external to the entity (macroeconomic factors). There are different estimation techniques using different variables, in this case microeconomic factors provided by the Superintendency of Banks and the Central Bank of Ecuador (BCE) in a quarterly period of time from 2017 to 2020, these indicators will be used for the construction of the data set and thus use statistical techniques such as regression analysis to estimate LGD. Currently, in Ecuador there are no public studies to estimate the LGD with statistical models, therefore, this project is an efficient contribution to the introduction of more statistical techniques to the financial system.

The panel data method, a technique that controls for effects at the individual and time levels, was used. The fixed effects model helped us to estimate the LGD and to be able to conclude that macroeconomic factors such as the CPI, CCI and real GDP influence financial entities to have a higher LGD, which causes lower recovery rates. Hence more losses. Including these variables allows us to fit a successful backtested model with short-term predictions of the dependent variable.

Keywords: Loss given default, panel data, common effects models, fixed effects, random effects.

Capítulo 1

Introducción

1.1. Antecedentes

El Sistema Financiero a nivel global desarrolla un rol muy importante para la economía, de manera eficiente, con un correcto manejo de sus recursos y la adecuada administración del riesgo, el sistema financiero tiene como propósito dinamizar la economía, principalmente con la captación de fondos de personas consideradas como ahorristas o agentes superactivos y la prestación de este a personas que lo demanden a través del otorgamiento de créditos conocidos como agentes deficitarios (Aguilar, Camargo y Saravia, 2004).

En el Ecuador, el sistema financiero cumple con este objetivo y está conformado por entidades públicas y privadas, bajo el control y supervisión de la Superintendencia de Bancos. Estas entidades podrían ver afectadas su estabilidad al conceder créditos con la incertidumbre del futuro pago de estos, aumentando su nivel de riesgo, por lo que es fundamental el análisis y evaluación para mitigar el riesgo de la entidad con el objeto de prevenir pérdidas ante una afectación. Actualmente, según las Normas de Control para las entidades de los sectores financieros, público y privado, las entidades controladas deben establecer esquemas eficientes de administración y control del riesgo de crédito al que se expone en el desarrollo del negocio. Cada entidad controlada tiene su propio perfil de riesgo según las características de los mercados en los que opera y de los productos que ofrece; por lo tanto, al no existir un modelo único de administración del riesgo de crédito, cada entidad debe desarrollar su propio esquema (Codificación de las Normas de la SB/ Libro Uno – Sistema Financiero, 2018). Así, cada entidad debe establecer, adecuar su capital y administrar los riesgos con terceros para cumplir sus obligaciones, es decir,

asegurar sus posibles pérdidas de acuerdo con su nivel de ingresos. Por lo tanto, las entidades financieras toman decisiones para el otorgamiento de créditos a los usuarios según sus necesidades, basándose en modelos matemáticos que predicen, dependiendo del comportamiento histórico de su cartera, mediante la probabilidad de incumplimiento de pago. Las entidades controladas dan préstamos a sus clientes con el objetivo de que estos puedan adquirirlos para un fin determinado y pagarlo en un tiempo acordado, para que las instituciones financieras mitiguen riesgo de crédito, estas deben establecer mecanismos que les permitan el control y aseguramiento de esta exposición, por lo que el incumplimiento genera pérdidas en la entidad, en el sistema financiero y en la economía del país.

La estabilidad financiera es el objetivo principal del Comité de Basilea; a largo de los años se crearon acuerdos para la regulación y supervisión bancaria, en junio de 2004 se crea el denominado “Convergencia internacional de medidas y normas de capital” o “Basilea II”, el cual establece a las entidades financieras de los países que determinen una revisión de dotaciones de capital para cubrir los riesgos. Esta norma establece el requerimiento de capital mínimo para cubrir el riesgo de crédito, mediante el método de calificaciones internas o IRBa (Internal Ratings Based approach) que se centra en las componentes que definen el riesgo de crédito: Probabilidad de Incumplimiento (PD), Exposición al incumplimiento (EAD) y Pérdida dado el incumplimiento (LGD) o Severidad de la pérdida. El problema en el que nos enfocaremos es en la estimación LGD definida como la pérdida que sufriría la institución después de haber realizado todas las gestiones para recuperar los créditos que han sido incumplidos, ejecutar las garantías o recibirlas como dación en pago. La severidad de la pérdida es igual a $1 - \text{Tasa de Recuperación}$ (Codificación de las Normas de la SB/ Libro Uno – Sistema Financiero, 2018).

En este trabajo nos centraremos en la modelización y previsión de la severidad de la pérdida del Sistema Financiero en el Ecuador, considerando para este estudio solamente a los bancos privados y basándonos en una clasificación adecuada de las instituciones controladas para una mejor estimación con el uso de variables macroeconómicas.

Es importante tener en cuenta, que la LGD es una estimación de parte de la exposición a la que está expuesta un banco al otorgar préstamos que probablemente no se recuperarán en caso de incumplimiento. Aslam (2020), expone de manera clara, si una inversión tiene una alta probabilidad de incumplimiento (PD) pero una LGD baja, el capital aún podría recuperarse. Por el contrario, si tiene una PD baja, pero una LGD alta, se podría perder dinero. Esto hace que sea importante conside-

rar LGD; por lo que dada esta explicación al aplicar a nuestro objetivo de estudio se podrá obtener una buena estimación de la LGD, ya sea por afecciones externas o internas en el sistema financiero.

Actualmente, en el Ecuador y en Latinoamérica hay escasos estudios acerca de la estimación de la severidad de pérdida a nivel macroeconómico, de hecho, Córdor (2020), es el único autor en el Ecuador que ha presentado estudios públicos de estimación para medir niveles de solvencia como para determinar nivel de liquidez en escenarios estresados e hipotéticos aplicando a su trabajo técnicas basadas en medias, más no, utilizando una metodología estadística robusta para mejor aserción de sus cálculos, en base a esto se pretende realizar un estudio que aporte más información con posibles medidas a tomar para que las entidades del sistema financiero del país puedan mitigar el riesgo tomando en cuenta los factores macroeconómicos que afectan en la recuperación de la deuda por atraso e incumplimiento de la misma.

Por otra parte, países más desarrollados han centrado sus estudios a la estimación de la severidad de pérdida, ignorando en gran medida los posibles efectos retardados de la macroeconomía sobre la LGD, por lo que, deja un vacío para poder demostrar, tanto las características microeconómicas de los deudores o de las organizaciones financieras, como las condiciones macroeconómicas influyen en el posible incumplimiento de pago al otorgar créditos. Por lo que, en países como República Checa e Italia han enfocado estudios a la importancia de la estimación LGD y la contribución directa del entorno macroeconómico a la variación de la severidad, factores como el PIB, consumo de hogares y tasa de desempleo son variables fundamentales de sus estudios ([18] y [25]). Con base en estas publicaciones, se podrá dar inicio para poder aplicar al sistema financiero del Ecuador, aportando información más detallada sobre la relación entre la severidad de pérdida en el sistema bancario y el ciclo económico de la economía ecuatoriana.

1.2. Justificación

1.2.1. Justificación Teórica

Puesto que, toda institución financiera está obligada a establecer y mantener un capital que le permita hacer frente a las posibles adversidades a las que está sujeto, por la falta de cumplimiento de pago al conceder un crédito. La estabilidad del sistema financiero respecto al riesgo de crédito depende de un eficiente proceso

de mitigación del riesgo; por lo cual, la estimación de una LGD tiene como objeto garantizar los requisitos establecidos por Basilea II, donde, los requisitos de capital reflejan adecuadamente la volatilidad sistemática de las pérdidas de crédito a lo largo del tiempo. Además, las tasas de recuperación de las exposiciones impagas pueden ser más bajas durante condiciones de recesión económica que en condiciones normales, una norma de capital diseñada para garantizar un capital suficiente para cubrir las pérdidas realizadas en circunstancias adversas debería reflejar esta tendencia (Sabato y Schmid, S.F.). Lo que constituye uno de los objetivos de este trabajo. Existen varios estudios que consisten en estimar la severidad de pérdida de una cartera. Aunque no es fácil de modelar, se puede construir modelos que proporcionan mejores estimaciones de la severidad y buenas previsiones de la severidad media de una cartera (Cook y Belloti, 2012). Varios autores presentan, además, el impacto de factores macroeconómicos mediante una modelización del riesgo de una cartera a través de modelos econométricos y estadísticos, usando variables financieras y macroeconómicas que afectan directamente el riesgo, por ende, la severidad de pérdida. Variables como: el Producto Interno Bruto (PIB), tasa de desempleo, consumo de hogares, etc., podrían ser variables endógenas o exógenas para la estimación de la severidad (Caselli, Gatti, y Querci, 2008). Para llevar a cabo este proyecto se propone analizar los diferentes modelos y métodos que involucren además variables macroeconómicas y que nos guíen en la modelización para entidades del sistema financiero del Ecuador, estimando la severidad de pérdida con una clasificación adecuada de instituciones según su dinámica financiera.

1.2.2. Justificación Metodológica

La estimación de la severidad de pérdida es una herramienta muy necesaria para prever un capital mínimo en caso de perturbaciones a una institución financiera, por lo que, al estimar la severidad a un nivel macroeconómico, mejorará la estabilidad del sistema financiero y la economía real del país.

En primer lugar, la estimación de la severidad de pérdida mejora la capacidad de los bancos para absorber pérdidas crediticias inesperadas, por lo que, los métodos más comunes para esta estimación son: “la regresión de respuesta fraccional, el árbol de regresión, el bosque aleatorio o random forest, el aumento de gradiente, la red neuronal artificial, y los mínimos cuadrados” (Hurlin, Leymarie y Patin, 2018). Por tanto, es posible analizar estos métodos porque podrían ser parte de la metodología para este proyecto.

Otras técnicas de los estudios más relevantes que fueron aplicados en República Checa e Italia, por lo que, tenerlos en cuenta para el inicio del proyecto es necesario para poder encontrar un modelo o los modelos que estimen la severidad de pérdida, el proyecto investigará la influencia de datos tanto a nivel microeconómico como macroeconómico, por lo que, un método flexible sobre los diferentes factores que afectan a la estimación de la severidad aplicada en los estudios mencionados es: el emplear un algoritmo de selección para modelos combinados y agregados. Entonces un modelo combinado permitirá el estudio a nivel micro, dentro de un marco paramétrico, posiblemente usando modelos de regresión lineal generalizada y regresión multinivel. Por otra parte, el modelo agregado se basa en un algoritmo automatizado de selección de modelos de general a específico (Gets) para la selección de variables a nivel macro dentro de un modelo ARIMA, puesto que mantiene la congruencia del modelo durante el proceso de selección, siendo un método viable para datos agregados. Por tanto, objetivo de ambos enfoques de modelado, combinados y agregados, es resaltar la importancia de las características macroeconómicas en la estimación de la LGD mediante el uso conjuntos de datos y empleando diferentes tipos de técnicas de modelado (Belyaev, Belyaeva, Konečný, Seidler, y Vojtek, 2012). Un procedimiento que podría ser útil es la metodología de datos en panel debido a que es un modelo que combina datos de tiempo e individuos, por otra parte, al utilizar datos de panel, se pueden incluir efectos fijos o aleatorios para capturar características individuales no observadas que pueden influir en las relaciones entre las variables de interés. Esto ayuda a reducir el sesgo causado por variables no controladas en estudios transversales o de series temporales simples.

1.2.3. Justificación Práctica

En primer lugar, la solución de este problema beneficiará al ente regulador y al sistema financiero. Al monitorear estos entornos se permite al regulador actuar de forma preventiva mitigando riesgos y costos, tanto para el sistema como para una entidad en particular y alertando sobre resultados adversos imprevistos causados por distintos riesgos e informando sobre el capital necesario para absorber pérdidas en caso de alteraciones mayores (Cóndor, 2020). Y, por otro lado, el estudio podrá contribuir a más procesos de investigación en el sector financiero del país o a la implementación de técnicas para que las entidades bancarias puedan calibrar el riesgo y anticipar las posibles pérdidas con condiciones macro que afectarían al cumplimiento de créditos otorgados.

1.3. Objetivos

1.3.1. Objetivo General

Estimar la severidad de pérdida en el sistema bancario del Ecuador (bancos privados) incluyendo factores macroeconómicos.

1.3.2. Objetivos Específicos

- Determinar el método estadístico que genere un mejor modelo a partir de criterios conceptuales y prácticos de modelamiento.
- Estimar y validar el modelo idóneo a través de pruebas estadísticas y realizar un backtesting de ajuste del modelo.
- Identificar e interpretar las variables macroeconómicas que inciden en la estimación de la severidad de pérdida.

Capítulo 2

Marco Teórico

2.1. Conceptos financieros

Según la Superintendencia de Bancos (2021), Libro I, Título IX, Capítulo II, Sección I, Artículo 2. - Se considerarán las siguientes definiciones para efecto de la aplicación de las disposiciones contenidas en este capítulo:

2.1 Riesgo de crédito.- Es la posibilidad de pérdida debido al incumplimiento del prestatario o la contraparte en operaciones directas, indirectas o de derivados que conlleva el no pago, el pago parcial o la falta de oportunidad en el pago de las obligaciones pactadas;

2.2 Incumplimiento.- Es no efectuar el pago pactado dentro del período predeterminado; o, efectuarlo con posterioridad a la fecha en que estaba programado, o, en distintas condiciones a las pactadas en el contrato;

2.3 Probabilidad de incumplimiento (PD).- Es la posibilidad de que ocurra el incumplimiento parcial o total de una obligación de pago o el rompimiento de un acuerdo del contrato de crédito, en un período determinado;

2.4 Nivel de exposición del riesgo de crédito (EAD).- Es el valor presente (al momento de producirse el incumplimiento) de los flujos que se espera recibir de las operaciones crediticias;

2.5 Tasa de recuperación (r).- Es el porcentaje de la recaudación realizada sobre las operaciones de crédito que han sido incumplidas;

2.6 Severidad de la pérdida (1 - r).- Es la medida de la pérdida que sufriría la institución controlada después de haber realizado todas las gestiones para recuperar los créditos que han sido incumplidos, ejecutar las garantías o recibirlas como dación en pago. La severidad de la pérdida es igual a (1 - Tasa de recuperación);

2.7 Pérdida esperada (PE).- Es el valor esperado de pérdida por riesgo crediticio en un horizonte de tiempo determinado, resultante de la probabilidad de incumplimiento, el nivel de exposición en el momento del incumplimiento y la severidad de la pérdida:

$$PE = EAD * PD * (1 - r) \quad (2.1)$$

A partir de estas cortas definiciones, se puede aclarar de mejor manera el tema puntual de este proyecto, la Severidad de Pérdida.

2.1.1. Severidad de Pérdida

La LGD, **Pérdida Dado el Incumplimiento** o **Severidad de la Pérdida** (en inglés Loss Given Default), es una pérdida económica que corresponde a la proporción de una deuda que no es recuperada en caso de un default del acreditado (Cuevas, 2019).

Cuevas (2019), ejemplifica el significado de la LGD a partir de la ecuación $1 - r$, en la figura 2.2 se muestra la historia de un crédito. El punto 1 corresponde al inicio de un crédito por un monto de 200. Luego de sucesivos pagos, el deudor cae en default en el punto 2, con una exposición (EAD) de 100. En el punto 3 el banco gasta 10 en acciones de cobranza, gracias a lo cual el deudor paga 50 de lo que debía.

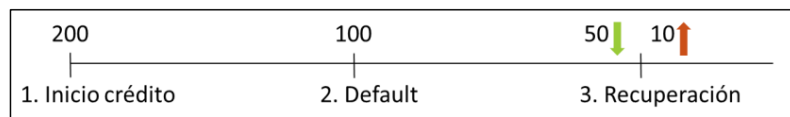


Figura 2.1: Ejemplo de la historia de un crédito. **Fuente:** (Cuevas, 2019)

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

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

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

A diferencia de otros modelos, en este proyecto se estimará la LGD con la ecuación (2.5) con el fin de implementar el modelo utilizando datos de acceso público.

Por tanto, la severidad de pérdida es:

$$LGD = \frac{PE}{EAD * PD} \quad (2.2)$$

2.1.2. Factores Microeconómicos

En esta sección, se presenta la interpretación y formulación de los indicadores microeconómicos relevantes para este estudio, obtenidos del Balance General para bancos privados, información proporcionada de la Superintendencia de Bancos:

<i>Indicador</i>	<i>Definición general</i>	<i>Definición técnica e Interpretación</i>	<i>Fórmula de cálculo</i>
GRADO DE SOLVENCIA	La solvencia financiera se refiere a la capacidad de una empresa o individuo para cumplir con sus obligaciones financieras a largo plazo, es decir, la capacidad de pagar sus deudas y mantener una situación financiera estable y sostenible en el tiempo.	Mide la relación entre el patrimonio técnico y los activos que suponen un riesgo. Un mayor valor de este índice, refleja una mejor situación de la entidad (SEPS, 2017).	Patrimonio Técnico Constituido / Activos y contingentes ponderados por riesgo
SUFICIENCIA PATRIMONIAL	La suficiencia patrimonial es el capital que es necesario para que un banco u otra entidad financiera pueda hacerle frente a los riesgos de su operación.	Mide la porción de los activos inmovilizados que estaría cubierta por el patrimonio. Una mayor cobertura patrimonial de activos, significa una mejor posición (SEPS, 2017).	(Patrimonio + Resultados) / Activo improductivos
ÍNDICE DE MOROSIDAD	El índice o ratio de morosidad es un indicador utilizado en el mundo de la banca. Este mide el volumen de operaciones que han sido consideradas morosas, con relación al volumen de préstamos y créditos concedidos por dicha entidad (Coll, 2021).	Mide la porción de los activos inmovilizados que estaría cubierta por el patrimonio. Una mayor cobertura patrimonial de activos, significa una mejor posición (SEPS, 2017).	Cartera Improductiva / Cartera Bruta
EFICIENCIA OPERATIVA	La eficiencia operativa se refiere a la capacidad de una empresa para utilizar sus recursos de manera óptima para producir bienes o servicios y generar beneficios de forma rentable y sostenible.	Miden la relación de los Gastos de Operación frente al Activo Promedio. Un menor valor de este índice, reflejan una mejor situación de la entidad (SEPS, 2017).	Gastos de Operación / Activo Total Promedio
GRADO DE ABSORCIÓN DEL MARGEN FINANCIERO	El grado de absorción del margen financiero es un indicador financiero que se utiliza para evaluar la capacidad de una entidad financiera para cubrir sus costos operativos y generar beneficios a través de los ingresos por intereses.	Miden la relación de los Gastos de Operación frente al Margen Neto Financiero. Un menor valor de este índice, reflejan una mejor situación de la entidad (SEPS, 2017).	Gastos de Operación / Margen Neto Financiero
EFICIENCIA OPERATIVA DEL PERSONAL	El grado de absorción del margen financiero es un indicador financiero que se utiliza para evaluar la capacidad de una entidad financiera para cubrir sus costos operativos y generar beneficios a través de los ingresos por intereses.	Miden la relación de los Gastos de Personal frente al Activo Promedio. Un menor valor de este índice, refleja una mejor situación de la entidad (SEPS, 2017).	Gastos de Personal Estiamdos / Activo Total Promedio
ROE	El ROE (Return on Equity) es un indicador financiero que mide la rentabilidad de una empresa en relación con el capital invertido por los accionistas. En otras palabras, el ROE mide la cantidad de beneficios que la empresa genera en relación con el dinero que los accionistas han invertido en ella.	Miden la relación de la Rentabilidad frente al Patrimonio Promedio. Un mayor valor de este índice, refleja una mejor situación de la entidad (SEPS, 2017).	Rendimiento / Patrimonio Promedio

Cuadro 2.1: Descripción de variables Microeconómicas

<i>Indicador</i>	<i>Definición general</i>	<i>Definición técnica e Interpretación</i>	<i>Fórmula de cálculo</i>
ROA	El ROA (Return on Assets) es un indicador financiero que mide la rentabilidad de una empresa en relación con sus activos totales. En otras palabras, el ROA mide la eficiencia con la que la empresa utiliza sus activos para generar ganancias.	Miden la relación de la Rentabilidad frente al Activo Promedio. Un mayor valor de este índice, refleja una mejor situación de la entidad (SEPS, 2017).	$\text{Rentabilidad} / \text{Activo Total Promedio}$
MARGEN DE INTERMEDIACIÓN SOBRE ACTIVO	El margen de intermediación sobre activo es un indicador financiero que mide la eficiencia de una empresa en la generación de ingresos por sus actividades de intermediación financiera en relación con sus activos totales.	Mide la rentabilidad de la gestión operativa en relación con los activos generadores de ingresos. La relación entre más alta es mejor. Mientras mayor es el indicador, significa que la ganancia por el proceso de intermediación (colocación de préstamos en función a la cantidad de depósitos) es más eficiente en relación al activo promedio (SEPS, 2017).	$\text{Margen de Intermediación} / \text{Activo Promedio}$
MARGEN DE INTERMEDIACIÓN SOBRE PATRIMONIO	El margen de intermediación sobre patrimonio es una medida que se utiliza en el sector financiero para evaluar la eficiencia y rentabilidad de una entidad bancaria. Se calcula dividiendo el margen de intermediación (es decir, la diferencia entre los intereses que se cobran por los préstamos y los intereses que se pagan por los depósitos) entre el patrimonio neto de la entidad.	Mide la rentabilidad de la gestión operativa en relación al patrimonio promedio. La relación entre más alta es mejor. Mientras mayor es el indicador, significa que la ganancia por el proceso de intermediación (colocación de préstamos en función a la cantidad de depósitos) es más eficiente en relación al patrimonio promedio (SEPS, 2017).	$\text{Margen de Intermediación} / \text{Patrimonio Promedio}$
ÍNDICE DE LIQUIDEZ	El índice de liquidez es una medida financiera que se utiliza para evaluar la capacidad de una empresa para cumplir con sus obligaciones de corto plazo (es decir, aquellas que vencen en un plazo de 12 meses o menos) utilizando sus activos corrientes (es decir, aquellos que se pueden convertir en efectivo en un plazo de un año o menos).	Mide la relación de los Fondos Disponibles frente a los Depósitos a Corto Plazo. Un mayor valor de este índice, refleja una mejor situación de la entidad (SEPS, 2017).	$\text{Fondos Disponibles} / \text{Depósitos a Corto Plazo}$
ÍNDICE DE CAPITALIZACIÓN NETO	El índice de capitalización neto es una medida financiera que se utiliza para evaluar la solvencia de una empresa y su capacidad para cumplir con sus obligaciones a largo plazo. Este índice compara la deuda a largo plazo de la empresa con su patrimonio neto.	Mide la relación entre el capital neto y los activos sujetos a riesgo, es una medición de la solvencia. Mientras más alto el indicador, representa que la entidad está teniendo una mayor ganancia a pesar de la variabilidad que podrían presentar los activos en riesgo (SEPS, 2017).	FK / FI
ÍNDICE DE APALANCAMIENTO	El índice de apalancamiento es una medida financiera que se utiliza para evaluar la cantidad de deuda que una empresa ha utilizado para financiar sus operaciones en comparación con su capital propio o patrimonio neto.	Mide la relación entre los Activos Totales y el Patrimonio, este valor entre más elevado indica un riesgo más bajo. Este indicador mide la calidad de los activos que puede generar el Patrimonio Invertido (SEPS, 2017).	$\text{Activos Totales} / \text{Patrimonio}$
ÍNDICE DE COBERTURA	El índice de cobertura es una medida financiera que se utiliza para evaluar la capacidad de una empresa para cumplir con sus obligaciones financieras a largo plazo. Este índice compara la cantidad de ingresos que la empresa genera con sus gastos e intereses sobre la deuda a largo plazo.	Mide el nivel de protección que la entidad asume ante el riesgo de cartera morosa. La relación mientras más alta es mejor (SEPS, 2017).	$\text{Provisiones} / \text{Cartera Improductiva}$

Cuadro 2.2: Descripción de variables Microeconómicas

2.1.3. Factores Macroeconómicos

Los factores macroeconómicos se refieren a las tendencias y condiciones económicas a nivel nacional o global, que tienen un impacto significativo en la economía en general. Algunos de los factores macroeconómicos más importantes son:

Índice de desempleo urbano

El índice de desempleo urbano es un indicador económico que mide la proporción de personas en edad de trabajar que se encuentran sin empleo en zonas urbanas.

$$\frac{\text{número de personas desempleadas en áreas urbanas}}{\text{población activa urbana}}$$

Este índice es una medida importante de la salud económica de una ciudad o área urbana, ya que, indica la cantidad de personas que no están ganando un ingreso y, por lo tanto, pueden estar experimentando dificultades financieras. Un índice de desempleo urbano alto puede indicar que hay una falta de oportunidades de empleo en una ciudad o que la economía está en una recesión.

El índice de desempleo urbano también es útil para los gobiernos y los responsables de la política económica, ya que, pueden utilizarlo para tomar decisiones informadas sobre cómo estimular la creación de empleo y mejorar la situación económica en áreas urbanas.

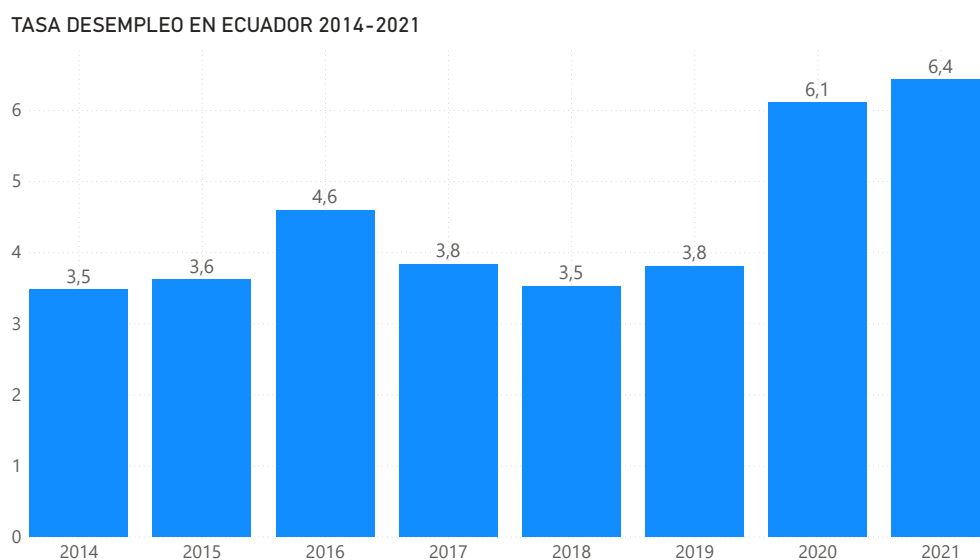


Figura 2.2: Tasa de desempleo en el Ecuador 2014-2021. **Fuente:** https://es.theglobaleconomy.com/ecuador/unemployment_rate/

Índice de precios al consumidor (IPC)

El IPC es un indicador económico mensual, nacional, que mide los cambios en el tiempo del nivel general de los precios, correspondientes al consumo final de bienes y servicios de los hogares de estratos de ingreso: alto, medio y bajo, residentes en el área urbana del país (INEC, 2022). El IPC se utiliza para medir la inflación, que es el aumento sostenido y generalizado del nivel de precios de bienes y servicios en una economía.

Para calcular el IPC, se recopilan los precios de los bienes y servicios de la canasta en un período determinado y se los compara con los precios de la misma canasta en un período anterior. Luego, se calcula el porcentaje de cambio en el precio de la canasta, y este porcentaje se utiliza para calcular el IPC.

El IPC se utiliza ampliamente como una medida de la inflación y se utiliza para ajustar los salarios, los precios de los contratos y otros pagos que están vinculados a la inflación. También es una herramienta importante para los responsables de la política económica, ya que, pueden utilizarla para tomar decisiones informadas sobre las tasas de interés y otras medidas destinadas a controlar la inflación.

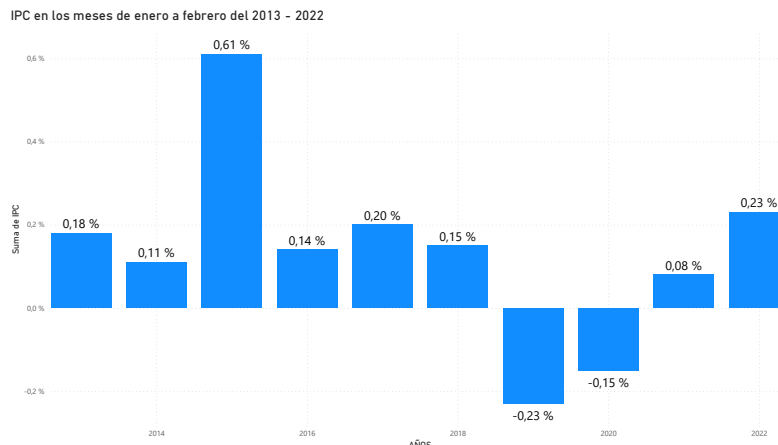


Figura 2.3: IPC en los meses de febrero en el Ecuador 2013-2022. **Fuente:** Ecuador en cifras. Boletín Técnico N.º 09-2022-IPC

Índice de confianza al consumidor

El índice de confianza del consumidor (ICC) es un indicador económico que mide la opinión de los consumidores sobre la salud de la economía y su propia situación financiera. El ICC se basa en encuestas que miden la opinión de los consumidores sobre factores como la situación económica actual, las expectativas futuras, el

empleo y la inflación.

El ICC se utiliza como una medida de la confianza del consumidor en la economía en general y en su propia situación financiera. Un ICC alto indica que los consumidores tienen una opinión positiva sobre la economía y que están dispuestos a gastar dinero, lo que puede ser un indicador de un aumento en el consumo y un crecimiento económico. Por otro lado, un ICC bajo indica que los consumidores tienen una opinión negativa sobre la economía y pueden ser menos propensos a gastar dinero, lo que puede indicar una disminución en el consumo y un menor crecimiento económico.

El ICC es utilizado por los gobiernos, los responsables de la política económica, las empresas y los inversores para evaluar la salud de la economía y tomar decisiones informadas sobre la inversión y el gasto.

Índice de actividad coyuntural (IDEAC)

El IDEAC es un indicador económico de periodicidad mensual, estructurado con variables físicas de producción que señalan la tendencia de la actividad económica coyuntural. Está elaborado como un indicador de cuántum de la producción y tiene una estructura matemática similar a un índice tipo Laspeyres, es decir, se fija el sistema de ponderaciones a un período base (Banco Central del Ecuador, 2012).

El objetivo de este indicador es presentar una visión instantánea de la coyuntura económica basándose en la evolución de las distintas actividades, así como prever su comportamiento inmediato y sugerir posibles correcciones. Este indicador está diseñado para reflejar las variaciones reales de la producción (Banco Central del Ecuador, 2012).

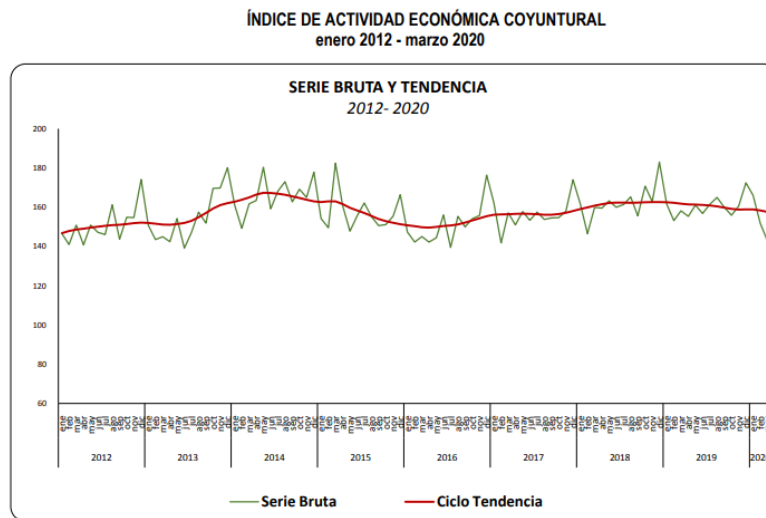


Figura 2.4: Índice de actividad económica coyuntural **Fuente:** Banco Central del Ecuador 2012-2020. <https://www.bce.fin.ec/index.php/component/k2/item/313-indice-de-actividad-econ%C3%B3mica-coyuntural-ideac>

WTI Precio del petróleo

El precio WTI del petróleo se refiere al precio del barril de petróleo crudo de tipo West Texas Intermediate (WTI), que es uno de los principales tipos de petróleo crudo que se negocian en el mercado internacional. El precio del WTI se utiliza como una referencia para determinar el precio de otros tipos de petróleo crudo en todo el mundo. **Fuente:** US. Energy Information Administration <https://www.eia.gov/>

El precio del WTI es determinado por la oferta y la demanda del mercado internacional de petróleo, y está influenciado por una variedad de factores, incluyendo la producción de petróleo de los principales productores, la demanda global de petróleo, los cambios en la política energética de los gobiernos y los eventos geopolíticos que pueden afectar la oferta y la demanda de petróleo.

El precio del petróleo WTI es seguido de cerca por los inversores, los responsables de la política energética y las empresas del sector energético, ya que, tiene un impacto significativo en la economía global y puede afectar a los precios de otros productos y servicios, como el combustible para vehículos y la calefacción de hogares y empresas.

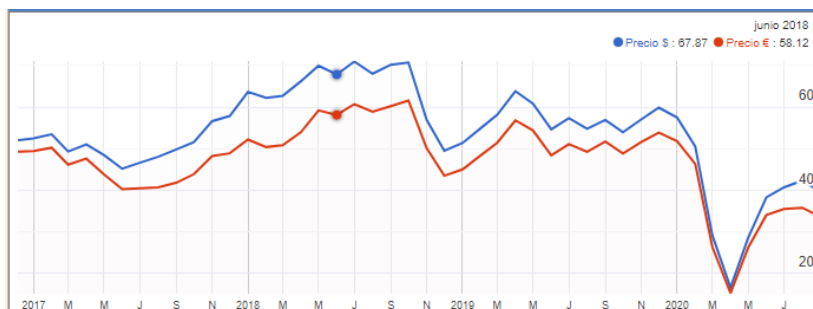


Figura 2.5: Evolución del Precio petróleo West Texas Intermediate **Fuente:** <https://datosmacro.expansion.com/materias-primas/petroleo-wti>

2.1.4. Riesgo País

El Índice de Riesgo País es un indicador financiero que mide la percepción del riesgo de invertir en un país determinado. Se calcula tomando en cuenta una serie de factores económicos, políticos y financieros que afectan la estabilidad económica del país y la capacidad del gobierno para pagar su deuda externa y es elaborado por el banco estadounidense J.P. Morgan.

El Índice de Riesgo País de J.P. Morgan es utilizado por los inversores y los gobiernos como una herramienta para evaluar el riesgo de invertir en un país determinado y para tomar decisiones informadas sobre la inversión y la financiación de proyectos en el país.

Es importante tener en cuenta que el índice de riesgo país es una medida subjetiva y no siempre refleja la situación económica real de un país. Además, los factores que influyen en el índice pueden cambiar rápidamente, lo que significa que el índice puede fluctuar en función de la percepción de los inversores y no necesariamente de la realidad económica del país en cuestión.

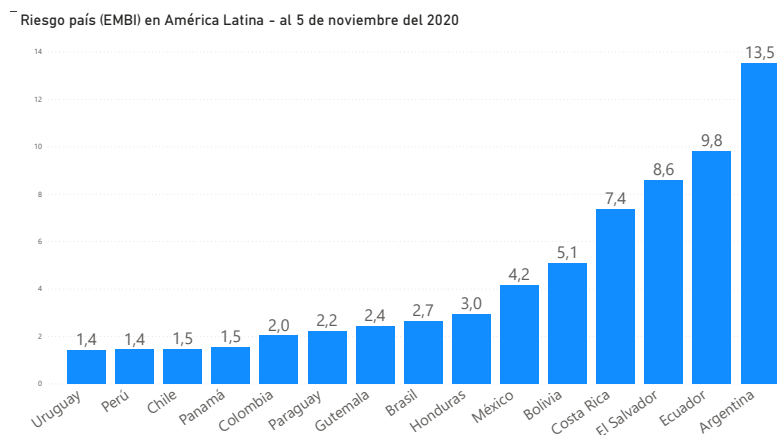


Figura 2.6: Riesgo País en América latina 2020. **Fuente:** Revista Panorámica 2020.

El índice se basa en la diferencia entre las tasas de interés de los bonos del Tesoro de los Estados Unidos y los bonos emitidos por el país en cuestión. Cuanto mayor sea la brecha entre estas dos tasas, mayor será el riesgo percibido de invertir en ese país. Así, hasta 2020 Ecuador mantenía un riesgo país alto en relación con América Latina.

PIB real

PIB real, también conocido como Producto Interno Bruto real, es una medida de la producción económica de un país que tiene en cuenta la inflación a diferencia del PIB nominal, entonces el PIB real proporciona una medida más precisa de la producción económica en términos de poder adquisitivo.

El PIB real se utiliza comúnmente como una medida del crecimiento económico a largo plazo de un país. También se utiliza para comparar la producción económica entre países y para analizar los cambios en la producción económica dentro de un país a lo largo del tiempo.

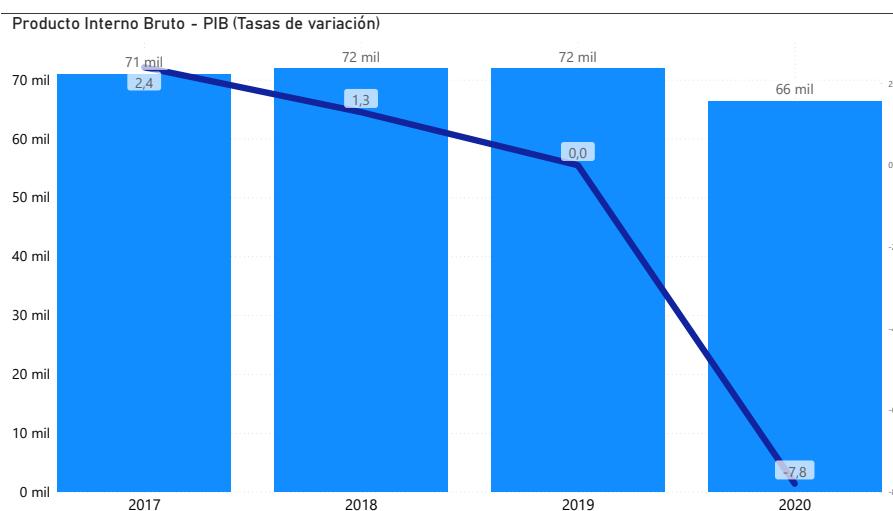


Figura 2.7: PIB 2017-2020 Fuente: (Banco Central del Ecuador, 2020)

2.2. Estructuras de representación de datos

Las bases de datos pueden ser de diferentes tipos, [56] define las características especiales de algunas bases de datos que se suelen encontrar.

2.2.1. Datos de Series de tiempo

Una **serie temporal** se define como una sucesión de observaciones que toma una variable a través del tiempo en orden cronológico. En este caso, en el campo de la Economía, las series de tiempo nos permiten visualizar el comportamiento de valores con frecuencia diaria, semanal, mensual, trimestral y anual de factores micro y macroeconómicos como: el Producto Interno Bruto (PIB), índice de desempleo, Indicadores de Solvencia, índice de Morosidad de una cartera, entre otros. Así como en diferentes ciencias, donde se desee medir el comportamiento de un fenómeno de estudio. Por otra parte, los eventos pasados pueden influir sobre los eventos futuros y los comportamientos rezagados son frecuentes en las ciencias sociales, el tiempo es una dimensión importante en las bases de datos de series de tiempo (Wooldrige, 2013).

2.2.2. Sección cruzada o datos de corte transversal

Una base de **datos de corte transversal** consiste en una muestra de individuos, hogares, empresas, ciudades, estados, países u otras unidades, tomada en algún punto dado en el tiempo. Una característica importante de los datos de corte transversal es que a menudo puede suponerse que han sido obtenidos de la población mediante un muestreo aleatorio (Wooldrige, 2013).

2.2.3. Datos en Panel

Los **datos en panel**, son una colección de datos que aprovecha la información tanto de los datos de corte transversal, como la de los datos de tiempo, en donde se asocia una serie de tiempo para cada elemento de la sección cruzada. Por lo tanto, se puede decir que los datos en panel presentan datos más grandes e informativos y genera un mayor grado de libertad (Achmad, Yoyon y Achmad, 2019).

La forma general de los modelos de regresión de datos en panel se pueden ver en la siguiente ecuación:

$$Y_{it} = \alpha_{it} + \beta' X_{it} + \varepsilon_{it} \quad (2.3)$$

Donde:

- Y_{it} : variable respuesta para el individuo i y periodo de tiempo t .
- α_{it} : valor constante/intercepción del individuo i y periodo de tiempo t .
- β' : $(\beta_1, \beta_2, \dots, \beta_k)$ variable de regresión.
- X_{it} : el valor de la variable predictora del i -individuo y t -periodo de tiempo.
- ε_{it} : variable inobservable (término de error).
- $i:1,2,\dots,N$ significa la i -ésima unidad transversal.
- $t:1,2,\dots,T$ periodo de tiempo.

2.3. Metodología para la estimación del modelo de regresión de datos en panel

Al realizar la estimación de parámetros, el modelo depende de las suposiciones sobre las intersecciones y los coeficientes de pendiente. El uso de datos en panel

permite que ocurran diferentes intersecciones y coeficientes de pendiente en cada individuo y cada período de tiempo. En este concepto, hay tres enfoques que se pueden utilizar: Modelo de Efectos Comunes, Modelo de Efectos Aleatorios y Modelo de Efectos Fijos (Widarjono, 2013).

2.3.1. Modelo de Efectos Comunes

El enfoque con el modelo de Efectos Comunes asume que las intersecciones y los coeficientes de pendiente tienen el mismo valor para todos los individuos en todos los períodos de tiempo. En otras palabras, este modelo ignora las dimensiones individuales y el tiempo, se supone que el comportamiento de los datos es el mismo en varios periodos (Achmad, et al, 2019). Este método puede utilizar el enfoque de mínimos cuadrados ordinarios (MCO u OLS) o la técnica de Mínimos Cuadrados para estimar el modelo de datos en panel.

$$Y_{it} = \alpha + \beta' X_{it} + \varepsilon_{it} \quad (2.4)$$

Para $i = 1, \dots, N$ y $t = 1, \dots, T$. Donde N = número de individuos o sección transversal y T = número de periodos de tiempo.

El modelo de efectos comunes es útil cuando se desea estimar el efecto medio de una variable independiente en la población en lugar de estimar los efectos individuales de cada individuo. Sin embargo, este modelo no es adecuado cuando se desea modelar las diferencias individuales en los efectos de las variables independientes, ya que, asume que los efectos son los mismos para todos los individuos en la muestra.

Es importante tener en cuenta que el modelo de efectos comunes se utiliza principalmente en el análisis de datos longitudinales o de panel y no se aplica a datos transversales, ya que, en los datos transversales no hay información sobre la evolución temporal de la variable de respuesta y las variables independientes.

2.3.2. Modelo de Efectos Fijos

Este enfoque FE (fixed effect) cree que las diferencias en las características entre los individuos están representadas en las intersecciones. Por lo tanto, las intersecciones para cada individuo serán diferentes, pero los coeficientes de pendiente permanecerán constantes en todos los períodos de tiempo (Achmad, et al, 2019). Por otra

parte, el supuesto de modelización que produce un intercepto constante para cada sección transversal y tiempo se considera menos realista, por lo que, se necesitan más modelos para captar la diferencia. Los efectos fijos suponen que las diferencias entre individuos (sección transversal) pueden acomodarse a los diferentes interceptos. Para estimar el modelo de efectos fijos con un intercepto diferente entre individuos, se utiliza la técnica de las variables ficticias, este modelo de estimación suele denominarse también técnica de la variable ficticia de mínimos cuadrados (LSDV) (Zulfikar, 2018).

La ecuación de regresión de datos en panel del modelo de efectos fijos es la siguiente:

$$Y_{it} = \alpha_i + \beta' X_{it} + \varepsilon_{it} \quad (2.5)$$

Para $i = 1, \dots, N$ y $t = 1, \dots, T$. Donde N = número de individuos o sección transversal y T = número de periodos de tiempo.

El modelo de efectos fijos es útil cuando se desea modelar las diferencias individuales en los efectos de las variables independientes, ya que, permite estimar los efectos de las variables independientes para cada individuo en la muestra. Sin embargo, este modelo no puede estimar efectos promedio para la población, ya que, el efecto de cada individuo se considera constante y específico para ese individuo.

2.3.3. Modelo de Efectos Aleatorios

En este enfoque RE (random effects) las diferencias en las características individuales se acomodan mediante términos de error. Los términos de error pueden correlacionarse entre individuos y entre tiempos. La ventaja de utilizar un modelo de Efecto Aleatorio es para eliminar la heteroscedasticidad. Este modelo también se lo denomina como modelo de componentes de error (ECM) o técnica de mínimos cuadrados generalizados (GLS). A diferencia de los modelos anteriores, es que el RE no usa el principio de mínimos cuadrados ordinarios, sino que usa el principio de máxima verosimilitud. En el modelo de efectos aleatorios, los residuos pueden estar interconectados entre el tiempo y entre individuos o secciones transversales. Por lo tanto, este modelo asume que existe una diferencia de intercepto para cada individuo y el intercepto es una variable aleatoria. Así que en el modelo de efectos aleatorios hay dos componentes residuales. El primero es el residual como un todo, donde el residual es una combinación de sección transversal y serie de tiempo. El segundo residuo es un residuo individual que es una característica aleatoria de la

i-ésima unidad de observación y permanece en todo momento (Zulfikar, 2018).

La ecuación de regresión de datos en panel del modelo de efectos aleatorios es la siguiente:

$$Y_{it} = \alpha + \beta' X_{it} + W_{it} \quad (2.6)$$

α es la media de la intercepción de la sección transversal y la serie temporal de los términos de error. $W_{it} : \mu_i + \varepsilon_{it}$, donde μ_i es la desviación de la sección transversal del error aleatorio, que explica las diferencias entre los individuos y entre sí (Achmad, et al, 2019).

El modelo de efectos aleatorios es útil cuando se desea modelar tanto las diferencias individuales en los efectos de las variables independientes como los efectos promedio de las variables independientes para la población. Además, este modelo permite estimar la varianza de los efectos aleatorios y la covarianza entre los efectos aleatorios y las variables independientes.

Finalmente, estos modelos RE Y FE sirven para controlar la heterogeneidad inobservable, en particular cuando esta es constante en el tiempo y está correlacionada con las variables independientes. Esta constante puede ser eliminada de los datos a través de la diferenciación, por ejemplo, teniendo una primera diferencia con la cual se eliminarán los componentes del modelo invariables en el tiempo. [55]

2.4. Método de selección del modelo de regresión de Datos en Panel

Ahora bien, dados los conceptos de los distintos modelos de regresión de datos en panel, debemos seleccionar el modelo más adecuado a trabajar, para esto existen pruebas que deben ser aplicadas. Estas son:

- Test de Chow
- Test de Hausman
- Test de Multiplicadores de Lagrange

2.4.1. Test de Chow

El test de Chow es un estadístico de prueba desarrollado por Gregory Chow, que se utiliza para probar si los coeficientes en dos modelos de regresión diferentes en

diferentes conjuntos de datos son iguales. Para el presente proyecto, esta prueba se utiliza para elegir el mejor modelo entre los modelos de Efectos Comunes y Efectos Fijos (Greene, 2000).

$$H_0 : \alpha_1 = \alpha_2 = \dots = \alpha_n = 0 \text{ (MCO)}$$

$$H_1 : \alpha_i \neq 0 \text{ (FE)}$$

2.4.2. Test de Hausman

La prueba de Hausman se utiliza para elegir un mejor modelo entre FE y RE. El estadístico de contraste es:

$$(\hat{\beta}_{FE} - \hat{\beta}_{RE})' [Var \hat{\beta}_{FE} - \hat{\beta}_{RE}]^{-1} (\hat{\beta}_{FE} - \hat{\beta}_{RE}) \sim \chi_{k-1}^2$$

donde k es el número de parámetros del modelo de efectos fijos. Si el valor del estadístico es superior al de las tablas, se rechaza la hipótesis nula de efectos aleatorios y, por tanto, elegiríamos el modelo de efectos fijos (Greene, 2000).

2.4.3. Test de Multiplicadores de Lagrange

La prueba LM se utiliza para elegir un mejor modelo entre MCO y RE. La prueba del multiplicador de Lagrange tiene la función de determinar la mejor estimación, ya sea utilizando un efecto aleatorio o no. (Zulfikar, 2018) No es necesario realizar la prueba LM si se determina que FE es el mejor en la prueba Chow y la prueba Hausman. Esta prueba solamente se realiza cuando el mejor modelo encontrado en el Chow Test es FE mientras que en el Hausman Test el mejor modelo es RE (Achmad, et al, 2019)

$$H_0 : \sigma_\varepsilon^2 = 0 \text{ (CE)}$$

$$H_1 : \sigma_\varepsilon^2 \neq 0 \text{ (RE)}$$

donde:

σ_ε^2 : varianza de los residuos

Se resume en el siguiente gráfico las pruebas de selección del mejor modelo de estimación para datos en panel:

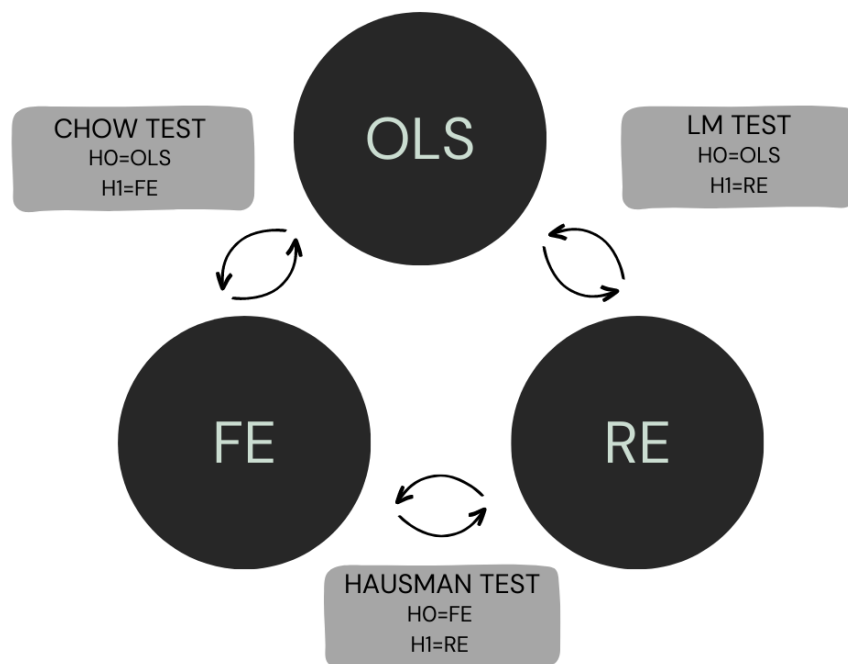


Figura 2.8: Elección de estimación para la regresión de datos en panel. **Fuente:** (Zulfikar, 2018)

2.5. Supuestos de prueba e idoneidad del modelo

Para obtener el estimador del modelo de parámetros en el enfoque CE y FE, se utiliza el método de mínimos cuadrados ordinarios (MCO u OLS). La suposición clásica se realiza para obtener el mejor estimador OLS conocido como BLUE (Best Linear Unbiased Estimator/Mejor estimador lineal imparcial).

Las pruebas de suposición clásicas incluyen:

1. Prueba de multicolinealidad
2. Prueba de normalidad
3. Prueba de heteroscedasticidad
4. Prueba de autocorrelación

De las cuatro pruebas, solo se deben cumplir las pruebas 3 y 4 para obtener el mejor estimador (BLUE) (Gujarati, 2004).

2.5.1. Prueba de Multicolinealidad

A la hora de estimar un modelo económico, los datos disponibles sobre las variables explicativas o regresores pueden presentar un alto grado de correlación, especialmente en un contexto de series temporales y con series macroeconómicas. Por ejemplo, la población y el PIB en general suelen estar altamente correlacionados. A este fenómeno se le conoce como multicolinealidad. En algún caso puede que los datos de una variable se obtengan como resultado de una identidad contable o de una combinación lineal exacta entre otros regresores (Widarjono, 2013).

Es importante tener en cuenta que la multicolinealidad no siempre es un problema en un modelo de regresión y depende del objetivo del análisis y de las variables involucradas. Uno de los supuestos utilizados en OLS es la ausencia de una fuerte relación entre las variables predictoras, por lo que la presencia de multicolinealidad hace que el modelo tenga una gran variante (Widarjono, 2013). Sin embargo, la multicolinealidad todavía puede producir estimadores BLUE. Una forma de detectar la multicolinealidad es utilizar la muestra de coeficiente de correlación (Gujarati, 2004).

2.5.2. Prueba de normalidad

Como se mencionó anteriormente, esta prueba es opcional, debido a que la regresión asume que los errores ε_{it} siguen una distribución normal (Gujarati, 2004). La regresión OLS no requiere que el término de error siga una distribución normal para producir estimaciones imparciales con la varianza mínima. Sin embargo, cumplir con esta suposición le permite realizar pruebas de hipótesis estadísticas y generar intervalos de confianza e intervalos de predicción confiables (Frost, 2020).

H_0 : (errores de los residuos están distribuidos normalmente)

H_1 : (errores de los residuos, no está distribuidos normalmente)

La prueba de normalidad estadística adoptó la fórmula de Jarque-Bera como podemos ver en la siguiente ecuación: (Gujarati, 2004).

$$JB = NT \left(\frac{S_k^2}{6} + \frac{(K_r - 3)^2}{24} \right)$$

Donde,

S_k : coeficiente de asimetría

K_r : altura(curtosis)

Si el valor de $JB < \chi_{tabla}^2 = \chi_{(a,2)^2}$ o el $p - valor > 0,05$, se acepta H_0 esto aclara que el error se distribuye normalmente.

2.5.3. Prueba de heteroscedasticidad

La suposición de que la varianza de error tiene una naturaleza constante (homoscedástica) es muy importante para producir el mejor estimador de MCO. (Achmad, et al, 2019)

H_0 : El error no tiene heteroscedasticidad

H_1 : El error tiene heteroscedasticidad

El método Glejser se utiliza para estadístico de prueba de heteroscedasticidad, que consiste en hacer una regresión del valor absoluto del error con su variable independiente (Widarjono, 2013). Si β_1 no es significativo en la prueba t o el $p - valor > 0,05$, entonces se acepta la hipótesis H_0 . Esto explica que el error no contenía heteroscedasticidad o era constante.

2.5.4. Prueba de autocorrelación

Esta prueba de suposición se usa para ver si no hay una correlación serial en el error. Esta prueba es crucial para obtener un estimador BLUE OLS. Una forma de detectar la autocorrelación es utilizar un método de prueba de Durbin Watson (Achmad, et al, 2019).

$H_0 : Cov(\varepsilon_{it}, \varepsilon_{i,t-1}) = 0$ (no hay autocorrelación)

$H_1 : Cov(\varepsilon_{it}, \varepsilon_{i,t-1}) \neq 0$ (existe autocorrelación)

El estadístico de la prueba de autocorrelación de Durbin Watson como podemos ver a continuación:

$$DW = \frac{\sum_{i=1}^N \sum_{i=2}^2 (\varepsilon_{it} - \varepsilon_{i,t-1})^2}{\sum_{i=1}^N \sum_{i=2}^2 \varepsilon_{it}^2}$$

En donde en la siguiente tabla se presenta el valor crítico del test de autocorrelación:

$n < d_W < d_L$:	Existe autocorrelación positiva
$d_L \leq d_W \leq d_u$:	No se puede decidir
$d_u \leq d_W \leq 4 - d_u$:	No existe autocorrelación
$4 - d_u \leq d_W \leq 4 - d_L$:	No se puede decidir
$4 - d_L \leq d_W \leq 4$:	Existe correlación negativa

Cuadro 2.3: Valor crítico de autocorrelación por Durbin Watson

2.6. Prueba de significación de los parámetros de regresión

La prueba de significación de los parámetros se utiliza para averiguar el nivel de influencia que tiene la variable de predicción en la variable de respuesta. Se realizan dos pruebas, a saber, la prueba simultánea y la prueba parcial. (Achmad, et al, 2019)

2.6.1. Prueba simultánea

La prueba simultánea se aplica para determinar el efecto de las variables predictoras sobre la variable de respuesta simultánea. La hipótesis de la prueba simultánea es la siguiente.

$$H_0 : \beta_1 = \beta_2 = \dots = \beta_k = 0$$

(todas las variables predictoras no tienen ningún efecto sobre la variable de respuesta)

$$H_1 : \text{mínimo de una } \beta_k \neq 0$$

(hay al menos una variable predictora que afecta a la variable de respuesta)

Por lo que el estadístico F se define:

$$F_{estadstico} = \frac{R^2 / (n + K - 1)}{(1 - R^2) / (nT - n - K)}$$

Donde, R^2 es el coeficiente de determinación. Si el valor de $F_{estadstico} > F_{tabla} = F_{(a, n+K-1, nT-n-K)}$ o $p\text{-valor} \leq 0,05$, entonces se rechaza la hipótesis nula H_0 , lo que significa que en el modelo hay al menos una variable de predicción que tiene un efecto significativo sobre la variable de respuesta (Achmad, et al, 2019).

2.6.2. Prueba parcial

Se utiliza una prueba parcial para determinar el efecto de cada variable de predicción sobre la variable de respuesta. La hipótesis de la prueba parcial es

$$H_0 : \beta_k = 0$$

(la variable predictora k no influye significativamente en la variable de respuesta)

$$H_1 : \beta_k \neq 0$$

(la variable predictora k influye significativamente en la variable de respuesta)

El estadístico de la prueba parcial se formula de la siguiente manera:

$$t_{valor} = \frac{\beta_k}{se(\beta_k)}$$

Donde, si el valor de $t_{valor} > t_{tabla} = t_{(\alpha/2, nT-n-k)}$ o el $p - valor \leq 0,05$, entonces rechaza la hipótesis nula H_0 . Esto describe que la variable k-predictora tiene un efecto significativo sobre la variable de respuesta (Achmad, et al, 2019).

2.7. Backtesting en datos en panel

En primer lugar, el backtesting es un término utilizado en el modelado para referirse a probar un modelo predictivo en datos históricos. El backtesting implica retroceder en el tiempo, paso a paso, en tantas etapas como sea necesario. Por lo tanto, es un tipo especial de validación cruzada aplicada a períodos anteriores. Debido a que se conforma por una serie temporal de datos ordenados cronológicamente, en principio, igualmente espaciados en el tiempo; el pronóstico es el proceso de predecir valores futuros de una serie de tiempo en función de sus patrones observados previamente (autoregresivos), o incluyendo variables externas.(Amat, s.f)

2.7.1. Backtesting de ajuste del modelo

El backtesting en datos en panel se refiere al proceso de probar la precisión y eficacia de un modelo estadístico o econométrico utilizando datos históricos. Como se mencionó anteriormente, los datos en panel son un tipo de datos que contienen información sobre varias personas, empresas o países durante un período de tiem-

po.

Para realizar backtesting en datos en panel, el primer paso es especificar el modelo a probar. Esto implica seleccionar la variable dependiente y las variables independientes, así como cualquier variable de control que pueda ser relevante. Una vez que se especifica el modelo, se utilizan datos históricos para estimar los parámetros del modelo utilizando técnicas estadísticas como mínimos cuadrados ordinarios, efectos fijos o efectos aleatorios.

Después de estimar los parámetros del modelo, el siguiente paso es probar la precisión y eficacia del modelo comparando sus valores pronosticados con los valores reales en los datos históricos. Normalmente, esto se hace calculando varias medidas estadísticas, como la suma residual de cuadrados (RSS), el R^2 y el $R^2 - \text{ajustado}$. Estas medidas brindan información sobre qué tan bien se ajusta el modelo a los datos históricos y se pueden usar para evaluar su poder predictivo.

Es importante tener en cuenta que las pruebas retrospectivas en datos en panel pueden ser complejas debido a la gran cantidad de datos involucrados y la necesidad de tener en cuenta los efectos a nivel individual y específicos del tiempo. Además, es importante seleccionar cuidadosamente los datos históricos utilizados para la prueba retrospectiva para garantizar que sean representativos de la población subyacente y el período de tiempo de interés.

Capítulo 3

Implementación

En este capítulo se detallará la metodología antes explicada para la modelización de la Severidad de Pérdida, considerando datos de acceso público se construyó una base de periodos trimestrales desde 2017 hasta el primer trimestre de 2020, los cuales se obtuvieron de la página web de la Superintendencia de Bancos y del Banco Central.

3.1. Descripción de las variables

En el presente estudio se utilizará una base de datos en panel, es decir, variables de corte transversal y a su vez datos en el tiempo que representan indicadores microeconómicos y macroeconómicos. En donde, los indicadores microeconómicos nos muestran el desarrollo de los bancos del sector privado, y los indicadores macroeconómicos nos muestran el estado a través del tiempo de la economía de un país.

Inicialmente, la base de datos cuenta con 40 variables, de las cuales se describen las más relevantes para este proyecto.

- **Variable Dependiente:**

Para el presente proyecto, la variable dependiente es la Severidad de Pérdida. Ya que no se dispone de información de acceso público para esta variable, lo que se realizará es un variable Proxy lo cual nos permitirá utilizarla como indicador de la variable de objeto de estudio.

- **Variable Proxy de la Severidad de Pérdida** En este trabajo se opta por la siguiente forma de cálculo para la Severidad de Pérdida expuestas en las ecua-

ciones (2.1) y (2.2)

$$LGD = \frac{PE}{EAD * PD}$$

En donde, $\frac{PE}{EAD}$ será explicada por las $Provisiones_i / CarteraBruta_i$. [31]. Por tanto,

$$LGD = \frac{Provisiones_i}{CarteraBruta_i * PD}$$

Para cada banco i .

- **Estructura de la base de datos**

DATE	BANCOS GRANDES, MEDIANOS O PEQUEÑOS	PD	PROVI_DOL	CARTERA_BRUTA_DOL	PROXY_PROVISIONES/CARTERABRUTA	LGD	INDICADORES	
							MICROECONÓMICOS	MACROECONÓMICOS
mar-17	x_11							
jun-17	x_12							
sep-17	x_13							
:	:							
T	x_3T							

Figura 3.1: Estructura de la base de datos. **Fuente:** Propia

En donde:

- **Date.** Periodo de tiempo en trimestres desde marzo de 2017 hasta septiembre de 2019.
- **Bancos.** Se realizará un estudio para estimar la severidad de pérdida para Bancos Grandes, Medianos y Pequeños, clasificación establecida en la Superintendencia de Bancos.
- **Variables para la construcción de la LGD.** Como se explicó en la Variable Proxy Provisiones/Cartera bruta, los datos que se obtuvieron de los boletines mensuales del balance general
- **Variables Exógenas.** Se describieron todos sus conceptos en el capítulo II.

3.1.1. Descripción Estadística de los Datos

En esta sección se tiene como objetivo definir los indicadores económicos que han aportado a la construcción del modelo.

Indicadores Microeconómicos

<i>Indicador</i>	<i>Variable</i>
GRADO DE SOLVENCIA	gra_sol
SUFICIENCIA PATRIMONIAL	COB_PATRI_ACT_INM
ÍNDICE DE MOROSIDAD	IND_MOROSIDAD
EFICIENCIA OPERATIVA	EFICI_OPE
GRADO DE ABSORCIÓN DEL MARGEN FINANCIERO	GRADO_ABS_MARG_FINAN
EFICIENCIA OPERATIVA DEL PERSONAL	EFIC_ADM_PERS
ROE	ROE
ROA	ROA
MARGEN DE INTERMEDIACIÓN SOBRE ACTIVO	MARG_INTER_EST_PAT
MARGEN DE INTERMEDIACIÓN SOBRE PATRIMONIO	MARG_INTER_EST_ACT
ÍNDICE DE LIQUIDEZ	IND_LIQUI
ÍNDICE DE CAPITALIZACIÓN NETO	IND_CAP
ÍNDICE DE APALANCAMIENTO	IND_APALANCAMIENTO
ÍNDICE DE COBERTURA	IND_COBER

Cuadro 3.1: Variables microeconómicas

Indicadores Macroeconómicos

<i>Indicador</i>	<i>Variable</i>
Índice de desempleo urbano	IND_DESEM_URB
Índice del precio al consumidor	IPC
Índice de confianza al consumidor	ICC
Índice de actividad económica coyuntural	IDEAC
Precio del Petróleo	WTI
Riesgo País	JP_MORGAN
PIB real	VALOR_REAL_DEL_PIB

Cuadro 3.2: Variables macroeconómicas

3.1.2. Análisis de Estacionariedad

La estructura de datos en panel ya descrita, se forma de una serie temporal que se define como una secuencia ordenada de observaciones, cada una de las cuales está asociada a un momento de tiempo[35]. Por lo que el objetivo del análisis de series de tiempo es poder estudiar el comportamiento de la serie para poder realizar futuras predicciones a través del uso de los procesos estocásticos, por ello para obtener predicciones consistentes se requiere series temporales estables. Esta estabilidad para los procesos estocásticos que generan series temporales se la llama estacionariedad. El análisis de estacionariedad se lo realizará por medio de la prueba de Dickey- Fuller aumentada, la cual nos permitirá observar si una variable es estacionaria o no. Por consiguiente debemos entender que son las series estacionarias y no estacionarias.

- **Serie Estacionaria**

Una serie es estacionaria cuando es estable, esto significa que la media, la varianza y la autocorrelación de la serie no cambian a lo largo del tiempo. Esto se refleja gráficamente en que los valores de la serie tienden a oscilar alrededor de una media constante y la variabilidad con respecto a esa media también permanece constante en el tiempo. Es una serie básicamente estable a lo largo del tiempo, sin que se aprecien aumentos o disminuciones sistemáticos de sus valores (Marin, (s.f)).

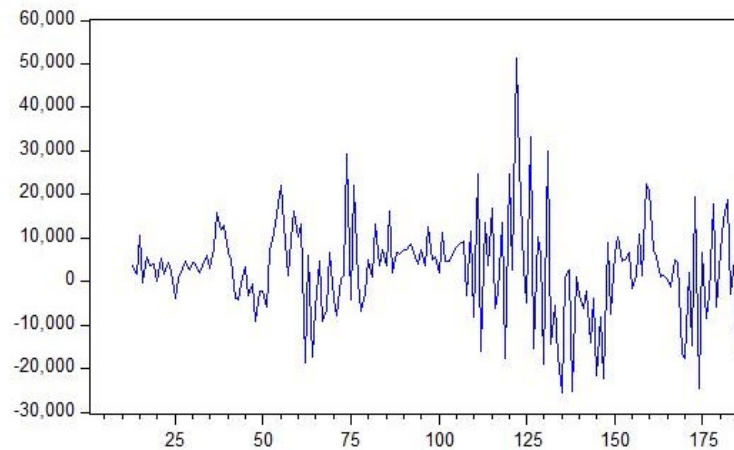


Figura 3.2: Serie de tiempo Estacionaria. **Fuente:** Propia

- **Serie No Estacionaria**

Una serie no estacionaria es una serie temporal cuyas propiedades estadísticas cambian a lo largo del tiempo. Esto significa que la media, la varianza y/o la autocorrelación de la serie pueden variar en el tiempo, por lo que los cambios en la media determinan una tendencia a crecer o decrecer a largo plazo, por lo que la serie no oscila alrededor de un valor constante (Marin, (s.f)).

Una serie no estacionaria puede presentar tendencias, estacionalidad, ciclos y/o cambios abruptos en la media o la varianza. Por lo tanto, el análisis de una serie no estacionaria es más complejo que el de una serie estacionaria, ya que los métodos estadísticos convencionales pueden no ser válidos debido a la presencia de las propiedades variables en el tiempo. Por ejemplo, la regresión lineal puede dar resultados engañosos si se aplica a una serie no estacionaria con una tendencia, ya que la relación entre las variables puede ser distorsionada por los cambios en la media.

Para analizar una serie no estacionaria, es necesario aplicar técnicas específicas de análisis de series temporales, como la diferenciación, la descomposición, la transformación o modelos ARIMA pues permiten transformar la serie no estacionaria en una serie estacionaria, lo que facilita el análisis y la modelización de la serie temporal.

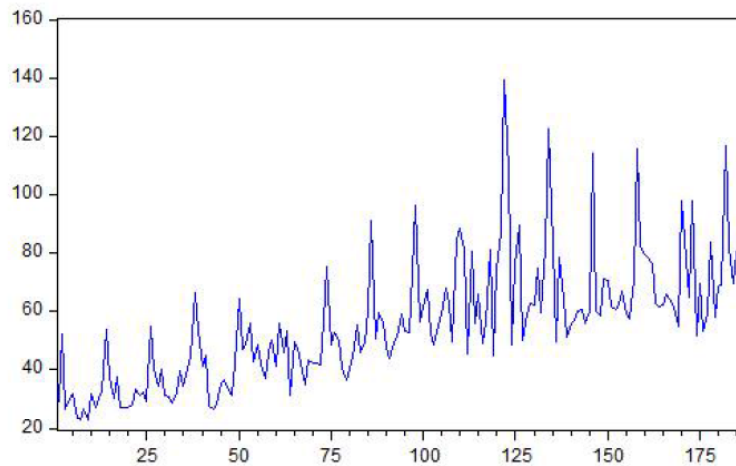


Figura 3.3: Serie de tiempo No Estacionaria. **Fuente:** Propia

- **Prueba de Dickey Fuller aumentada** La prueba de Dickey Fuller aumentada (DFA) es una prueba de raíz unitaria para la estacionariedad por lo que las raíces unitarias pueden causar resultados impredecibles al realizar el análisis de series de tiempo. Por tanto, al aplicar esta prueba a las variables de la base se puede determinar que series son estacionarias y cuáles no, así poder transformarlas a series estacionarias por medio del método de Diferenciación.

- **Diferenciación de una serie**

La diferenciación de una serie es una técnica comúnmente utilizada en el análisis de series temporales para transformar una serie no estacionaria en una serie estacionaria.

El método más general para eliminar la tendencia consiste en suponer que la tendencia evoluciona lentamente en el tiempo, de manera que en el instante t la tendencia debe estar próxima a la tendencia en el instante $t - 1$. De esta forma, si restamos a cada valor de la serie el valor anterior, la serie resultante estará aproximadamente libre de tendencia. Esta operación se denomina diferenciación de la serie y consiste en pasar de la serie original x_t a la serie y_t mediante:

$$y_t = x_t - x_{t-1}$$

Donde, x_t es el valor de la serie en el tiempo t , y y_t es el valor de la serie diferenciada en el tiempo t .

De este modo, la serie diferenciada resulta ser estacionaria (Marin, (s.f)).

3.2. Construcción del modelo de datos en panel

Para la construcción del modelo, se selecciona las variables con las que se va a trabajar y se construirán los paneles de la base dependiendo del tamaño del banco, es decir, grandes, medianos y pequeños, y se aplicará el análisis de estacionariedad para cada una de las bases. Una vez realizado el análisis se unirán las nuevas bases. Se utilizará como base de prueba que va desde el segundo trimestre del año 2017 hasta el tercer trimestre del año 2019.

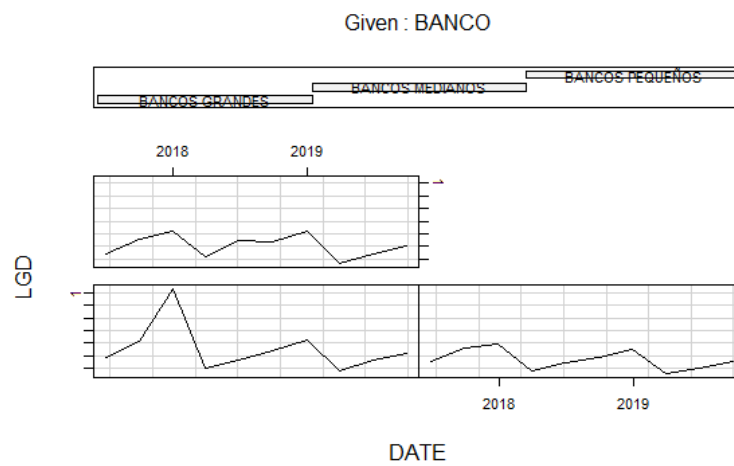


Figura 3.4: Comportamiento de la Severidad de Pérdida. **Fuente:** Propia

A simple vista se puede observar la severidad de pérdida para los tres tipos de panel, que tiene un comportamiento decreciente y creciente con un ciclo de 2 periodos, mostrando que a la mitad de un año la LGD empieza a decrecer.

3.2.1. Estructura

Se realizará la regresión basándose en lo especificado en el capítulo 2 construyendo el modelo para efectos comunes, efectos fijos y efectos aleatorios, y así aplicar la metodología para la selección del mejor modelo. Con la ayuda del software R que proporciona un paquete para la teoría de datos en panel, se puede construir un modelo de manera más sencilla y con resultados óptimos, este paquete se lo puede encontrar como `plm`.

Para esto se debe trabajar con una base balanceada, es decir, que el número de periodos T es igual para todos los individuos de cada grupo N .

DATE	BANCO		
	BANCOS GRANDES	BANCOS MEDIANOS	BANCOS PEQUEÑOS
2017-06-30	1	1	1
2017-09-30	1	1	1
2017-12-31	1	1	1
2018-03-31	1	1	1
2018-06-30	1	1	1
2018-09-30	1	1	1
2018-12-31	1	1	1
2019-03-31	1	1	1
2019-06-30	1	1	1
2019-09-30	1	1	1

Figura 3.5: Base de datos Balanceada **Fuente:** Propia

Como podemos observar en la figura 3.5 para cada período de tiempo existe un solo dato para cada tipo de banco. Por otra parte, vale recalcar las ventajas del uso de la metodología de datos en panel, pues al tener datos que varían a lo largo del tiempo y/o entre los individuos, nuestro modelo tendrá más información que en los modelos de corte transversal, obteniendo así estimadores más eficientes. Los datos en panel permitirán controlar por medio de las heterogeneidades existentes entre las unidades que no varían en el tiempo, y reducir el sesgo de las variables omitidas, así como testear la hipótesis del comportamiento de las observaciones en el tiempo (Urdinez y Cruz, 2021). Así, teniendo el enfoque de datos en panel podemos hacer uso del paquete `plm` para la selección del modelo más óptimo.

3.2.2. Modelo de efectos comunes

El modelo de efectos comunes es una regresión estándar de mínimos cuadrados ordinarios, es decir, sin ningún efecto transversal o de tiempo.

Cuadro 3.3: Modelo MCO

	<i>Dependent variable:</i>
	LGD
SOLVENCIA	-13.8600*** (4.4424)
INTERME_FIN	-1.0317*** (0.2571)
MARG_INTER_EST_PAT	-12.2041*** (2.1611)
PORC_ENCAJE	12.8852*** (2.5022)
IPC	-0.7290*** (0.2374)
ICC	0.1283*** (0.0379)
IDEAC	-0.1503*** (0.0264)
VALOR_REAL_PIB	0.0043*** (0.0005)
Constant	22.5519 (28.3075)
Observations	30
R ²	0.8636
Adjusted R ²	0.8116
F Statistic	16.6165*** (df = 8; 21)
p-value	1.7907e-07

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

Al construir el modelo de efectos comunes se obtuvieron las variables más significativas para evaluar el mismo, por los resultados estadísticos el modelo es válido dado que el *p – valor* es menor que 0.05.

A continuación, se va a interpretar a nivel financiero y macroeconómico todas las variables significativas que influyen en la severidad de pérdida (LGD).

Interpretación

Este modelo tiene como variables significativas la *solventia, intermediación financiera, margen de intermediación sobre patrimonio, porcentaje de encaje bancario, índice del precio al consumidor (IPC), índice de confianza del consumidor (ICC), índice de actividad económica coyuntural (IDEAC), y el PIB real.*

Las variables para que la severidad de pérdida (LGD) sea menor dentro de este modelo son:

- La solventia es el patrimonio técnico constituido, que es la cantidad disponible que debe tener una entidad bancaria para cubrir las operaciones diarias y pérdidas inesperadas (Junta de Regulación Monetaria y Financiera, 2015), sobre activos y contingentes ponderados por riesgo. Esto da, como resultado, la capacidad que tiene el capital del banco de hacerse a cargo del riesgo que asume. Esta variable tiene una afectación negativa, debido a que permite a la entidad tener una inmediata capacidad de respuesta hacia los ahorristas y también a que depende de una buena gestión financiera (Carrión, 2020), dando paso a que la entidad tenga mayores herramientas para poder hacer frente al incumplimiento de la cartera crediticia, reduciendo el riesgo de pérdida en un crédito.
- La intermediación financiera es la cartera bruta, que es el total de créditos otorgados por un banco, sobre la suma de los depósitos a la vista y a plazos, que son las obligaciones que tiene la entidad, esto incluye, depósitos monetarios, de ahorros, y todos los depósitos que retiran antes y después de los treinta días (Superintendencia de bancos, 2022). En otras palabras, mide cuánto dinero la entidad bancaria presta de los depositantes a un grupo de personas para la compra de un bien o servicio (ASOBANCA, 2022). En este modelo es negativo, pero lo más lógico sería que esta variable sea positiva, debido a que, si una entidad presta a más personas el dinero de sus depositantes, va a existir una mayor probabilidad de que estos incumplan con su crédito, dando paso a un mayor riesgo de pérdida. Según Villano (2013) explica que, mientras más dinero se desembolse para un crédito, habrá no solo mayores costos, sino más cargas financieras a lo largo del tiempo, como también el incremento del

riesgo.

- El margen de intermediación sobre activos es un ratio que mide la eficiencia de una entidad bancaria al prestar dinero o por intermediación financiera, servicios o comisiones. Es el porcentaje de gastos administrativos y recursos humanos que se absorbe de la utilidad operacional bruta (Calderón, 2019). Esto, dentro de nuestro modelo, significa que, como tenemos un mayor margen de intermediación, debido a que nuestros gastos administrativos y de personal son los menores posibles, la entidad bancaria es más eficiente. Por lo tanto, cuando el banco es eficiente en su gestión, los costos implicados para recuperar la cartera crediticia van a ser los más bajos posibles. Cuando los bancos, son más eficientes y están enfocados en ciertos segmentos, se traduce en menores costos y sobre todo menor riesgo (Carrasco Tineo, 2014).
- El IPC o inflación nos permite analizar la variación de los precios a lo largo del tiempo de los bienes y servicios más importantes de la canasta de consumo de los hogares (INEC, 2022). Esta variable es muy importante analizarla, debido a que mientras más alta es la inflación, se reduce la eficiencia de todas las tareas de intermediación (Uribe, 1999). Según Azariadis y Smith (1993) sostienen que, si un país presenta alza de precios, esto va a generar asimetría de información, provocando que se otorgue una mala o ineficiente asignación de un crédito (Uribe, 1999). En este caso, hay que tomar en cuenta que Ecuador es un país dolarizado, es decir, las tasas de inflación siempre van a tender a igualarse a las de Estados Unidos, no necesariamente van a ser iguales, porque la velocidad de convergencia va a depender de otras variables como la flexibilidad laboral, los desajustes de partida de los precios relativos, etc (BCE, 2021). De la misma manera, la tasa de interés nominal va a ser equivalente a la tasa de interés real, creando un sistema económico de confianza y provocando la reducción de las tasas de interés (BCE, 2021). Es importante mencionar que, un país al dolarizarse pierde su política cambiaria y monetaria, dando paso a que la oferta monetaria sea endógena, es decir, se determine de una manera interna, según la balanza de pagos o el gasto gubernamental (Onur Tas y Togay, 2014). Por lo tanto, como el IPC es medida en una economía dolarizada, su incremento se debe al aumento de la oferta monetaria, dado por el ingreso de más dinero al país, ya sea por mayor exportación o por gasto gubernamental, dando paso a que la economía ecuatoriana crezca. En consecuencia, cuando se da el aumento de precios en una economía dolarizada, el LGD es menor, ya que, el sistema económico al estar dinamizado por gasto público o ingreso de

divisas, los agentes económicos van a tener más alternativas de pago para un crédito incumplido o simplemente no van a incumplir sus préstamos.

- El IDEAC nos indica la tendencia que va a seguir la economía en el corto plazo, describiendo la variación del volumen (producción física) de las actividades económicas (Banco Central, 2014) de catorce industrias seleccionadas y que tienen una cobertura mayor al 60 % del PIB ecuatoriano, estas son “Acuicultura y pesca de camarón, Flores, Cereales, Electricidad, gas y agua, Pesca, Banano, café y cacao, Servicios financieros imputados, Otros cultivos agrícolas, Servicios gubernamentales, Transporte, Construcción y obras públicas, Comercio, Petróleo y Manufactura” (BCE, 2021, p. 1). Por tanto, el IDEAC en nuestro modelo disminuye el LGD, debido a que, si se tiene una buena tendencia de producción en las catorce industrias que dan una cobertura mayor al 60 % del PIB, significa que tenemos una economía más dinamizada, por lo que hay más probabilidad de que los agentes económicos paguen sus créditos incumplidos.

Por el contrario, las variables para que el LGD sea mayor dentro del modelo son:

- El Encaje bancario es un porcentaje de dinero que las entidades financieras depositan en el Banco Central, cuyo objetivo es garantizar el retiro de dinero de los depositantes, mediante la restricción de que los bancos no presten toda su base monetaria (García Díaz, Cárdenas, y Molina, 2011). En este caso, es positiva, pero debería ser negativa, debido a que, en la lógica financiera, si el Banco Central ajusta el encaje bancario, es para justamente precautelar que la entidad financiera pueda pagar sus obligaciones, pero sobre todo al reducir su base monetaria, van a dar menos créditos, y así van a tener menor riesgo de pérdida.
- El ICC mide el nivel de optimismo de los hogares ecuatorianos sobre la situación actual del país, como el de las expectativas acerca del consumo, situación económica del hogar y escenario económico a nivel nacional (BCE, 2012). La interpretación de esta variable en este modelo, es que mientras mayor optimismo, confianza o expectativas tengan los agentes sobre el escenario económico del país, dará paso al aumento del LGD. Esto se puede dar, debido a que, al tener una buena expectativa económica, las personas tienden a sobre consumir, provocando que su crédito se vea vulnerable al incumplimiento y, por lo tanto, mayor riesgo de pérdida para la entidad.

- El PIB real es aquella que mide la actividad económica de un país (López, Yaelga, y Espinosa, 2021). En otras palabras, es la constitución del número de bienes y servicios finales que produjeron ecuatorianos y residentes en el país en un periodo de tiempo, tomando como base los precios de un determinado año (Camargo y Blanco, 2010). Según Coyle (2017), la mejor manera de medir el PIB es a través del gasto, y su ecuación es:

$$PIB = C + I + G + (X - M)$$

donde:

- C = consume de los agentes económicos privados u hogares
- I = inversión de las empresas
- G = gasto público del estado en bienes o servicios
- X = exportaciones
- M = importaciones
- (X-M) = exportaciones netas, en la cual puede haber un superávit o déficit comercial

Cuando se da un crecimiento económico o el incremento del PIB, se da una mayor propensión de que un agente económico tome un crédito sin haber analizado adecuadamente su real capacidad de pago. En otras palabras, es una mayor probabilidad de que se incumpla un crédito, sobre todo los préstamos de consumo (Aparicio, Gutiérrez, Jaramillo, y Moreno, 2013). Adicionalmente, se puede deducir que, una estabilidad financiera, promovida por el crecimiento económico, va a producir el otorgamiento de un mayor volumen de créditos, con mayor riesgo, generando alza en precios y liquidez en el sistema económico; como es dado por la especulación financiera, provocará un estallido de una burbuja con una cartera vencida, llegando a una crisis financiera (Ochoa, Leano, y Agudelo, 2010). Esto da paso a que al aumentar el PIB o al tener un crecimiento económico, la severidad de pérdida va a ser mayor.

Debido a este análisis el modelo se descarta y más adelante se hará el correspondiente análisis estadístico que nos permitirá corroborar el análisis.

3.2.3. Modelo de efectos aleatorios

Un modelo de efectos aleatorios, por lo general se usará cuando cuanto la población N es grande y T pequeño y por consiguiente se trabaja tomando una muestra (Gutierrez, 2020), por lo que es la primera razón para excluir a este modelo, puesto que no se adapta al panel de datos con el que se está trabajando, pero aún se realizará el análisis respectivo.

Cuadro 3.4: Modelo efectos aleatorios

	<i>Dependent variable:</i>
	LGD
PROP_ACT_IMPRD	15.5284*** (4.3951)
IND_COBER	0.3125* (0.1597)
IPC	-0.7769** (0.3462)
IDEAC	-0.1436*** (0.0384)
VALOR_REAL_PIB	0.0037*** (0.0008)
Constant	36.7627 (40.8973)
Observations	30
R ²	0.6633
Adjusted R ²	0.5932
F Statistic	47.2814***
p-value	4.9783e-09

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

Evaluamos el modelo con las variables que aportaron significativamente al mis-

mo, con una precisión del modelo 59 % y un $p - valor < 0,05$ el modelo es válido, ahora se analiza las variables según la lógica financiera.

Interpretación

Este modelo tiene como variables significativas *el porcentaje de activos improductivos, cobertura de la cartera problemática, IPC, IDEAC y PIB real*. Las variables para que el LGD sea menor dentro del modelo son:

- El IPC en una economía dolarizada, como su incremento se debe al aumento de la oferta monetaria, produce el aumento de precios, produciendo que la severidad de pérdida sea menor, ya que, los agentes económicos van a tener más alternativas de pago para un crédito incumplido o simplemente no van a incumplir sus préstamos.
- La otra variable es el IDEAC, ya que, si se tiene una buena tendencia de producción, se da una economía más dinamizada, por lo que hay más probabilidad de que los agentes económicos paguen sus créditos incumplidos.

En cambio, las variables para que el LGD sea mayor dentro del modelo son:

- El porcentaje de activos improductivos, es aquel que mide la participación de los bienes cuya rentabilidad financiera es cero o no generan un ingreso, en relación con los activos totales; mientras menor sea, este indicador es mejor, debido a que el banco es más eficiente colocando recursos en activos productivos (SEPS, 2017). Por lo tanto, al tener más activos improductivos, significa que la entidad es menos eficiente, provocando que los costos implicados sean más altos para recuperar la cartera crediticia y por consecuencia mayor riesgo de pérdida.
- La otra variable es la cobertura de la cartera problemática, la división de las provisiones de cartera de créditos, que son las reservas para proteger la cartera de crédito que tiene problemas de recuperación, sobre la cartera improductiva bruta, aquella cartera que no genera ingresos al banco. En otras palabras, es el porcentaje de protección que asume el banco ante el riesgo de tener una cartera morosa (SEPS, 2017). Lo más lógico sería que esta variable sea negativa, debido a que, si tenemos una alta cobertura de provisiones de cartera de crédito, significa que hay una mayor cantidad de reserva para asumir el riesgo

de tener créditos irrecuperables, provocando que la severidad de pérdida sea menor.

- El PIB real influye de manera positiva, debido a que su aumento se traduce en un crecimiento económico, provocando una mayor propensión de que una persona tome un crédito sin haber analizado adecuadamente su capacidad real de pago, provocando que una entidad bancaria tenga un mayor porcentaje de LGD.

En conclusión, este análisis es el segundo factor para descartar el uso de este modelo para la base utilizada.

3.2.4. Modelo de efectos fijos

A diferencia del modelo de efectos aleatorios, se utiliza el modelo de efectos fijos cuando tenemos los datos de todos los miembros de la población. (Gutierrez, 2020)

Cuadro 3.5: Modelo de efectos fijos. Parte 1

	<i>Dependent variable:</i>
	LGD
SOLVENCIA	-104.2672*** (11.9034)
COB_PATRI_ACT_INM	0.4664*** (0.1038)
IND_MOROSIDAD	56.6337** (17.6932)
EFICI_OPE	-362.6708*** (52.1154)
GRADO_ABS_MARG_FINAN	-8.9859*** (2.0609)
EFIC_ADM_PERS	-632.1457*** (94.8193)
ROE	55.9813*** (9.1343)
ROA	-55.7343*** (9.0097)
MARG_INTER_EST_PAT	-73.7895*** (7.2866)
MARG_INTER_EST_ACT	-163.3752*** (27.2923)

Note:

*p<0.1; **p<0.05; ***p<0.01

Cuadro 3.6: Modelo de efectos fijos. Parte 2

	<i>Dependent variable:</i>
	LGD
IND_LIQUI	15.1106*** (2.8603)
IND_CAP	10.0379*** (2.7764)
IND_APALANCAMIENTO	-36.1243** (12.9735)
IND_COBER	-1.2413** (0.4054)
IPC	-1.0138*** (0.1872)
ICC	0.2213*** (0.0330)
IDEAC	-0.2223*** (0.0307)
VALOR_REAL_PIB	0.0088*** (0.0007)
Observations	30
R ²	0.9824
Adjusted R ²	0.9433
F Statistic	27.9260*** (df = 18; 9)
p-value:	9.1648e-06

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

El modelo presentado es un modelo válido con un R^2 – *ajustado* de 90 %, es decir, que la variable dependiente está explicada por el 90 % de las variables independientes, además, dado un p – *valor* < 0,05 tenemos un modelo significativo, por lo que se procede a realizar el correspondiente análisis económico.

Interpretación

Este modelo tiene como variables significativas *la solvencia, cobertura patrimonial de activos improductivos, índice de morosidad, eficiencia operativa, grado de absorción del margen financiero, eficiencia administrativa de personal, return on equity (ROE), return on assets (ROA), margen de intermediación sobre patrimonio, margen de intermediación sobre activo, índice de liquidez, índice de capitalización neto, índice de apalancamiento, cobertura de la cartera problemática, IPC, ICC, IDEAC y PIB real*. Las variables para que el LGD sea menor son:

- La solvencia mide la capacidad que tiene el capital del banco de hacerse a cargo del riesgo que asume. Esta variable tiene una afectación negativa, debido a que permite a la entidad tener una inmediata capacidad de respuesta hacia los ahorristas y también a que depende de una buena gestión financiera (Carrión, 2020), dando paso a que la entidad tenga mayores herramientas para poder hacer frente al incumplimiento de la cartera crediticia, reduciendo el riesgo de pérdida en un crédito.
- La eficiencia operativa es la división de los gastos de operación, costos que se realizan para el desarrollo de las actividades del banco, sobre el activo total promedio. Este índice representa el porcentaje de gastos operativos que se utiliza en la administración de los activos, un mayor índice, significa que el banco está destinando mayor gasto para la administración del activo total (SEPS, 2017). Dentro de nuestro modelo, el que un banco realice un mayor gasto en salarios, alquiler de locales, compra de suministros, va a provocar un menor LGD, debido a que hay más recursos para la administración de los activos.
- La eficiencia administrativa de personal, es la división de los gastos de personal, costos para la administración de personal, sobre el activo total promedio. Este índice representa el porcentaje de gastos de operación que se utiliza en la administración de los activos, un mayor índice, significa que el banco está destinando mayor gasto en personal para gestionar sus activos totales (SEPS, 2017). Dentro de nuestro modelo, el que un banco realice un mayor gasto en salarios, alquiler de locales, compra de suministros, va a provocar un menor LGD, debido a que hay más recursos para la administración de los activos.
- El grado de absorción del margen financiero neto, es la división de los gastos de operación, costos que se realizan para el desarrollo de las actividades del banco, sobre margen financiero neto, diferencia del margen bruto financiero y

las provisiones. Este indicador mide la absorción de los costos operacionales en los ingresos que vienen de la gestión operativa, mientras mayor es el índice, quiere decir que, el banco no genera los suficientes ingresos para cubrir sus gastos operativos (SEPS, 2017). Dentro de nuestro modelo, el que los gastos de operación sean mayores al margen financiero neto, provoca que el LGD sea menor, esto debido a que hay más recursos para el cobro de un crédito incumplido.

- El ROA mide la eficacia que tiene una entidad sobre el manejo de sus recursos, mientras menor es este índice, significa que el banco no puede generar los suficientes ingresos para fortalecer su patrimonio (SEPS, 2017). Por lo tanto, si una entidad financiera tiene mayor eficacia en la administración de sus recursos, va a provocar que su LGD disminuya, debido a que hay mayores ingresos que egresos que permite la generación de más herramientas para el cobro de un crédito incumplido.
- El margen de intermediación sobre el patrimonio, es la división de las ganancias netas generadas por la intermediación financiera, sobre el patrimonio promedio. Este índice mide la rentabilidad que tiene la entidad bancaria en relación al patrimonio; mientras mayor es el indicador, es mejor, debido a que, la ganancia obtenida de la intermediación es más eficiente en relación al patrimonio (SEPS, 2017). Por lo tanto, si un banco tiene ganancias más eficientes en relación al patrimonio, da paso a que su LGD sea menor.
- El margen de intermediación sobre el activo, es la división de las ganancias netas generadas por la intermediación financiera, sobre el activo promedio. Este índice mide la rentabilidad que tiene la entidad bancaria en relación al activo promedio; mientras mayor es el indicador, es mejor, debido a que, la ganancia obtenida de la intermediación es más eficiente en relación al activo promedio (SEPS, 2017). Por lo tanto, si un banco tiene ganancias más eficientes en relación al activo, da paso a que su LGD sea menor.
- El índice de apalancamiento, mide el endeudamiento que tiene una entidad bancaria, es la división del pasivo sobre el patrimonio. Este indicador tiene dos objetivos, limitar el exceso de deuda que un banco puede asumir, y el ser una medida para el reforzamiento del capital independientemente del riesgo (BBVA, 2016). El apalancamiento es una herramienta que puede afectar positiva o negativamente una entidad, debido a que, si obtiene una deuda, se está asumiendo un riesgo, por eso esta decisión siempre va a estar en función a una

buena gestión del equipo operativo y capacidad de pago (Flórez, 2011). Por lo tanto, si un banco tiene un índice alto de apalancamiento, quiere decir que tiene una buena capacidad de pago, como un buen gestionamiento operativo, provocando que el LGD sea menor.

- La cobertura de la cartera problemática, es un indicador que nos permite medir el porcentaje de protección que asume el banco ante el riesgo de tener una cartera morosa (SEPS, 2017). A comparación del modelo de efectos aleatorios, esta variable es negativa, comprobando que, si tenemos una alta cobertura de provisiones de cartera de créditos, significa que hay una mayor cantidad de reserva para asumir el riesgo de tener carteras de crédito irrecuperable, provocando que el LGD sea menor.
- El IPC, como su incremento se debe al aumento de la oferta monetaria, produce el aumento de precios en una economía dolarizada, produciendo que la severidad de pérdida sea menor, ya que, los agentes económicos van a tener más alternativas de pago para un crédito incumplido o simplemente no van a incumplir sus préstamos.
- Otra variable es el IDEAC, ya que, si se tiene una buena tendencia de producción, se da una economía más dinamizada, por lo que hay más probabilidad de que los agentes económicos paguen sus créditos incumplidos.

En cambio, las variables para que el LGD sea mayor son:

- La cobertura patrimonial de activos improductivos, es la medición del grado de cobertura que tienen los activos improductivos, bienes cuya rentabilidad financiera es cero o no generan un ingreso como una cartera vencida, del patrimonio efectivo, mientras más alto es el índice, la pérdida que se de en activos podrá ser cubierta de mejor manera con el patrimonio (Unidad de Gestión del Riesgo del Sistema Financiero, 2015). Por lo tanto, cuando el patrimonio efectivo de una entidad tiene mayor cobertura sobre los activos improductivos como una cartera vencida, cartera que no devenga intereses, equipos, o cuentas por cobrar, el LGD va a aumentar, es decir, una entidad va a tener una mayor pérdida en un crédito, porque la pérdida producida por esa cartera incumplida va a estar cubierta por más patrimonio de la entidad.
- El índice de morosidad, es la división de la cartera improductiva, préstamos que no generan renta financiera al banco, como por ejemplo carteras vencidas

y que no devenguen interés, sobre el total de la cartera de crédito, mientras mayor es el indicador, significa que tiene problemas con la recuperación de la cartera (SEPS, 2017). Por lo tanto, cuando la entidad bancaria tiene problemas con la recuperación de su cartera, el LGD es mayor.

- El ROE, es un indicador que mide el nivel de retorno producido por el patrimonio invertido por los accionistas del banco, mientras menor sea este índice, significa que la entidad no tiene los suficientes recursos para cubrir la remuneración correspondiente de cada accionista (SEPS, 2017). Por lo tanto, mientras mayor retorno del patrimonio tenga cada accionista, el LGD va a ser mayor, dado que, cada accionista puede extraer su parte invertida, dejando vulnerable al banco para enfrentar las carteras incumplidas.
- El índice de liquidez mide la capacidad que tiene un banco para responder a obligaciones inmediatas, es la división de los fondos disponibles, recursos que tiene la entidad para realizar pagos con disposición inmediata, sobre los depósitos que pueden ser retirados en cualquier momento (depósitos a corto plazo); mientras más alto es el indicador, se tendrá mayor capacidad de responder a obligaciones inmediatas (SEPS, 2017). Por lo tanto, mientras la entidad tenga mayor capacidad de hacer frente a obligaciones inmediatas, el LGD será mayor, debido a que según estudios, mayor liquidez en el sistema financiero genera mayor morosidad (Altuve y Briceño, 2018).
- El índice de capitalización neto, es un indicador de solvencia que mide la capacidad que tiene una entidad bancaria para enfrentar pérdidas inesperadas (Comisión Nacional Bancaria y de Valores, 2017). Es la división entre el capital neto y los activos sujetos a riesgo, mientras mayor sea este índice, significa que el banco está teniendo mayores ganancias a pesar de los cambios que se podrían dar en los activos en riesgo (SEPS, 2017). Esta variable tiene una afectación negativa, debido a que permite a la entidad tener una inmediata capacidad de respuesta hacia pérdidas inesperadas (Carrión, 2020), dando paso a que la entidad tenga mayores herramientas para poder hacer frente al incumplimiento de la cartera crediticia, reduciendo el riesgo de pérdida en un crédito.
- El ICC es aquel que mide el nivel de expectativas acerca del consumo, por lo tanto, mientras mayor sea esta expectativa, mayor será el LGD, esto se puede dar, debido a que, al tener una buena expectativa económica, las personas

tienden a sobre consumir, provocando que su crédito se vea vulnerable al incumplimiento, por lo tanto, el LGD será mayor.

- El PIB real influye de manera positiva, debido a que su aumento se traduce en un crecimiento económico, provocando una mayor propensión de que una persona tome un crédito sin haber analizado adecuadamente su capacidad de pago real, provocando que una entidad bancaria tenga un mayor porcentaje de LGD.

Después de la interpretación se tiene un modelo válido, así como el resto de modelos MCO Y RE, por lo que se procede a evaluar los 3 modelos mediante pruebas estadísticas de selección para datos en panel.

3.3. Método de selección

3.3.1. Test de Hausman

El Test de Hausman suele utilizarse para decidir si es preferible estimar el modelo con efectos fijos o con efectos aleatorios. En este test se asume que el modelo de efectos fijos es consistente para los parámetros verdaderos y el modelo de efectos aleatorios es una especificación eficiente de los efectos individuales bajo el supuesto de que son aleatorios y siguen una distribución normal. Se supone que el modelo de efectos fijos calcula siempre el estimador consistente, mientras que el modelo de efectos aleatorios calculará el estimador que es consistente y eficiente bajo H_0 (Urdinez y Cruz, 2021).

H_0 : modelo efectos aleatorios

H_1 : modelo de efectos fijos.

Con la función `phtest` del paquete `plm` entre los modelos de efectos fijos y aleatorios se obtiene el p – *valor* $< 2,2e - 16$, por tanto, se rechaza la hipótesis nula, así el modelo seleccionado es el modelo de efectos fijos.

3.3.2. Test de Chow

Por otro lado, el test de Chow se utilizará para decidir si el modelo se estima con efectos comunes o con efectos fijos.

H_0 : modelo efectos comunes

H_1 : modelo de efectos fijos.

Con la función `pooltest` del paquete `plm` se obtiene el $p - valor = 0,0075$ entonces se rechaza H_0 . Por tanto, el modelo de efectos fijos sigue siendo el más eficiente realizando la prueba de Hausman y de Chow.

Con estas pruebas se obvia el test de los multiplicadores de Lagrange, debido a que este elige un mejor modelo entre MCO y RE.

En conclusión, debido a estos parámetros de selección, el modelo que se ajusta a los datos es el **modelo de efectos fijos**

3.4. Idoneidad del modelo

Una vez seleccionado el modelo de efectos fijos, se debe validar los supuestos de prueba. Es decir, se debe verificar que los residuos sigan una distribución normal e independiente, que sean homocedásticos (varianza de los residuos constante) y que no exista multicolinealidad (González, 2017).

Test de Glejser para la heteroscedasticidad

Este test detecta heteroscedasticidad para ver si la variante de error tiene una naturaleza constante o no (Achmad, Yoyon y Achmad, 2019). A su vez, si el $p - valor < 0,05$ se rechaza H_0 : la varianza del error no tiene heteroscedasticidad, es decir, que a partir del resultado del test se puede corregir la heteroscedasticidad con la transformación el modelo.

En el modelo de efectos fijos se obtiene el $p - valor = 0,1573$, por lo que el supuesto de heteroscedasticidad no está presente en el modelo y no se rechaza H_0 .

Test de Durbin-Watson para la autocorrelación

Para modelos en panel se ha modificado el test de Durbin Watson para paneles balanceados y consecutivos, puesto que, el test de Durbin-Watson permite diagnosticar la heteroscedasticidad para regresión lineal simple o múltiple. Así, el paquete `plm` establece la función `pdwftest` para modelos de efectos fijos:

```
Durbin-watson test for serial correlation in panel models  
  
data:  FORMULAR_LGD  
DW = 2.6715, p-value = 0.9793
```

Figura 3.6: Test de Durbin Watson **Fuente:** propia

El estadístico nos arroja el $p - valor > 0,05$ por lo que se rechaza la H_0 de que no hay autocorrelación entre los residuos.

Test de Jarque-Bera para la Normalidad

El test de Jarque-Bera es una prueba de bondad de ajuste que se utiliza para verificar si una muestra de datos sigue una distribución normal, que utiliza un estadístico de prueba que involucra la curtosis y la asimetría. Es decir, compara estos coeficientes con los que se obtendría en una muestra de datos distribuidos normalmente. Si ambos coeficientes de la muestra no se desvían significativamente de estos valores, entonces, el test concluye que no se puede rechazar la hipótesis de normalidad (Borgnino, 2014).

H_0 : los residuos siguen una distribución normal.

```
Jarque Bera Test  
  
data:  residuals(reg.plm.f)  
x-squared = 2.2061, df = 2, p-value = 0.3319
```

Figura 3.7: Test de Jarque Bera **Fuente:** propia

Como se observa en la figura 3.7 el $p - valor$ obtenido sobre el modelo de efectos fijos no rechaza la hipótesis de normalidad, por lo tanto, los residuos siguen una distribución normal.

Dadas las pruebas que cumplen con los supuestos establecidos, se ha encontrado entonces el mejor estimador lineal imparcial (BLUE) para el modelo de enfoque FE, así, este modelo de efectos fijos es un modelo idóneo.

3.5. Pruebas de significación

Una vez que el modelo ha superado todas las pruebas de suposición clásicas, el siguiente paso es realizar una prueba simultánea y una prueba parcial para determinar el efecto de las variables predictoras en la variable de respuesta de forma simultánea y parcial (Achmad, et al, 2019).

3.5.1. Prueba simultánea

Como se mencionó en el capítulo anterior, sección 2.5, la prueba simultánea se aplica para determinar el efecto de las variables predictoras sobre la variable respuesta. A continuación, se muestran los resultados de las pruebas del modelo de efectos fijos seleccionado.

R ²	0.90241
R ² -ajustado	0.86332
RSS	0.19292
F-estadístico	27.926, df1=18 y df2= 9
p-valor	9.1648e-06

Cuadro 3.7: Resultados de la prueba simultánea

Con los resultados que se muestran el cuadro 3.7, el valor $F = 27,926 > F(0,05; 18; 9) = 2,960$ (Tablas estadísticas), el $p - valor < 0,05$ y $R^2 = 90\%$ con un nivel de confianza del 95 %, se rechaza H_0 , así, hay al menos una variable predictora que afecta simultáneamente a la variable respuesta, es decir, las 18 variables predictoras pueden afectar a la LGD en los diferentes paneles establecidos con características similares.

3.5.2. Prueba Parcial

La prueba parcial determina el efecto de cada variable de predicción sobre la variable respuesta. A continuación, se muestran los resultados de cada variable con el correspondiente estadístico de prueba t-valor.

<i>Estimador</i>	<i>t-valor</i>	<i>Pr(> t)</i>
SOLVENCIA	8.7594	1.07E-05
COB_PATRI_ACT_INM	4.4929	0.0015042
IND_MOROSIDAD	3.2009	0.0108161
EFICI_OPE	-6.959	6.62E-05
GRADO_ABS_MARG_FINAN	-4.3603	0.0018227
EFIC_ADM_PERS	6.6668	9.20E-05
ROE	6.1287	0.0001731
ROA	-6.186	0.0001616
MARG_INTER_EST_PAT	-10.1267	3.22E-06
MARG_INTER_EST_ACT	5.9861	0.000206
IND_LIQUI	5.283	0.0005051
IND_CAP	3.6155	0.0056112
IND_APALANCAMIENTO	-2.7845	0.0212452
IND_COBER	-3.0616	0.0135367
IPC	-5.4159	0.0004241
ICC	6.6987	8.87E-05
IDEAC	-7.2481	4.83E-05
VALOR_REAL_PIB	12.5102	5.40E-07

Cuadro 3.8: Resultados del t-valor de las variables predictoras

Como se muestra en el cuadro 3.8 todas las variables son significativas a un nivel de confianza del 95 %, esto se debe a la construcción del modelo, ya que se buscó las variables más significativas para el mismo.

Comparación empírica

La LGD ha sido un tema de estudio en varios países, lo cual va a permitir comparar empíricamente el estudio de este trabajo. Caselli, Gatti y Querci (2008) examinaron 11.649 préstamos en el mercado de Italia, concluyendo que existe una relación entre la LGD y ciertas variables macroeconómicas. Ellos determinaron que la LGD tiene una relación positiva cuando se da una variación anual de mora y un aumento en la tasa de desempleo; mientras que una relación negativa con el aumento del PIB real, la cantidad de préstamos bancarios, el total de personas ocupadas, el total bruto de inversiones, aumento de la renta bruta disponible, y del consumo de los hogares.

Por otra parte, otro estudio realizado en una cartera de préstamos a pymes de un banco comercial en Eslovenia nos explica que la LGD tiene una relación positiva cuando tienen garantías financieras, y una relación negativa cuando tiene una asignación de cuentas por cobrar y préstamos a largo plazo (Kosak y Poljšak, 2010). El PIB anual y el año de incumplimiento no fueron variables estadísticamente significativas, debido a que el conjunto de datos era muy pequeño, y a que hubo un crecimiento favorable en ese periodo (Kosak y Poljšak, 2010).

Según Konečný, Seidler, Belyaeva y Belyaev (2017), mediante un estudio al 15 % del mercado minorista de crédito al consumo en República Checa, los factores que contribuyen a que la LGD aumente es la tasa de interés, crecimiento en precios inmobiliarios y crecimiento del PIB real; mientras que las variables que hacen que la LGD disminuya es la tasa de desempleo, crecimiento del consumo real en hogares y pymes, crecimiento de la inversión, crecimiento de los préstamos, y salarios.

Cabe recalcar que, la diferencia dada entre uno y otro modelo se puede dar debido a que cada país tiene diferentes factores que intervienen en la variación del LGD. En este caso, comparando el modelo de efectos fijos de este trabajo con los otros, se puede decir que, la variable de intermediación financiera comparte la misma relación con el modelo de (Caselli, Gatti y Querci, 2008) y de (Konečný, Seidler, Belyaeva, y Belyaev, 2017). La variable del PIB real y del IDEAC solo con el de (Konečný, Seidler, Belyaeva y Belyaev, 2017). Por el contrario, las variables que difiere es el ICC y la inflación con (Konečný, Seidler, Belyaeva y Belyaev, 2017). Por último, las variables aporte de este trabajo son: la solvencia, el patrimonio técnico constituido, el margen de intermediación sobre activos, el IDEAC y el encaje bancario.

3.6. Predicción

Una vez cumpliendo el objetivo de determinar el método estadístico para la selección de un modelo idóneo, se realizará el backtesting del mismo.

Se han identificado las variables microeconómicas y macroeconómicas que influyen en la estimación de la severidad de pérdida para el periodo 2017 al tercer trimestre del año 2019. Por lo que es importante tener en cuenta que la predicción para datos en panel puede ser compleja, pero gracias a las funciones del paquete plm se tiene una predicción válida, ya que cumple con los parámetros necesarios para los datos en panel.

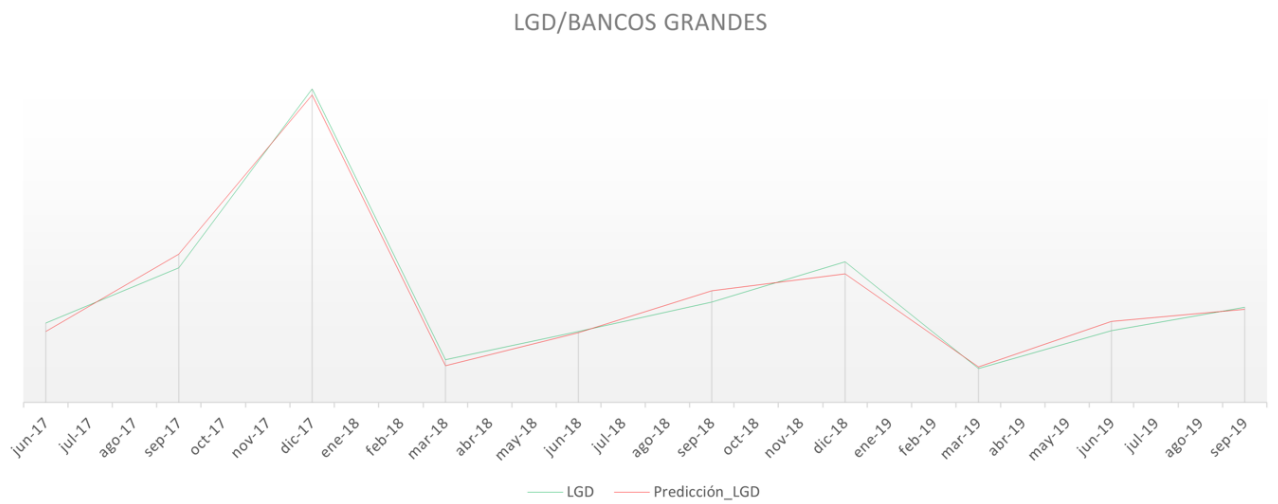


Figura 3.8: Estimación de la LGD, Bancos Grandes **Fuente:** Propia

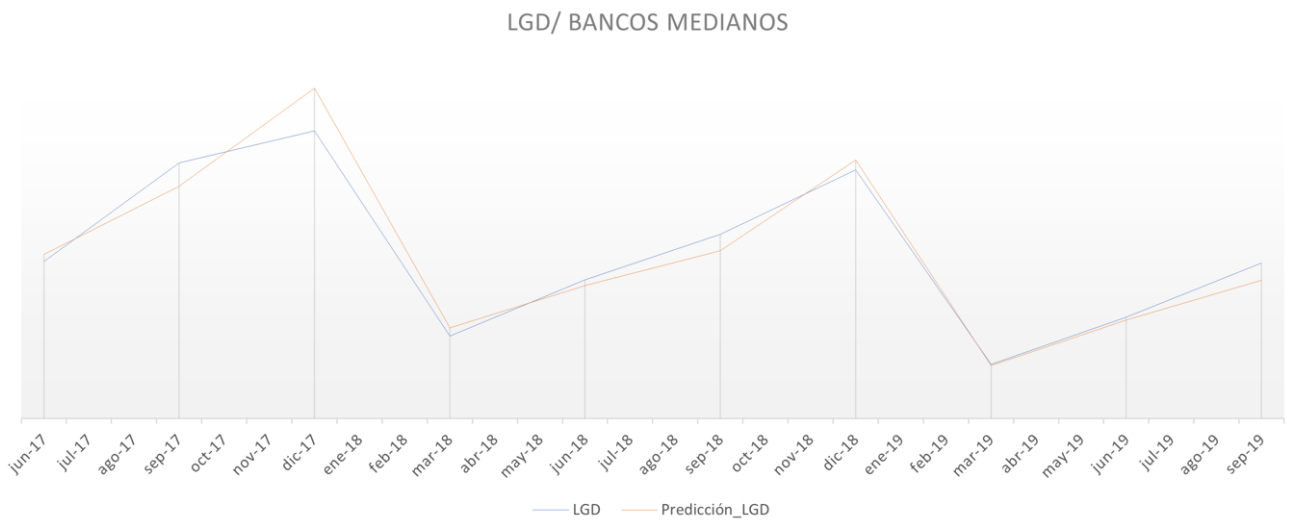


Figura 3.9: Estimación de la LGD, Bancos Medianos **Fuente:** Propia

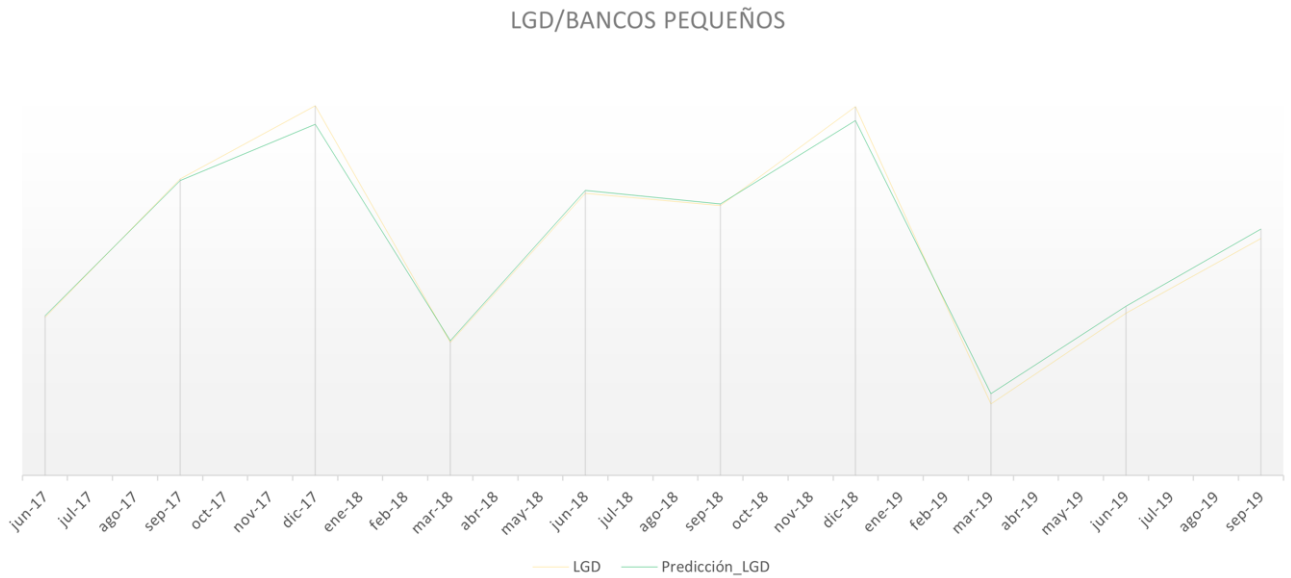


Figura 3.10: Estimación de la LGD, Bancos Pequeños **Fuente:** Propia

Predicciones en Base de Prueba

Se ha considerado la base de prueba en el periodo de diciembre de 2019 a marzo de 2020, antes de la Pandemia, debido a que esto podría hacer un modelo inestable debido a una recesión muy fuerte.

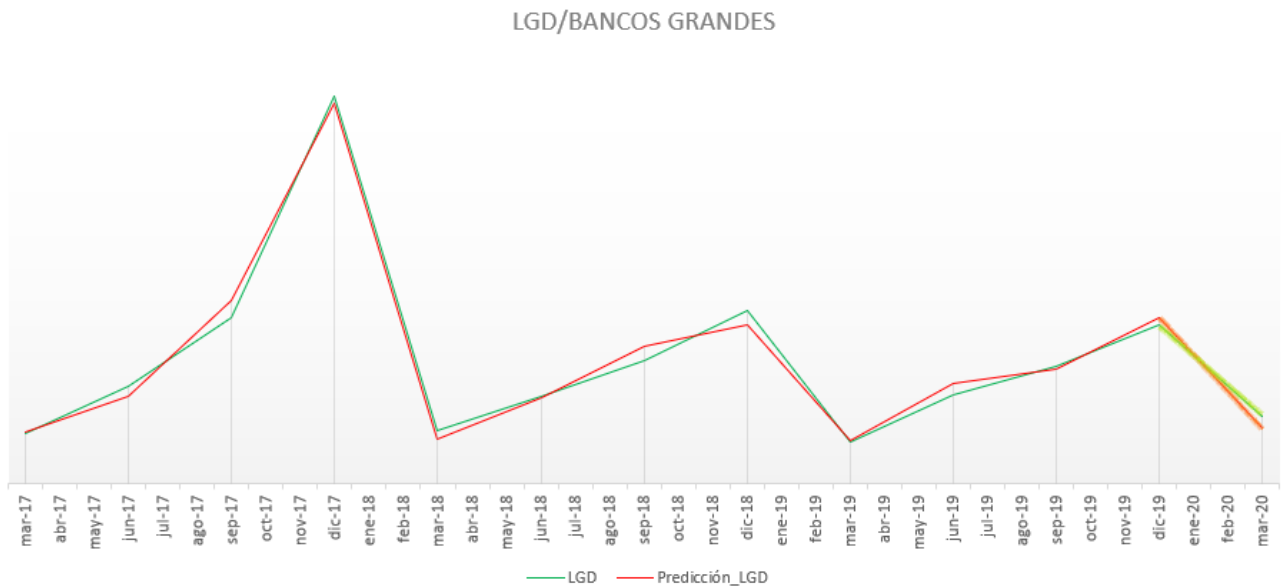


Figura 3.11: Predicción de la LGD, Bancos Grandes **Fuente:** Propia

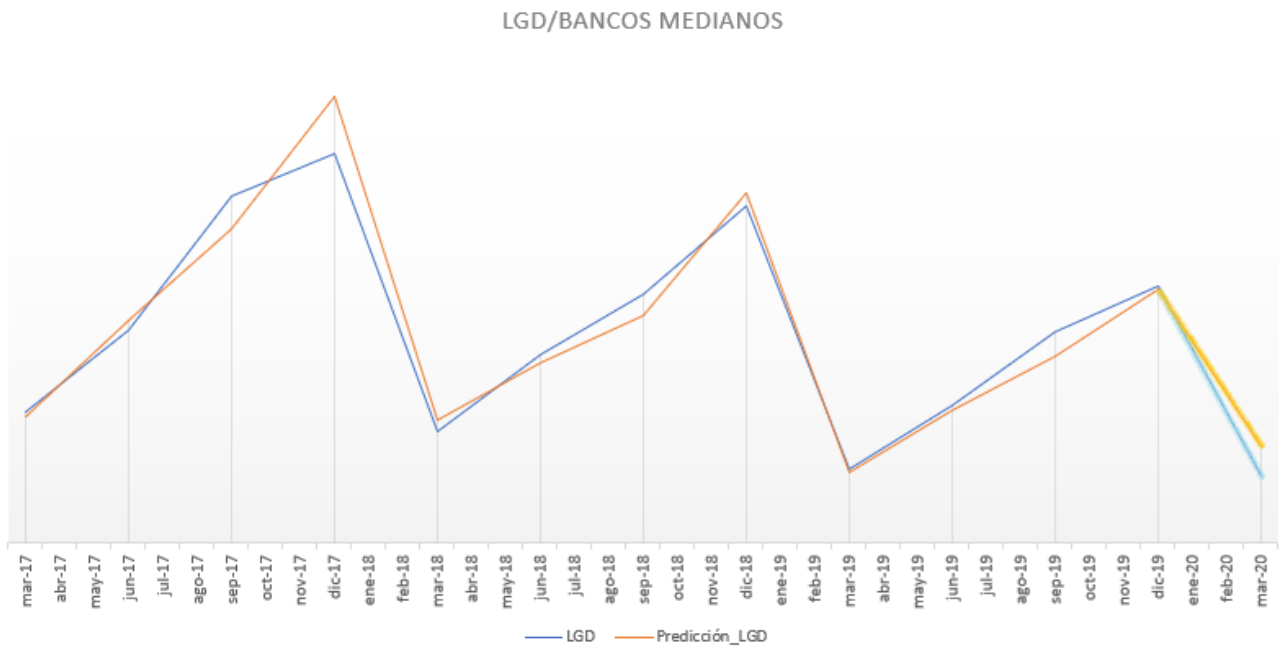


Figura 3.12: Predicción de la LGD, Bancos Medianos **Fuente:** Propia

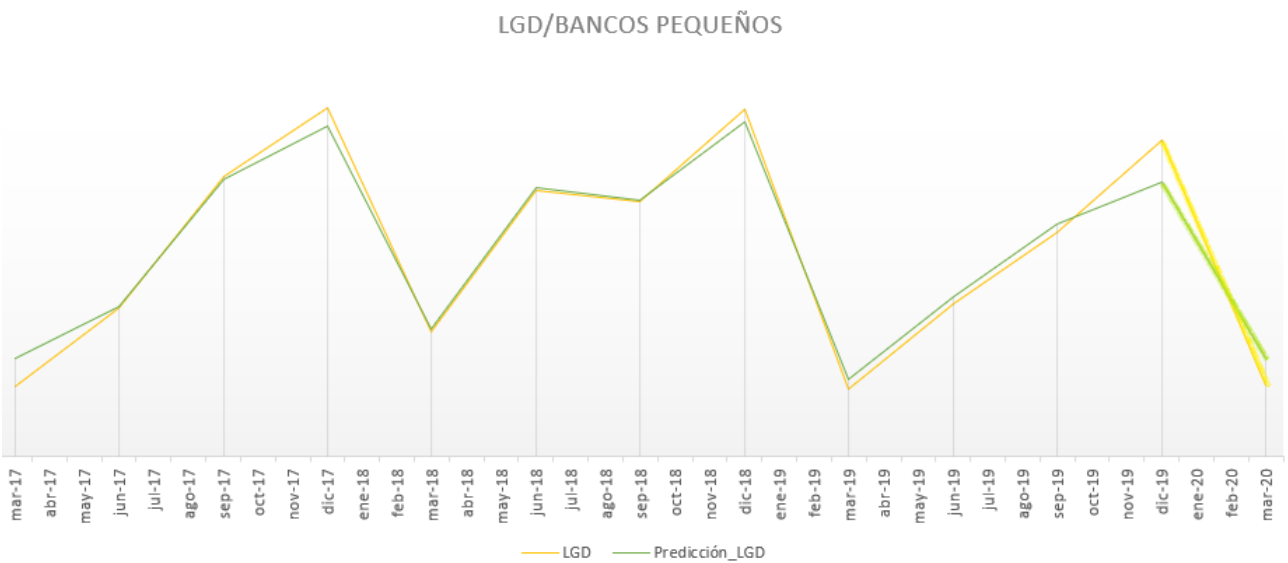


Figura 3.13: Predicción de la LGD, Bancos Pequeños **Fuente:** Propia

Se aplicó el modelo, y como se puede observar en las figuras 3.11, 3.12 y 3.13, los valores obtenidos son bastante acertados, ya que se asemejan al porcentaje de la LGD que se utilizó en el presente estudio, demostrando que el método seleccionado es válido.

Capítulo 4

Conclusiones y Recomendaciones

La severidad de pérdida se refiere a una proporción de pérdida financiera no recuperable, que una organización experimenta debido al incumplimiento de créditos o eventos inesperados, por lo que toda institución financiera debe tener en cuenta esta proporción para hacer frente al riesgo y costos de estos imprevistos. Existen diferentes métodos de estimación de la LGD, pero este proyecto se enfocó en una selección histórica y transversal con base en indicadores financieros disponibles públicamente.

4.1. Conclusiones

Así, después del estudio se presentan las conclusiones relevantes que se pudieron encontrar sobre la severidad de pérdida y la metodología seleccionada:

1. De manera general, la severidad de pérdida puede ser muy alta en algunos casos: Si una institución experimenta un evento inesperado grave, la LGD puede ser muy alta y, en algunos casos, incluso podría ser suficiente para poner en peligro la supervivencia de la entidad.
2. En conclusión, en el presente trabajo se ha podido cumplir con la estimación objetivo de la Severidad de pérdida del sistema bancario del Ecuador mediante indicadores microeconómicos de acceso público y la influencia de factores macroeconómicos, ya que los cambios a nivel nacional o internacional afectan directamente a las entidades financieras, puesto que son los entes que manejan la economía del país.

3. El método estadístico propuesto *datos en panel*, fue sin duda un modelo bastante acertado debido a su versatilidad para poder trabajar con series de tiempo y datos transversales. Los efectos de la metodología proponen iniciar con modelos ignorando dimensiones individuales y de tiempo, para poder comparar con los efectos que si los incluyen, pues nos muestran las inconsistencias y sobre todo la interpretación económica que sustenta la invalidez del modelo. Vale recalcar que los modelos son explicados a través de datos reales y esto permite identificar la relación lógica que estiman la severidad de pérdida.

Por otra parte, los efectos fijos fue el modelo idóneo que se ajustó estadísticamente para la estimación y el backtesting, debido a la estructura de la base y cantidad de datos, además de las especificaciones del modelo, se rechazó el trabajar con un aumento de factores de tiempo, ya que el modelo FE permite remover el sesgo que se genera por variables omitidas, especialmente de las variables que no varían en el tiempo. Finalmente, el modelo proporciona predicciones cuantitativamente similares, por lo que es idóneo y válido

4. Es importante señalar que la estimación de LGD con datos de panel puede ser compleja y requiere una selección cuidadosa de las variables independientes y de control en el modelo. Además, es importante asegurarse de que los datos históricos utilizados para la estimación de LGD sean representativos de la población subyacente y el período de tiempo de interés.
5. Se concluye que, los factores macroeconómicos son efectivamente importantes para la estimación, por ejemplo, para que disminuya la LGD con el modelo de efectos fijos, es necesario el aumento del IPC, debido a que Ecuador, al ser un país dolarizado, su variación de precios va a estar en función al ingreso de dinero por superávit comercial o la adquisición de una gran deuda pública, traduciéndose en una dinamización económica. Esto va a provocar que, los agentes económicos tengan más alternativas de pago para un crédito incumplido o simplemente no incumplan con sus préstamos.

Por otra parte, para un incremento en el LGD, es necesario el aumento del ICC, es decir, que el jefe de hogar tenga un alto optimismo, confianza o expectativa del entorno económico, en consecuencia, esto va a provocar que los agentes sobre consuman y los créditos se vean vulnerables a ser incumplidos. Otra variable macroeconómica relevante es el PIB real, dado que su aumento significa que el país está teniendo un crecimiento económico, y esto da como resultado que las personas tomen un crédito sin haber analizado adecuadamente su ca-

pacidad real de pago y tengan una mayor probabilidad de incumplimiento de un crédito.

4.2. Recomendaciones

- La evaluación de la severidad de pérdida es un paso importante en la gestión de riesgos, ya que permite a una entidad entender, evaluar y asignar recursos para manejar los riesgos a los que está expuesta y así tomar medidas para reducirlos. Por lo que, una evaluación con variables propias que guardan las instituciones financieras permitirán que modelos como este funcionen de forma más óptima, es decir, trabajar con información más real y actualizada podrá dar paso a que la estimación de la LGD se la realice con la importancia que implica este factor. Se recomienda que las organizaciones construyan un data warehouse con información fiable y necesaria para un estudio de LGD sea tomado en cuenta y no trabajarlo de manera implícita como se mostró en este proyecto como medida alternativa, y así poder personalizarlo para cada institución financiera.

Bibliografía

- [1] *Junta de Regulación Monetaria y Financiera*. Recuperado de : <https://bit.ly/31TMYLU>, 2015.
- [2] *Comisión Nacional Bancaria y de Valores*. Recuperado de : [https://www.gob.mx/cnbv/acciones-y-programas/alertas-tempranas-banca-multiple#:~:text=El%20%20%C3%8Dndice%20de%20Capitalizaci%C3%B3n%20\(ICAP,riesgos%20en%20los%20que%20incurre.,2017](https://www.gob.mx/cnbv/acciones-y-programas/alertas-tempranas-banca-multiple#:~:text=El%20%20%C3%8Dndice%20de%20Capitalizaci%C3%B3n%20(ICAP,riesgos%20en%20los%20que%20incurre.,2017).
- [3] *Comisión Nacional Bancaria y de Valores*. Recuperado de : [https://www.gob.mx/cnbv/acciones-y-programas/alertas-tempranas-banca-multiple#:~:text=El%20%20%C3%8Dndice%20de%20Capitalizaci%C3%B3n%20\(ICAP,riesgos%20en%20los%20que%20incurre.,2017](https://www.gob.mx/cnbv/acciones-y-programas/alertas-tempranas-banca-multiple#:~:text=El%20%20%C3%8Dndice%20de%20Capitalizaci%C3%B3n%20(ICAP,riesgos%20en%20los%20que%20incurre.,2017).
- [4] *CODIFICACIÓN DE LAS NORMAS DE LA SB/ LIBRO UNO – SISTEMA FINANCIERO, título IX, cap. II (Ecuador)*. Recuperado de https://www.superbancos.gob.ec/bancos/wp-content/uploads/downloads/2022/02/L1_IX_cap_II.pdf, 2021.
- [5] Achmad, J. Yoyon, K. y Achmad, M. Data panel modelling with fixed effect model (fem) approach to analyze the influencing factors of dhf in pasuruan regency. *Science and Technology Publications*, 2019.
- [6] Aguilar, G. Camargo G. y Saravia, R. Análisis de la morosidad en el sistema bancario peruano. *Instituto de Estudios Peruanos*, 2004.
- [7] Altuve, J. y Briceño, A. Análisis de los factores que influyen en la morosidad del sistema bancario venezolano. *Revista venezolana de análisis de coyuntura*, 2018.
- [8] Amat, J. *Introduction to forecasting*. Recuperado de: <https://joaquinamatrodrigo.github.io/skforecast/0.4.3/quick-start/introduction-forecasting.html>, (s.f).

- [9] Aparicio, C., G. J. J. M. y Moreno, H. Indicadores alternativos de riesgo de crédito en el Perú: matrices de transición crediticia condicionadas al ciclo económico. *Documentos de Trabajo*, 2013.
- [10] Aslam, T. When sovereigns default: a global and regional look. *SP Global Market Intelligence.*, 2020.
- [11] ASOBANCA. El rol de la intermediación financiera para cumplir objetivos de empresas y personas. 2022.
- [12] BBVA. ¿Qué es el ratio de apalancamiento y cómo se calcula? Recuperado de : <https://www.bbva.com/es/ratio-apalancamiento-bancario-la-calidad-tan-importante-la-cantidad/>, 2016.
- [13] BCE. Reporte anual del índice de confianza del consumidor - icc. *Revisión anual*, 2012.
- [14] BCE. RRAAE. Recuperado de : https://rraae.cedia.edu.ec/Record/BCE_4d9dd32b7602060147f0f64b811ab640, 2014.
- [15] BCE. *METODOLOGÍA PARA MEDIR LA VULNERABILIDAD*. Recuperado de : <https://contenido.bce.fin.ec/documentos/PublicacionesNotas/Catalogo/Cuadernos/Cuad135.pdf>, 2015.
- [16] BCE. Dolarización. 2021.
- [17] BCE. Índice de actividad económica coyuntural (ideac). *Revisión anual*, 2021.
- [18] Belyaev, K. Belyaeva, A. K. T. S. J. y Vojtek, M. Macroeconomic factors as drivers of Igd prediction: Empirical evidence from the czech republic. *Czech National Bank.*, 2012.
- [19] Bhargava, A. Franzini, L. y. N. W. Serial correlation and the fixed effects model. *The Review of Economic Studies*, 1982.
- [20] Borgnino, J. *Testeando simetría y kurtosis sobre los residuos de un modelo de datos en panel agrícola*. Universidad de San Andrés. Recuperado de: www.ugr.es , 2014.
- [21] Calderón, A. *Eficiencia bancaria: del mito a la realidad*. Recuperado de : <https://www.elfinanciero.cr/opinion/eficiencia-bancaria-del-mito-a-la-realidad/GE5PVPQJIJCNBCCMMKR7ROHCSA/story/>, 2019.

- [22] Camargo, B. y Blanco, G. Macroeconomía. *Bogotá: Fundación Politécnico Granacolombiano.*, 2010.
- [23] Carrasco, D. Análisis comparativo de rentabilidad, eficiencia y concentración entre bancos y microfinancieras durante el período 2005–2011. *USAT*, 74-75, 2014.
- [24] Carrión, M. *Evaluación del patrimonio técnico en las instituciones del sector financiero popular y solidario del Ecuador.* Recuperado de : <https://www.cienciamatriarevista.org.ve/index.php/cm/article/view/361/460>, 2020.
- [25] Caselli, S. Gatti, S. y Querci, F. . The sensitivity of the loss given default rate to systematic risk: New empirical evidence on bank loans. *Journal of Financial Services Research.*, 2008.
- [26] Coll, F. *Índice de morosidad.* Recuperado de : <https://www.janushenderson.com/en/row/documents/>, 2021.
- [27] Coyle, D. Producto interno bruto. *Ciudad de México: International Editors Co. y Princeton University Press.*, 2017.
- [28] Crook, J. y Belloti, T. Loss given default models incorporating macroeconomic variables for credit cards. *International Journal of Forecasting*, 28(1),pp-171-182, 2012.
- [29] Cuevas, V. *Desafíos de la Industria Bancaria en la estimación de Loss Given Default.* Universidad de Chile, 2019.
- [30] Córdor, J. *Reporte de pruebas de tensión.* Recuperado de https://estadisticas.superbancos.gob.ec/portalestadistico/portalestudios/wp-content/uploads/sites/4/downloads/2020/06/pruebas_tension_mar_20.pdf , 2020.
- [31] End, J., Hoerberichts, M., y Tabbæ, M. Modelling scenario analysis and macro stress-testing. *Netherlands Central Bank, Research Department, DNB Working Papers*, 2006.
- [32] Flórez, S. El apalancamiento financiero y operativo: un gran desafío para los negocios en Colombia. *Unilibre*, 2011.
- [33] Frost, J. *Regression Analysis: An Intuitive Guide for Using and Interpreting Linear Models.* Recuperado de : <https://statisticsbyjim.com/regression/ols-linear-regression-assumptions/>, 2020.

- [34] García Díaz, C., C. G. y Molina, C. Análisis de la intermediación financiera en el escenario de las crisis de los siglos xx y xxi. *Redalyc*, 109-110., 2011.
- [35] Gonzáles, M. *Análisis de series temporales : modelos ARIMA*. Universidad del País Vasco. Recuperado de: es.slideshare.net, 2009.
- [36] González, C. *Stress testing de la morosidad del crédito de consumo en el sistema financiero ecuatoriano*. Recuperado de: <https://bibdigital.epn.edu.ec/>, 2017.
- [37] Gujarati, D. *Basic Econometrics*. The McGraw-Hill Companies, New York, 4th edition, 2004.
- [38] Gutierrez, J. *El modelo de Datos en Panel*. Recuperado de: <https://todoeconometria.com/paneldata1/>, 2020.
- [39] Hurlin, C. Leymarie, J. y Patin, A. Loss functions for loss given default model comparison. *European Journal of Operational Research.*, 2018.
- [40] INEC. Boletín técnico. *Ecuador en cifras*, 2022.
- [41] Khieu, H. Mullineaux, D. y H.-C. Y. The determinants of bank loan recovery rates. *Journal of Banking Finance.*, 2011.
- [42] Konečný, T. Seidler, J. B. A. y. B. K. *The time dimension of the links between loss given default and the macroeconomy*. 2017.
- [43] Kosak, M. y Poljšak, J. *Loss given default determinants in a commercial bank lending: an emerging market case study*. 2010.
- [44] López, F. Yaselga, E. y Espinosa, F. Modelo nowcast con factores dinámicos para la estimación trimestral del pib real para el ecuador. *Revista Cuestiones Económicas.*, 2021.
- [45] Marin, J. *Series Temporales*. Recuperado de: www.slideshare.net , (s.f).
- [46] Ochoa, J. Leano, W. y Agudelo, L. *Construcción de un modelo de scoring para el otorgamiento de crédito en una entidad financiera*. Recuperado de : http://www.scielo.org.co/scielo.php?script=sci_arttext&pid=S1657-42142010000200010, 2010.
- [47] Onur, B. y Togay, S. Efectos de la dolarización oficial en una pequeña economía abierta: el caso de ecuador. *Investigación económica*, 51-86, 2014.

- [48] Sabato G. y Schmid, M. *Estimating conservative loss given default*. Recuperado de <http://financedocbox.com/Investing/72424846-Estimating-conservative-loss-given-default.html> , S.F.
- [49] SB. *Glosario de Términos*. Recuperado de : <https://www.superbancos.gob.ec/bancos/glosario-de-terminos/#:~:text=Los%20dep%C3%B3sitos%20a%20la%20vista,mecanismos%20de%20pago%20y%20registro.> , 2022.
- [50] SEPS. *Fichas Metodológicas de Indicadores Financieros*. Recuperado de : <https://www.seps.gob.ec/> , 2017.
- [51] Urdinez, F. y Cruz, A. *Analizar Datos Políticos*. Recuperado de: <https://arcruz0.github.io/libroadp/panel.html> , 2021.
- [52] Uribe, J. *Inflación y crecimiento económico en Colombia*. *Borradores de Economía*, 1999.
- [53] Villano, F. *Cuantificación del riesgo de incumplimiento en créditos de libre inversión: un ejercicio econométrico para una entidad bancaria del municipio de Popayán, Colombia*. *Estudios Gerenciales*, 416-427, 2013.
- [54] Widarjono, A. *Ekonometrika Teori dan Aplikasi Untuk Ekonomi dan Bisnis*. Recuperado de: addi.ehu.es , 2013.
- [55] Wikipedia. *Modelo de efectos fijos* — *Wikipedia, La enciclopedia libre*. Recuperado de: https://es.wikipedia.org/w/index.php?title=Modelo_de_efectos_fijos&oldid=145126875 , 2022.
- [56] Wooldridge, J. *Introductory econometrics: A modern approach*. Cengage learning, 2013.
- [57] Zulfikar, R. *Estimation Model And Selection Method Of Panel Data Regression : An Overview Of Common Effect, Fixed Effect, And Random Effect Model*. Recuperado de: https://www.researchgate.net/publication/326142125_Estimation_Model_And_Selection_Method_Of_Panel_Data_Regression_An_Overview_Of_Common_Effect_Fixed_Effect_And_Random_Effect_Model , 2018.

Anexos

Apéndice A

A.1. Funciones de R

Paquetes de R

```
library(readxl)
library(tidyverse)
library(ggplot2)
library(stats)
library(tseries)
library(ggfortify)
library(forecast)
library(corrplot)
```

Funciones de Estacionariedad

```
gra_solv<-ts(grandes$SOLVENCIA, start = c(2017,1),
end = c(2020,1), frequency = 4)
autoplot(acf(gra_solv, plot = FALSE))
autoplot(pacf(gra_solv, plot = FALSE))
#test de dickey fuller aumentado
#ho: La serie no es estacionaria
#ha: La serie es estacionaria
adf.test(gra_solv, alternative = "stationary",k=0)
# no es estacionaria
# diferenciacion
adf.test(diff((gra_solv)), alternative="stationary", k=0)
```

A.1.1. Funciones de datos de panel

```
#MCO
reg.mco<- plm(FORMULAR_MCO, data = BANCOS_EST,
index=c("BANCO", "DATE"), model = "pooling")

#EFECTOS FIJOS
reg.plm.f=plm(FORMULAR_LGD, data = BANCOS_EST,
index=c("BANCO", "DATE"), model="within")

#EFECTOS ALEATORIOS
reg.plm=plm(FORMULAR_LGD, data = BANCOS_EST,
index=c("BANCO", "DATE"), model="ramdon")
```

A.1.2. Pruebas

Factor de tiempo

```
#EFECTOS FIJOS
reg.plm.ft=plm(FORMULAR_LGD+factor (DATE), data=BANCOS_EST,
index=c("BANCO", "DATE"), model="within"))

#Prueba para integrar factores de tiempo
#Multiplicadores de Lagrange
plmtest(reg.plm.f, c("time"), type=("bp"))

##
## Lagrange Multiplier Test - time effects (Breusch-Pagan)
## for balanced panels
##
## data: FORMULAR_LGD
## chisq = 0.53831, df = 1, p-value = 0.4631
## alternative hypothesis: significant effects

# en este caso no se usan efectos de tiempo puesto
que el p-valor es >0.05 no se rechaza H_0:
```

Test de Hausman FE vs RE

```
phtest (reg .plm . f , reg .plm)
```

Test de Chow FE vs MCO

```
pooltest (reg .mco , reg .plm . f)
```