

**Modelo estadístico para establecer la relación  
entre el indicador de cartera vencida total, la  
pérdida neta de crédito y el costo de cartera de  
un crédito de consumo libre inversión otorgado  
por una entidad bancaria de Bogotá, Colombia**

Oscar Mauricio Medina Monroy  
Henry Cárdenas Pinto

**Fundación Universitaria Los Libertadores**  
Departamento de Ciencias Básicas  
Especialización en Estadística Aplicada  
Bogotá  
2017

**Modelo estadístico para establecer la relación  
entre el indicador de cartera vencida total, la  
pérdida neta de crédito y el costo de cartera de  
un crédito de consumo libre inversión otorgado  
por una entidad bancaria de Bogotá, Colombia**

Oscar Mauricio Medina Monroy  
Henry Cárdenas Pinto

Asesor  
MSc. Juan Camilo Santana Contreras

**Fundación Universitaria Los Libertadores**  
Departamento de Ciencias Básicas  
Especialización en Estadística Aplicada  
Bogotá  
2017

Nota de aceptación

---

---

---

---

---

---

Firma del presidente del jurado

---

Firma del jurado

---

Firma del jurado

*Bogotá, D.C., 02 de 02 de 2017*

Las Directivas de la Universidad de los  
Libertadores, los jurados calificadores y el  
cuerpo docente, no son responsables por los  
criterios e ideas expuestas en el presente  
documento, éstos corresponden únicamente a  
los autores.

---

# Resumen

En este documento se presenta un modelo estadístico para establecer la relación entre el indicador de cartera vencida total, la pérdida neta de crédito y el costo de cartera, para un crédito de consumo de libre inversión otorgado por una entidad bancaria de Bogotá (Colombia). Los datos contienen información del periodo comprendido entre junio de 2011 hasta enero de 2015. Las diferentes pruebas estadísticas que se aplicaron al modelo propuesto sugieren que éste es consistente. La prueba de raíz unitaria de Dickey-Fuller aumentada permitió establecer que las series pérdida neta de crédito, costo de crédito e indicador de cartera vencida total no son estacionarias y tienen el mismo orden de integración. La prueba de cointegración de Engle-Granger proporcionó suficiente evidencia estadística sobre la relación de equilibrio a largo y corto plazo entre la pérdida neta de crédito y el indicador de cartera vencida total y entre el costo de crédito y la pérdida neta de crédito. Utilizando los datos históricos del indicador de cartera vencida total en el modelo estadístico propuesto junto con la información contable suministrada en el estado de pérdidas y ganancias, se realizó una simulación de Monte Carlo que permitió pronosticar el costo de cartera y el índice de rentabilidad sobre los activos. Con los resultados obtenidos es posible indicar a la entidad bancaria cual es la probabilidad de obtener utilidades mínimas para el producto de crédito de consumo de libre inversión.

**Palabras clave:** Indicador de cartera vencida total, pérdida neta de crédito, costo de cartera, rentabilidad sobre los activos, raíz unitaria, cointegración, regresión, simulación Monte Carlo.

---

# Abstract

This paper presents a statistical model to establish the relationship between the total overdue portfolio indicator, the net loss of credit and the cost of credit, for a credit of consumption of free investment granted by a bank of Bogotá (Colombia). The data contain information for a period from June 2011 to January 2015. The different statistical tests that were applied to the proposed model suggest that it is consistent. The augmented Dickey-Fuller unit root test allows us to establish that the series net loss of credit, credit cost and total overdue portfolio indicator are not stationary and have the same order of integration. The Engle-granger cointegration test provided sufficient statistical evidence on the long-run and short-run equilibrium relationship between net credit loss and the total overdue portfolio indicator and between the cost of credit and the net loss of credit. Using historical data from the total overdue portfolio indicator in the proposed statistical model together with the accounting information contained in the profit and loss statement, a Monte Carlo simulation was carried out, which allowed the forecast of the cost of credit and the rate of return on assets. With the results obtained it is possible to indicate to the bank what the probability of obtaining minimum profits for the product of consumer credit of free investment is.

**Key words:**Total overdue portfolio indicator, net loss of credit, cost of credit, rate of return on assets, unit root, cointegration, regresion, Monte Carlo simulation.

---

# Tabla de Contenido

<b>1. Introducción</b>	<b>1</b>
1.1. Planteamiento del problema . . . . .	3
1.2. Formulación del problema . . . . .	5
1.3. Justificación . . . . .	5
1.4. Objetivos . . . . .	6
1.4.1. Objetivo general . . . . .	6
1.4.2. Objetivos específicos . . . . .	6
<b>2. Marco de referencia</b>	<b>7</b>
<b>3. Marco metodológico</b>	<b>14</b>
3.1. Tipo de estudio . . . . .	14
3.2. Variables de estudio . . . . .	14
3.3. Unidad de análisis . . . . .	15
3.4. Consideraciones éticas . . . . .	15
3.5. Diseño estadístico . . . . .	15
3.5.1. Autocorrelación . . . . .	17
3.5.2. Función de autocorrelación y autocorrelación parcial . . . . .	18

---

3.5.3. Prueba de Durbin-Watson . . . . .	19
3.5.4. Prueba de Ljung-Box . . . . .	21
3.5.5. Pruebas de raíz unitaria . . . . .	22
3.5.6. Prueba de hipótesis para subconjuntos . . . . .	25
3.5.7. Prueba de cointegración de Engle-Granger . . . . .	27
3.5.8. Modelo de corrección de error . . . . .	28
3.5.9. Prueba de Jarque-Bera . . . . .	30
3.5.10. Prueba de Shapiro-Wilk . . . . .	31
3.5.11. Heteroscedasticidad . . . . .	31
3.5.12. Prueba de White . . . . .	32
3.5.13. Método de Monte Carlo . . . . .	33
3.5.14. Prueba de Anderson-Darling . . . . .	37
<b>4. Resultados y discusión</b>	<b>40</b>
4.1. Análisis descriptivo de las series temporales. . . . .	40
4.2. Prueba de Dickey-Fuller aumentada. . . . .	46
4.3. Regresiones del modelo . . . . .	55
4.4. Prueba de cointegración de Engle-Granger . . . . .	56
4.5. Pruebas de diagnóstico de los residuos de las regresiones del modelo	59
4.6. Modelo de corrección de error . . . . .	66
4.7. Pruebas de diagnóstico de los residuos del MCE . . . . .	68
4.8. Relación entre el CoCre y el ROA . . . . .	73
<b>5. Conclusiones</b>	<b>79</b>

---

---

A. Estadística descriptiva de las series temporales en R	82
B. Prueba de raíz unitaria de Dickey-Fuller aumentada en EViews 8	85
C. Prueba de cointegración de Engle-Granger y modelo de corrección de error en EViews 8	88
D. Prueba de cointegración de Engle-Granger y modelo de corrección de error en R	94
Referencias	98

---

# Índice de Tablas

2.1. Estado del cliente de acuerdo a su rango de mora. . . . .	8
3.1. Características de las variables de estudio. . . . .	14
4.1. Resumen estadística descriptiva. . . . .	40
4.2. Prueba de normalidad de las series de tiempo. . . . .	42
4.3. Contraste $\phi_3$ para la serie del ICVT. . . . .	47
4.4. Contraste $\phi_2$ para la serie del ICVT. . . . .	47
4.5. Prueba de Dickey-Fuller aumentada de la serie del ICVT y su primera diferenciación para regresión auxiliar con intercepto y tendencia. . . . .	47
4.6. Contraste $\phi_3$ para la serie de la PNC. . . . .	50
4.7. Contraste $\phi_2$ para la serie de la PNC. . . . .	50
4.8. Prueba de Dickey-Fuller aumentada de la serie del PNC y su primera diferenciación para regresión auxiliar sin intercepto ni tendencia. . . . .	50
4.9. Contraste $\phi_3$ para la serie del CoCre. . . . .	52
4.10. Contraste $\phi_2$ para la serie del CoCre. . . . .	53
4.11. Prueba de Dickey-Fuller aumentada de la serie del CoCre y su primera diferenciación para regresión auxiliar con intercepto y tendencia. . . . .	54
4.12. Regresión entre la PNC y el ICVT . . . . .	55

---

4.13. Regresión entre el CoCre y la PNC. . . . .	56
4.14. Prueba de cointegración y raíz unitaria para los residuos de la regresión entre el PNC y el ICVT. . . . .	57
4.15. Prueba de cointegración y raíz unitaria para los residuos de la regresión entre el CoCre y el PNC. . . . .	58
4.16. Resumen de las pruebas de diagnóstico de los residuos de la regresión del ICVT contra el PNC. . . . .	62
4.17. Resumen de las pruebas de diagnóstico de los residuos de la regresión del CoCre contra el PNC. . . . .	65
4.18. Resultado de la regresión por MCO del MCE para la relación entre la PNC y el ICVT. . . . .	67
4.19. Resultado de la regresión por MCO del MCE para la relación entre el CoCre y la PNC. . . . .	67
4.20. Resumen de las pruebas de diagnóstico de los residuos del MCE para la relación entre el ICVT y la PNC. . . . .	70
4.21. Resumen de las pruebas de diagnóstico de los residuos del MCE para la relación entre el CoCre y la PNC. . . . .	73
4.22. Estado de resultados publicado por la vicepresidencia financiera de la entidad bancaria para el año 2015. . . . .	73
4.23. Resultados de la prueba de bondad de ajuste de Anderson-Darling para los datos del ICVT. . . . .	75
4.24. Resumen estadística descriptiva del ajuste a una distribución gamma de los datos del ICVT. . . . .	76
4.25. Nivel de certidumbre del ROA para utilidades mínimas entre 0.5 % y 2.0 %. . . . .	78

---

# Índice de figuras

1.1. Cartera de consumo por tipo de producto. Fuente: Superintendencia Financiera de Colombia . . . . .	2
3.1. Arbol de decisión para elegir el contraste en la prueba ADF. . . . .	25
4.1. Histograma y boxplot de la serie índice de cartera vencida total. . . . .	41
4.2. Histograma y boxplot de la serie pérdida neta de crédito. . . . .	41
4.3. Histograma y boxplot de la serie costo de crédito. . . . .	42
4.4. Serie del índice de cartera vencida total (ICVT) y su primera diferenciación. . . . .	43
4.5. Función de autocorrelación y autocorrelación parcial del ICVT. . . . .	43
4.6. Serie de la pérdida neta de crédito (PNC) y su primera diferenciación. . . . .	44
4.7. Función de autocorrelación y autocorrelación parcial del PNC. . . . .	44
4.8. Serie del costo de crédito (CoCre) y su primera diferenciación. . . . .	45
4.9. Función de autocorrelación y autocorrelación parcial del CoCre. . . . .	45
4.10. Correlograma de los residuos de la regresión con intercepto y tendencia con cero rezagos para la serie del ICVT. . . . .	46
4.11. Correlograma de los residuos de la regresión con intercepto y tendencia con un rezago para la serie del ICVT. . . . .	48

---

4.12. Correlograma de los residuos de la regresión con intercepto y tendencia con cero rezagos para la serie de la PNC. . . . .	49
4.13. Correlograma de los residuos de la regresión sin intercepto y tendencia con dos rezagos para la serie de la PNC. . . . .	51
4.14. Correlograma de los residuos de la regresión con intercepto y tendencia con cero rezagos para la serie del CoCre. . . . .	52
4.15. Correlograma de los residuos de la regresión con intercepto y tendencia con cuatro rezagos para la serie del CoCre. . . . .	53
4.16. Gráfico de los residuos de la regresión PNC vs. ICVT . . . . .	57
4.17. Gráfico de los residuos de la regresión CoCre vs. PNC . . . . .	58
4.18. Gráfico de probabilidad normal y boxplot de los residuos de la regresión de la PNC contra el ICVT. . . . .	60
4.19. Histograma de los residuos de la regresión de la PNC contra el ICVT. . . . .	60
4.20. Correlograma de los residuos de la regresión de la PNC contra el ICVT. . . . .	61
4.21. Gráfico de los residuos y el cuadrado de los residuos contra la PNC estimada. . . . .	61
4.22. Gráfico de la serie pérdida neta de crédito y pérdida neta de crédito estimada. . . . .	62
4.23. Gráfico de probabilidad normal y boxplot de los residuos de la regresión del CoCre contra el PNC. . . . .	63
4.24. Histograma de los residuos de la regresión del CoCre contra la PNC. . . . .	63
4.25. Correlograma de los residuos de la regresión del CoCre contra el PNC. . . . .	64
4.26. Gráfico de los residuos y el cuadrado de los residuos contra el CoCre. . . . .	64
4.27. Gráfico de la serie costo de crédito y costo de crédito estimada. . . . .	66
4.28. Gráfico de probabilidad normal y boxplot de los residuos de la regresión por MCO con MCE para la relación entre el ICVT y la PNC . . . . .	68

---

---

4.29. Histograma de los residuos de la regresión por MCO con MCE para la relación entre el ICVT y la PNC. . . . .	69
4.30. Correlograma de los residuos de la regresión por MCO con MCE para la relación entre el el ICVT y la PNC. . . . .	69
4.31. Gráfico de los residuos y el cuadrado de los residuos del MCE contra la primera diferenciación del PNC. . . . .	70
4.32. Gráfico de probabilidad normal y boxplot de los residuos de la regresión por MCO con MCE para la relación entre el CoCre y la PNC. . . . .	71
4.33. Histograma de los residuos de la regresión por MCO con MCE para la relación entre el CoCre y la PNC. . . . .	71
4.34. Correlograma de los residuos de la regresión por MCO con MCE para la relación entre el CoCre y la PNC. . . . .	72
4.35. Gráfico de los residuos y el cuadrado de los residuos de MCE contra la primera diferenciación del CoCre. . . . .	72
4.36. PNC observado y pronosticado con intervalos de confianza del 95 %.	74
4.37. CoCre observado y pronosticado con intervalos de confianza del 95 %.	74
4.38. Datos del ICVT ajustados a una distribución gamma según el criterio de Anderson-Darling. . . . .	76
4.39. Pronóstico del CoCre para crédito de consumo de libre inversión. . .	77
4.40. Pronóstico del ROA para crédito de consumo de libre inversión. . .	77
4.41. Pronóstico del ROA para utilidades mínimas entre 0.5 % y 2.0 % . .	78
B.1. Prueba ADF a nivel intercepto y tendencia de la serie ICVT con cero rezagos . . . . .	86
C.1. Regresión por MCO de la serie PNC contra ICVT . . . . .	89

---

---

C.2. Prueba de cointegración de Engle-Granger para los residuos de la regresión. . . . .	90
C.3. Prueba de heroscedasticidad de White para los residuos de la regresión. 91	
C.4. . . . .	93
C.5. Resultados de la estimación por MCO del modelo de corrección de error para la relación entre la PNC y el ICVT. . . . .	93

---

# Capítulo 1

## Introducción

De acuerdo al último reporte de estabilidad financiera presentado por el Banco de la República (*Reporte de estabilidad financiera*, 2014) el crecimiento real de la actividad económica mundial durante el 2014 fue más débil de lo esperado, explicado, en parte, por la fuerte caída de la cotización internacional del petróleo, la incertidumbre frente a la situación fiscal de Grecia, la gradual desaceleración de la economía China y de las economías emergentes. En el caso de América Latina se presentó una fuerte desaceleración económica como consecuencia de la reducción de las exportaciones, especialmente de materias primas, la disminución en precios de los bienes básicos y de las dificultades socio-políticas internas de algunos países de la región. A pesar de este difícil entorno externo Colombia presentó el mejor desempeño regional con un crecimiento en su actividad económica del 4.6 % en 2014.

El adecuado crecimiento económico que se presentó en Colombia en el periodo 2010-2014, impulsado por actividades como la construcción, los servicios financieros y el incremento de la inversión extranjera, vinculada principalmente a la industria minero-energética, provocó un desempeño positivo de la cartera de créditos durante estos años. Sin embargo, la dinámica económica del país en el 2015 se ha visto afectada principalmente por la fuerte caída de la cotización internacional del petróleo y la devaluación del peso frente al dolar, ocasionando un deterioro en el balance externo del país debido a la disminución de la exportaciones de bienes básicos y materias primas (*Reporte de estabilidad financiera*, 2014).

El ritmo de crecimiento y la calidad de la cartera de créditos en el mercado colombiano se ha mantenido estable en los últimos años gracias a factores como: una mayor demanda de créditos en las modalidades de comercio, consumo, vivienda y microcrédito, el fortalecimiento patrimonial de las instituciones financieras, un mayor nivel en la profundización de la banca, los ajustes en las políticas de otorgamiento de crédito y la implementación de mecanismos adecuados para el cubrimiento de los créditos mediante un sistema de provisiones.

De acuerdo a cifras reportadas por el Banco de la República (*Reporte de estabilidad financiera*, 2014), en diciembre del 2014 se observó que el ritmo de crecimiento de la cartera bruta fue superior que en junio, explicado por la dinámica de la cartera comercial y el comportamiento estable de la cartera de consumo en los últimos meses del año, lo cual compensó la desaceleración en las modalidades de vivienda y microcrédito. Se aprecia que la cartera comercial sigue siendo la más representativa del mercado (56.3%), seguida de la de consumo (29.6%), vivienda (11.1%) y microcrédito (3%).

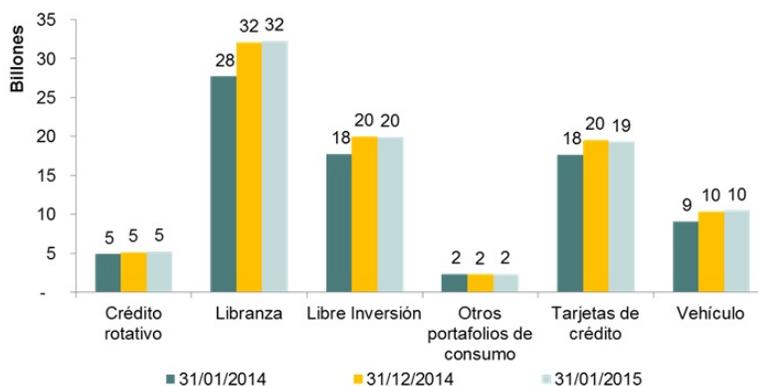


Figura 1.1: Cartera de consumo por tipo de producto. Fuente: Superintendencia Financiera de Colombia

Como se puede observar la cartera de consumo ocupa un reglón muy importante de la cartera de créditos en Colombia y se ha caracterizado por presentar un crecimiento considerable y un comportamiento relativamente estable en años recientes, lo anterior resultado del aumento en la rentabilidad de los establecimientos financieros en la colocación de recursos en la cartera de consumo, la disminución en

---

los costos de financiación producto de la baja en la tasa de interés de referencia y el crecimiento en la clase media y la mejora en sus ingresos. Sin embargo, esta disminución en las tasas de interés, que hace que la demanda de crédito de consumo crezca, puede también influir de manera significativa en el riesgo de crédito de consumo y en general en el sistema crediticio (Martinez, 2014).

Según cifras reportadas por la Superintendencia Financiera de Colombia (*Evolución del sistema financiero colombiano*, 2015), para diciembre del 2014, dentro de los diferentes productos que componen la cartera de consumo en Colombia, se destaca la participación en el mercado del crédito de libranza (32 %), libre inversión (20 %) y las tarjetas de crédito (20 %). Las estrategias utilizadas por algunas compañías de financiamiento comercial y algunos bancos para incrementar el otorgamiento de crédito al público en general, en especial, en productos como las tarjetas de crédito y el crédito de libre inversión, generan altos niveles de rentabilidad a corto plazo, gracias a las elevadas tasas de colocación y las comisiones por servicios en tarjetas de crédito. Sin embargo, esta situación contribuye al aumento del capital en riesgo de los establecimientos financieros, ante un eventual incumplimiento de sus deudores, y esto puede afectar la rentabilidad y sostenibilidad de las instituciones de crédito (Martinez, 2014), (Salazar, 2013).

## 1.1. Planteamiento del problema

En Colombia todas las instituciones autorizadas a otorgar crédito, y cuya actividad esta supervisada por la Superintendencia Financiera de Colombia, están obligadas a diseñar e implementar un sistema administrativo de riesgo de crédito (SARC), en conformidad con lo establecido en el acuerdo de capitales de Basilea II, que tiene como propósito principal servir como un mecanismo de alerta temprana, que le permita a las instituciones de crédito detectar una disminución en sus activos ante un eventual incumplimiento en el pago de sus obligaciones por parte de sus deudores.

El sistema administrativo de riesgo de crédito (SARC) conlleva la aplicación de modelos de referencia para estimar las pérdidas esperadas que permitan establecer el nivel de provisiones que deben crear las instituciones de crédito para cubrir el riesgo de crédito de sus deudores. Esto implica que los establecimientos crediticios

---

---

deben realizar un gasto en provisiones (GPR) que proviene directamente de sus utilidades y determinará el costo de cartera o de crédito (CoCre). El gasto en provisiones (GPR) junto con el indicador de cartera vencida total (ICVT) son dos de las medidas más utilizadas para evaluar el riesgo crediticio (Sagner, 2011).

Un manejo inapropiado del riesgo de crédito afectará las provisiones, y por ende, las utilidades de las instituciones de crédito. Asimismo, las instituciones de crédito necesitan establecer cuánto del dinero desembolsado ya no es recuperable después de un determinado periodo de tiempo, ya que esto tiene un efecto sobre sus activos crediticios y en sus utilidades, lo que se logra mediante el indicador de pérdida neta de crédito (PNC). El efecto del riesgo de crédito sobre la calidad de cartera se manifiestan sobre la capacidad que tienen las instituciones de crédito para generar utilidades a partir de sus activos crediticios y se suele medir mediante el indicador de rentabilidad sobre los activos (ROA). Mediante información del indicador de cartera vencida total (ICVT) y la pérdida neta de crédito (PNC) es posible establecer una relación entre el costo de crédito (CoCre) y el índice de rentabilidad sobre los activos (ROA).

La utilización de modelos para la estimación del riesgo crediticio y la calidad de cartera se ha incrementado en los últimos años debido a los pronunciamientos dados en el Acuerdo de Capitales de Basilea II (*Aplicación de Basilea 2-Apectos prácticos*, 2004) y a los efectos de la crisis financiera mundial del periodo 2008-2011. En el caso colombiano se han realizado estudios de carácter descriptivo para estimar el riesgo crediticio y su efecto sobre el índice de rentabilidad (ROA)(Martinez, 2014), también se han desarrollado algunos modelos econométricos basados en series de tiempo que utilizan modelos de vectores autoregresivos (VAR) para examinar la relación causal entre la cartera de las entidades financieras y del sistema, y la cartera vencida considerando variables macroeconómicas como: el índice de producción manufacturera, la tasa de interés real y la demanda nacional de energía y variables microeconómicas como: el saldo de la cartera respecto de la cartera del sistema financiero, los gastos administrativos y monto real de los préstamos sin garantía (Giraldo, 2010).

Además, se han realizado modelos discriminantes logit y modelos log-log para estimar el riesgo de incumplimiento y el índice de cartera vencida, respectivamente, en créditos de libre inversión en una entidad bancaria de Colombia considerando, en el primer caso, información de los clientes y sus créditos, y en el segundo caso,

---

información de variables macro como: el indicador de actividad económica local, la tasa de interés activa del banco, la tasa de desempleo local y la quiebra de corredores financieros no autorizados como variable *dummy* (Salazar, 2013).

## 1.2. Formulación del problema

Por las razones expuestas anteriormente se plantea la siguiente pregunta de investigación: ¿ existe una relación estadística entre las variables pérdida neta de crédito (PNC), indicador de cartera vencida total (ICVT) y costo de cartera (CoCre) como insumo para proponer un índice de rentabilidad sobre los activos (ROA) de un crédito de libre inversión otorgado a personas naturales por un establecimiento bancario de Bogotá, Colombia?

## 1.3. Justificación

La crisis económica mundial de 2008 a 2011, que fue iniciada por el colapso de la burbuja inmobiliaria en Estados Unidos en el año 2006 y que luego provocaría la crisis de las hipotecas *subprime* en 2007, ha sido una de las razones principales por la cual en los últimos años se ha incrementado la utilización de diferentes modelos para estimar el riesgo crediticio al que están expuestos los establecimientos financieros.

Por ésta razón, y siguiendo las recomendaciones hechas en el Acuerdo de Capitales de Basilea II, se hizo necesario que los organismos de regulación de la actividad financiera en cada país, desarrollaran modelos de referencia basados en indicadores internos, con el propósito de estimar a futuro el nivel de pérdidas esperadas de las entidades financieras, para que estas pueden disponer el nivel de provisiones adecuado que les permita ajustar sus políticas de otorgamiento de crédito a sus clientes, ya que una sobreestimación de sus provisiones puede reducir eventualmente la rentabilidad de sus activos crediticios, mientras que su subestimación puede generar mayores pérdidas.

Es por esta razón que se deben explorar modelos que utilizan el comportamiento

---

---

histórico de variables e indicadores financieros de los establecimientos de crédito, como el indicador de cartera vencida total (ICVT), la pérdida neta de crédito (PNC) y el costo de cartera (CoCre), que permitan determinar que tan eficiente es una entidad de crédito para generar utilidades a partir de sus activos representados en su inversión crediticia, en particular, en el portafolio de libre inversión, ya que esto permite reflejar, en cierta medida, la calidad de la cartera y proporciona la información necesaria para que la institución de crédito constituya las provisiones necesarias para cada periodo, que garanticen su solvencia y sostenibilidad.

## 1.4. Objetivos

### 1.4.1. Objetivo general

Evaluar estadísticamente la existencia de una relación entre las variables pérdida neta de crédito (PNC), indicador de cartera vencida total (ICVT) y costo de cartera (CoCre) para proponer un índice de rentabilidad sobre los activos (ROA) de un crédito de libre inversión otorgado a personas naturales por un establecimiento bancario de Bogotá, Colombia.

### 1.4.2. Objetivos específicos

- Determinar el orden de integración de las series de tiempo pérdida neta de crédito (PNC), indicador de cartera vencida total (ICVT) y costo de cartera o crédito (CoCre).
  - Analizar la estacionariedad de las series de tiempo bajo estudio.
  - Identificar si existe una relación de equilibrio a largo plazo entre las series de la PNC y el ICVT y entre el CoCre y la PNC.
  - Analizar los efectos del corto plazo sobre la dinámica de largo plazo de las relaciones entre la PNC y el ICVT y entre el CoCre y la PNC.
  - Identificar el modelo más adecuado basado en series de tiempo para pronosticar la pérdida neta de crédito (PNC) y el costo de cartera (CoCre).
-

# Capítulo 2

## Marco de referencia

La cartera de créditos en Colombia se compone de las siguientes modalidades: comercial, consumo, vivienda y microcrédito. De acuerdo a la circular básica contable y financiera número 100 de 1995, publicada por la Superintendencia Financiera de Colombia (*Capítulo 2: Reglas Relativas a la Gestión de Riesgo Crediticio. Circular Externa 035 de 2009*, 2009), cada una de estas modalidades de crédito se define como sigue:

- Los créditos comerciales son los concedidos a personas naturales o jurídicas para el desarrollo de actividades económicas organizadas.
- Los créditos de consumo son los otorgados a personas naturales para financiar la adquisición de bienes de consumo o el pago de servicios para fines no comerciales o empresariales.
- Los créditos de vivienda, independientemente del monto, son aquellos concedidos a personas naturales destinados a la adquisición de vivienda nueva o usada o a la construcción de vivienda individual.
- El microcrédito es el constituido por las operaciones activas de crédito a las cuales se refiere el artículo 39 de la Ley 590 de 2000, o las normas que lo modifiquen, sustituyan o adicionen, así como las realizadas con microempresas en las cuales la principal fuente de pago de la obligación provenga de los ingresos derivados de su actividad.

---

Cada una de estas modalidades puede subdividirse en diferentes productos o portafolios dependiendo de la metodología interna de cada establecimiento de crédito.

En lo concerniente a las organizaciones o instituciones involucradas en la gestión y control del crédito en Colombia se ha establecido que: el Banco de la República, fija la tasa de intervención de la política monetaria que afecta al sistema de crédito, la Superintendencia Financiera de Colombia, debe supervisar y controlar el sistema en forma eficaz y oportuna realizando monitoreo permanente al comportamiento de la cartera de crédito y las instituciones otorgantes de créditos, son responsables de permitir el acceso a personas naturales y jurídicas a los recursos (Martínez, 2014).

La oferta de crédito de consumo en Colombia proviene de diferentes sectores como los establecimientos bancarios, corporaciones financieras, cooperativas financieras, cajas de compensación familiar, fondos de empleados y en gran medida de las compañías de financiamiento comercial, considerando que el crédito de consumo se constituye en el principal estimulador de la demanda para las empresas productoras y comerciantes. La cartera de consumo esta constituida por los siguientes portafolios o productos: libranza, crédito de libre inversión, tarjetas de crédito, crédito de vehículo y crédito rotativo. De acuerdo a cifras reportadas por la Superintendencia Financiera de Colombia, para diciembre del 2014, la libranza, el crédito de libre inversión y las tarjetas de crédito fueron los portafolios que tuvieron mayor intervención en el mercado (*Evolución del sistema financiero colombiano*, 2015).

Estado del cliente	Rango de mora
Bucket 0 ( $B_0$ )	0 días (al corriente)
Bucket 1 ( $B_1$ )	1-30 días
Bucket 2 ( $B_2$ )	31-60 días
Bucket 3 ( $B_3$ )	61-90 días
Bucket 4 ( $B_4$ )	91-120 días
Bucket 5 ( $B_5$ )	121-150 días
Bucket 6 ( $B_6$ )	151-180 días
Bucket 7 ( $B_7$ )	Mayor o igual a 181 días

Tabla 2.1: Estado del cliente de acuerdo a su rango de mora.

---

Por otra parte, la cartera de crédito se puede clasificar, según el comportamiento de pago de los deudores, en las categorías indicadas en la tabla 2.1. Cada *bucket* se define como la cartera que esta en un rango de mora determinado y caracteriza el estado del cliente. En el momento en que una entidad financiera le otorga a los clientes un crédito les asigna una altura de mora correspondiente al *bucket 0*, es decir, los clientes no tienen retraso con la mora y se les considera clientes al día o al corriente.

Cuando un cliente adquiere un bien de consumo, por ejemplo, a través de un crédito de libre inversión o una tarjeta de crédito, el establecimiento de crédito la asigna generalmente un periodo de un mes para liquidar su deuda sin cobrarle intereses. En teoría se espera que el cliente cuente con el dinero para pagar su compra y se mantenga al corriente. Si el cliente no cuenta con el capital suficiente para cubrir el monto total puede cubrir una cantidad que va desde el mínimo establecido por la institución de crédito hasta el saldo total antes de la fecha límite de pago, y quedará como saldo el monto de la deuda, sin que se cobren intereses moratorios ni cargos por cobranza (Nieto, 2010).

La entidad que otorga el crédito establece los intereses mensuales que se aplicaran por falta de pago de la deuda o intereses moratorios, la *fecha límite de pago*, la *fecha pago mínimo*, que se considera como el pago anticipado de intereses, y la *fecha de corte o de facturación*, que es un día de cada mes en que el establecimiento de crédito revisa el estado de cuenta del cliente y donde se calcula el saldo entre la fecha de corte actual y la fecha de corte anterior. Este intervalo de tiempo se denomina *periodo de corte* y suele ser de 30 días (Nieto, 2010).

Para determinar el estado de los clientes se realizan los estados de cuenta en la fecha de corte de cada mes. Dependiendo de los pagos que realicen los clientes pasan de un estado a otro. Si en un determinado mes del año el establecimiento de crédito realiza un desembolso de dinero a los clientes, esta deuda pasa a ocupar el *bucket 0*, que corresponde a la cartera con 0 días de mora, y en ese estado se clasifican los clientes. Si al siguiente mes los clientes pagan parcialmente o la totalidad de su deuda se quedan en el *bucket 0* ocasionando una amortización natural de su deuda. Sin embargo, si no pagan la totalidad de su deuda ni realizan el pago mínimo pasan al *bucket 1*, que corresponde a la cartera con 1 a 30 días de mora, y si la tendencia de no pago se mantiene continuaran descendiendo hasta el *bucket*

---

6, provocando un *rodamiento* de una parte del dinero desembolsado hacia estos *bucket de mora*. No obstante, se sabe por experiencia que gran parte del *rodamiento* ocurre entre el *bucket 0* y el *bucket 1*. El *rodamiento* del dinero desembolsado hacia los *bucket de mora* se puede determinar mediante la *tasa de rodamiento* o *roll rate* que es un indicador que permite comparar dos alturas de mora en dos instantes de tiempo diferentes.

Los clientes pueden retornar uno o mas *bucket* hacia atrás, cuando reestructuran su deuda, dependiendo del pago que realicen, sin embargo, por experiencia se sabe que cuando los clientes caen en el *bucket 4* ya es muy difícil que regresen al *bucket 0*. Finalmente, a partir de los 181 días de mora o *bucket 7*, los clientes pasan al grupo de pérdida, un estado del que ya no saldrán, donde se considera que la deuda ya es incobrable. Para determinar cuanto del dinero desembolsado por la entidad de crédito ya no es recuperable se utiliza un indicador conocido como la pérdida neta de crédito (PNC), definido como la proporción de cartera al día que alcanza los 181 días de mora y que se calcula como el producto de las *tasas de rodamiento* entre franjas de mora desde el *bucket 0* hasta el *bucket 7*.

La situación que se ha descrito anteriormente corresponde al desembolso de un solo mes, pero, la institución de crédito realizará desembolsos durante todos los meses del año y esto ocasionará que los *bucket de mora*, correspondientes al desembolso de un mes en particular, se solapen con los *bucket de mora* asociados a los desembolsos de meses anteriores y posteriores. Teniendo esto en cuenta, el establecimiento de crédito necesita establecer, cada mes, que porcentaje de la cartera total está en alturas de moras superiores a 31 días. Para esto se utiliza el indicador de cartera vencida total (ICVT) que mide la proporción entre el volumen de cartera con mora mayor a 31 días y el volumen total de cartera disponible en un momento del tiempo (Sagner, 2011).

Por otra parte, todas aquellas entidades vigiladas por la Superintendencia Financiera de Colombia que dentro de su objeto social principal se encuentren autorizadas para otorgar crédito deben evaluar permanentemente el riesgo incorporado en sus activos crediticios, tanto en el momento de otorgar créditos como a lo largo de la vida de los mismos, incluidos los casos de reestructuraciones. Para tal efecto, de conformidad con la circular básica contable número 100 de 1995 emanada por la Superintendencia Financiera de Colombia, las entidades de crédito deben

---

diseñar e implementar un sistema administrativo de riesgo de crédito (SARC), que debe contener políticas y procedimientos claros y precisos que permitan evitar que las instituciones incurran en pérdidas o disminución de sus activos debido al incumplimiento de sus clientes en el pago de sus obligaciones (*Capítulo 2: Reglas Relativas a la Gestión de Riesgo Crediticio. Circular Externa 032 de 2014, 2014*). Este sistema administrativo de riesgo de crédito (SARC) debe contener los siguientes componentes básicos:

- Modelos internos o de referencia para estimación de pérdidas esperadas.
- Sistema de provisiones para cubrir el riesgo crediticio.
- Políticas de administración del riesgo crediticio.
- Procesos de administración del riesgo crediticio.
- Procesos de control interno.

La Superintendencia Financiera de Colombia como entidad supervisora del sistema ha diseñado, en concordancia con lo establecido en el acuerdo de capitales de Basilea II, los modelos de referencia para la estimación o cuantificación de la pérdida esperada en cada modalidad de crédito y sus diversos portafolios (*Aplicación de Basilea 2-Apectos prácticos, 2004*). El propósito del supervisor es determinar un mecanismo que permita establecer la componente de pérdida esperada  $PE$  en función del monto de crédito otorgado a los clientes y la probabilidad de deterioro que sufra la cartera a través del tiempo, para lo cual dispone la siguiente fórmula:

$$PE = (PI)(VEA)(PDI)$$

donde  $PI$  es la probabilidad de incumplimiento de los deudores,  $VEA$  es el valor expuesto del activo y  $PDI$  es la pérdida dado el incumplimiento. La probabilidad de incumplimiento  $PI$  es la probabilidad de que en un lapso de doce meses los clientes incurran en incumplimiento debido al no pago de sus obligaciones. Para el caso particular del crédito de consumo el incumplimiento ocurre cuando una operación de crédito se encuentra en mora mayor a 90 días. El valor expuesto del activo  $VEA$  es el saldo de la obligación al momento del cálculo de la pérdida esperada y cubre el saldo, intereses, cargos fijos, entre otros. La pérdida dado el incumplimiento  $PDI$  es el deterioro económico en que incurriría la entidad crediticia en el caso en que se materialice el incumplimiento. Las pérdidas esperadas disminuirán

---

entre más alta sea la tasa de recuperación esperada y aumentarán en función del monto del crédito y de la probabilidad de deterioro de cada activo (*Capítulo 2: Reglas Relativas a la Gestión de Riesgo Crediticio. Circular Externa 027 de 2011, 2014*). La probabilidad de incumplimiento y la pérdida dado el incumplimiento son cantidades que ya se han determinado en los modelos de referencia mediante la aplicación de técnicas estadísticas como modelos Logit y Probit, análisis discriminante y matrices de transición (Salazar, 2013), (Nieto, 2010), (Aristizabal, Támara, y Velásquez, 2010), (Lando, 2004).

Las pérdidas esperadas obtenidas de la aplicación del modelo de referencia establecen el nivel de provisiones que debe constituir la institución de crédito para cubrir el riesgo de crédito de sus deudores y eso determina el nivel de caja con el que puede contar la institución para poder ejercer su labor de colocación de recursos en el público. El gasto en provisiones junto con el indicador de cartera vencida total son dos de las medidas más utilizadas para evaluar el riesgo de crédito y son indicadores elaborados a partir de información contable (Sagner, 2011).

El gasto en provisiones es una medida basada en pérdidas esperadas ya que corresponde al reconocimiento anticipado por parte de las instituciones de crédito de las pérdidas asociadas a la materialización del evento de no pago por parte de los deudores. En términos contables, el gasto en provisiones significa una disminución del patrimonio de las instituciones debido al incremento del stock de provisiones que afecta directamente las utilidades de la entidad de crédito. El costo de cartera o de crédito (CoCre) corresponde al gasto en provisiones que debe constituir la entidad con respecto al stock de cartera que se tuvo en promedio durante un periodo de tiempo, que por lo general es de un año (Sagner, 2011). Un manejo inadecuado del riesgo de crédito afectará las provisiones y pondrá en riesgo la utilidad de la entidad ya que el gasto en provisiones proviene directamente de las utilidades del establecimiento de crédito. Un indicador que permite determinar que tan eficiente es una institución crediticia para generar utilidades a partir de sus activos, representados en su inversión crediticia, es el índice de rentabilidad sobre los activos (ROA) que mide la proporción entre la suma de utilidades mensuales y el activo total (Martinez, 2014).

El costo de crédito (CoCre) y el índice de rentabilidad sobre los activos (ROA) se relacionan a través de la información contable presentada en el estado de resultados

---

o estados de pérdidas y ganancias (PyG) elaborado por el departamento financiero y contable de la entidad bancaria. El estado de resultados es un reporte financiero anual en el que se reflejan los ingresos y gastos, y en consecuencia, el beneficio o pérdida que ha generado la entidad en ese periodo de tiempo en relación con su actividad crediticia.

Los elementos de un estado financiero se agrupan de la siguiente manera: ingresos, costos y gastos. En base a estos tres rubros principales se estructura el reporte. Las principales cuentas que lo conforman son las siguientes:

- Ingresos: comprende los intereses devengados por todos los activos financieros con rendimiento.
  - Costo de los fondos: es un porcentaje anual establecido por el Banco de la República para los intereses generados por los préstamos interbancarios, con el objetivo de hacer frente a nuevos créditos u obligaciones, como retiros bancarios, cuando la entidad bancaria carece de depósitos o reservas suficientes.
  - Utilidad financiera: se refiere a la diferencia que se obtiene al restar los ingresos y el costo de los fondos.
  - Utilidad después del CoCre: hace referencia a la diferencia que se obtiene al restar la utilidad financiera y el costo del crédito, e indica la ganancia o pérdida de la entidad bancaria en función de su actividad crediticia.
  - Gastos: en este rubro se incluyen todos aquellos gastos que están directamente involucrados con la operación de crédito.
  - Utilidad antes de impuestos: este concepto se refiere a la ganancia o pérdida de la entidad bancaria después de cubrir sus compromisos operacionales y financieros.
  - Impuestos: son las contribuciones sobre las utilidades que la entidad bancaria paga al gobierno.
-

# Capítulo 3

## Marco metodológico

### 3.1. Tipo de estudio

Se propone un estudio de tipo correlacional dado que se describirá la relación existente entre los indicadores financieros: indicador de cartera vencida total (ICVT), pérdida neta de crédito (PNC) y costo de crédito (CoCre), para proponer un índice de rentabilidad sobre los activos (ROA) de un crédito de libre inversión otorgado a personas naturales por un establecimiento bancario de Bogotá, Colombia.

### 3.2. Variables de estudio

Nombre de la variable	Tipo de variable	Unidad de medida
Índice de cartera vencida total (ICVT)	Cuantitativa	%, mensual
Pérdida neta de crédito (PNC)	Cuantitativa	%, mensual
Costo de crédito (CoCre)	Cuantitativa	%, mensual

Tabla 3.1: Características de las variables de estudio.

---

### 3.3. Unidad de análisis

Las series temporales consideradas comprenden datos mensuales desde junio de 2011 hasta enero de 2015 de los siguientes indicadores financieros: indicador de cartera vencida total (ICVT), pérdida neta de crédito (PNC) y costo de cartera o crédito (CoCre). Los datos fueron obtenidos de la Superintendencia Financiera de Colombia para un producto de crédito de consumo en su modalidad de libre inversión otorgado a personas naturales por un establecimiento bancario de Bogotá, Colombia.

### 3.4. Consideraciones éticas

La información de los indicadores financieros se obtuvo de las bases de datos suministradas por la Superintendencia Financiera de Colombia y son de libre acceso al público.

### 3.5. Diseño estadístico

1. Se estableció el orden de integración de las series de tiempo: indicador de cartera vencida total (ICVT), pérdida neta de crédito (PNC) y costo de crédito (CoCre), a ser incluidas en el modelo. Al determinar el orden de integración pueden presentarse tres situaciones (Mata, 2004):
    - Si las series son integradas de orden  $I(0)$  se aplica el procedimiento estándar de mínimos cuadrados ordinarios (MCO).
    - Si las series resultan integradas de diferente orden, como  $I(0)$ ,  $I(1)$ ,  $I(2)$ , es posible concluir que las mismas no estén cointegradas, y la aplicación de mínimos cuadrados ordinarios (MCO) producirá resultados espurios, es decir, valores altos del coeficiente de correlación  $R^2$  y los estadísticos  $t$  y  $F$  y un valor muy bajo del estadístico  $d$  de Durbin-Watson.
    - Por último, puede ocurrir que las series tengan el mismo orden de integración, por ejemplo  $I(1)$ . En general, cualquier combinación lineal de series de orden  $I(1)$  será también de orden  $I(1)$ . Sin embargo, puede haber alguna combinación lineal particular entre las series de orden  $I(1)$
-

---

que sea una serie de orden  $I(0)$ , es decir, estacionaria, y en este caso las series serán cointegradas.

Muchas series de tiempo, y en especial las series económicas, no son estacionarias, porque no oscilan alrededor de una media constante o la varianza no es constante en el tiempo, a este tipo de procesos se les considera procesos integrados y el orden de integración  $I(d)$  de una serie de tiempo se refiere al número  $d$  de diferenciaciones que hay que realizar sobre la serie no estacionaria para convertirla en un ARMA estacionario (Montenegro, 2011).

Una serie integrada de orden cero  $I(0)$  es una serie que no hay que diferenciarla por que ya es estacionaria y tiene una representación ARMA. Una serie no estacionaria será integrada de orden uno, es decir  $I(1)$ , si al diferenciarla una vez se vuelve estacionaria. Si la integración se logra después de diferenciarla dos veces la serie es  $I(2)$ . De la experiencia se encuentra que la mayoría de las series económicas son hasta  $I(2)$  (Montenegro, 2011).

2. Para analizar la estacionariedad de las series temporales, indicador de cartera vencida total (ICVT), pérdida neta de crédito (PNC) y costo de crédito (CoCre), se desarrollaron las siguientes pruebas informales y estadísticas:
    - a) Se efectuó una representación gráfica y un análisis descriptivo de las series temporales, lo que permitió obtener información inicial sobre la naturaleza de las series, por ejemplo, indicios sobre posible tendencia, varianza dependiente del tiempo y ciclo estacional.
    - b) Se hizo el correlograma de cada una las series de tiempo.
    - c) Se realizó la transformación de las series en primeras diferencias para filtrar la tendencia y se efectuó su correspondiente representación gráfica. Esta transformación permite convertir las series de tiempo en series estacionarias.
    - d) Se desarrolló la prueba  $Q$  de Ljung-Box para contrastar la hipótesis nula de *ruido blanco* de los residuos.
    - e) Se efectuaron las pruebas de raíces unitarias de Dickey-Fuller (DF) y Dickey-Fuller Aumentada (ADF) que permiten determinar formalmente el número de raíces unitarias de la serie de tiempo.
-

### 3.5.1. Autocorrelación

La autocorrelación se produce cuando las perturbaciones del modelo presentan correlaciones entre ellas. La autocorrelación supone que la matriz de varianzas y covarianzas de las perturbaciones presenta valores distintos de cero en los elementos que están fuera de la diagonal principal. Esto significa que se viola la hipótesis de que  $E(\epsilon_t \epsilon_s) = 0$  para todo  $t \neq s$ , y por tanto, existen relaciones lineales entre distintas perturbaciones (Pérez y Fernández, 2009).

La consecuencia fundamental es que, al aplicar mínimos cuadrados ordinarios (MCO), el estimador de  $\beta$  es lineal e insesgado, pero el de  $\sigma^2$  es sesgado, por lo que el estimador de la matriz de varianzas y covarianzas de  $\beta$  utilizado usualmente también es sesgado e incorrecto. Por este motivo, el estimador deja de ser eficiente y los contrastes de hipótesis quedan invalidados. En estos casos, es posible encontrar otro estimador, denominado estimador por mínimos cuadrados generalizados, que recoja la información sobre las correlaciones entre las perturbaciones y que sea más eficiente (Pérez y Fernández, 2009).

En la forma más simple, la presencia de autocorrelación en las perturbaciones considera la posibilidad que exista una dependencia temporal entre valores de la perturbación, de manera que, si la influencia sobre el instante  $t$  es sólo el instante anterior  $t - 1$ , la perturbación seguiría la siguiente expresión:

$$\epsilon_t = \rho \epsilon_{t-1} + \mu_t \quad (3.1)$$

donde  $\mu_t$  es una variable aleatoria ruido blanco. A este tipo de dependencia se le denomina proceso autorregresivo de orden 1, o AR(1). En general, si la dependencia temporal se establece hasta  $p$  rezagos, el proceso será un proceso autorregresivo de orden  $p$ , o AR(p):

$$\epsilon_t = \rho_1 \epsilon_{t-1} + \rho_2 \epsilon_{t-2} + \dots + \rho_p \epsilon_{t-p} + \mu_t \quad (3.2)$$

Otra posibilidad consiste en considerar que la autocorrelación viene generada por una combinación de variables aleatorias: ruido blanco, contemporáneas y rezagadas, es decir, un proceso de medias móviles. El orden de este proceso vendrá dado por el rezago  $q$  hasta el que depende de la perturbación. El

proceso más general será:

$$\epsilon_t = \mu_t + \theta_1\mu_{t-1} + \theta_2\mu_{t-2} + \dots\theta_q\mu_{t-q} \quad (3.3)$$

La última posibilidad sería una combinación de ambos procesos, construyendo un proceso mixto ARMA(p,q), es decir:

$$\epsilon_t = \rho_1\epsilon_{t-1} + \rho_2\epsilon_{t-2} + \dots + \rho_p\epsilon_{t-p} + \mu_t + \theta_1\mu_{t-1} + \dots\theta_q\mu_{t-q} \quad (3.4)$$

### 3.5.2. Función de autocorrelación y autocorrelación parcial

Uno de los procedimientos utilizados para detectar correlaciones entre las perturbaciones es el correlograma de las funciones de autocorrelación simple (ACF) y la función de autocorrelación parcial (PACF). El correlograma es un gráfico de los coeficientes de autocorrelación en función del orden de rezago. Su interpretación puede permitir, si la serie es estacionaria, identificar el tipo de proceso de autocorrelación, ya que cada proceso tiene una representación distinta y única (Pérez y Fernández, 2009).

- La función de autocorrelación (ACF) mide la correlación entre dos variables separadas por  $k$  periodos y esta definida como (Cryer y Chan, 2008)

$$\begin{aligned} \rho(j) &= \text{corr}(x_j, x_{j-k}) \\ &= \frac{\text{cov}(x_j, x_{j-k})}{\sqrt{\text{Var}(x_j)}\sqrt{\text{Var}(x_{j-k})}} \end{aligned} \quad (3.5)$$

y presenta las siguientes propiedades:

$$\begin{aligned} \rho_0 &= 1 \\ -1 &< \rho_j < 1 \\ \rho_j &= \rho_{-j} \end{aligned}$$

- La función de autocorrelación parcial (PACF) mide la correlación entre dos variables separadas por  $k$  periodos cuando no se considera la dependencia creada por los retardos intermedios existentes entre ambas y se define por (Cryer y Chan, 2008)

$$\begin{aligned}\pi_j &= \text{corr}(x_j, x_{j-k} / x_{j-1}x_{j-2}\dots x_{j-k-1}) \\ &= \frac{\text{cov}(x_j - \hat{x}_j, x_{j-k} - \hat{x}_{j-k})}{\sqrt{\text{Var}(x_j - \hat{x}_j)}\sqrt{\text{Var}(x_{j-k} - \hat{x}_{j-k})}}\end{aligned}\quad (3.6)$$

La representación gráfica de la función de autocorrelación contra los rezagos, o correlograma, es una prueba simple de estacionariedad de las series temporales que proporciona información sobre la presencia de tendencia o ciclo estacional en la serie. Si el correlograma empieza en un valor muy alto y decae muy lentamente hacia cero las series no son estacionarias. Si el correlograma decae muy rápidamente después del primer rezago la series son estacionarias (Pérez y Fernández, 2009).

### 3.5.3. Prueba de Durbin-Watson

Si el término de error del modelo econométrico tiene autocorrelación de primer orden:  $\epsilon_t = \rho\epsilon_{t-1} + \mu_t$ , donde  $\mu_t$  no tiene autocorrelación, entonces el estadístico  $d$  de Durbin-Watson permite contrastar las siguientes hipótesis nula  $H_o$  y alternativa  $H_a$ , respectivamente:

- $H_o : \rho = 0$ ,: no existe autocorrelación de primer orden.
- $H_a : 0 < |\rho| < 1$  existe autocorrelación  $AR(1)$ .

El estadístico  $d$  se define como la razón entre la suma de las diferencias al cuadrado de residuos sucesivos sobre la suma de cuadrados de los residuos, es decir:

$$d = \frac{\sum_{t=2}^n (\hat{\epsilon}_t - \hat{\epsilon}_{t-1})^2}{\sum_{t=1}^n \hat{\epsilon}_t^2}\quad (3.7)$$

Este estadístico se basa en los siguientes supuestos (Gujarati, 2010):

- El modelo de regresión incluye el término del intercepto.

- Las variables explicativas,  $X$ , son no estocásticas, es decir, son fijas en muestreo repetido.
- Los residuos  $\epsilon_t$  se generan mediante el esquema autorregresivo de primer orden. Por tanto, no se pueden utilizar para detectar esquemas autorregresivos de orden superior.
- El término de error  $\epsilon_t$  está normalmente distribuido.
- El modelo de regresión no incluye valores rezagados de la variable dependiente como variable explicativa.
- No hay observaciones faltantes en los datos.

Una de las ventajas del estadístico  $d$  de Durbin-Watson es que se basa en los residuos estimados que se calculan usualmente en los análisis de regresión y es frecuente su inclusión junto con otros estadísticos como  $R^2$ ,  $R^2$  ajustado  $t$  y  $F$ . Sin embargo, como los residuos  $\hat{\epsilon}_t$  dependen de los valores de la variable explicativa  $X$  en una muestra dada, la distribución de probabilidad del estadístico  $d$ , dada en 3.7, también tendrá una dependencia compleja con los  $X$  y, como demostraron Durbin y Watson, será difícil determinar dicha distribución de probabilidad. Por tanto, a diferencia de las pruebas  $t$ ,  $F$  o  $\chi^2$ , no existe un valor crítico único que conduzca al rechazo o no rechazo de la hipótesis nula de que no hay autocorrelación de primer orden en los residuos (Pérez y Fernández, 2009).

A pesar de esto, Durbin y Watson lograron obtener cotas inferiores y superiores para los niveles de significación del estadístico  $d$ , sobre el conjunto de todas sus posibles distribuciones de probabilidad. Dichos valores críticos se presentan en tablas estadísticas (Dickey y Fuller, 1981). Si el valor  $d$  calculado de 3.7 cae por fuera de estos valores críticos, puede tomarse una decisión respecto de la presencia de autocorrelación positiva o negativa de primer orden (Pérez y Fernández, 2009).

El estadístico  $d$  puede aproximarse por:

$$d = 2(1 - \hat{\rho}) \quad \text{donde} \quad \hat{\rho} = \frac{\sum_{t=2}^n \hat{\epsilon}_t \hat{\epsilon}_{t-1}}{\sum_{t=1}^n \hat{\epsilon}_t^2} \quad (3.8)$$

Como los valores permitidos del parámetro  $\hat{\rho}$  están entre -1 y 1, entonces el estadístico  $d$  oscilará entre 0 y 4. Si  $\hat{\rho} = -1$ , entonces  $d \approx 4$ , y existe autocorrelación negativa de primer orden en los residuos. Si  $\hat{\rho} = 1$ , entonces

$d \approx 0$  y hay autocorrelación negativa de primer orden. Cuando el término de error es independiente a lo largo del tiempo,  $\hat{\rho} = 0$ , y el valor del estadístico  $d$  estará próximo a 2, y se puede suponer que no hay autocorrelación positiva o negativa de primer orden (Pérez y Fernández, 2009),(Gujarati, 2010).

### 3.5.4. Prueba de Ljung-Box

El estadístico  $Q$  de Ljung-Box sirve para contrastar esquemas generales de autocorrelación de la perturbación permitiendo probar en forma conjunta que todos los coeficientes de autocorrelación  $\hat{\rho}_k$  son simultáneamente iguales a cero, esto es, que son independientes.

El estadístico  $Q$  se define como:

$$Q = n(n+2) \sum_{k=1}^m \left( \frac{\hat{\rho}_k^2}{n-k} \right) \sim \chi_{r-p-q}^2 \quad (3.9)$$

y se distribuye asintóticamente como una chi-cuadrado con  $r - p - q$  grados de libertad, siendo  $n$  el número de observaciones utilizadas en la estimación de la ecuación con la que se está trabajando,  $p$  y  $q$  el orden del proceso de autocorrelación propuesto,  $\hat{\rho}_k$  el coeficiente de autocorrelación muestral de orden  $k$  de los residuos mínimos cuadráticos y  $m$  es la longitud del rezago (Pérez y Fernández, 2009).

Para esta prueba la hipótesis nula  $H_o$  y alternativa  $H_a$  son, respectivamente:

- $H_o : \rho_k = 0, k = 1, 2, \dots, m$ : las autocorrelaciones son independientes.
- $H_a : \rho_k \neq 0, k = 1, 2, \dots, m$ : las autocorrelaciones no son independientes.

En esta prueba la regla de decisión establece que si el  $Q$  calculado excede el valor crítico de la tabla  $\chi^2$  al nivel de significación seleccionado, no se acepta la hipótesis nula de que todos los coeficientes de autocorrelación son iguales a cero, por lo menos algunos de ellos deben ser diferentes de cero (Gujarati, 2010).

La elección del valor de  $r$  en el estadístico  $Q$  es uno de los inconvenientes de este contraste, ya que si es demasiado pequeño puede fallar en la detección

---

de procesos donde las correlaciones de retardos altos son importantes, y si es demasiado grande puede tener poca potencia al incluir correlaciones de retardos que en realidad son insignificantes (Pérez y Fernández, 2009).

### 3.5.5. Pruebas de raíz unitaria

El procedimiento para estudiar la cointegración consiste, primero, en determinar si las series individualmente vistas son  $I(1)$ ; segundo, en examinar si hay cointegración, y tercero, en caso afirmativo, incorporar la cointegración en las ecuaciones de regresión. Las pruebas estadísticas para determinar el grado de integración de una serie, en particular si son  $I(1)$ , se conocen como pruebas de raíz unitaria (Montenegro, 2011).

#### Prueba de Dickey-Fuller

Cuando una serie temporal no es integrada de orden cero  $I(0)$ , es decir, no es estacionaria, presenta al menos una raíz unitaria. Cuando esto sucede es posible convertir la serie original en una serie estacionaria a través de una transformación adecuada como una diferenciación. Las pruebas de Dickey-Fuller (DF) y Dickey-Fuller Aumentada (ADF) permiten determinar el número de raíces unitarias que la serie original presenta y que corresponde al número de diferencias que habrá que tomar en la serie para hacerla estacionaria (López, Rodríguez, y Pulido, 2004).

Dickey y Fuller proponen un modelo autorregresivo AR(1) como:

$$Y_t = \rho Y_{t-1} + \epsilon_t \quad (3.10)$$

donde  $\rho$  es el parámetro a estimar y  $\epsilon_t$  es un término de error de ruido blanco. Como  $\rho$  es un coeficiente de autocorrelación, por tanto, toma valores entre:  $-1 < \rho < 1$ . Si  $\rho = 1$  la ecuación 3.10 se convierte en un modelo de caminata aleatoria sin intercepto y la serie  $Y_t$  será no estacionaria. Sin embargo, no es posible estimar la ecuación 3.10 por mínimos cuadrados ordinarios (MCO) y probar la hipótesis por medio de la prueba  $t$  usual, ya que tiene un sesgo importante en el caso de una raíz unitaria (Gujarati, 2010). Por tanto, es necesario transformar la serie  $Y_t$  mediante la aplicación de diferencias para

---

obtener:

$$\Delta Y_t = \delta Y_{t-1} + \epsilon_t \quad (3.11)$$

donde  $\delta = \rho - 1$  y  $\Delta$  es el operador de primeras diferencias. De esta manera, la prueba de Dickey y Fuller plantea probar la hipótesis nula de que  $\delta = 0$ , entonces  $\rho = 1$ ; es decir, se tiene una raíz unitaria, lo que significa que la serie de tiempo es estacionaria, frente a la hipótesis alterna de que  $\delta < 0$ .

Al analizar la naturaleza del proceso de raíz unitaria Dickey y Fuller sugieren las siguientes ecuaciones para determinar la presencia o no de raíces unitarias

$$\Delta Y_t = \delta Y_{t-1} + \epsilon_t \quad (3.12a)$$

$$\Delta Y_t = \beta_1 + \delta Y_{t-1} + \epsilon_t \quad (3.12b)$$

$$\Delta Y_t = \beta_1 + \beta_2 t + \delta Y_{t-1} + \epsilon_t \quad (3.12c)$$

La primera es un modelo de caminata aleatoria, la segunda incorpora un intercepto y la tercera incluye intercepto y un término de tendencia (Gujarati, 2010). En cada caso, las hipótesis nula y alternativa son, respectivamente:

- $H_0 : \delta = 0$ , existe una raíz unitaria, la serie de tiempo es no estacionaria o tiene tendencia estocástica.
- $H_a : \delta \neq 0$ , la serie de tiempo es estacionaria posiblemente alrededor de una tendencia determinista.

Aunque la prueba calcula un estadístico  $t$  sobre el valor estimado de  $\delta$ , éste estadístico no tiene la distribución convencional t-Student. Por ello, Dickey y Fuller y posteriormente MacKinnon construyeron las tablas que se usan para determinar los valores críticos del estadístico  $\tau$  de la prueba DF. La regla de decisión compara el estadístico  $\tau$  con los valores críticos de MacKinnon, si el estadístico  $\tau$  no excede los valores críticos se rechaza la hipótesis nula de existencia de una raíz unitaria, y por tanto, la serie es estacionaria (Gujarati, 2010).

### Prueba de Dickey-Fuller Aumentada

El problema de la prueba de Dickey-Fuller simple es que asume que el proceso estocástico es AR(1) y que el término de error no está correlacionado.

Por tanto, cuando el proceso siga otro esquema, la estimación de la regresión auxiliar de la prueba producirá un incumplimiento de la condición de *ruido blanco* de los residuos, debido a una mala especificación. Por esto, aunque el primer paso para detectar raíces unitarias es una prueba DF simple a continuación se debe contrastar estadísticamente las características de *ruido blanco* de los residuos (López y cols., 2004).

Para estos casos se dispone del test ampliado de Dickey-Fuller (ADF) que permite considerar otros modelos de autocorrelación y desarrollar una prueba cuando el término de error  $\epsilon_t$  sí está correlacionado. La estrategia para la prueba Dickey-Fuller Aumentada (ADF) consiste en determinar cuál de los siguientes tres casos determina una mejor aproximación a la serie original,  $Y_t$ .

- Caso 1: Modelo  $AR(p)$  con media cero

$$\Delta Y_t = \sum_{j=1}^{p-1} \delta_{j+1} \Delta Y_{t-j} + \epsilon_t \quad (3.13)$$

- Caso 2: Modelo  $AR(p)$  con media diferente de cero

$$\Delta Y_t = \beta_1 + \delta Y_{t-1} + \sum_{j=2}^p \delta_j \Delta Y_{t-j+1} + \epsilon_t \quad (3.14)$$

- Caso 3: Modelo  $AR(p)$  con tendencia lineal

$$\Delta Y_t = \beta_1 + \beta_2 t + \delta Y_{t-1} + \sum_{j=2}^p \delta_j \Delta Y_{t-j+1} + \epsilon_t \quad (3.15)$$

donde el número de rezagos de  $\Delta Y_t$  puede elegirse efectuando la regresión por mínimos cuadrados ordinarios (MCO) inicialmente con muchos rezagos y luego disminuyéndolos hasta que el último sea significativo y asegurándose que  $\epsilon_t$  sea ruido blanco, según la función de autocorrelación ACF, la prueba  $Q$  de Ljung-Box, la prueba Durbin-Watson, etc. En la prueba ADF se sigue probando la misma hipótesis nula que en la prueba DF simple, y además esta prueba sigue la misma distribución que el estadístico DF por lo que sirven los mismos valores críticos (Gujarati, 2010).

### 3.5.6. Prueba de hipótesis para subconjuntos

Se considera el siguiente modelo de regresión lineal múltiple o modelo no reducido:

$$Y_i = \beta_0 + \beta_1 X_1 + \dots + \beta_l X_l + \beta_{l+1} X_{l+1} + \dots + \beta_k X_k + \epsilon_i \quad (3.16)$$

frente al modelo reducido:

$$Y_i = \beta_0 + \beta_{l+1} X_{l+1} + \dots + \beta_k X_k + \epsilon_i \quad (3.17)$$

y se establece el siguiente contraste de hipótesis:

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

$$H_a : \text{al menos un } \beta_j = 0, \quad j = 1, 2, \dots, l$$

No rechazar la hipótesis nula significa que el modelo reducido es significativo mientras que rechazarla significa que al menos una de las primeras  $l$  variables es significativa (Pérez y Fernández, 2009). Con base en esta prueba se estableció el árbol de decisión, mostrado en la figura 3.1, para poder determinar cual de los tres contrastes de la prueba de Dickey-Fuller aumentada debía interpretarse.

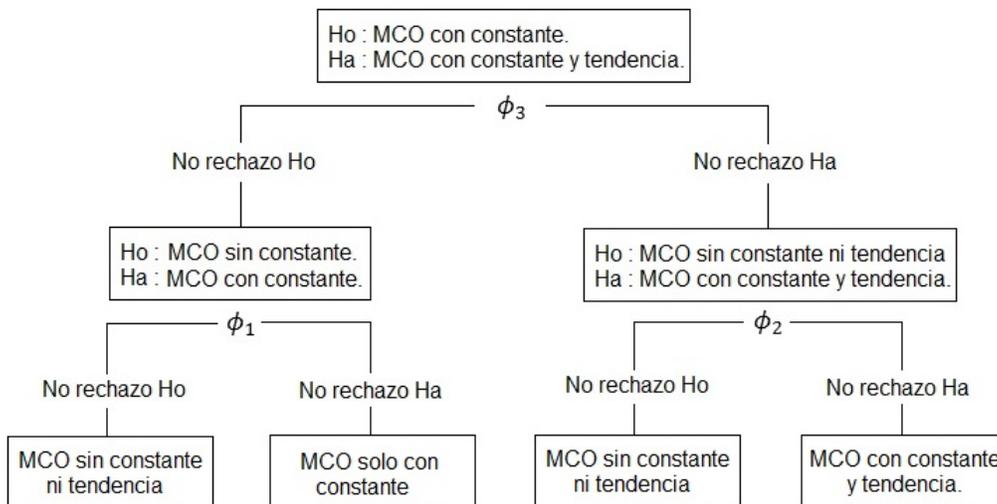


Figura 3.1: Árbol de decisión para elegir el contraste en la prueba ADF.

En este caso el estadístico de prueba se define como:

$$\phi_i = \frac{(SSE_r - SSE_{nr})/l}{SSE_{nr}/(N - k)} \quad (3.18)$$

donde  $SSE_r$  es la suma de cuadrados de los residuos del modelo reducido,  $SSE_{nr}$  es la suma de cuadrados de los residuos del modelo no reducido,  $N$  es el número total de observaciones,  $l$  es el número de restricciones y  $k$  es el número de coeficientes del modelo no reducido, que en este caso, consideran el número de rezagos, el coeficiente no diferenciado y la presencia o no de intercepto y tendencia. En este caso se rechaza la hipótesis nula si el estadístico  $\phi_i$  es mayor que el valor crítico reportado en las tablas de Dickey y Fuller para la pruebas de raíz unitaria basadas en el estadístico  $F$  (Dickey y Fuller, 1981).

3. Se establecieron las siguientes relaciones entre las series temporales pérdida neta de crédito (PNC) e indicador de cartera vencida total (ICVT) y entre costo de crédito (CoCre) y pérdida neta de crédito (PNC):

$$PNC_t = \beta_0 + \beta_1 ICVT_t + z_{1,t} \quad (3.19a)$$

$$CoCre_t = \alpha_0 + \alpha_1 PNC_t + z_{2,t} \quad (3.19b)$$

donde  $\beta$  y  $\alpha$  son constantes y  $z_t$  es un término de error. Se sabe que realizar la regresión de una serie no estacionaria sobre otra no estacionaria puede conducir a una regresión espuria. Sin embargo, si dos series temporales son individualmente  $I(1)$ , es decir, tienen tendencias estocásticas, y su combinación lineal es  $I(0)$ , entonces la regresión de una serie sobre otra no necesariamente será espuria y en este caso las dos variables estarán cointegradas.

Regresiones como las indicadas en las ecuaciones 3.19a y 3.19b son plausibles sólo si existe algún mecanismo que mantengan las dos variables cerca una de la otra a través del tiempo, esto es, que no permita que se separen demasiado aunque ambas individualmente sean no estacionarias. Por tanto, debe existir algún mecanismo de equilibrio a largo plazo entre las dos. Aunque en el corto plazo las series pueden diferir un poco, algo las obliga a acercarse en caso de que se estén alejando demasiado, de modo que su diferencia sea estacionaria (Montenegro, 2011).

Cuando existe este mecanismo de equilibrio de largo plazo, la combinación lineal de dos variables no estacionarias puede producir una estacionaria  $I(0)$ . Cuando esto sucede se llama *cointegración* y se dice que las dos variables no estacionarias están *cointegradas*. En general, para que haya cointegración entre dos variables, éstas deben tener el mismo orden de integración; de otra manera, una cambia más rápido que la otra y será imposible que su combinación lineal sea  $I(0)$ . Por tanto, dos variables  $x_t$  y  $y_t$  están cointegradas si ambas son  $I(1)$  y se tiene que:

$$y_t - \beta x_t = z_t \quad (3.20)$$

donde  $z_t$  es  $I(0)$  y representa el residuo de la regresión de las dos variables. Es importante resaltar que si bien los residuos de la regresión entre las dos variables deben ser  $I(0)$ , para que exista equilibrio, esto no implica que deban ser *ruido blanco* (Montenegro, 2011).

Si se logra verificar que los residuos de regresiones como las indicadas en las ecuaciones 3.19a y 3.19b son  $I(0)$ , la metodología tradicional de regresión por mínimos cuadrados ordinarios (MCO) es aplicable a las series de tiempo. En tal caso éstas regresiones serán *regresiones cointegrantes* y los parámetros  $\beta_1$  y  $\alpha_1$  serán los *parámetros cointegrantes* (Montenegro, 2011).

### 3.5.7. Prueba de cointegración de Engle-Granger

Para verificar si los residuos de las regresiones establecidas en las ecuaciones 3.19a y 3.19b son estacionarios se efectuó la prueba de cointegración de Engle-Granger. En el contexto de la teoría de cointegración las pruebas de raíz unitaria de Dickey-Fuller (DF) y Dickey-Fuller Aumentada (ADF) se conocen como las pruebas de Engle-Ganger (EG) y Engle-Ganger Aumentada (AEG). Por tanto, solo se necesita estimar las regresiones 3.19a y 3.19b, obtener los residuos y usar las pruebas DF y ADF (Gujarati, 2010). Sin embargo, los valores críticos de significación DF y ADF no son apropiados y deben utilizarse los valores críticos para la prueba de raíz unitaria y cointegración calculados a partir de la ecuación:

$$CV = b + b_1\left(\frac{1}{N}\right) + b_2\left(\frac{1}{N^2}\right) \quad (3.21)$$

donde los coeficientes  $b$ ,  $b_1$  y  $b_2$  han sido estimados en las tablas para los valores críticos de cointegración reportadas por MacKinnon (Mackinnon, 2010).

- 
4. Una vez se estableció que los residuos de las regresiones propuestas en las ecuaciones 3.19a y 3.19b son estacionarios y, por tanto, existe una relación de equilibrio a largo plazo entre las series temporales pérdida neta de crédito (PNC) e indicador de cartera vencida total (ICVT) y costo de crédito (Co-Cre) y pérdida neta de crédito, se procedió a aplicar el esquema tradicional de regresión por mínimos cuadrados ordinarios (MCO). Como es usual en esta metodología se realizó el siguiente procedimiento:
- Se obtuvieron los estadísticos  $t$  y  $F$  y sus correspondientes  $p$  valores.
  - Se interpretó el coeficiente de determinación  $R^2$ .
  - Se interpretaron los coeficientes  $\beta$  de la regresión y se estableció si eran estadísticamente significativos.
  - Se realizaron las pruebas informales de normalidad de los residuos de las regresiones a través del gráfico de probabilidad, el diagrama de caja y el histograma de los residuos y se efectuó la prueba de normalidad de Jarque-Bera.
  - Para detectar posibles problemas de heteroscedasticidad en el modelo se efectuó el estudio gráfico sobre el comportamiento del cuadrado de los residuos contra la variable dependiente estimada y se realizó el contraste de White.
5. Se realizó la estimación por mínimos cuadrados ordinarios (MCO) de un modelo de corrección de error (MCE) para vincular el comportamiento de corto plazo con la dinámica de largo plazo de las relaciones establecidas entre la pérdida neta de crédito (PNC) y el índice de cartera vencida total (ICVT) y entre el costo de crédito (CoCre) y la pérdida neta de crédito (PNC). Posteriormente, se realizó el diagnóstico del modelo mediante las pruebas de autocorrelación, normalidad y heteroscedasticidad de los residuos de la regresión por MCO.

### 3.5.8. Modelo de corrección de error

Si dos series  $y_t$  y  $x_t$  están cointegradas entonces la relación entre ellas se expresa como un modelo de corrección de error, es decir, debe existir algún mecanismo de corrección de error que preserve el equilibrio. Las series  $y_t$  y  $x_t$  pueden mantener una relación de largo plazo estable y presentar desequilibrios en el corto plazo, el modelo de corrección de errores permite determinar

---

la magnitud del ajuste de la variable hacia su tendencia de largo plazo. Este resultado se conoce como *teorema de representación de Granger* (Granger y Engle, 1987). El modelo de corrección de error consiste en una regresión que acopla la información de largo plazo contenida en la cointegración o relación de equilibrio,  $z_t = y_t - \beta x_t$ , con la información de corto plazo contenida en las variaciones de  $y_t$  y  $x_t$  (Montenegro, 2011).

Si las variables están cointegradas se pueden utilizar los residuos para corregir los errores y estimar también los efectos a corto plazo. El modelo a estimar se denomina de corrección de error y esta dado por:

$$\Delta y_t = \alpha \Delta x_t + \gamma z_{t-1} + \epsilon_t \quad (3.22)$$

donde  $\epsilon_t$  es un término de error de ruido blanco,  $z_{t-1} = y_{t-1} - \beta x_{t-1}$  es el valor rezagado del término de error o de cointegración,  $\beta$  es el efecto a largo plazo de  $x_t$  sobre  $y_t$  y  $\gamma$  es el efecto a corto plazo de  $x_t$  sobre  $y_t$  (Pfaff, 2008).

La expresión de un modelo de corrección de error es similar a un vector auto-regresivo (VAR) donde hay una ecuación para cada una de las variables. La variable dependiente en cada ecuación es la primera diferencia de la variable correspondiente, ya sea  $y_t$  o  $x_t$ , expresada en función de sus propios rezagos en diferencias, los rezagos de la otra variable en diferencias y el término de error o de cointegración rezagado un periodo. Para el caso bivariado se tiene el sistema de ecuaciones:

$$\Delta y_t = \alpha_0 + \gamma_1 z_{t-1} + \sum_{i=1}^k \alpha_{1,i} \Delta x_{t-i} + \sum_{i=1}^l \alpha_{2,i} \Delta y_{t-i} + \epsilon_{1t} \quad (3.23a)$$

$$\Delta x_t = \phi_0 + \gamma_1 z_{t-1} + \sum_{i=1}^k \phi_{1,i} \Delta x_{t-i} + \sum_{i=1}^l \phi_{2,i} \Delta y_{t-i} + \epsilon_{1t} \quad (3.23b)$$

Todas las variables explicativas utilizadas en éstas regresiones, incluido el término de error o de cointegración son  $I(0)$ . Se pueden agregar más rezagos de ambas variables para incrementar el poder explicativo o para asegurarse de que  $\epsilon_t$  sea ruido blanco (Montenegro, 2011). El coeficiente  $\gamma$  se conoce como coeficiente de ajuste y su magnitud indica el grado de ajuste con respecto a desviaciones del equilibrio a largo plazo. Mientras más cerca este de uno, más rápido será el ajuste hacia el equilibrio y siempre debe ser de signo negativo,

---

de otra manera el sistema puede diverger de su equilibrio a largo plazo (Pfaff, 2008).

### 3.5.9. Prueba de Jarque-Bera

La prueba de normalidad de Jarque-Bera es una prueba asintótica basada en los residuos de MCO. Esta prueba calcula primero los coeficientes de asimetría y curtosis de los residuos de MCO, con el siguiente estadístico de prueba:

$$JB = n \left[ \frac{S^2}{6} + \frac{(K - 3)^2}{24} \right] \quad (3.24)$$

donde  $n$  es el tamaño de la muestra,  $S$  es el coeficiente de asimetría y  $K$  es el coeficiente de curtosis. Para este contraste la hipótesis nula  $H_o$  y alternativa  $H_a$  son, respectivamente:

- $H_o$ : los residuos están normalmente distribuidos.
- $H_a$ : los residuos no están normalmente distribuidos.

Para una variable distribuida normalmente, la asimetría debe ser cero y la curtosis debe ser tres. Por tanto, la prueba de normalidad de Jarque-Bera constituye una prueba de hipótesis conjunta de que  $S$  y  $K$  son cero y tres respectivamente. En este caso se espera que el valor del estadístico  $JB$  sea igual a cero (Gujarati, 2010).

De acuerdo con la hipótesis nula, la cual afirma que los residuos están normalmente distribuidos, Jarque y Bera mostraron que en muestras grandes el estadístico  $JB$  dado en 3.18 sigue una distribución chi-cuadrada, con 2 gl. Si el p valor calculado del estadístico  $JB$  es lo bastante bajo, lo cual sucederá si el valor del estadístico difiere en gran medida de cero, se puede rechazar la hipótesis de que los residuos están normalmente distribuidos. Pero si el p valor es razonablemente alto, lo que sucede cuando el valor del estadístico está cerca de cero, no rechazamos la suposición de normalidad (Gujarati, 2010).

---

### 3.5.10. Prueba de Shapiro-Wilk

La prueba de Shapiro-Wilk se utiliza para determinar si una muestra aleatoria tiene una distribución normal. Se considera uno de los test más potentes para el contraste de normalidad, sobre todo para muestras pequeñas ( $n < 50$ ). La prueba se basa en la correlación de una muestra de estadísticos de orden con aquellos de una distribución normal. El uso de estadísticos de orden implica que los datos de la muestra deben estar ordenados de manera creciente. La muestra ordenada se representa por la transpuesta de un vector:  $x' = (x_1, \dots, x_N)$ . El estadístico  $W$  de la prueba se define como:

$$W = \frac{(\sum_{i=1}^N a_i x_i)^2}{\sum_{i=1}^N (x_i - \bar{x})^2} \quad (3.25)$$

donde  $\bar{x} = \frac{1}{N} \sum_{i=1}^N x_i$  es la media muestral y las variables  $a_i$  se calculan como:

$$a = (a_1, \dots, a_n)' = \frac{mV^{-1}}{(m^T V^{-1} V^{-1} m)^{1/2}} \quad (3.26)$$

donde  $m = (m_1, \dots, m_n)$ , con  $m_1, \dots, m_n$  como los valores medios del estadístico ordenado de variables aleatorias independientes e idénticamente distribuidas muestreadas de distribuciones normales y  $V$  es la matriz de covarianzas de ese estadístico de orden. El valor de  $W$  puede oscilar entre 0 y 1. Para este contraste la hipótesis nula  $H_o$  y alternativa  $H_a$  son, respectivamente:

- $H_o$ : la muestra aleatoria tiene una distribución normal.
- $H_a$ : la muestra no aleatoria tiene una distribución normal.

La hipótesis nula se rechazará si  $W$  es demasiado pequeño o si el p-valor es menor al nivel de significación  $\alpha$ . Si el p-valor es mayor a  $\alpha$ , no se rechaza la hipótesis y se concluye que los datos siguen una distribución normal (Shapiro y Wilk, 1965).

### 3.5.11. Heteroscedasticidad

Uno de los problemas estadísticos más frecuentes en los modelos econométricos y que genera resultados no deseables en la estimación paramétrica es la

heteroscedasticidad. No hay reglas generales que permitan detectar la presencia de heteroscedasticidad en un modelo. Las dificultades que surgen para detectar la heteroscedasticidad están asociadas con el hecho de que las perturbaciones estocásticas no son producto de la observación, de manera que la detección se basa principalmente en los errores que resultan por la estimación realizada al aplicar los mínimos cuadrados ordinarios (MCO). En general, se asume que el modelo a considerar puede tener problemas de heteroscedasticidad y luego se procede a estimar el modelo por MCO. Una vez obtenidos los residuos se utilizan métodos informales y formales, basados fundamentalmente en el comportamiento de los residuos de MCO (Pérez y Fernández, 2009).

Un supuesto básico del modelo clásico de regresión lineal es que la varianza del término de perturbación estocástica  $\epsilon_i$  es la misma en todas las observaciones. Este supuesto se llama *supuesto de homocedasticidad* y se expresa como  $V(\epsilon_i) = \sigma^2$ . Este supuesto establece entonces que la varianza de cada  $\epsilon_i$  es algún número positivo constante e igual a  $\sigma^2$ , cuando éste supuesto se viola se tiene un problema de heteroscedasticidad. Simbólicamente  $V(\epsilon_i) = \sigma_i^2$ , donde el subíndice de  $\sigma^2$  indica que las varianzas condicionales de  $\epsilon_i$  ya no son constantes. Hay, entonces, problemas de heteroscedasticidad cuando el error muestra diferentes dispersiones para cada observación o, de igual manera, cuando la probabilidad de que el error tome un valor determinado es diferente para cada observación (Pérez y Fernández, 2009).

### 3.5.12. Prueba de White

Uno de los contrastes estadísticos más usuales para detectar heteroscedasticidad es la prueba general de heteroscedasticidad de White que no depende estrictamente de la normalidad de los términos de error del modelo de regresión:

$$Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \dots + \beta_k X_{ik} + \epsilon_i \quad (3.27)$$

donde White propone que:

$$\epsilon_i = \alpha_0 + \alpha_1 Z_{i1} + \alpha_2 Z_{i2} + \dots + \alpha_m Z_{im} + \mu_i \quad (3.28)$$

White sugiere que si hay presencia de heteroscedasticidad relacionada con alguna variable  $X$ , entonces  $Z$  se sustituye por los valores originales de las

$X$ , por sus cuadrados y por sus productos cruzados dos a dos, para evitar así la linealidad (Pérez y Fernández, 2009).

White propone, en su prueba, la aplicación de los siguientes pasos:

- Calcular los residuales,  $\hat{\epsilon}_i$  mediante la aplicación de MCO del modelo original

$$Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \dots + \beta_k X_{ik} + \epsilon_i \quad (3.29)$$

- Efectuar la regresión auxiliar:

$$\hat{\epsilon}_i^2 = \alpha_0 + \alpha_1 X_{i1} + \alpha_2 X_{i2} + \alpha_3 X_{i1}^2 + \alpha_4 X_{i2}^2 + \alpha_5 X_{i1} X_{i2} + \dots + \mu_i \quad (3.30)$$

- Bajo la hipótesis nula de heteroscedasticidad se verifica, para muestras grandes, que  $nR^2 \sim \chi_m^2$ , donde  $R^2$  es el coeficiente de determinación ajustado de la regresión auxiliar,  $n$  es el tamaño de la muestra y  $m$  es el número de grados de libertad, igual al número de variables regresoras en la regresión auxiliar.
  - Si  $nR^2 > \chi^2$  para un nivel de significación definido, se rechaza la hipótesis nula de homoscedasticidad.
6. Con los datos históricos del indicador de cartera vencida total (ICVT) se determinó su distribución de probabilidad teórica con la prueba de bondad de ajuste de Anderson-Darling. Luego, con el modelo estadístico obtenido a partir de la aplicación de mínimos cuadrados ordinarios a las regresiones propuestas en las ecuaciones 3.19a y 3.19b, y considerando un estado de pérdidas y ganancias (PyG) para un producto de crédito de consumo de libre inversión, se realizó una simulación de Monte Carlo que permitió pronosticar el costo de cartera (CoCre) y el índice de rentabilidad sobre los activos (ROA) para una determinada calidad de cartera.

### 3.5.13. Método de Monte Carlo

La simulación de Monte Carlo es un método estadístico no determinista que permite la realización de experimentos con muestreo de números aleatorios en una computadora, para simular, mediante modelos matemáticos, el

---

comportamiento de sistemas reales complejos no dinámicos y proporcionar soluciones aproximadas a éstos (Raychaudhuri, 2008). El fundamento teórico del método de Monte Carlo se encuentra en la ley de los grandes números y en el teorema del límite central de la teoría de probabilidades (Canavos, 1988).

### Metodología

Los pasos que se desarrollan en una simulación de Monte Carlo son los siguientes (Raychaudhuri, 2008):

- *Generación del modelo estático.* Toda simulación de Monte Carlo comienza con el desarrollo de un modelo determinista del sistema, denominado *escenario básico*, similar al escenario real. En este modelo determinista se introducen los parámetros de entrada. Los parámetros de entrada del modelo dependen de varios factores externos. Debido a estos factores, los modelos realistas están sujetos a la variación aleatoria de las variables de entrada. El escenario básico no considera estas variaciones, por tanto, se eligen los valores más probables como los valores de las variables de entrada. Luego, se aplican las relaciones matemáticas, donde se usan los valores de los parámetros de entrada, y se obtienen los parámetros de salida.
  - *Identificación de la distribución de probabilidad de entrada.* Cuando se está satisfecho con el modelo determinista se adicionan las componentes de aleatoriedad al modelo. Como la aleatoriedad se origina en la naturaleza estocástica de las variables de entrada, se intenta identificar las distribuciones de probabilidad que gobierna los parámetros de entrada. Este paso necesita los datos históricos para las variables de entrada. Existen diferentes procedimientos estadísticos para identificar las distribuciones de probabilidad de los parámetros de entrada.
  - *Generación de variables aleatorias.* Después de que se han identificado las distribuciones de probabilidad de los parámetros de entrada, se genera un conjunto de números aleatorios o pseudo-aleatorios (denominados así ya que los números aleatorios generados por este método no son reales sino simulados) de estas distribuciones de probabilidad. Un conjunto de números aleatorios, consistente de un valor por cada una
-

---

de las variables de entrada, será usado en el modelo determinista para proporcionar un conjunto de valores de salida. Luego, se repite este proceso generando más conjuntos de números aleatorios, uno por cada distribución de entrada, y se recolectan diferentes conjuntos de posibles valores de salida. Esta parte es el núcleo de la simulación de Monte Carlo.

- *Análisis y toma de decisiones.* Después de que se ha recolectado la muestra de valores de salida de la simulación, se realiza el análisis estadístico de estos valores. Este paso proporciona la confianza estadística necesaria para la toma de decisiones y la acción a desarrollar.

### Identificación de la distribución de entrada

El procedimiento para identificar la distribución de entrada para el modelo de simulación de Monte Carlo se denomina ajuste de distribuciones. Cuando existen datos históricos de un parámetro de entrada particular, se usan métodos numéricos para ajustar los datos a una distribución teórica de probabilidad continua o discreta. Los métodos de ajuste proporcionan una forma de determinar la distribución de probabilidad más adecuada para un conjunto dado de datos. Cada distribución de probabilidad puede identificarse por su conjunto de parámetros, de modo que, el ajuste de distribuciones es esencialmente igual a encontrar los parámetros de la distribución que generará los datos. Desde ésta perspectiva, los métodos de ajuste no son más que problemas de optimización no lineal, donde las variables son parámetros de las distribuciones (Raychaudhuri, 2008). Los procedimientos comúnmente usados para el ajuste de datos a distribuciones de probabilidad son los siguientes:

- a) Métodos de ajuste de distribuciones
    - Método de máxima verosimilitud
    - Método de momentos
    - Optimización no lineal
  - b) Estadísticos de bondad de ajuste
    - Prueba Chi-cuadrado
    - Prueba de Kolmogorov-Smirnov
    - Prueba Cramer-von Mises
-

- Prueba Anderson-Darling

### Generación de números aleatorios

Después de que se han identificado las distribuciones de probabilidad para los parámetros de entrada en el modelo de simulación, se generan los números aleatorios de éstas distribuciones. Los números aleatorios generados representan valores específicos de la variable. El método de transformación inversa proporciona la forma más común para generar una muestra aleatoria de una distribución. En éste método, se utiliza la función densidad de probabilidad (PDF) para distribuciones continuas o la función de masa de probabilidad (PMF) para distribuciones discretas, y se convierte un número aleatorio entre 0 y 1 en un valor aleatorio para la distribución de entrada (Raychaudhuri, 2008). Una descripción detallada del método de transformación inversa se encuentra en (Fishman, 1995), (Law y Kelton, 2000).

El método de Monte Carlo está fundamentado en la generación de una adecuada secuencia de números aleatorios. Cada número aleatorio debe ser totalmente independiente de los otros números de la secuencia. Además, dos generadores aleatorios independientes deben proporcionar estadísticamente el mismo valor promedio de salida. Comúnmente las secuencias de números se obtienen de algún algoritmo y se denominan números pseudoaleatorios. Esto hace que los generadores tengan que cumplir una serie de características (Gómez, 2010):

- a)* Se requiere una buena distribución, es decir, los números obtenidos deben estar uniformemente distribuidos en el intervalo en el que se obtienen: (0, 1). Si tomamos un subintervalo cualquiera, la fracción de números aleatorios que aparece respecto del total tiene que ser la misma para todo subintervalo de la misma amplitud.
- b)* Al ser generados mediante un algoritmo, siempre tienen un ciclo más o menos largo. En el caso de simulaciones en que se usa una gran cantidad de números aleatorios es importante que éstos no se repitan para evitar las correlaciones.
- c)* Es importante que se pueda reproducir la sucesión de números usados. Si se repite la simulación en las mismas condiciones el resultado debe ser el mismo.

---

Una discusión detallada sobre la generación de números aleatorios uniformes en el intervalo  $(0,1)$  se encuentra en (Fishman, 1995), (Law y Kelton, 2000).

### 3.5.14. Prueba de Anderson-Darling

La prueba de bondad de ajuste de Anderson-Darling (AD) es una prueba no paramétrica que permite determinar si una muestra de datos proviene de una distribución de probabilidad específica (Anderson y Darling, 1954). Esta prueba es una forma de estimación de mínima distancia y uno de los estadísticos más potentes para detectar discrepancia con respecto a la normalidad cuando se analizan distribuciones que provienen de muestras pequeñas ( $n \leq 30$ ). Estudios recientes recomiendan el uso de muestras con tamaño de por lo menos  $n = 50$  para tener un buen desempeño de la prueba AD (Flowers, Flowers, y Trejo, 2014).

La prueba de Anderson-Darling (AD) es una modificación de la prueba de Kolmogorov-Smirnov (KS) y da mayor peso a las colas de la distribución que la prueba KS. Por tanto, la prueba AD es relevante cuando las colas de la distribución teórica seleccionada son de importancia práctica. Al calcular los valores críticos para un determinado nivel de significación la prueba AD hace uso de la distribución teórica propuesta, esto tiene la ventaja de permitir una prueba más sensible y la desventaja de tener que calcular los valores críticos para cada tipo de distribución (Ang y Tang, 2007).

En su forma básica, la prueba AD asume que no existen parámetros a estimar en la distribución que se está probando, en cuyo caso la prueba y su conjunto de valores críticos siguen una distribución libre. Sin embargo, la prueba AD se utiliza con mayor frecuencia en situaciones donde se está probando una familia de distribuciones de probabilidad, en cuyo caso deben ser estimados los parámetros de esa familia y debe tenerse estos en cuenta a la hora de ajustar la prueba estadística y sus valores críticos.

El procedimiento para aplicar el método de Anderson-Darling es el siguiente (Ang y Tang, 2007):

- a) Ordenar los datos observados en orden ascendente:  $x_1, x_2, \dots, x_i, \dots, x_n$  con  $x_n$  como el valor más grande.
-

- b) Evaluar la función de distribución acumulada de la distribución propuesta  $F_X(x_i)$  en  $x_i$ , para  $i = 1, 2, \dots, n$
- c) Calcular el estadístico AD de Anderson-Darling:

$$A^2 = \sum_{i=1}^n \frac{2i-1}{n} \{\ln F_X(x_i) + \ln[1 - F_X(x_{n+1-i})]\} - n \quad (3.31)$$

El estadístico AD se utiliza para calcular el p-valor para la prueba de bondad de ajuste, lo que permite determinar si una muestra de datos proviene de una población con una distribución específica. Por lo general, mientras mejor se ajusta la distribución a los datos, menor será el estadístico AD. Las hipótesis para la prueba de Anderson-Darling son las siguientes:

- $H_o$  : Los datos siguen una distribución especificada.
- $H_a$  : Los datos no siguen una distribución especificada.

Si el p-valor para la prueba de Anderson-Darling es inferior al nivel de significación  $\alpha$  seleccionado, se concluye que los datos no siguen la distribución especificada.

- d) Calcular el estadístico AD ajustado  $A^*$  para considerar el efecto del tamaño  $n$  de la muestra. Este ajuste depende de la forma de la distribución de probabilidad seleccionada.
- e) Seleccionar un nivel de significación  $\alpha$  y determinar el correspondiente valor crítico  $c_\alpha$  para el tipo de distribución especificada. Los valores críticos  $c_\alpha$  para las distribuciones más comunes se encuentra reportados en tablas.
- f) Para una distribución dada se compara el estadístico ajustado  $A^*$  con el valor crítico  $c_\alpha$  apropiado. Si  $A^*$  es menor que  $c_\alpha$ , la distribución propuesta es aceptable al nivel de significación  $\alpha$ .

Para la distribución normal, los valores críticos  $c_\alpha$  están dados por:

$$c_\alpha = a_\alpha \left( 1 + \frac{b_0}{n} + \frac{b_1}{n^2} \right) \quad (3.32)$$

donde los valores  $a_\alpha$ ,  $b_0$  y  $b_1$  están dados en tablas (Ang y Tang, 2007) para un nivel de significación dado y el estadístico ajustado  $A^*$  para un tamaño

---

de muestra  $n$  es:

$$A^* = A^2 \left( 1.0 + \frac{0.75}{n} + \frac{2.25}{n^2} \right) \quad (3.33)$$

Para la distribución exponencial, los valores críticos  $c_\alpha$  para una nivel de significación dado están reportados en tablas (Ang y Tang, 2007) y el correspondiente estadístico ajustado  $A^*$  está dado por:

$$A^* = A^2 \left( 1.0 + \frac{0.6}{n} \right) \quad (3.34)$$

Para la distribución gamma, los valores críticos  $c_\alpha$  dependen del parámetro  $k$ , reportado en las tablas (Ang y Tang, 2007), y el estadístico ajustado  $A^*$  también depende de  $k$  como:

$$A^* = \begin{cases} A^2 \left( 1.0 + \frac{0.6}{n} \right) & \text{para } k = 1 \\ A^2 + \frac{1}{n} \left( 0.2 + \frac{0.3}{k} \right) & \text{para } k \geq 2 \end{cases}$$

# Capítulo 4

## Resultados y discusión

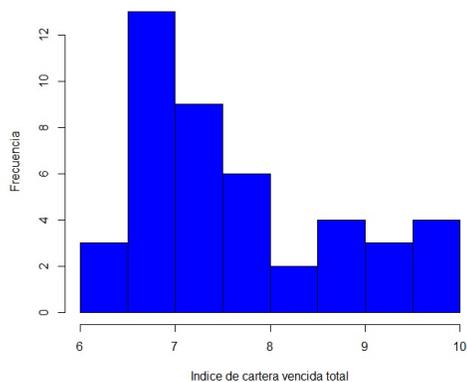
En este capítulo se muestran y analizan los resultados obtenidos para las pruebas estadísticas desarrolladas de acuerdo a la metodología propuesta. El análisis descriptivo de las series se realizó usando el software R, las pruebas de raíz unitaria se efectuaron con el software EViews 8 y las pruebas de cointegración según el método de Engle-Granger se realizaron con el software Eviews 8 y R. La prueba de bondad de ajuste de Anderson-Darling y la simulación de Monte Carlo se realizaron con la aplicación Oracle Crystal Ball. En los apéndices A, B, C y D se muestra como implementar las pruebas.

### 4.1. Análisis descriptivo de las series temporales.

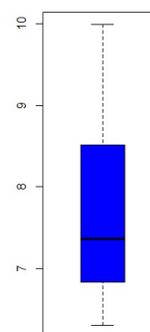
Variable	Mínimo	Mediana	Media	Máximo	Sesgo	Curtosis	Desviación
ICVT	6.308	7.361	7.676	9.987	0.7648	-0.6552	1.0708
PNC	0.4467	0.6092	0.6459	0.8984	0.5099	-0.8047	0.1225
CoCre	0.1351	0.5668	0.5563	0.8035	-0.6886	0.4963	0.1469

Tabla 4.1: Resumen estadística descriptiva.

La tabla 4.1 proporciona un resumen de la estadística descriptiva del conjunto de datos de cada serie temporal. Como se observa, en la variable CoCre, la media y la mediana no están tan alejadas entre sí, lo que indica que posiblemente esta serie es ligeramente simétricas.

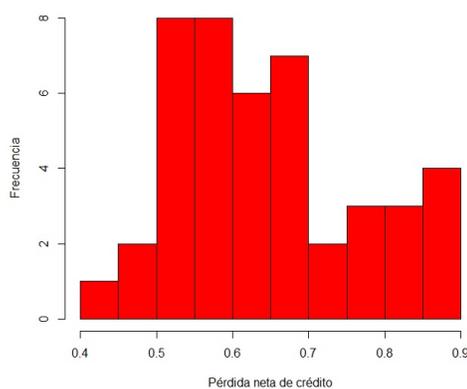


(a) Histograma de la serie ICVT.

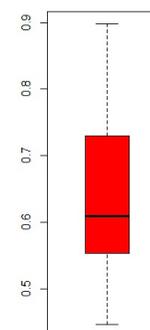


(b) Boxplot de la serie ICVT.

Figura 4.1: Histograma y boxplot de la serie índice de cartera vencida total.



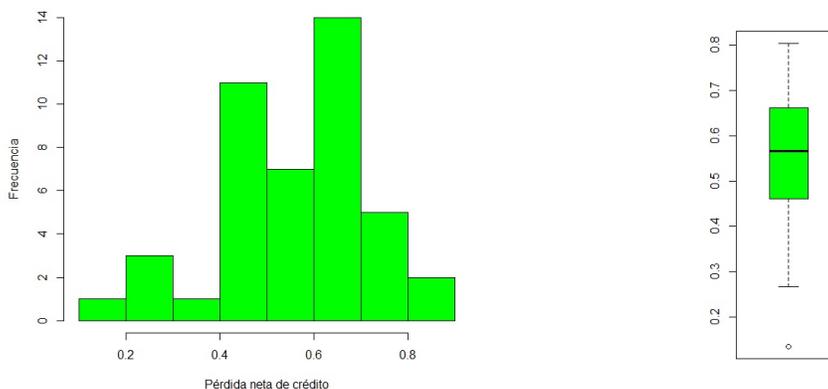
(a) Histograma de la serie PNC.



(b) Boxplot de la serie PNC.

Figura 4.2: Histograma y boxplot de la serie pérdida neta de crédito.

Como se observa en los boxplot de las figuras 4.1(b) y 4.2(b), la mediana está desplazada hacia el primer cuartil y el bigote de arriba es más largo, lo que indica, en concordancia con los resultados del resumen estadístico, mostrados en la tabla 4.1, y los histogramas de las figuras 4.1(a) y 4.2(a), que las series del ICVT y la PNC están sesgadas a la derecha. La altura de ambos boxplot sugiere que los datos están bastante dispersos alrededor de la mediana y en las dos series la curtosis es negativa por lo que las distribuciones son platicúrticas.



(a) Histograma de la serie COCRE.

(b) Boxplot de la serie COCRE.

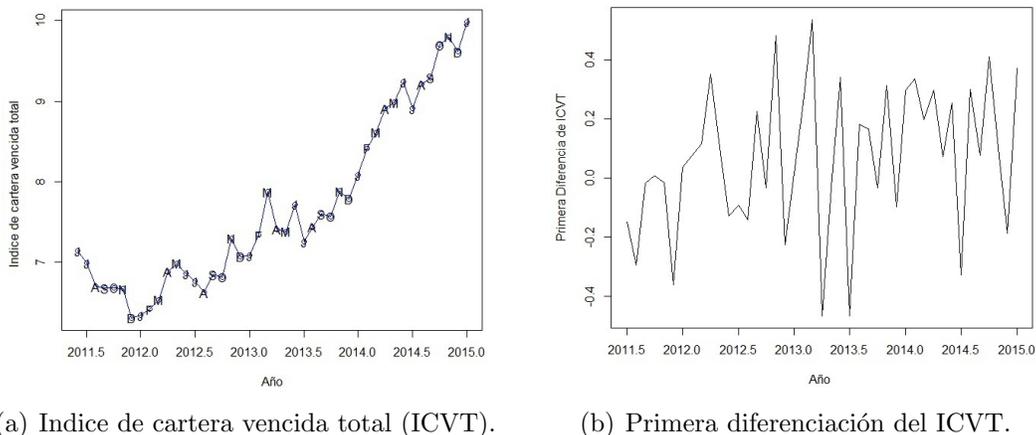
Figura 4.3: Histograma y boxplot de la serie costo de crédito.

En el boxplot de la figura 4.3(b), se observa que la mediana está levemente desplazada hacia el tercer cuartil y el bigote de abajo es un poco más largo, lo que sugiere, en coincidencia con el histograma de la figura 4.3(a) y los resultados del resumen estadístico, que la serie del CoCRE está ligeramente sesgada a la izquierda. La altura del boxplot indica que los datos están dispersos alrededor de la mediana y la curtosis positiva muestra que la distribución es leptocúrtica.

Variable	Shapiro-Wilk	p-valor
ICVT	0.8962	0.000837
PNC	0.9416	0.02714
CoCre	0.9610	0.1426

Tabla 4.2: Prueba de normalidad de las series de tiempo.

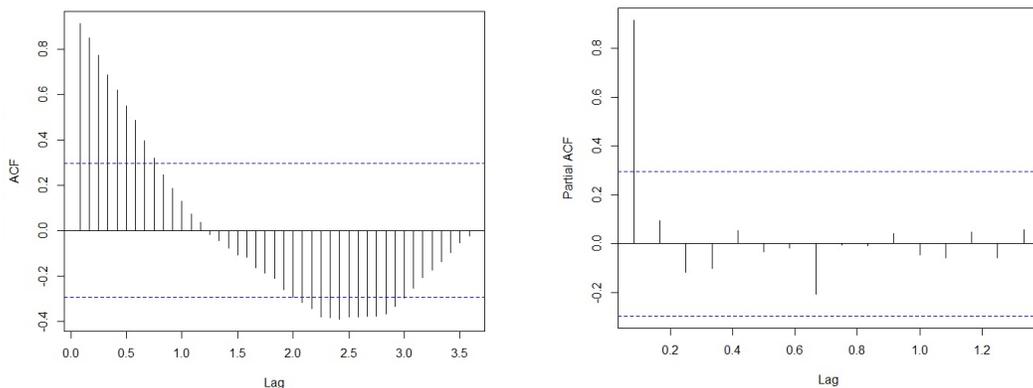
Los p-valores de las series ICVT y PNC son menores al nivel de significación  $\alpha = 0.05$  entonces no hay evidencias estadísticas de que el proceso generador de datos provenga de una distribución normal. Por otra parte, el p-valor de la serie del CoCre es mayor al nivel de significación  $\alpha$  y entonces hay suficiente evidencia estadística de que el mecanismo generador de datos provenga de una distribución normal.



(a) Índice de cartera vencida total (ICVT).

(b) Primera diferenciación del ICVT.

Figura 4.4: Serie del índice de cartera vencida total (ICVT) y su primera diferenciación.



(a) Función de autocorrelación del ICVT.

(b) Función de autocorrelación parcial del ICVT.

Figura 4.5: Función de autocorrelación y autocorrelación parcial del ICVT.

En la figura 4.4(a) se observa que la serie del índice de cartera vencida total (ICVT) presenta tendencia, y esto se confirma con su función de autocorrelación, mostrada en la figura 4.5(a), donde se aprecia claramente una caída suave. La figura 4.5(a) también nos indica que la serie no presenta un ciclo estacional. En la gráfica de la primera diferenciación del ICVT, mostrada en la figura 4.4(b), se observa que la serie fluctúa alrededor de la media cero y que la varianza se mantiene aproximadamente constante. Por tanto, la serie del ICVT no es estacionaria.

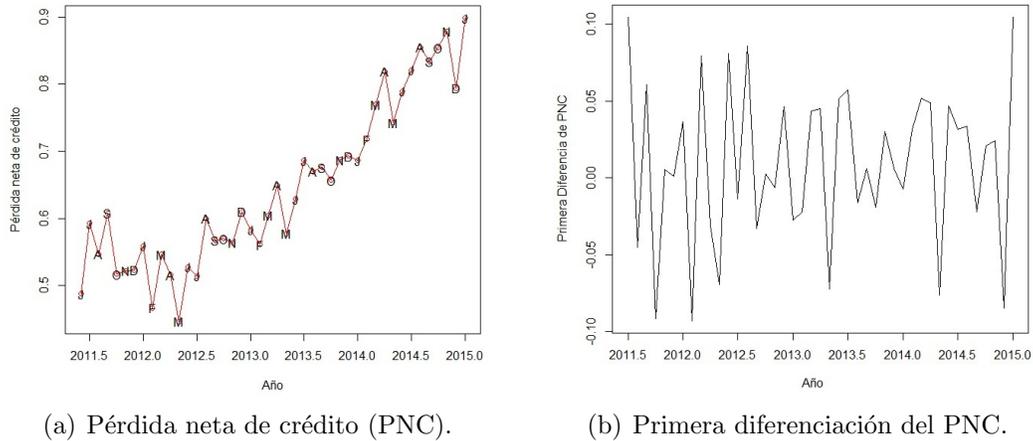


Figura 4.6: Serie de la pérdida neta de crédito (PNC) y su primera diferenciación.

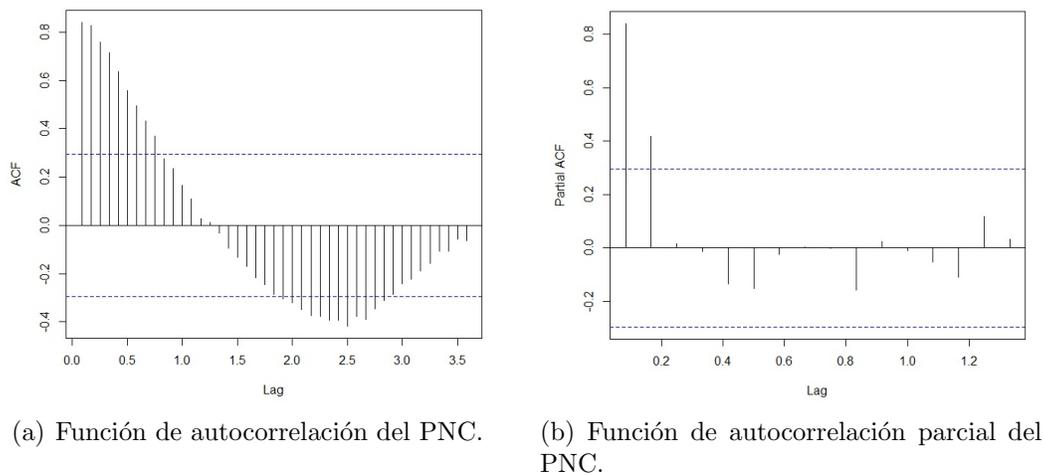
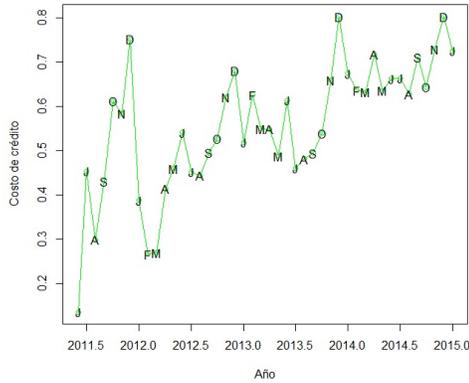
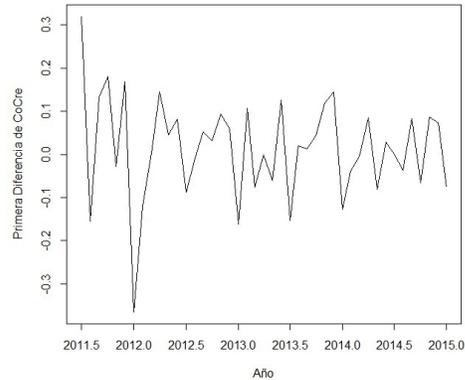


Figura 4.7: Función de autocorrelación y autocorrelación parcial del PNC.

La serie pérdida neta de crédito (PNC) también presenta tendencia, como se puede apreciar en la figura 4.6(a), y esto se corrobora con su función de autocorrelación, donde se presenta una caída suave, como se puede observar en la figura 4.7(a). Además, tampoco hay presencia de un ciclo estacional. La primera diferenciación del PNC, mostrada en la figura 4.6(b), indica que la serie se mantiene fluctuando en una región delimitada alrededor de la media cero. En consecuencia, la serie del PNC tampoco es estacionaria.

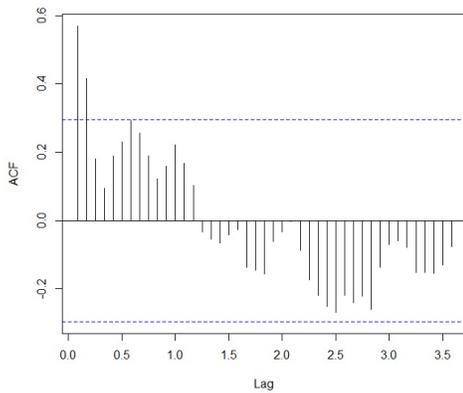


(a) Costo de crédito COCRE.

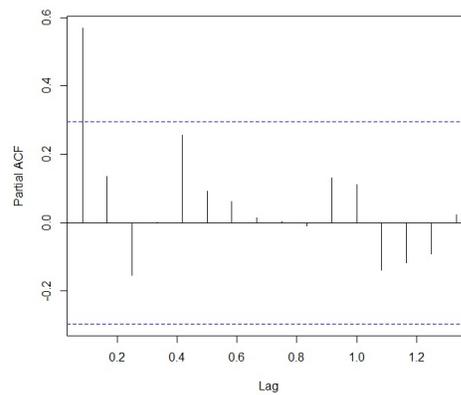


(b) Primera diferenciación del CoCRE.

Figura 4.8: Serie del costo de crédito (CoCre) y su primera diferenciación.



(a) Función de autocorrelación del COCRE.



(b) Función de autocorrelación parcial del COCRE.

Figura 4.9: Función de autocorrelación y autocorrelación parcial del CoCre.

Por último, la figura 4.8(a) muestra la serie del costo de crédito (CoCre). Esta serie exhibe cierta volatilidad y aparentemente presenta una tendencia, sin embargo, al observar la figura 4.9(a) se ve que la función de autocorrelación cae de manera muy abrupta, lo que indicaría la ausencia de una tendencia. En este caso tampoco se observa la existencia de un ciclo estacional. La primera diferenciación del CoCre, mostrada en la figura 4.8(b), muestra que la serie fluctúa alrededor de la media cero y que la varianza se mantiene aproximadamente constante. Por tanto, la serie del CoCre tampoco es estacionaria.

## 4.2. Prueba de Dickey-Fuller aumentada.

Al realizar la prueba de raíz unitaria de Dickey-Fuller aumentada (ADF), para la regresión con intercepto y tendencia y usando cero rezagos, para la serie del ICVT, se encontró que los residuos de la regresión son ruido blanco, como se puede observar en el correlograma de la figura 4.10, donde todos los rezagos se encuentran dentro de las bandas y los p-valores están por encima del 5 %, sin embargo, el estadístico de Durbin-Watson fue de 2.414, que esta por fuera del intervalo aceptable establecido entre 1.85 y 2.15, por tanto, hay presencia de autocorrelación. Para corregir la existencia de autocorrelación se añadió más dinámica al modelo incorporando un rezago y se realizó de nuevo la prueba ADF para la regresión con intercepto y tendencia obteniendo un valor de 2.037 para el estadístico DW.

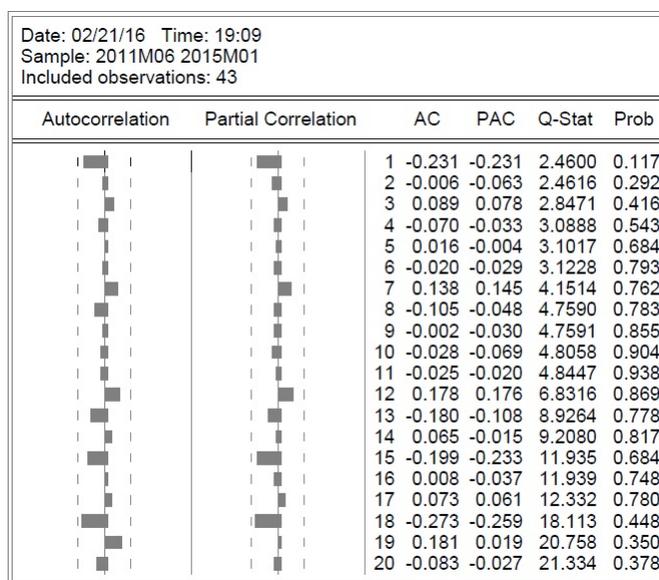


Figura 4.10: Correlograma de los residuos de la regresión con intercepto y tendencia con cero rezagos para la serie del ICVT.

Luego, se realizó la prueba ADF para las regresiones con intercepto y sin intercepto ni tendencia incluyendo un rezago. Con los valores obtenidos de la suma de los cuadrados de los residuos, de cada una de las tres regresiones, se realizó la validación del modelo, aplicando las pruebas de hipótesis para subconjuntos, para poder establecer cual de los tres contrastes debe interpretarse. Los resultados se muestran en las tablas 4.3 y 4.4.

$SSE_r$	$SSE_{nr}$	$N$	$l$	$k$	$\phi_3$	Valor crítico n.s. 5 %
2.3792	1.8778	42	1	4	10.14	6.73

Tabla 4.3: Contraste  $\phi_3$  para la serie del ICVT.

Como se puede observar en la tabla 4.3 el valor del estadístico  $\phi_3 = 10.14$  cae a la derecha del valor crítico 6.73, en la región de rechazo. Se concluye, con un nivel de significación del 5 %, que hay evidencia suficiente para rechazar la hipótesis nula de una regresión por MCO con intercepto, por tanto, se debe continuar a la siguiente prueba para establecer si el contraste a interpretar es el que se realiza con la regresión por MCO con intercepto y tendencia o una regresión por MCO sin intercepto ni tendencia.

$SSE_r$	$SSE_{nr}$	$N$	$l$	$k$	$\phi_2$	Valor crítico n.s. 5 %
2.4262	1.8778	42	2	4	5.54	5.13

Tabla 4.4: Contraste  $\phi_2$  para la serie del ICVT.

En la tabla 4.4 se aprecia que el valor del estadístico  $\phi_2 = 5.54$  cae a la derecha del valor crítico 5.13, en la región de rechazo. Se concluye, con un nivel de significación del 5 %, que hay suficiente evidencia para rechazar la hipótesis nula de una regresión por MCO sin intercepto ni tendencia, por tanto, se debe interpretar el contraste que se efectúa con la regresión por MCO con intercepto y tendencia.

Regresión auxiliar con intercepto y tendencia				
Variable	Coefficiente	Desviación estándar	t	P-valor
ICVT(-1)	-0.221191	0.090945	-2.432146	0.0198
D(ICVT(-1))	-0.233515	0.147272	-1.585602	0.1211
C	1.232744	0.541260	2.277545	0.0285
TREND	0.024042	0.007548	3.185352	0.0029
Estadístico de Durbin-Watson		2.037527	Rezagos	1
Prueba ADF a nivel		Valor crítico n.s. 5 %	t	P-valor
		-3.520787	-2.432146	0.3586
Prueba ADF primera diferenciación		Valor crítico n.s. 5 %	t	P-valor
		-3.523623	-5.956496	0.0001

Tabla 4.5: Prueba de Dickey-Fuller aumentada de la serie del ICVT y su primera diferenciación para regresión auxiliar con intercepto y tendencia.

En la tabla 4.5 se muestran los resultados de la prueba de Dickey-Fuller aumentada, para el contraste con intercepto y tendencia utilizando un rezago, para la serie del ICVT y su primera diferencia. Como se puede observar el estadístico Durbin-Watson es 2.037, por tanto, no hay presencia de autocorrelación, y el correlograma de los residuos, mostrado en la figura 4.11, es ruido blanco, los rezagos están dentro de las bandas y los p-valores están por arriba del 5 %. El p-valor del último rezago de la regresión es muy grande y en consecuencia no es significativo, sin embargo, el problema de la no significación es menos grave que el de la autocorrelación.

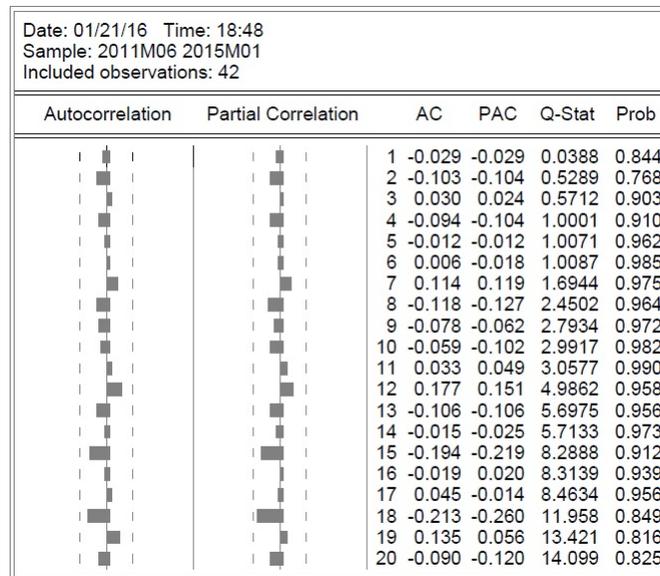


Figura 4.11: Correlograma de los residuos de la regresión con intercepto y tendencia con un rezago para la serie del ICVT.

El estadístico t de la prueba ADF a nivel da un valor de -2.43 que cae a la derecha del valor crítico de -3.52 en la región de no rechazo, por tanto, se concluye, con un nivel de significación del 5 %, que no hay evidencia suficiente para rechazar la hipótesis nula y por tanto la serie del ICVT tiene una raíz unitaria y no es estacionaria. Finalmente, al realizar la prueba ADF para la primera diferenciación se obtiene que el estadístico t es -5.95 y se encuentra a la izquierda del valor crítico de -3.52, en la región de rechazo, por tanto, se concluye, con un nivel de significación del 5 %, que hay suficiente evidencia para rechazar la hipótesis nula y por tanto la primera diferenciación de la serie de ICVT no tiene raíz unitaria, es decir, es estacionaria y por consiguiente la serie del ICVT es integrada de orden uno.

Cuando se efectuó la prueba ADF para la regresión con intercepto y tendencia utilizando cero rezagos, para la serie de la pérdida neta de crédito (PNC), se halló que los residuos de la regresión son ruido blanco, como se puede apreciar en el correlograma de la figura 4.12, donde todos los rezagos se encuentran dentro de las bandas y los p-valores están por encima del 5 %. En este caso el estadístico de Durbin-Watson fue de 2.147 y se encuentra dentro del rango aceptable, sin embargo, al realizar los contrastes con intercepto y sin intercepto ni tendencia se encontró que el valor del estadístico de Durbin-Watson era de 2.69 y 2.82, respectivamente, lo cual puede conducir a un problema de autocorrelación en el momento de realizar la validación del modelo. Por tal motivo se decidió incrementar el número de rezagos a dos y realizar de nuevo las regresiones con intercepto y tendencia, intercepto y sin intercepto ni tendencia, obteniendo respectivamente los valores 2.086, 2.091 y 2.079 para el estadístico de Durbin-Watson, con lo que se corrigió el problema de posible presencia de autocorrelación.

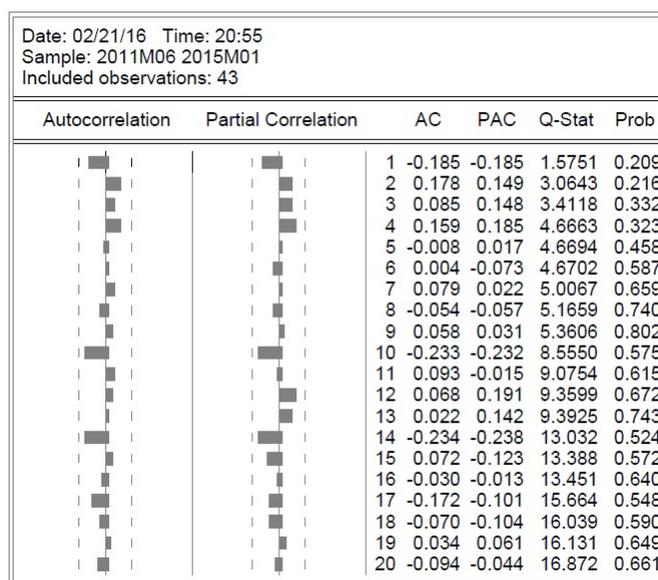


Figura 4.12: Correlograma de los residuos de la regresión con intercepto y tendencia con cero rezagos para la serie de la PNC.

Los resultados de la validación del modelo se indican en las tablas 4.6 y 4.7. Como se puede apreciar, el valor del estadístico  $\phi_3 = 8.13$  cae a la derecha del valor crítico 6.73, en la región de rechazo. Se concluye, con un nivel de significación del 5 %, que hay evidencia suficiente para rechazar la hipótesis nula, por tanto, se

procede a la siguiente prueba para establecer si el contraste a interpretar es el que se realiza con la regresión con intercepto y tendencia o una regresión sin intercepto ni tendencia.

$SSE_r$	$SSE_{nr}$	$N$	$l$	$k$	$\phi_3$	Valor crítico n.s. 5 %
0.069891	0.057007	41	1	5	8.13	6.73

Tabla 4.6: Contraste  $\phi_3$  para la serie de la PNC.

El valor del estadístico  $\phi_2 = 4.10$  cae a la izquierda del valor crítico 5.13, en la región de no rechazo. Se concluye, con un nivel de significación del 5 %, que no hay suficiente evidencia para rechazar la hipótesis nula, por tanto, se debe interpretar el contraste que se realiza con la regresión sin intercepto ni tendencia.

$SSE_r$	$SSE_{nr}$	$N$	$l$	$k$	$\phi_2$	Valor crítico n.s. 5 %
0.069993	0.057007	41	2	5	4.10	5.13

Tabla 4.7: Contraste  $\phi_2$  para la serie de la PNC.

Regresión auxiliar sin intercepto ni tendencia				
Variable	Coeficiente	Desviación estándar	t	P-valor
PNC(-1)	0.023008	0.010861	2.118453	0.0407
D(PNC(-1))	-0.677799	0.161117	-4.206879	0.0002
D(PNC(-2))	-0.279976	0.159832	-1.751693	0.0879
Estadístico de Durbin-Watson		2.079632	Rezagos	2
Prueba ADF a nivel		Valor crítico n.s. 5 %	t	P-valor
		-1.949097	2.118453	0.9907
Prueba ADF primera diferenciación		Valor crítico n.s. 5 %	t	P-valor
		-1.949319	-5.031370	0.0000

Tabla 4.8: Prueba de Dickey-Fuller aumentada de la serie del PNC y su primera diferenciación para regresión auxiliar sin intercepto ni tendencia.

En la tabla 4.8 se indican los resultados de la prueba de Dickey-Fuller aumentada, para el contraste sin intercepto ni tendencia utilizando dos rezagos, para la serie de la PNC y su primera diferencia. Como se puede apreciar el estadístico Durbin-Watson es de 2.079, por tanto, no hay autocorrelación. Además el correlograma de los residuales, mostrado en la figura 4.13, es ruido blanco, y todos los p-valores

están por arriba del 5 %. El p-valor del último rezago de la regresión es muy grande y en consecuencia no es significativo. El p-valor de la prueba ADF a nivel es de 0.9907, que es mayor al nivel de significación del 5 %, por tanto, se concluye que no hay evidencia suficiente para rechazar la hipótesis nula y entonces la serie de la pérdida neta de crédito (PNC) tiene una raíz unitaria y no es estacionaria. Finalmente, al realizar la prueba ADF para la primera diferenciación se obtiene que el estadístico t es -5.03 y se encuentra a la izquierda del valor crítico de -1.94, en la región de rechazo, por tanto, se concluye, con un nivel de significación del 5 %, que hay evidencia suficiente para rechazar la hipótesis nula y en consecuencia la primera diferenciación de la serie de la pérdida neta de crédito no tiene raíz unitaria, luego la serie es estacionaria y por tanto, la serie de la PNC también es integrada de orden uno.

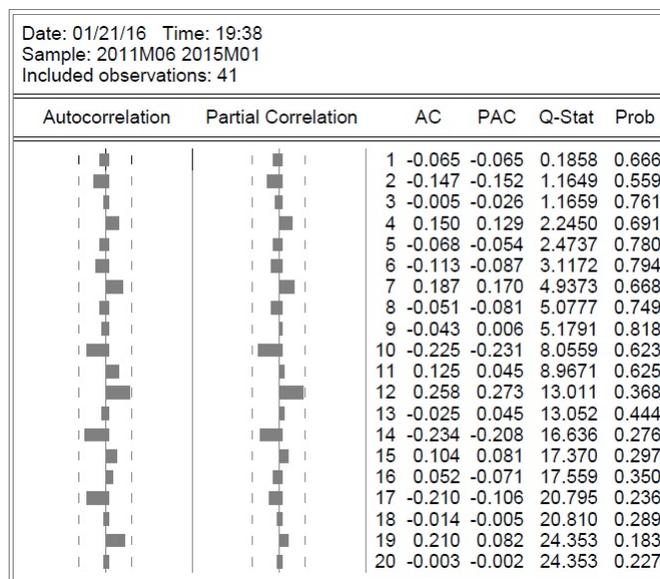


Figura 4.13: Correlograma de los residuos de la regresión sin intercepto y tendencia con dos rezagos para la serie de la PNC.

Cuando se efectuó la prueba de Dickey-Fuller aumentada (ADF), para la regresión con intercepto y tendencia y utilizando cero rezagos, para la serie del costo de crédito (CoCre), se determinó que los residuos de la regresión son ruido blanco, como se puede observar en el correlograma de la figura 4.14, donde todos los rezagos se encuentran dentro de las bandas y las probabilidades están por encima del 5 %, En este caso el estadístico de Durbin-Watson fue de 1.776, que esta por fuera del

rango admisible, por tanto, hay presencia de autocorrelación. Se estableció que era necesario incorporar cuatro rezagos al realizar la prueba ADF para la regresión con intercepto y tendencia para corregir el problema de autocorrelación, obteniendo para el estadístico de Durbin-Watson un valor de 2.049.

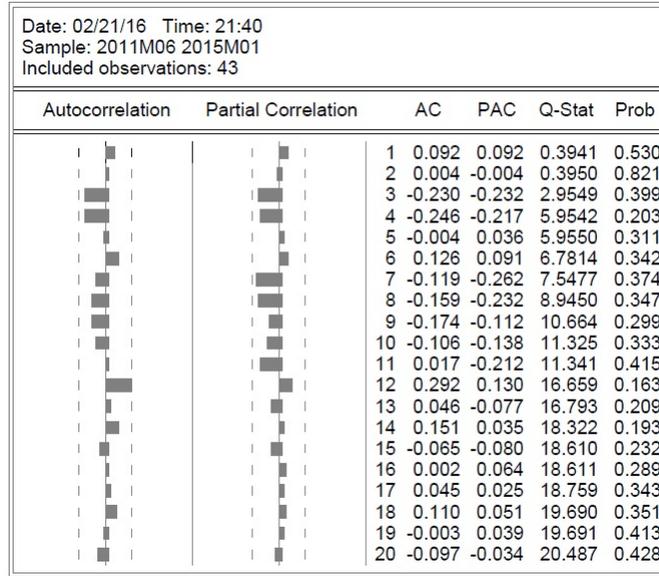


Figura 4.14: Correlograma de los residuos de la regresión con intercepto y tendencia con cero rezagos para la serie del CoCre.

Los resultados de la validación del modelo se indican en las tablas 4.9 y 4.10. Como se puede observar, el valor del estadístico  $\phi_3 = 8.74$  cae a la derecha del valor crítico 6.73, en la región de rechazo. Se concluye, con un nivel de significación del 5 %, que hay evidencia suficiente para rechazar la hipótesis nula, por consiguiente, se procede a la siguiente prueba para determinar si el contraste a interpretar es el que se efectúa con la regresión con intercepto y tendencia o con la regresión sin intercepto ni tendencia.

$SSE_r$	$SSE_{nr}$	$N$	$l$	$k$	$\phi_3$	Valor crítico n.s. 5 %
0.300015	0.235615	39	1	7	8.74	6.73

Tabla 4.9: Contraste  $\phi_3$  para la serie del CoCre.

En la tabla 4.10 se observa que el valor del estadístico  $\phi_2 = 5.29$  cae a la derecha del valor crítico 5.13, en la región de rechazo. Se concluye, con un nivel de significación del 5 %, que hay suficiente evidencia para rechazar la hipótesis nula de una regresión sin intercepto ni tendencia, por tanto, se debe interpretar el contraste que se efectúa con la regresión con intercepto y tendencia.

$SSE_r$	$SSE_{nr}$	$N$	$l$	$k$	$\phi_2$	Valor crítico n.s. 5 %
0.313633	0.235615	39	2	7	5.29	5.13

Tabla 4.10: Contraste  $\phi_2$  para la serie del CoCre.

En la tabla 4.11 se muestran los resultados obtenidos para la prueba de Dickey-Fuller aumentada, para el contraste con intercepto y tendencia utilizando cuatro rezagos, para la serie del CoCre y su primera diferencia. Como se observa el estadístico Durbin-Watson es de 2.049, por tanto, no hay autocorrelación. Además, el correlograma de los residuales, mostrado en la figura 4.15, es ruido blanco, y todos los p-valores están por arriba del 5 %.

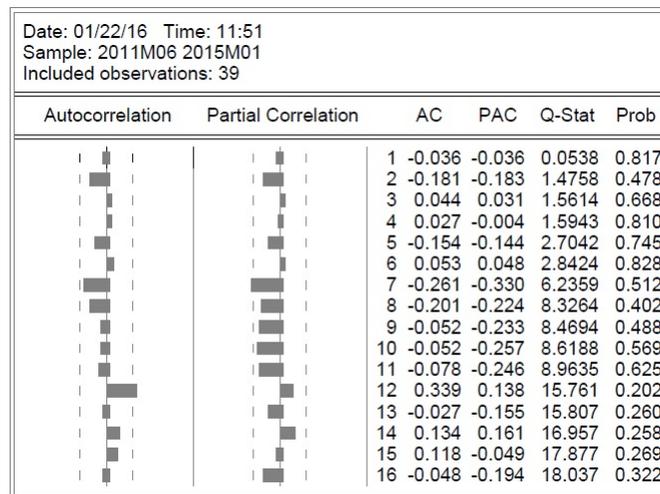


Figura 4.15: Correlograma de los residuos de la regresión con intercepto y tendencia con cuatro rezagos para la serie del CoCre.

Por otra parte, el p-valor del último rezago de la regresión es muy grande y por consiguiente no es significativo. El estadístico t de la prueba ADF a nivel da un valor de -3.17 que cae a la derecha del valor crítico de -3.52, en la región de no rechazo, por tanto, se concluye, con un nivel de significación del 5 %, que no hay

evidencia suficiente para rechazar la hipótesis nula y por consiguiente, la serie del costo de crédito (CoCre) tiene una raíz unitaria, es decir, no es estacionaria. Finalmente, al realizar la prueba ADF para la primera diferenciación se obtiene que el estadístico  $t$  es  $-4.60$  y se encuentra a la izquierda del valor crítico de  $-3.53$ , en la región de rechazo, por tanto, se concluye, con un nivel de significación del 5 %, que hay evidencia suficiente para rechazar la hipótesis nula y en consecuencia la primera diferenciación de la serie del costo de crédito no tiene raíz unitaria, es decir, la serie es estacionaria y por consiguiente la serie del CoCre es integrada de orden uno.

Regresión auxiliar con intercepto y tendencia				
Variable	Coficiente	Desviación estándar	$t$	P-valor
COSTO(-1)	-0.961867	0.303070	-3.173750	0.0033
D(COSTO(-1))	0.337457	0.238152	1.416983	0.1662
D(COSTO(-2))	0.414296	0.203295	2.037899	0.0499
D(COSTO(-3))	0.082727	0.196404	0.421211	0.6764
D(COSTO(-4))	-0.034362	0.148995	-0.230624	0.8191
C	0.373763	0.122013	3.063297	0.0044
TREND	0.007279	0.002461	2.957459	0.0058
Estadístico de Durbin-Watson		2.049161	Rezagos	4
Prueba ADF a nivel		Valor crítico n.s. 5 %	$t$	P-valor
		-3.529758	-3.173750	0.1045
Prueba ADF primera diferenciación		Valor crítico n.s. 5 %	$t$	P-valor
		-3.533083	-4.607637	0.0037

Tabla 4.11: Prueba de Dickey-Fuller aumentada de la serie del CoCre y su primera diferenciación para regresión auxiliar con intercepto y tendencia.

La prueba de raíz unitaria de Dickey-Fuller aumentada permitió establecer que las series indicador de cartera vencida total (ICVT), pérdida neta de crédito (PNC) y costo de crédito (CoCre) son integradas de orden uno, es decir, son series no estacionarias cada una con una raíz unitaria. Por tanto, es posible que exista una relación de equilibrio a largo plazo entre estas variables, es decir, que estén cointegradas.

### 4.3. Regresiones del modelo

A partir de los resultados indicados en la tabla 4.12 se obtuvo la siguiente regresión por MCO de la pérdida neta de crédito (PNC) contra el indicador de cartera vencida total (ICVT):

$$PNC_t = -0.173051 + 0.106678ICVT_t \quad (4.1)$$

Se puede apreciar que existe una relación positiva de largo plazo entre la PNC y el ICVT. En este caso si el ICVT se incrementa en un punto porcentual, la PNC aumenta alrededor de 0.11 puntos porcentuales. El valor de  $r^2$  de 0.8691 indica que alrededor del 86.91 % de la variación de la PNC se explica por la variación del ICVT, que es un porcentaje alto, por consiguiente, los datos se ajustan bastante bien a la línea de regresión. El valor del estadístico t asociado al coeficiente de la pendiente  $\beta_1$  es 16.70255 y el p-valor de éste estadístico es muy pequeño, por tanto,  $\beta_1$  es muy significativo. Entonces, se concluye, al nivel de significación del 5 %, que hay evidencia suficiente para rechazar la hipótesis nula de que  $\beta_1 = 0$ , por consiguiente, la pérdida neta de crédito si es explicada por el índice de cartera vencida total.

Variable	Coefficiente	Desviación estándar	t	P-valor
C	-0.173051	0.049493	-3.396451	0.0011
ICVT	0.106678	0.006387	16.70255	0.0000
N	$r^2$	Estadístico F	P-valor (estadístico F)	
44	0.869149	278.9753	0.000000	

Tabla 4.12: Regresión entre la PNC y el ICVT

Con los resultados indicados en la tabla 4.13 se determinó la siguiente regresión por MCO entre el costo de crédito (CoCre) y la pérdida neta de crédito (PNC):

$$CoCre_t = 0.082806 + 0.421377PNC_t + 0.379426CoCre_{t-1} \quad (4.2)$$

Para que el modelo fuera consistente se decidió incluir un término de la variable dependiente retardada un periodo, lo que sugiere que el pasado influye en el comportamiento de la serie del costo de crédito, y el modelo adquiere una estructura autoregresiva. Sin embargo, en la práctica no es común que se disponga de información del indicador para periodos de tiempo anteriores.

Variable	Coefficiente	Desviación estándar	t	P-valor
C	0.082806	0.080741	1.025579	0.3113
PNC	0.421377	0.151585	2.779804	0.0082
COSTO(-1)	0.379426	0.125863	3.014598	0.0045
N	$r^2$	Estadístico F	P-valor (estadístico F)	
43	0.512400	21.01719	0.000001	

Tabla 4.13: Regresión entre el CoCre y la PNC.

Con este modelo el valor que se obtuvo para  $r^2$  de 0.5124 muestra que alrededor del 51.24 % de la variación del CoCre se explica mediante la variación del PNC y el CoCre retardado un periodo. El coeficiente de 0.4213 indica que si el PNC se incrementa en un punto porcentual, manteniendo constante la influencia del CoCre retardado un periodo, el costo de crédito aumenta en 0.4213 puntos porcentuales y el coeficiente de 0.3794 señala que si el efecto de la PNC se mantiene constante el CoCre aumenta en 0.3794 puntos porcentuales, si el CoCre retardado un periodo sube un punto porcentual.

Los valores del estadístico  $t$  asociado a los coeficientes de pendiente  $\beta_1$  y  $\beta_2$  son de 2.7798 y 3.0145, respectivamente, y sus correspondientes p-valores son muy pequeños, por consiguiente, individualmente los coeficientes  $\beta_1$  y  $\beta_2$  son muy significativos. Además, el valor del estadístico  $F$  fue de 21.017 y su p-valor fue de 0.000001, que es una probabilidad muy pequeña, lo que implica que, al nivel de significación del 5 %, hay suficiente evidencia para rechazar la hipótesis nula de que todos los coeficientes de pendiente son simultáneamente cero y, por tanto, se concluye que la PNC y el CoCre retardado un periodo si afectan a el CoCre.

## 4.4. Prueba de cointegración de Engle-Granger

Para determinar si es posible una regresión cointegrante entre la serie de la pérdida neta de crédito (PNC) y el indicador de cartera vencida total (ICVT) y entre el costo de crédito (CoCre) y la pérdida neta de crédito (PNC) se realizó la prueba de cointegración de Engle-Granger. Primero se realizó el análisis gráfico de las series de residuos y luego se aplicó la prueba de raíz unitaria en el contexto de la prueba de cointegración para establecer si los residuos son estacionarios.

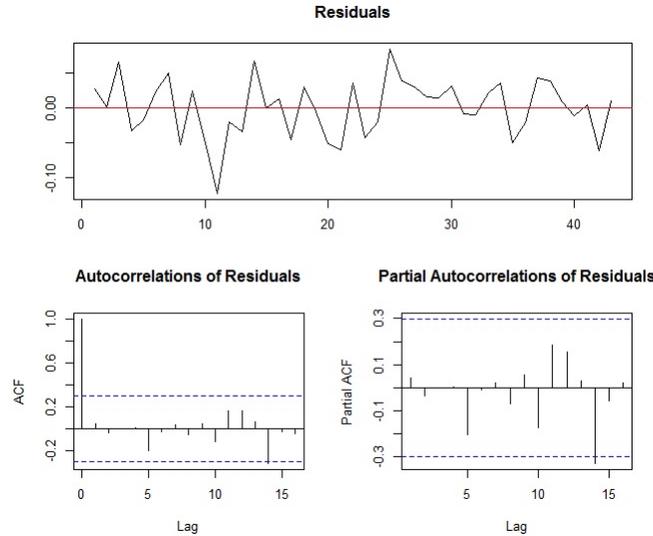


Figura 4.16: Gráfico de los residuos de la regresión PNC vs. ICVT

Regresión auxiliar con intercepto				
Variable	Coficiente	Desviación estándar	t	P-valor
RESID01(-1)	-0.897785	0.145412	-6.174090	0.0000
C	0.002358	0.006444	0.365914	0.7163
Estadístico de Durbin-Watson		1.901711	Rezagos	0
Prueba ADF		Valor crítico n.s. 5 % -3.48	t	-6.174090

Tabla 4.14: Prueba de cointegración y raíz unitaria para los residuos de la regresión entre el PNC y el ICVT.

En la figura 4.16 se observa la gráfica de la serie de residuos de la regresión entre la pérdida neta de crédito (PNC) y el índice de cartera vencida total (ICVT), como se observa la serie de residuos se mueve alrededor de la media cero y no presenta una tendencia determinista.

En la tabla 4.14 se muestran los resultados para la prueba de cointegración y raíz unitaria para los residuos de la regresión entre el PNC y el ICVT. Como se aprecia el estadístico de Durbin-Watson tiene un valor de 1.90, que se encuentra dentro del rango admisible, por tanto, no hay autocorrelación. El estadístico t de los residuos retardados un periodo es significativo y el estadístico t de la prueba de Dickey-Fuller aumentada fue -6.17, que esta en la región de rechazo a la izquierda del valor

crítico -3.48, obtenido según la tabla de MacKinnon para una regresión con dos variables, sin tendencia y al nivel de significación del 5 %. Según estos resultados podemos concluir, con un nivel de significación del 5 %, que hay evidencia suficiente para rechazar la hipótesis nula, por tanto, los residuos no tienen una raíz unitaria, es decir, son estacionarios. Luego, las series pérdida neta de crédito (PNC) e índice de cartera vencida total (ICVT) están cointegradas y existe una relación de largo plazo entre ellas.

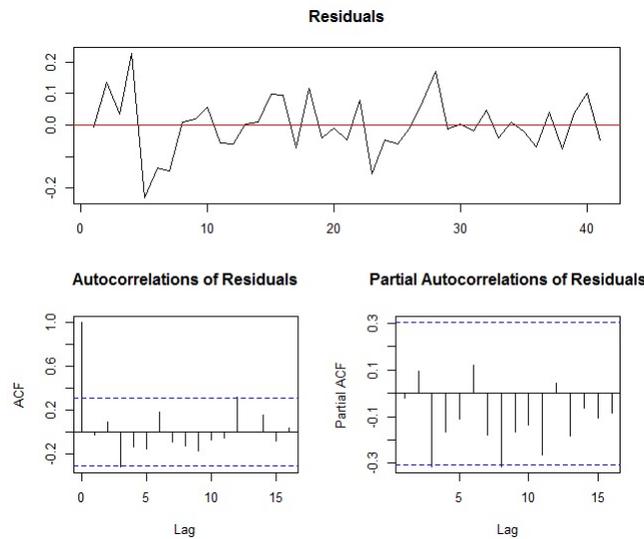


Figura 4.17: Gráfico de los residuos de la regresión CoCre vs. PNC

Regresión auxiliar con intercepto				
Variable	Coefficiente	Desviación estándar	t	P-valor
RESID02(-1)	-0.980484	0.213679	-4.588575	0.0000
D(RESID02(-1))	0.052793	0.155446	0.339624	0.7360
C	0.002829	0.014369	0.196913	0.8449
Estadístico de Durbin-Watson		2.035560	Rezagos	1
Prueba ADF		Valor crítico n.s. 5 % -3.48	t -4.588575	

Tabla 4.15: Prueba de cointegración y raíz unitaria para los residuos de la regresión entre el CoCre y el PNC.

En la figura 4.17 se muestra la gráfica de la serie de residuos de la regresión entre el costo de crédito (CoCre) y la pérdida neta de crédito (PNC), como se puede observar la serie de residuos se mueve alrededor de la media cero y la varianza se mantiene aproximadamente constante en el tiempo.

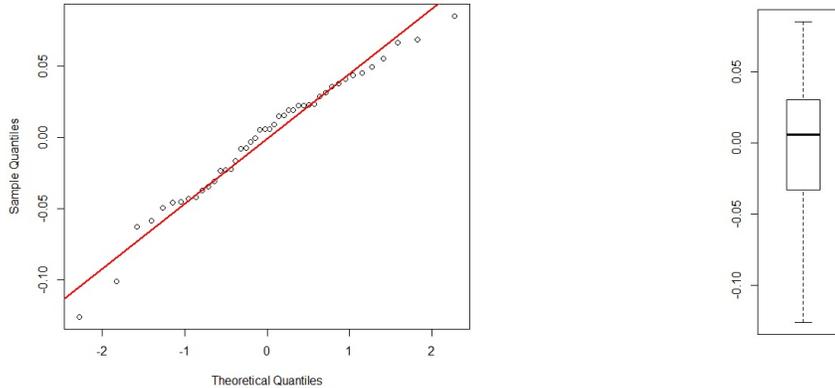
Los resultados para la prueba de cointegración y raíz unitaria para los residuos de la regresión entre el CoCre y el PNC se indican en la tabla 4.15. El valor del estadístico de Durbin-Watson fue 2.03, que se encuentra dentro del intervalo 1.85 y 2.15, por lo cual, no hay presencia de autocorrelación. El último retardo de los residuos tiene un p-valor igual a 0.73, que es mayor al 5 %, por consiguiente, no es significativo. El estadístico t de la prueba de Dickey-Fuller aumentada fue -4.58, que esta en la región de rechazo a la izquierda del valor crítico -3.48, obtenido según la tabla de MacKinnon para una regresión con dos variables, sin tendencia y al nivel de significación del 5 %. Según esto podemos concluir, con un nivel de significación del 5 %, que hay evidencia suficiente para rechazar la hipótesis nula, por lo cual, los residuos no tienen una raíz unitaria, es decir, son estacionarios. Por consiguiente, las series costo de crédito (CoCre) y pérdida neta de crédito (PNC) también están cointegradas.

## 4.5. Pruebas de diagnóstico de los residuos de las regresiones del modelo

Como se aprecia en el gráfico de probabilidad normal de los residuos de la regresión cointegrante entre el ICVT y la PNC, mostrado en la figura 4.18(a), una línea recta se ajusta de manera razonable a los datos, por tanto, los residuos parecen distribuirse normalmente. Sin embargo, al observar el boxplot de los residuos, mostrado en la figura 4.18(b), se aprecia que la mediana esta ligeramente desplazada hacia el tercer cuartil y el bigote de abajo es un poco más largo, lo que indica que hay un ligero sesgo hacia la izquierda.

El histograma de los residuos de la regresión se muestra en la figura 4.19 y exhibe un leve sesgo a la izquierda. Como se observa, el coeficiente de asimetría de -0.56 nos confirma que hay un ligero sesgo a la izquierda y la kurtosis es un poco mayor a tres, por lo que la distribución de los residuos no parece ser perfectamente normal.

---



(a) Gráfico de probabilidad de los residuos.

(b) Boxplot de los residuos.

Figura 4.18: Gráfico de probabilidad normal y boxplot de los residuos de la regresión de la PNC contra el ICVT.

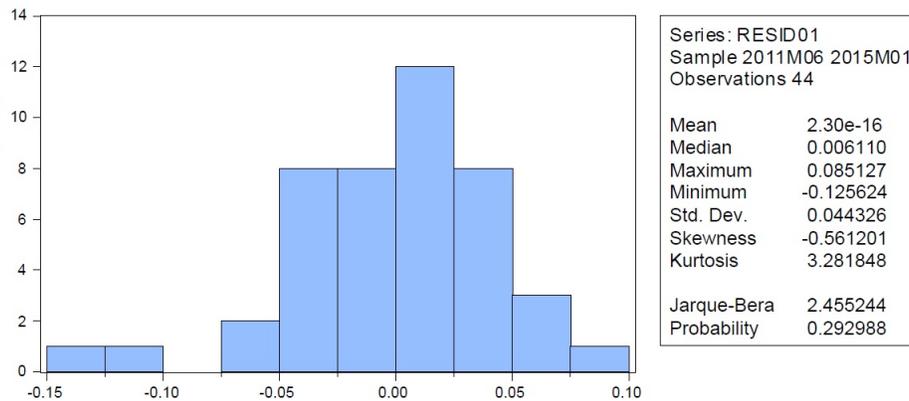
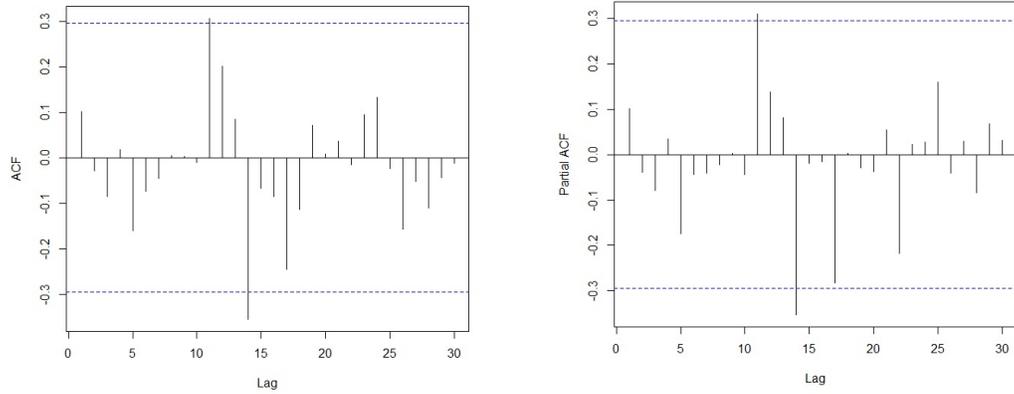


Figura 4.19: Histograma de los residuos de la regresión de la PNC contra el ICVT.

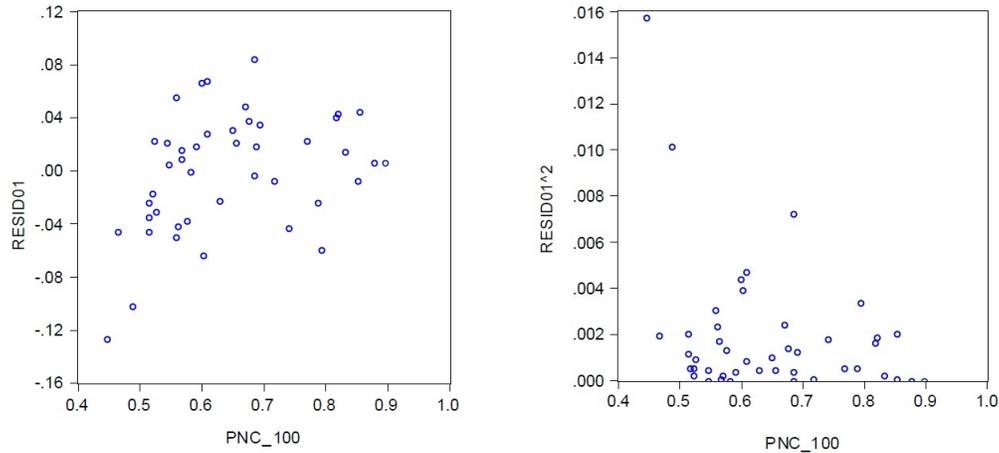
En la figura 4.20 se muestra el correlograma de los residuos de la regresión de la PNC contra el ICVT. Se observa que los rezagos 11 y 14 están ligeramente fuera de las bandas, lo que parece indicar la posible presencia de correlación serial. De otro lado, en la gráfica de los residuos y del cuadrado de los residuos contra la pérdida neta de crédito estimada, mostrada en la figura 4.21, se observa que no hay un patrón claramente definido entre las variables, por consiguiente, es posible que no exista un problema de heteroscedasticidad.



(a) Función de autocorrelación de los residuos.

(b) Función de autocorrelación parcial de los residuos.

Figura 4.20: Correlograma de los residuos de la regresión de la PNC contra el ICVT.



(a) Residuos vs. PNC.

(b) Cuadrado de los residuos vs. PNC.

Figura 4.21: Gráfico de los residuos y el cuadrado de los residuos contra la PNC estimada.

En la tabla 4.16 se resumen los resultados de las pruebas de diagnóstico para los residuos de la regresión entre el ICVT y la PNC. Los p-valores en la tabla 4.16 se compararon con el nivel de significación del 0.05. El estadístico Jarque-Bera estimado para la regresión de la PNC y el ICVT es 2.45, que es menor a 5.99 y se encuentra en la región de no rechazo, por consiguiente, la hipótesis nula de

que los residuos están normalmente distribuidos no puede rechazarse. Además, la prueba de Shapiro-Wilk proporcionó un p-valor de 0.4029, que es mayor al nivel de significación  $\alpha$ , y entonces se concluye que los residuos se distribuyen normalmente.

Prueba	Estadístico	p-valor	Conclusión
Jarque-Bera	2.4564	0.2928	Normalmente distribuidos
Shapiro-Wilk	0.9736	0.4029	Normalmente distribuidos
Ljung-Box	29.412	0.205	No hay correlación serial
White	1.5958	0.2065	No hay heteroscedasticidad

Tabla 4.16: Resumen de las pruebas de diagnóstico de los residuos de la regresión del ICVT contra el PNC.

La prueba de autocorrelación de Ljung-Box calculada con 24 rezagos suministró un p-valor de 0.205, que está por arriba del 5 %, por tanto, no se rechaza la hipótesis nula de autocorrelaciones independientes y se puede concluir que los residuos son ruido blanco. Al realizar la prueba de White se encontró que el estadístico  $\chi^2$  es 1.5958 con un p-valor de 0.2065, en este la probabilidad es mayor que el nivel de significación del 5 % y, por lo tanto, no se rechaza la hipótesis nula, por lo que la varianza es constante y homocedástica. Por último, en la figura 4.22, se aprecia que la pérdida neta de crédito estimada con el modelo de regresión, dado en la ecuación 4.1, describe adecuadamente el comportamiento de la serie.

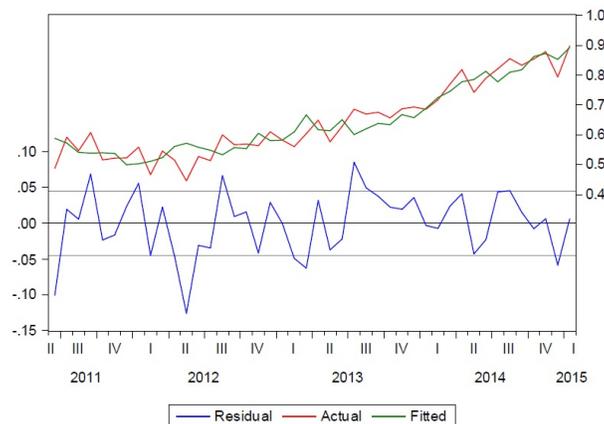
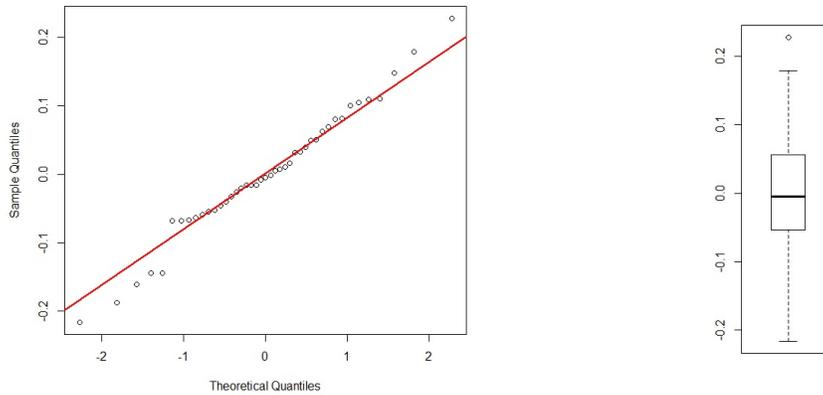


Figura 4.22: Gráfico de la serie pérdida neta de crédito y pérdida neta de crédito estimada.



(a) Gráfico de probabilidad de los residuos.

(b) Boxplot de los residuos.

Figura 4.23: Gráfico de probabilidad normal y boxplot de los residuos de la regresión del CoCre contra el PNC.

El gráfico de probabilidad normal de los residuos de la regresión del CoCre contra el PNC, mostrado en la figura 4.23(a), indica que la mayoría de los datos se ubican sobre la recta y al observar el boxplot de los residuos, mostrado en la figura 4.23(b), se encuentra que hay un punto atípico, la mediana está aproximadamente en el centro de la caja y el bigote inferior es levemente más largo que el superior. Por consiguiente, los residuos posiblemente se distribuyen normalmente.

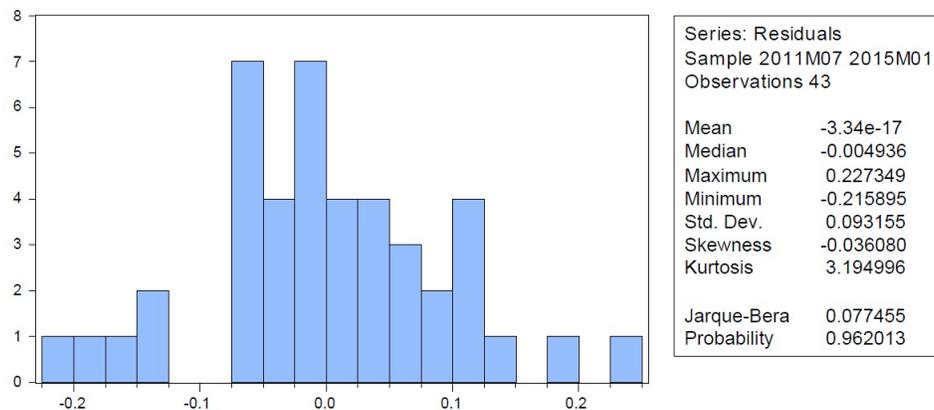
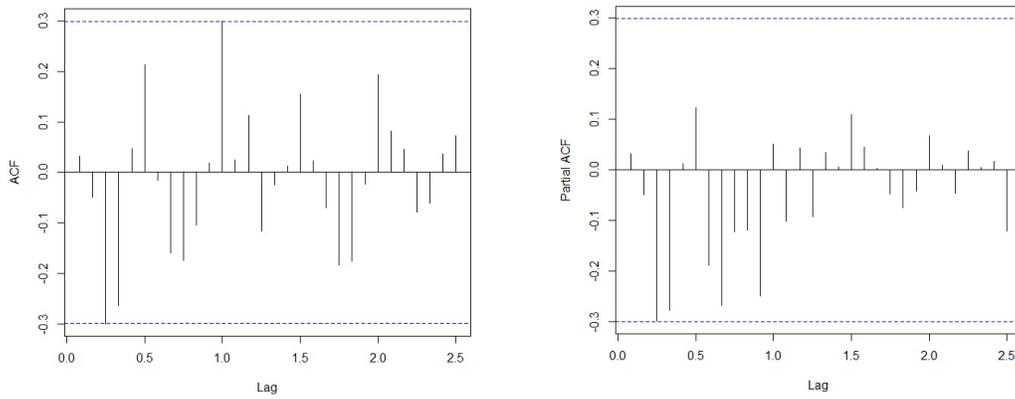


Figura 4.24: Histograma de los residuos de la regresión del CoCre contra la PNC.

Se determinó que el coeficiente de asimetría tiende a cero y la curtosis es cercana a tres, como se indica en los resultados en la figura 4.24, lo que da más indicios de

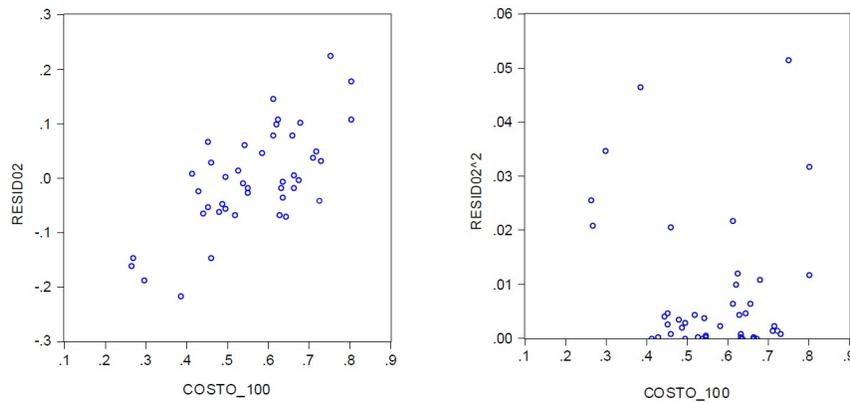
que los residuos tiene una distribución normal. Por otra parte, se encontró que los residuos de la regresión del CoCre contra la PNC no están correlacionados, como se puede observar en el correlograma mostrado en la figura 4.24, donde todos los rezagos se encuentran dentro de las bandas.



(a) Función de autocorrelación de los residuos.

(b) Función de autocorrelación parcial de los residuos.

Figura 4.25: Correlograma de los residuos de la regresión del CoCre contra el PNC.



(a) Residuos vs. Cocre.

(b) Cuadrado de los residuos vs. Co-Cre.

Figura 4.26: Gráfico de los residuos y el cuadrado de los residuos contra el CoCre.

En la figura 4.26 se muestran los residuos y el cuadrado de los residuos contra el costo de crédito estimado. Como se observa para valores altos del costo de crédito también se presentan valores altos en los residuos y se aprecia un patrón definido entre las variables, lo que sugiere la presencia de un posible problema de varianza no constante ni homocedastica.

Prueba	Estadístico	p-valor	Conclusión
Jarque-Bera	0.07746	0.9620	Normalmente distribuidos
Shapiro-Wilk	0.9835	0.7842	Normalmente distribuidos
Ljung-Box	33.846	0.0874	No hay correlación serial
White	7.4353	0.0243	Hay heteroscedasticidad

Tabla 4.17: Resumen de las pruebas de diagnóstico de los residuos de la regresión del CoCre contra el PNC.

En la tabla 4.17 se resumen los resultados de las pruebas de diagnóstico para los residuos de la regresión entre el CoCre y la PNC. Los p-valores en la tabla 4.17 se compararon con el nivel de significación del 0.05. La prueba de Jarque-Bera proporcionó un p-valor de 0.9620, por consiguiente, existe una probabilidad de 96.20 %, que es bastante alta, de no rechazar la hipótesis nula de normalidad, mientras que, la prueba de Shapiro-Wilk arrojó un p-valor de 0.7842, que es mayor al nivel de significación  $\alpha$ . Por tanto, se concluye, al nivel de significación del 5 %, que no hay evidencia suficiente para rechazar la hipótesis nula de normalidad de los residuos.

La prueba de autocorrelación de Ljung-Box suministró un p-valor de 0.087, que es mayor al nivel de significación  $\alpha$ , por tanto, no se rechaza la hipótesis y se puede concluir que los residuos son ruido blanco. Al realizar la prueba de White se encuentra que el estadístico  $\chi^2$  es 7.4353 con un p-valor de 0.0243, en este caso la probabilidad es menor que el nivel de significación del 5 % y, por tanto, se rechaza la hipótesis nula, por lo que hay suficientes indicios que confirman la existencia de un problema de varianza no constante ni homocedastica. Finalmente, en la figura 4.27, se muestra la comparación entre el costo de crédito y el costo de crédito estimado a partir de la regresión 4.2. Como se observa el modelo logra capturar el comportamiento de la serie aunque presenta algunos problemas en las regiones donde la serie es más volátil.

---

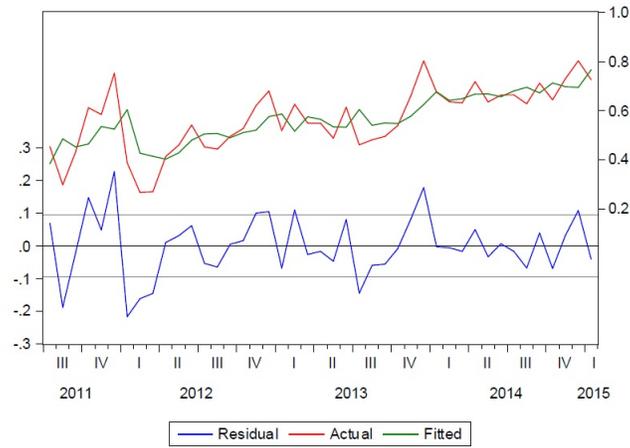


Figura 4.27: Gráfico de la serie costo de crédito y costo de crédito estimada.

## 4.6. Modelo de corrección de error

Con los resultados obtenidos mediante la estimación por MCO del modelo de corrección de error (MCE), indicados en la tabla 4.18, se obtuvo la siguiente ecuación en primeras diferencias:

$$\Delta PNC_t = 0.070297\Delta ICVT_{t-5} - 0.660736\epsilon_{t-1} \quad (4.3)$$

Se observa que el signo negativo asociado al término MCE es consistente con lo predicho con la teoría, ya que si la variable se encuentra por encima (debajo) de su valor de largo plazo debe esperarse un descenso (aumento) para que converja hacia dicho valor. En relación con el coeficiente del término MCE, éste es estadísticamente significativo y se puede interpretar que mensualmente se produce una corrección, de alrededor del 66 %, en la desviación del  $PNC$  respecto a su nivel de equilibrio de largo plazo.

Se aprecia que existe una relación positiva de corto plazo entre la variación intermensual de la pérdida neta de crédito ( $\Delta PNC$ ) y la variación del indicador de cartera vencida total ( $\Delta ICVT$ ), siendo esta relación estadísticamente significativa para el quinto rezago. Este resultado indica que los efectos de la variación en el ICVT sobre el comportamiento del PNC desaparecen hasta cinco meses después de ocurrida.

Variable	Coefficiente	Desviación estándar	t	P-valor
DICVT(-5)	0.070297	0.023900	2.941282	0.0057
RESID01(-1)	-0.660736	0.142565	-4.634647	0.0000
$r^2$	0.438471	Estadístico Durbin-Watson		2.307339

Tabla 4.18: Resultado de la regresión por MCO del MCE para la relación entre la PNC y el ICVT.

En la tabla 4.19 se indican los resultados de la estimación por MCO del modelo de corrección de error (MCE) para la relación entre el costo de crédito (CoCre) y la pérdida neta de crédito (PNC), de donde, se determinó el siguiente modelo:

$$\Delta CoCre_t = 1.363991\Delta CoCre_{t-1} - 0.373741\Delta CoCre_{t-3} + 0.755215\Delta PNC_{t-3} - 0.893940\epsilon_{t-1} \quad (4.4)$$

Se aprecia que el coeficiente del término del MCE es estadísticamente significativo y presenta el signo negativo correcto, siendo esto último consistente con la teoría, ya que actúa para disminuir el desequilibrio en el próximo periodo, en este caso, mensualmente. Además, la desviación del CoCre respecto a su valor de equilibrio de largo plazo se corrige mensualmente en un 89 % aproximadamente.

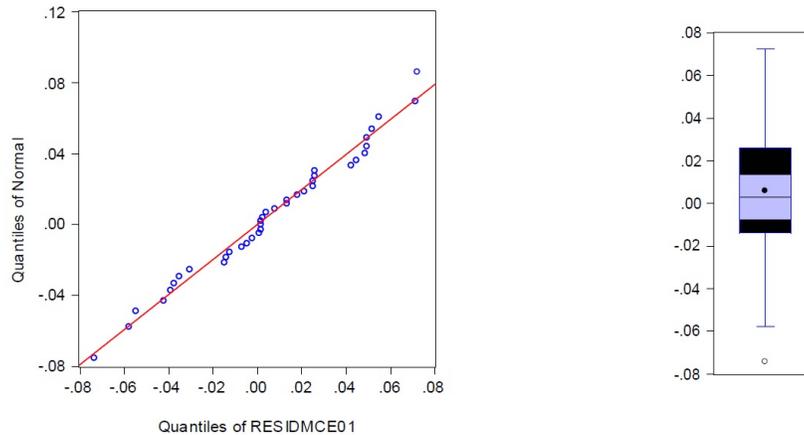
Variable	Coefficiente	Desviación estándar	t	P-valor
DCOSTO(-1)	1.363991	0.150923	9.037645	0.0000
DCOSTO(-3)	-0.373741	0.153199	-2.439577	0.0198
DPNC(-3)	0.755215	0.283425	2.664602	0.0115
RESID02(-1)	-0.893940	0.234251	-3.816169	0.0005
$r^2$	0.573966	Estadístico Durbin-Watson		2.092162

Tabla 4.19: Resultado de la regresión por MCO del MCE para la relación entre el CoCre y la PNC.

Se observa que existe una relación directa significativa de corto plazo entre la variación en el costo de crédito (CoCre) y la variación en el costo de crédito rezagado un periodo. Además, los resultados de la prueba de cointegración proporcionaron evidencia de una relación estable y significativa de largo plazo entre el CoCre y el CoCre retardado un periodo. Por tanto, desde el punto de vista empírico, parece ser que son importantes los efectos de la variación del costo de crédito sobre su propia dinámica. También se aprecia que los impactos de las variaciones del CoCre

y la PNC sobre la dinámica del CoCre persisten hasta tres meses después de ocurridas. Resulta interesante que este periodo de tres meses coincide con el periodo de tiempo para el cual se considera que un cliente de un crédito de consumo incurre en incumplimiento de acuerdo a la normativa establecida por la Superintendencia Financiera de Colombia.

## 4.7. Pruebas de diagnóstico de los residuos del MCE



(a) Gráfico de probabilidad de los residuos.

(b) Boxplot de los residuos.

Figura 4.28: Gráfico de probabilidad normal y boxplot de los residuos de la regresión por MCO con MCE para la relación entre el ICVT y la PNC

En las figuras 4.28(a) y 4.28(b) se muestran, respectivamente, el gráfico de probabilidad normal y el boxplot de los residuos del modelo de corrección de error para la pérdida neta de crédito. En el boxplot la media está levemente desplazada hacia el primer cuartil pero los bigotes son bastante simétricos y se observa un dato atípico, mientras que, en el gráfico de probabilidad normal se observa que la mayoría de los datos se localizan sobre la recta. Se encontró que la kurtosis es menor a tres y que el coeficiente de asimetría es  $-0.18$ , como se indica en la figura 4.29, por consiguiente, hay un leve sesgo a la izquierda y la distribución no es completamente normal.

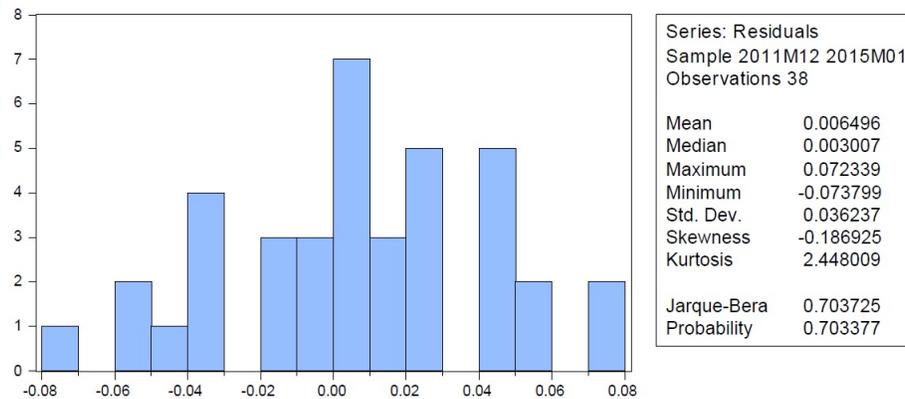


Figura 4.29: Histograma de los residuos de la regresión por MCO con MCE para la relación entre el ICVT y la PNC.

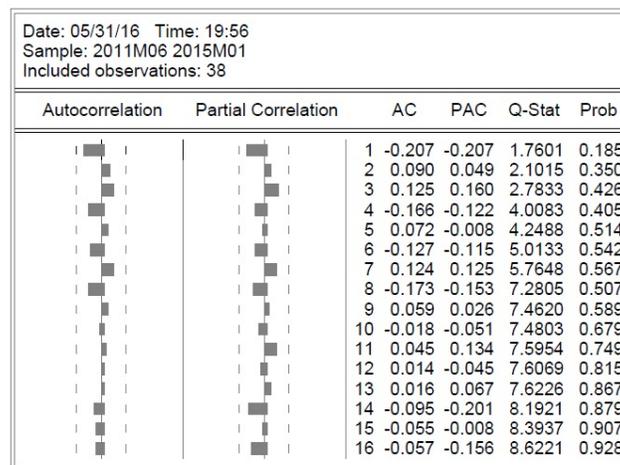


Figura 4.30: Correlograma de los residuos de la regresión por MCO con MCE para la relación entre el el ICVT y la PNC.

Se estableció que los residuos del modelo de corrección de error para la relación entre la pérdida neta de crédito y el índice de cartera vencida total son ruido blanco, como se puede apreciar en el correlograma de la figura 4.30, donde los rezagos se encuentran dentro de la bandas y todas las probabilidades ligadas a los estadísticos de Ljung-Box están por encima del 5 %. Al observar la figura 4.31 se aprecia la posible existencia de un patrón definido entre las variables lo que sugiere la presencia de un problema de heteroscedasticidad.

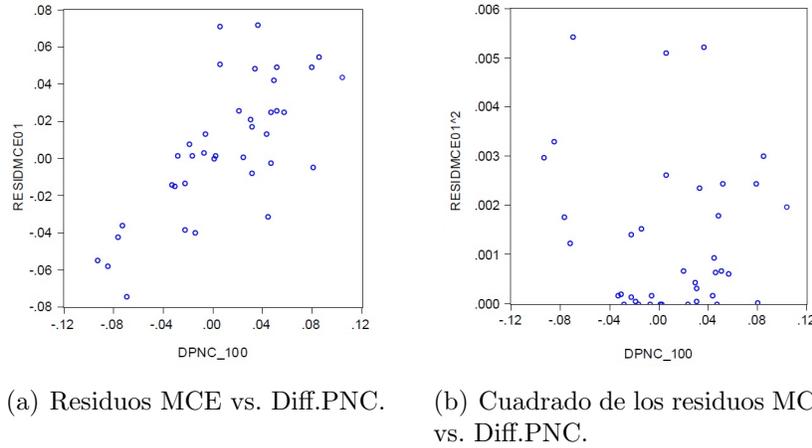


Figura 4.31: Gráfico de los residuos y el cuadrado de los residuos del MCE contra la primera diferenciación del PNC.

Prueba	Estadístico	p-valor	Conclusión
Jarque-Bera	0.7035	0.7034	Normalmente distribuidos
Ljung-Box	13.977	0.9472	No hay correlación serial
White	3.0257	0.2203	No hay heteroscedasticidad

Tabla 4.20: Resumen de las pruebas de diagnóstico de los residuos del MCE para la relación entre el ICVT y la PNC.

En la tabla 4.20 se resumen los resultados de las pruebas de diagnóstico para los residuos del modelo de corrección de error (MCE) para la relación entre el ICVT y la PNC. Los p-valores en la tabla se compararon con el nivel de significación del  $\alpha = 5\%$ . El estadístico Jarque-Bera es 0.7035, que es menor a 5.99 y se encuentra en la región de no rechazo, por tanto, no puede rechazarse la hipótesis nula de que los residuos se distribuyen normalmente. La prueba de White permitió determinar un p-valor de 0.2203, Como se observa, la probabilidad es mayor que el nivel de significación  $\alpha$  y, por tanto, no se rechaza la hipótesis nula, por lo que la varianza es constante y homocedastica.

En la figura 4.32 se observa en el boxplot que la mediana se encuentra en el centro de la caja, los bigotes son simétricos y hay un punto atípico. Adicionalmente, en el gráfico de probabilidad normal se aprecia que una recta se ajusta bien a los datos indicando que los residuos parecen distribuirse normalmente.

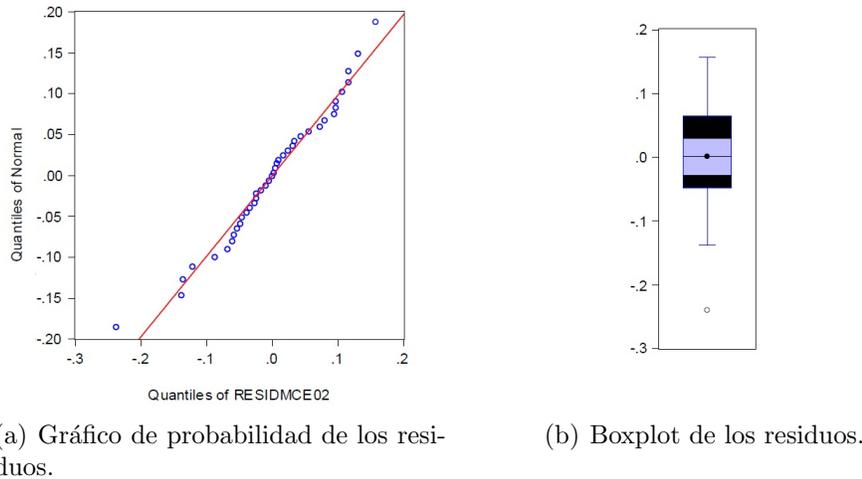


Figura 4.32: Gráfico de probabilidad normal y boxplot de los residuos de la regresión por MCO con MCE para la relación entre el CoCre y la PNC.

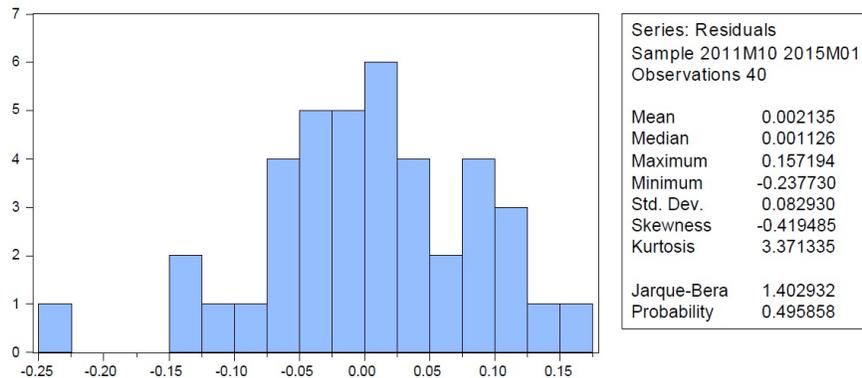


Figura 4.33: Histograma de los residuos de la regresión por MCO con MCE para la relación entre el CoCre y la PNC.

El correlograma de los residuos del modelo de corrección de error para la relación entre el CoCre y la PNC, mostrado en la figura 4.34, indica que todos los p-valores son mayores al 5 %, por consiguiente, los residuos son ruido blanco. El gráfico de los residuos al cuadrado del modelo de corrección de error, mostrado en la figura 4.35 parece indicar la existencia de una tendencia y de un posible problema de heterocedasticidad.

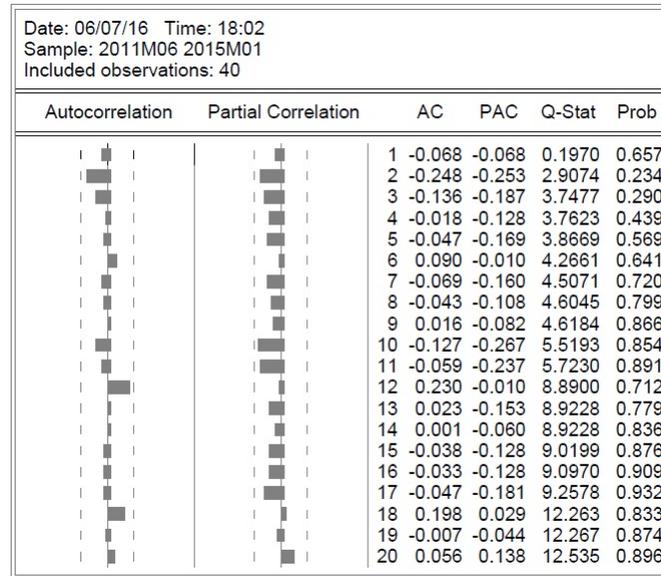


Figura 4.34: Correlograma de los residuos de la regresión por MCO con MCE para la relación entre el CoCre y la PNC.

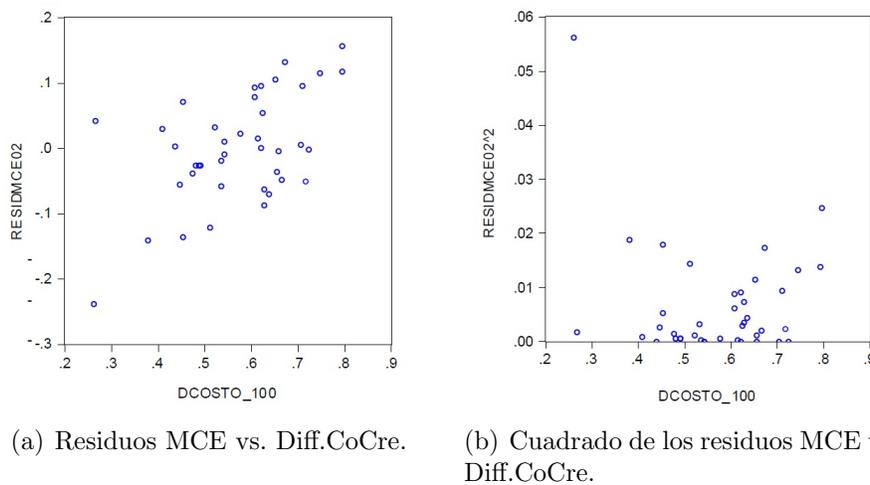


Figura 4.35: Gráfico de los residuos y el cuadrado de los residuos de MCE contra la primera diferenciación del CoCre.

En la tabla 4.21 se resumen los resultados de las pruebas de diagnóstico para los residuos del modelo de corrección de error (MCE) para la relación entre el CoCre

y la PNC. Los p-valores en la tabla se compararon con el nivel de significación del  $\alpha = 5\%$ . Se confirma con la prueba de JB y con un nivel de significación del 5 %, que no hay suficiente evidencia para rechazar la hipótesis nula de normalidad de los residuos. La prueba de White, que proporciona un p-valor de 0.0286, que es menor a  $\alpha$  y, por tanto, hay suficiente evidencia para ratificar la existencia de un problema de varianza no constante ni homocedastica.

Prueba	Estadístico	p-valor	Conclusión
Jarque-Bera	1.4029	0.4958	Normalmente distribuidos
Ljung-Box	22.529	0.548	No hay correlación serial
White	10.825	0.0286	Hay heteroscedasticidad

Tabla 4.21: Resumen de las pruebas de diagnóstico de los residuos del MCE para la relación entre el CoCre y la PNC.

## 4.8. Relación entre el CoCre y el ROA

PyG	
Ingreso	21.43 %
Costo de fondos	4.87 %
Utilidad Financiera	16.56 %
CoCre	7.68 %
Utilidad después de CoCre	8.88 %
Otros Ingresos	0.97 %
Gastos	6.38 %
Utilidad Antes de Impuestos	3.42
Impuestos	1.30 %
ROA	2.17 %

Tabla 4.22: Estado de resultados publicado por la vicepresidencia financiera de la entidad bancaria para el año 2015.

Para establecer la relación entre el costo de crédito (CoCre) y la rentabilidad sobre los activos (ROA) para una determinada calidad de cartera, se consideró un estado de resultados o estado de pérdidas y ganancias (PyG) para un periodo de un año, comprendido entre enero y diciembre de 2015, para un producto de

crédito de consumo de libre inversión otorgado por una entidad bancaria de Bogotá, Colombia. El estado de resultados se muestra en la tabla 4.22. El valor de todos los rubros es el promedio anual extraído del estado de pérdidas y ganancias publicado por la vicepresidencia financiera de la entidad bancaria.



Figura 4.36: PNC observado y pronosticado con intervalos de confianza del 95 %.



Figura 4.37: CoCre observado y pronosticado con intervalos de confianza del 95 %.

En la figura 4.36 se muestra la serie original del PNC y los resultados obtenidos, con el modelo planteado en la ecuación 4.1, para el pronóstico del PNC con intervalos de confianza del 95 %. El pronóstico se realizó para el periodo comprendido entre febrero de 2015 y enero de 2016. Como se observa, para el periodo pronosticado, indicado por la línea vertical en la figura 4.36, la serie original del PNC permanece dentro de los límites del intervalo de confianza y la serie pronosticada no logra describir completamente el comportamiento de la serie original del PNC, debido posiblemente a la dinámica particular que exhibió el ICVT durante el año 2015.

Con el modelo propuesto en la ecuación 4.2 se determinaron los intervalos de confianza del 95 % para el CoCre. El pronóstico se realizó para el periodo comprendido entre febrero de 2015 y enero de 2016. El CoCre es un indicador muy volátil por que es muy sensible a factores como la calificación del crédito, sin embargo, el comportamiento del estado de resultados es bastante estable, por tanto, para el término asociado al costo de crédito rezagado un periodo ( $\text{CoCre}_{t-1}$ ) se eligió el valor promedio anual del CoCre reportado en el PyG. Los resultados del pronóstico se indican con la línea vertical en la figura 4.37. Se observa que la serie pronosticada del CoCre trata de seguir el comportamiento de la serie original y la serie permanece dentro de los límites del intervalo de confianza, sin embargo, la serie pronosticada no se ajusta bien en los periodos donde la serie original del CoCre es más volátil.

Distribución	Estadístico AD	P-valor	Distribución	Estadístico AD	P-valor
Gamma	1.2639	0.000	Extremo mínimo	2.4259	0.000
Uniforme	1.4602	0.143	Pareto	2.4453	-
Logarítmico nor	1.4629	0.000	Triangular	4.1861	-
Extremo máximo	1.8484	0.000	t de Student	4.6307	-
Weibull	1.9006	0.027	Beta	14.2897	-
Logística	2.0477	0.000	Exponencial	18.3616	0.000
Normal	2.1523	0.000	Beta PERT	45.9249	-

Tabla 4.23: Resultados de la prueba de bondad de ajuste de Anderson-Darling para los datos del ICVT.

Por medio del software Oracle Crystal Ball se realizó la prueba de bondad de ajuste de Anderson-Darling a los datos del indicador de cartera vencida total. Se utilizó la información del ICVT para el periodo comprendido entre junio de 2011 y enero de 2016, que constituye una muestra de 56 observaciones. La prueba se

efectuó para diferentes distribuciones de probabilidad y se obtuvo que el valor más pequeño para el estadístico AD fue 1.2369 con un p-valor de 0.000, como se indica en la tabla 4.23. Por tanto, se concluye con un nivel de significación del 5 % que los datos del ICVT se ajustan a una distribución gamma. Con base en ésta distribución de probabilidad para el ICVT y utilizando el modelo propuesto en las ecuaciones 4.1 y 4.2, se realizó una simulación de Monte Carlo con 100000 iteraciones para pronosticar el CoCre y el ROA.

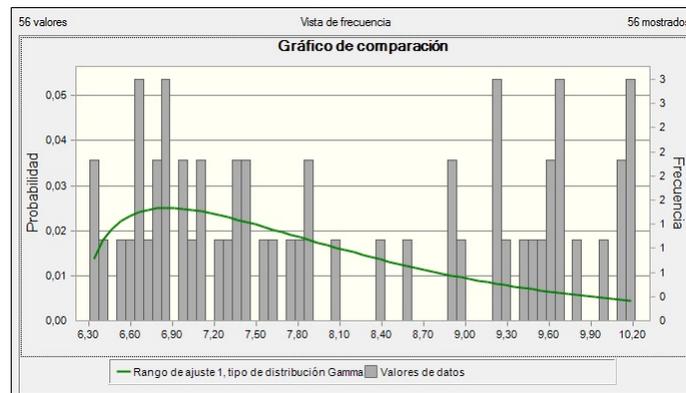


Figura 4.38: Datos del ICVT ajustados a una distribución gamma según el criterio de Anderson-Darling.

Mínimo	Mediana	Media	Máximo	Sesgo	Curtosis	Desviación
6.27	7.73	8.13	Infinito	1.65	7.08	1.54

Tabla 4.24: Resumen estadística descriptiva del ajuste a una distribución gamma de los datos del ICVT.

El cuadro de pronóstico para el CoCre, mostrado en la figura 4.39, nos indica que el valor mensual del CoCre varía entre 0.60 % y 0.82 % con un nivel de certidumbre del 95 %, que convertido a un valor efectivo anual, corresponde a una variación entre 7.44 % y 10.30 % con un nivel de certidumbre del 95 %. De acuerdo al PyG reportado por la vicepresidencia financiera de la entidad bancaria para el año 2015, indicado en la tabla 4.22, el CoCre observado fue 7.68 %, que se encuentra dentro del rango del intervalo.

Al analizar el cuadro de pronóstico para el ROA, mostrado en la figura 4.40, se observa que el rango de valores entre los indicadores de certidumbre muestra un

nivel de certidumbre del 92.20 %. Este resultado nos indica que la entidad bancaria tiene un 92.20 % de posibilidades de obtener utilidades a partir de su inversión en el producto de crédito de consumo de libre inversión, por tanto, existe una posibilidad de 7.8 % de experimentar pérdidas netas.

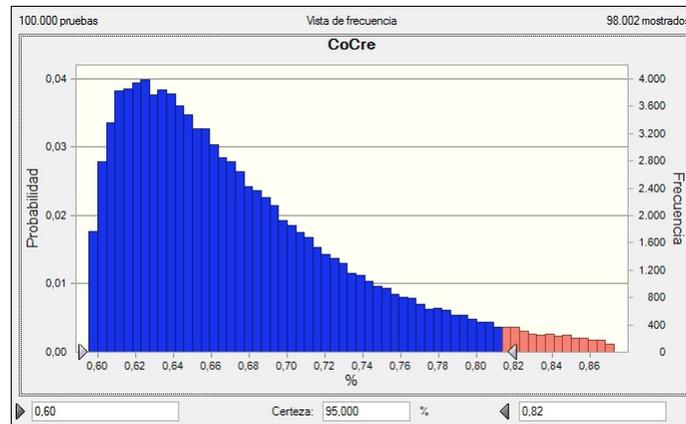


Figura 4.39: Pronóstico del CoCre para crédito de consumo de libre inversión.

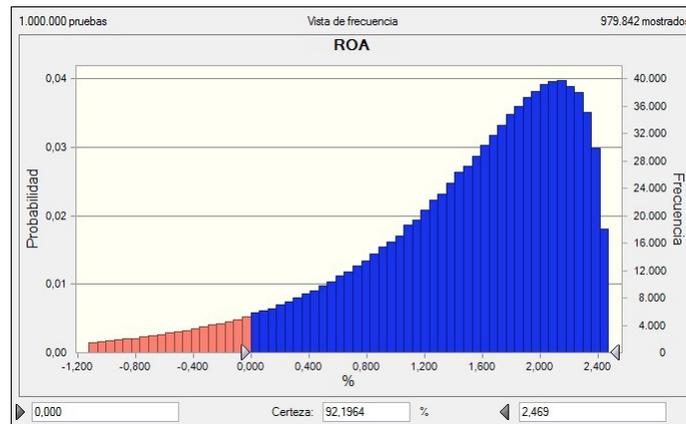


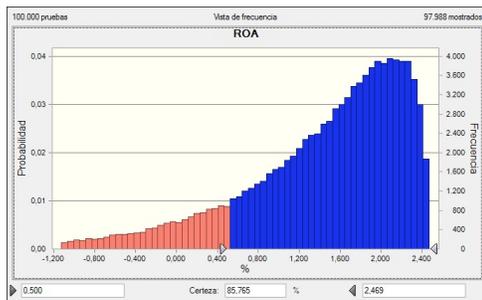
Figura 4.40: Pronóstico del ROA para crédito de consumo de libre inversión.

Además, los resultados de la simulación de Monte Carlo permiten establecer cuales son las certidumbres de que la entidad bancaria obtenga utilidades mínimas, en un rango entre 0.5 % y 2.0 %, como resultado de su inversión en el producto de

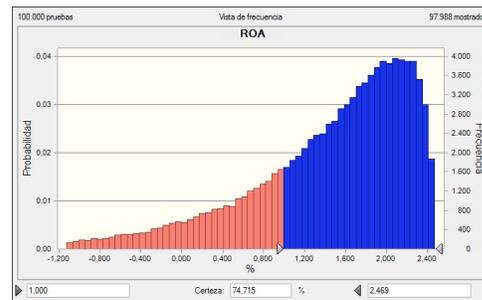
crédito de consumo de libre inversión, como se indica en la tabla 4.25 y en la figura 4.41. Como se puede observar, el producto de crédito de consumo de libre inversión resulta rentable para la entidad bancaria, ya que hay entre 85.74 % y 74.62 % de probabilidad de obtener utilidades mínimas entre 0.5 % y 1.0 %. Por otra parte, los resultados de la simulación nos indican que el ROA de 2.17 %, reportado por la vicepresidencia financiera de la entidad bancaria para el año 2015, como se indica en la tabla 4.22, se encuentra aproximadamente en el rango entre 0.19 % y 2.47 %, con un nivel de certidumbre del 90 %.

ROA	0.5 %	1.0 %	1.5 %	2.0 %
Nivel de certidumbre	85.74 %	74.62 %	56.11 %	27.77 %

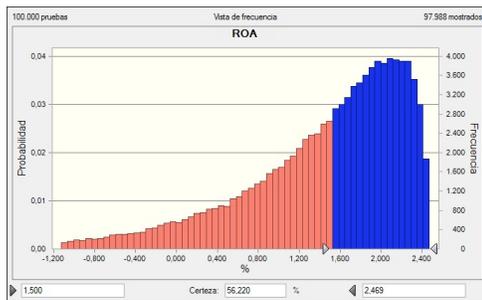
Tabla 4.25: Nivel de certidumbre del ROA para utilidades mínimas entre 0.5 % y 2.0 %.



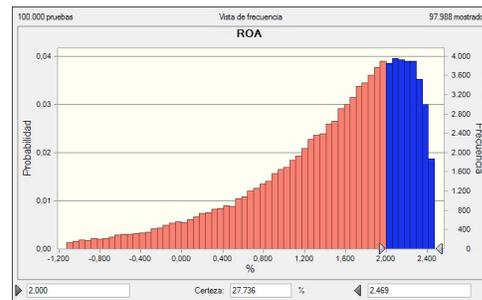
(a) Pronóstico ROA - mínimo 0.5 %.



(b) Pronóstico ROA - mínimo 1.0 %..



(c) Pronóstico ROA - mínimo 1.5 %..



(d) Pronóstico ROA - mínimo 2.0 %..

Figura 4.41: Pronóstico del ROA para utilidades mínimas entre 0.5 % y 2.0 %.

# Capítulo 5

## Conclusiones

La prueba de raíz unitaria de Dickey-Fuller aumentada aplicada a las series indicador de cartera vencida total (ICVT), pérdida neta de crédito (PNC) y costo de cartera o crédito (CoCre), permitió establecer que éstas series son no estacionarias, cada una con una raíz unitaria, es decir, son integradas de orden uno. Por consiguiente, es posible la existencia de relaciones de equilibrio a largo plazo entre estas series.

La aplicación de la prueba de cointegración de Engle-Granger a los residuos de la regresión entre la pérdida neta de crédito (PNC) y el indicador de cartera vencida total (ICVT) proporcionó suficiente evidencia estadística para rechazar la hipótesis nula de que los residuos de la regresión tienen una raíz unitaria. Por consiguiente, existe una relación de equilibrio a largo plazo entre las series pérdida neta de crédito (PNC) e indicador de cartera vencida total (ICVT).

Mediante la aplicación de mínimos cuadrados ordinarios se estableció la siguiente relación de largo plazo entre la serie pérdida neta de crédito (PNC) y la serie indicador de cartera vencida total (ICVT):

$$PNC_t = -0.173051 + 0.106678ICVT_t$$

Se encontró que el coeficiente de pendiente es estadísticamente significativo y, por tanto, el indicador de cartera vencida total sí afecta a la pérdida neta de crédito.

La prueba de cointegración de Engle-Granger efectuada a los residuos de la regresión entre el costo de crédito (CoCre) y la pérdida neta de crédito (PNC) suministró suficiente evidencia estadística para rechazar la hipótesis nula de que los residuos de la regresión son no estacionarios. Por tanto, las series costo de crédito (CoCre) y pérdida neta de crédito (PNC) están cointegradas.

Se determinó, mediante mínimos cuadrados ordinarios, la siguiente relación de largo plazo entre el costo de crédito (CoCre) y la pérdida neta de crédito (PNC):

$$CoCre_t = -0.082806 + 0.421377PNC_t + 0.379426CoCre_{t-1}$$

Se estableció que todos los coeficientes de pendiente son estadísticamente significativos, por tanto, se concluye que la pérdida neta de crédito (PNC) y el costo de crédito (CoCre) retardado un periodo si tienen influencia sobre el comportamiento del costo de crédito.

La estimación por mínimos cuadrados ordinarios del modelo de corrección de error para la relación entre la pérdida neta de crédito (PNC) y el indicador de cartera vencida total (ICVT) permitió establecer que existe una relación positiva de corto plazo entre la variación intermensual de la PNC y la variación de ICVT, siendo esta relación estadísticamente significativa para el quinto rezago.

La estimación por mínimos cuadrados ordinarios del modelo de corrección de error para la relación entre el costo de crédito (CoCre) y la pérdida neta de crédito (PNC) proporcionó evidencia estadística de que hay una relación directa significativa de corto plazo entre la variación en el CoCre y la variación en el CoCre rezagado un periodo. Además, los resultados de la prueba de cointegración proporcionaron evidencia de una relación estable y significativa de largo plazo entre el CoCre y el CoCre retardado un periodo. Por tanto, desde el punto de vista empírico, parece ser que son importantes los efectos de la variación del costo de crédito sobre su propia dinámica. También se aprecia que los impactos de las variaciones del CoCre y la PNC sobre la dinámica del CoCre persisten hasta tres meses después de ocurridas.

Con los datos históricos del indicador de cartera vencida total (ICVT) y con el modelo estadístico propuesto se realizó una simulación de Monte Carlo que per-

---

mitió pronosticar el costo de cartera (CoCre) y el índice de rentabilidad sobre los activos (ROA). A partir de los resultados obtenidos se logro establecer que:

- el valor efectivo anual del CoCre de 7.68 %, reportado en el PyG por la vicepresidencia financiera de la entidad bancaria para el año 2015, se encuentra en el rango del intervalo entre 7.44 % y 10.30 % con un nivel de certidumbre del 95 %.
  - el ROA de 2.17 %, reportado en el PyG por la vicepresidencia financiera de la entidad bancaria para el año 2015, se encuentra aproximadamente en el rango entre 0.19 % y 2.47 %, con un nivel de certidumbre del 90 %.
  - la entidad bancaria tiene un 92.20 % de posibilidades de obtener utilidades a partir de su inversión en el producto de crédito de consumo de libre inversión.
  - el producto de crédito de consumo de libre inversión resulta muy rentable para la entidad bancaria, si se desean obtener utilidades mínimas entre 0.5 % y 1.0 %, ya que la probabilidad de obtener estos rendimientos varía entre 85.74 % y 74.62 %.
-

# Apéndice A

## Estadística descriptiva de las series temporales en R

- Cargar las librerías para el análisis de series de tiempo

```
>library(TSA)
>library(faraway)
>library(tseries)
```
- Importar las bases de datos

```
>icvt<-read.table(file="C:\\Users\\Documents\\Estadistica
aplicada\\Calculos\\ICVT.txt",header=FALSE)
>pnc<-read.table(file="C:\\Users\\Documents\\Estadistica
aplicada\\Calculos\\PNC.txt",header=FALSE)
>cocre<-read.table(file="C:\\Users\\Documents\\Estadistica
aplicada\\Calculos\\COCRE.txt",header=FALSE)
```
- Convertir las bases de datos a formato de serie de tiempo

```
> icvt.t <- ts(icvt, start = c(2011, 6), freq = 12)
> pnc.t <- ts(pnc, start = c(2011, 6), freq = 12)
> cocre.t <- ts(cocre, start = c(2011, 6), freq = 12)
```
- Combinar la información de las series temporales ICVT, PNC y CoCre

```
> datos.consumo <- cbind(icvt.t, pnc.t, cocre.t)
```
- Realizar el resumen de la estadística descriptiva

```
> summary(datos.consumo)
```

---

En R-studio cargar la libreria:> library(Rcdmr) y ejecutar la orden  
>numSummary(datos.consumo,statistics=c("skewness","kurtosis",  
"sd"),type="2")

- Realizar las pruebas gráficas de normalidad de las series temporales  
>hist(icvt.t,col="blue",ylab="Frecuencia",xlab="Indice de  
cartera vencida total")  
>boxplot(icvt.t,col="blue")  
>hist(pnc.t,col="red",ylab="Frecuencia",xlab="Pérdida neta de  
crédito")  
>boxplot(pnc.t,col="red")  
>hist(cocre.t,col="green",ylab="Frecuencia",xlab="Costo de  
crédito")  
>boxplot(cocre.t,col="green")
  - Efectuar la prueba de normalidad de Shapiro-Wilk  
>shapiro.test(icvt.t)  
>shapiro.test(pnc.t)  
>shapiro.test(cocre.t)
  - Realizar las gráficas de las series de tiempo y sus primeras diferencias  
>plot(icvt.t,xlab="Año",ylab="Indice de cartera vencida total"  
,type="l",col="blue")  
> points(y = icvt.t,x = time(icvt.t),pch = as.vector(season(icvt.t)))  
>plot(diff(icvt.t),xlab="Año",ylab="Primera diferencia del  
ICVT")  
  
>plot(pnc.t,xlab="Año",ylab="Pérdida neta de crédito",type="l"  
,col="blue")  
> points(y = pnc.t,x = time(pnc.t),pch = as.vector(season(pnc.t)))  
>plot(diff(pnc.t),xlab="Año",ylab="Primera diferencia de la  
PNC")  
  
>plot(cocre.t,xlab="Año",ylab="Costo de crédito",type="l",col  
="blue")  
>points(y=cocre.t,x=time(cocre.t),pch=as.vector(season(  
cocre.t)))
-

```
>plot(diff(cocre.t),xlab="Año",ylab="Primera diferencia del  
CoCre")
```

- Realizar las gráficas de las funciones de autocorrelación y autocorrelación parcial

```
>acf(icvt.t,lag=50,main="Función de autocorrelación ICVT")  
>pacf(icvt.t,main="Función de autocorrelación parcial ICVT")
```

```
>acf(pnc.t,lag=50,main="Función de autocorrelación PNC")  
>pacf(pnc.t,main="Función de autocorrelación parcial PNC")
```

```
>acf(cocre.t,lag=50,main="Función de autocorrelación CoCre")  
>pacf(cocre.t,main="Función de autocorrelación parcial CoCre")
```

---

## Apéndice B

# Prueba de raíz unitaria de Dickey-Fuller aumentada en EViews 8

Para realizar la prueba de raíz unitaria de Dickey-Fuller Aumentada se debe realizar el siguiente procedimiento:

- Abrir la base de datos de la serie  $ICVT_{100}$  en el archivo de trabajo o Workfile.
- En la ventana de la serie hacer clic en el menú **View** y elegir la opción **Unit Root Test**.
- Correr la prueba ADF a nivel con intercepto y tendencia: En el cuadro de nombre **Test for unit root in** elegir la opción **level**, en el cuadro de nombre **Include test equation** elegir la opción **Trend and intercept**, en el cuadro de nombre **Lag length** elegir la opción **User specified** con cero rezagos. Luego, hacer clic en **OK** para cerrar la ventana de grupo. La salida se muestra en la figura B.1.
- En la ventana del archivo de trabajo abrir la base de datos de los residuos generados de la prueba ADF.
- Para generar el correlograma de los residuos, en la ventana de la serie de residuos hacer clic en el menú **View** y elegir la opción **Correlogram**, en el

cuadro de nombre **Correlogram of** elegir la opción **Level** y en el cuadro de nombre **Lags to include** dejar la opción por defecto y hacer clic en **OK**.

Null Hypothesis: ICVT_100 has a unit root Exogenous: Constant, Linear Trend Lag Length: 0 (Fixed)				
			t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic			-2.635606	0.2674
Test critical values:				
	1% level		-4.186481	
	5% level		-3.518090	
	10% level		-3.189732	
*MacKinnon (1996) one-sided p-values.				
Augmented Dickey-Fuller Test Equation Dependent Variable: D(ICVT_100) Method: Least Squares Date: 02/21/16 Time: 19:12 Sample (adjusted): 2011M07 2015M01 Included observations: 43 after adjustments				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
ICVT_100(-1)	-0.217935	0.082689	-2.635606	0.0119
C	1.242153	0.500301	2.482813	0.0173
@TREND("2011M06")	0.022060	0.006729	3.278472	0.0022
R-squared	0.223047	Mean dependent var		0.066200
Adjusted R-squared	0.184199	S.D. dependent var		0.249380
S.E. of regression	0.225244	Akaike info criterion		-0.076049
Sum squared resid	2.029397	Schwarz criterion		0.046825
Log likelihood	4.635059	Hannan-Quinn criter.		-0.030737
F-statistic	5.741577	Durbin-Watson stat		2.414113
Prob(F-statistic)	0.006425			

Figura B.1: Prueba ADF a nivel intercepto y tendencia de la serie ICVT con cero rezagos

- Si al observar el correlograma, los estadísticos  $Q$  de Ljung-Box o el estadístico  $DW$  de Durbin-Watson, se descubre un problema de autocorrelación, se debe incorporar más rezagos al modelo y correr de nuevo la prueba ADF a nivel con intercepto y tendencia hasta que los residuos sean ruido blanco y el estadístico  $DW$  este en el rango adecuado.

Una vez se ha determinado el número de rezagos que deben incluirse en el modelo para que no halla problemas de autocorrelación se realiza el siguiente proceso:

- Correr la prueba ADF a nivel sin intercepto ni tendencia: hacer clic en el menú **View-Unit Root Test**, elegir en el cuadro de nombre **Test**

**for unit root in-level-Include test equation-None** y en el cuadro de nombre **Lag length-User specified** colocar el número de rezagos necesario para que no halla autocorrelación.

- Correr la prueba ADF a nivel con intercepto: hacer clic en el menú **View-Unit Root Test**, elegir en el cuadro de nombre **Test for unit root in-level-Include test equation-Intercept** y en el cuadro de nombre **Lag length-User specified** colocar el número de rezagos necesario para que no halla autocorrelación.
  - Con los valores de la suma de los cuadrados de los residuos, obtenidos de las tres pruebas ADF, se determinan los estadísticos  $\phi$  y se establece cual de los tres contrastes debe interpretarse.
  - Una vez elegido el contraste a interpretar se corre la prueba ADF en primeras diferencias: hacer clic en el menú **View-Unit Root Test**, elegir en el cuadro de nombre **Test for unit root in-1st difference-Include test equation-Intercept** y en el cuadro de nombre **Lag length-User specified** colocar el número de rezagos necesario para que no halla autocorrelación.
-

## Apéndice C

# Prueba de cointegración de Engle-Granger y modelo de corrección de error en EViews 8

El procedimiento para estimar la relación de equilibrio a largo plazo es el siguiente:

- Hacer clic en el menú **Quick** y seleccionar **Estimate Equation**.
- En el cuadro de nombre **Equation specification** se escribe primero el nombre de la variable dependiente, luego la constante para calcular la ordenada en el origen y el nombre de la variable independiente, todo separado por espacios en blanco:  $PNC_{100} C ICVT_{100}$ .
- En el cuadro de nombre **Estimate settings** en la opción **Method** se deja la opción por defecto **LS-Least Squares**, luego clic en **Aceptar** y la salida que se obtiene se muestra en la figura C.1.
- En la ventana de la ecuación hacer clic en el menú **Name** y en el cuadro de nombre **Name to identify object** escribir el nombre del archivo para guardarlo en el archivo de trabajo y hacer clic en **OK** para cerrar la ventana del grupo.

- Para generar los residuos estimados de la regresión por MCO en la ventana de la ecuación hacer clic en el menú **Proc** y elegir la opción **Make Residual Series**.

Dependent Variable: PNC_100 Method: Least Squares Date: 01/28/16 Time: 16:23 Sample: 2011M06 2015M01 Included observations: 44				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-0.173051	0.049493	-3.496451	0.0011
ICVT_100	0.106678	0.006387	16.70255	0.0000
R-squared	0.869149	Mean dependent var	0.645864	
Adjusted R-squared	0.866033	S.D. dependent var	0.122537	
S.E. of regression	0.044850	Akaike info criterion	-3.326578	
Sum squared resid	0.084486	Schwarz criterion	-3.245479	
Log likelihood	75.18473	Hannan-Quinn criter.	-3.296503	
F-statistic	278.9753	Durbin-Watson stat	1.675351	
Prob(F-statistic)	0.000000			

Figura C.1: Regresión por MCO de la serie PNC contra ICVT

- En el cuadro de nombre **Residual type** elegir la opción **Ordinary**, en el cuadro de nombre **Name for resid series** escribir el nombre de la serie de residuos para guardarlo en el archivo de trabajo y clic en **OK** para cerrar la ventana del grupo.
- En la ventana del archivo de trabajo abrir la base de datos de los residuos hacer clic en el menú **View** y elegir la opción **Unit Root Test**.
- En el cuadro de nombre **Test for unit root in** elegir la opción **level**, en el cuadro de nombre **Include test equation** dejar la opción por defecto **Intercept**, en el cuadro de nombre **Lag length** elegir la opción **User specified** y colocar cero rezagos y luego hacer clic en **OK** para cerrar la ventana de grupo. La salida se muestra en la figura C.2. Es conveniente verificar que el estadístico DW esta en el rango adecuado, de lo contrario se pueden incorporar más rezagos al modelo.
- Para obtener el gráfico de los residuos de la regresión por MCO en la ventana de la serie hacer clic el menú **View** y elegir la opción **Graph-Line & Symbol** y hacer clic en **OK**.

Augmented Dickey-Fuller Unit Root Test on RESID01

Null Hypothesis: RESID01 has a unit root Exogenous: Constant Lag Length: 0 (Fixed)				
			t-Statistic	Prob.*
<b>Augmented Dickey-Fuller test statistic</b>			-6.174090	0.0000
Test critical values:				
	1% level		-3.592462	
	5% level		-2.931404	
	10% level		-2.603944	
*MacKinnon (1996) one-sided p-values.				
Augmented Dickey-Fuller Test Equation Dependent Variable: D(RESID01) Method: Least Squares Date: 01/27/16 Time: 20:28 Sample (adjusted): 2011M07 2015M01 Included observations: 43 after adjustments				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
RESID01(-1)	-0.897785	0.145412	-6.174090	0.0000
C	0.002358	0.006444	0.365914	0.7163
R-squared	0.481796	Mean dependent var		0.002484
Adjusted R-squared	0.469157	S.D. dependent var		0.057998
S.E. of regression	0.042257	Akaike info criterion		-3.444714
Sum squared resid	0.073211	Schwarz criterion		-3.362797
Log likelihood	76.06135	Hannan-Quinn criter.		-3.414506
F-statistic	38.11939	Durbin-Watson stat		1.901711
Prob(F-statistic)	0.000000			

Figura C.2: Prueba de cointegración de Engle-Granger para los residuos de la regresión.

Para realizar las pruebas de diagnóstico de normalidad, autocorrelación y heteroscedasticidad de los residuales de la regresión se efectúa el siguiente procedimiento:

- En la ventana del archivo de trabajo abrir la base de datos de la serie de los residuos.
- Para generar el gráfico de probabilidad normal en la ventana de la serie hacer clic en el menú **View** y elegir la opción **Graph-Quantile-Quantile** y clic en **OK**.
- Para obtener el boxplot, en la ventana de la serie hacer clic en el menú **View** y elegir la opción **Graph-Boxplot** y clic en **OK**.
- Para generar el histograma de los residuos y realizar la prueba de Jarque-Bera, en la ventana de la ecuación hacer clic en el menú **View** y elegir la opción **Residual Diagnostics-Histogram-Normality Test**.

- Para obtener el correlograma de los residuales y realizar la prueba  $Q$  de Ljung-Box, en la ventana de la serie hacer clic en el menú **View** y elegir la opción **Correlogram**.
- En el cuadro de nombre **Correlogram of** elegir la opción **Level** y clic en **OK**.
- Para generar el gráfico del cuadrado de los residuos contra la variable dependiente hacer clic en el menu **Quick**, elegir la opción **Graph**, en el cuadro de nombre **List of series** colocar el nombre de la variable dependiente y el nombre de los residuos al cuadrado separados por espacios en blanco:  $PNC_{100} \text{ RESID01}^2$ . Luego, en el cuadro de nombre **Graph type** elegir la opción **Scatter** y clic en **OK**.
- Para realizar la prueba de heteroscedasticidad en la ventana de la ecuación hacer clic en el menú **View**, elegir la opción **Residual Diagnostics-Heteroskedasticity Test**, en el cuadro de nombre **Specification** elegir en tipo de test **White** y no incluir términos cruzados y hacer clic en **OK**. La salida que se obtiene se muestra en la figura C.3.

Heteroskedasticity Test: White				
F-statistic	1.580644	Prob. F(1,42)	0.2156	
Obs*R-squared	1.595854	Prob. Chi-Square(1)	0.2065	
Scaled explained SS	1.658987	Prob. Chi-Square(1)	0.1977	
Test Equation:				
Dependent Variable: RESID^2				
Method: Least Squares				
Date: 01/30/16 Time: 14:26				
Sample: 2011M06 2015M01				
Included observations: 44				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.003859	0.001604	2.406552	0.0206
ICVT_100^2	-3.23E-05	2.57E-05	-1.257237	0.2156
R-squared	0.036269	Mean dependent var	0.001920	
Adjusted R-squared	0.013323	S.D. dependent var	0.002934	
S.E. of regression	0.002914	Akaike info criterion	-8.793901	
Sum squared resid	0.000357	Schwarz criterion	-8.712802	
Log likelihood	195.4658	Hannan-Quinn criter.	-8.763826	
F-statistic	1.580644	Durbin-Watson stat	1.989100	
Prob(F-statistic)	0.215614			

Figura C.3: Prueba de heroscedasticidad de White para los residuos de la regresión.

---

La estimación por mínimos cuadrados ordinarios (MCO) del modelo de corrección de error (MCE) se realiza de acuerdo al siguiente procedimiento:

- Se calcula la primera diferencia de las variables, PNC e ICVT, aplicando la siguiente orden en la ventana de comandos:

$$\begin{aligned}\text{Genr dPNC}_{100} &= \text{d(PNC}_{100}) \\ \text{Genr dICVT}_{100} &= \text{d(ICVT}_{100})\end{aligned}$$

o también, se puede hacer clic en el menú **Quick** y elegir la opción **Generate Series** y escribir:  $\text{dPNC}_{100} = \text{PNC}_{100} - \text{PNC}_{100}(-1)$  y clic en **OK**. Las nuevas variables aparecerán en el archivo de trabajo.

- Para la especificación del modelo de corrección de error se escribe la siguiente orden en la barra de comandos:

```

Ls dPNC100 C dPNC100(-1) dPNC100(-2) dPNC100(-3) dPNC100(-4)
  dPNC100(-5) dPNC100(-6) dICVT100 dICVT100(-1) dICVT100(-2)
  dICVT100(-3) dICVT100(-4) dICVT100(-5) dICVT100(-6) RESID01(-1)

```

donde se ha introducido la variable dependiente en primeras diferencias  $\text{dPNC}_{100}$ , la constante C para calcular la ordenada en el origen, los primeros seis retardos de la variable dependiente en primeras diferencias, la variable independiente en primeras diferencias  $\text{dICVT}_{100}$ , los primeros seis retardos de la variable independiente en primeras diferencias y los residuos de la regresión por MCO retardados un periodo  $\text{RESID01}(-1)$ .

- Al hacer clic en **Enter** se obtiene la salida que se muestra en la figura C.4.
  - Se debe comenzar a eliminar variables desde la menos significativa hasta la más significativa. Se continua el proceso hasta que todas, o al menos la mayoría, de las variables del modelo sean significativas y el coeficiente de los residuos retardados un periodo sea negativo y menor a la unidad. Además, el modelo debe aprobar las pruebas de diagnóstico de normalidad, autocorrelación y heteroscedasticidad. La continuación del proceso condujo al modelo que se muestra en la figura C.5.
-

Dependent Variable: DPNC_100				
Method: Least Squares				
Date: 06/01/16 Time: 16:50				
Sample (adjusted): 2012M01 2015M01				
Included observations: 37 after adjustments				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.011262	0.010715	1.050988	0.3047
DPNC_100(-1)	-0.567467	0.475281	-1.193961	0.2452
DPNC_100(-2)	-0.408213	0.407566	-1.001587	0.3274
DPNC_100(-3)	-0.186948	0.338280	-0.552645	0.5861
DPNC_100(-4)	0.124902	0.303019	0.412192	0.6842
DPNC_100(-5)	0.277910	0.272675	1.019199	0.3192
DPNC_100(-6)	0.170172	0.202301	0.841184	0.4093
DICVT_100	0.021281	0.038140	0.557978	0.5825
DICVT_100(-1)	-0.008692	0.041674	-0.208580	0.8367
DICVT_100(-2)	-0.035656	0.042429	-0.840361	0.4097
DICVT_100(-3)	-0.034648	0.042390	-0.817370	0.4225
DICVT_100(-4)	0.018006	0.036535	0.492852	0.6270
DICVT_100(-5)	0.092355	0.033157	2.785357	0.0108
DICVT_100(-6)	0.040071	0.034676	1.155600	0.2602
RESID01(-1)	-0.451220	0.473521	-0.952903	0.3510
R-squared	0.632827	Mean dependent var	0.010143	
Adjusted R-squared	0.399171	S.D. dependent var	0.049805	
S.E. of regression	0.038605	Akaike info criterion	-3.379928	
Sum squared resid	0.032788	Schwarz criterion	-2.726853	
Log likelihood	77.52866	Hannan-Quinn criter.	-3.149688	
F-statistic	2.708373	Durbin-Watson stat	2.003248	
Prob(F-statistic)	0.017767			

Figura C.4:

Dependent Variable: DPNC_100				
Method: Least Squares				
Date: 06/07/16 Time: 07:19				
Sample (adjusted): 2011M12 2015M01				
Included observations: 38 after adjustments				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
DICVT_100(-5)	0.070297	0.023900	2.941282	0.0057
RESID01(-1)	-0.660736	0.142565	-4.634647	0.0000
R-squared	0.438471	Mean dependent var	0.009901	
Adjusted R-squared	0.422873	S.D. dependent var	0.049150	
S.E. of regression	0.037338	Akaike info criterion	-3.686397	
Sum squared resid	0.050189	Schwarz criterion	-3.600208	
Log likelihood	72.04154	Hannan-Quinn criter.	-3.655732	
Durbin-Watson stat	2.307339			

Figura C.5: Resultados de la estimación por MCO del modelo de corrección de error para la relación entre la PNC y el ICVT.

## Apéndice D

# Prueba de cointegración de Engle-Granger y modelo de corrección de error en R

El procedimiento para estimar la relación de equilibrio a largo plazo es el siguiente:

- Realizar la regresión entre el ICVT y la PNC,  

```
> datos.01 <- cbind(pnc.t, icvt.t)
> regresion.01 <- lm(pnc.t ~ icvt.t, data = datos.01)
> eq.01 <- summary(regresion.01)
> eq.01
```
- Efectuar la prueba ADF sobre los residuos de la regresión,  

```
> library(urca)
> residuos.01 <- -(resid(eq.01))
> dfa.01 <- ur.df(residuos.01, lags=0, type="drift")
> summary(dfa.01)
> plot(dfa.01)
```

Para hacer las pruebas de diagnóstico de los residuos de la regresión se efectúa el siguiente procedimiento:

- Realizar el resumen de la estadística descriptiva,  

```
> summary(residuos.01)
```

---

En R-studio cargar la librería: `> library(Rcdmr)` y ejecutar la orden `> numSummary(residuos.01, statistics=c("skewness", "kurtosis", "sd"), type="2")`

- Efectuar las pruebas de normalidad,  
`> qqnorm(residuos.01)`  
`> qqline(residuos.01, col="red", lwd=2)`  
`> hist(residuos.01, col="yellow", ylab="Frecuencia", xlab="Residuos01")`  
`> boxplot(residuos.01)`  
`> jarque.bera.test(residuos.01)`  
`> shapiro.test(residuos.01)`
- Realizar las pruebas de incorrelación,  
`> acf(residuos.01, lag.max = 30)`  
`> pacf(residuos.01, lag.max = 30)`  
`> Box.test(residuos.01, lag=24, type="Ljung-Box")`
- Efectuar la prueba de heterocedasticidad de White  
`> R.01 < -summary(lm(residuos.01^2 ~ I(icvt.t^2)))$r.squared`  
`> LM.01 < -nrow(datos.01) * R.01`  
`> p.valor < -1 - pchisq(LM.01, 1)`  
`> p.valor`

El procedimiento para estimar la relación de equilibrio a corto plazo es el siguiente:

- Para poder realizar la regresión considerando términos rezagados se debe cargar la siguiente librería:  
`> library(lmtest)`
  - Determinar los residuos de la regresión,  
`> tce.01 < -resid(regresion.01)[2 : 44]`
  - Obtener las primeras diferencias de las series,  
`> dpnc.t = diff(pnc.t)`  
`> dicvt.t = diff(icvt.t)`
  - Combinar la información de las series en primeras diferencias y los residuos de la regresión,  
`> datos.mce.01 < -cbind(dpnc.t, dicvt.t, tce.01)`
-

- 
- Especificar el modelo de corrección de error,
 

```

>mce.01< -dynlm(dpnc.t~L(dpnc.t,1)+L(dpnc.t,2)+L(dpnc.t,3)
+L(dpnc.t,4)+L(dpnc.t,5)+L(dpnc.t,6)+L(dicvt.t,0)+L(dicvt.t,1)
+L(dicvt.t,2)+L(dicvt.t,3)+L(dicvt.t,4)+L(dicvt.t,5)+L(dicvt.t,6)
+L(tce.01,1),data=datos.mce.01)
> summary(mce.01)

```
  - Se debe comenzar a eliminar variables desde la menos significativa hasta la más significativa. Se continua el proceso hasta que todas, o al menos la mayoría, de las variables del modelo sean significativas y el coeficiente de los residuos retardados un periodo sea negativo y menor a la unidad. Además, el modelo debe aprobar las pruebas de diagnóstico de normalidad, autocorrelación y heteroscedasticidad. La continuación del proceso condujo al siguiente modelo:
 

```

>mce.f.01< -dynlm(dpnc.t~0+L(dicvt.t,5)+L(tce.01,1),data
=datos.mce.01)
> summary(mce.f.01)

```
  - Se obtiene los residuos del modelo de corrección de error,
 

```

> residuos.mce.f.01 < -resid(mce.f.01)

```
  - Se realizan las pruebas de diagnóstico de los residuos del modelo de corrección de error
    - Realizar el resumen de la estadística descriptiva,
 

```

> summary(residuos.mce.f.01)

```

 En R-studio cargar la libreria:> library(Rcdmr) y ejecutar la orden
 

```

>numSummary(residuos.mce.f.01,statistics=c("skewness", "kurtosis",
"sd"),type="2")

```
    - Efectuar las pruebas de normalidad,
 

```

> qqnorm(residuos.01)
> qqline(residuos.01,col="red",lwd=2)
> hist(residuos.01,col="yellow",ylab="Frecuencia",xlab=
"Residuos01")
> boxplot(residuos.mce.f.01)
> jarque.bera.test(residuos.mce.f.01)
> shapiro.test(residuos.mce.f.01)

```
    - Realizar las pruebas de incorrelación,
 

```

> acf(residuos.mce.f.01, lag.max = 30)

```
-

```
> pacf(residuos.mce.f.01, lag.max = 30)
> Box.test(residuos.mce.f.01, lag=24, type="Ljung-Box")
```

- Efectuar la prueba de heterocedasticidad

```
> library(lmtest)
> bptest(residuos.mce.f.01)
```

---

# Referencias

- Anderson, T. W., y Darling, D. A. (1954). A Test of Goodness of Fit. *Journal of the American Statistical Association*, 49(268), 765-769.
- Ang, A. H., y Tang, W. H. (2007). *Probability Concepts in Engineering: Emphasis on Applications to Civil and Environment Engineering*. N.Y., USA : John Wiley and Sons.
- Aplicación de basilea 2-apectos prácticos*. (2004). Banco de pagos internacionales. (Basilea: publications@bis.org.)
- Aristizabal, V. R., Támara, A. A., y Velásquez, C. E. (2010). *Modelación del riesgo crediticio como elemento fundamental en el cálculo de la pérdida esperada en una institución financiera*. Universidad Eafit. Medellín, Colombia. (Disponible en: [repository.eafit.edu.co/handle/10784/1187](http://repository.eafit.edu.co/handle/10784/1187))
- Canavos, G. C. (1988). *Probabilidad y Estadística Aplicaciones y Métodos*. México, D.F., México: McGraw-Hill/Interamericana de México, S.A. de C.V.
- Capítulo 2: Reglas relativas a la gestión de riesgo crediticio. circular externa 027 de 2011*. (2014). Superintendencia Financiera de Colombia. (Disponible en: [www.superfinanciera.gov.co](http://www.superfinanciera.gov.co).)
- Capítulo 2: Reglas relativas a la gestión de riesgo crediticio. circular externa 032 de 2014*. (2014). Superintendencia Financiera de Colombia. (Disponible en: [www.superfinanciera.gov.co](http://www.superfinanciera.gov.co).)
- Capítulo 2: Reglas relativas a la gestión de riesgo crediticio. circular externa 035 de 2009*. (2009). Superintendencia Financiera de Colombia. (Disponible en: [www.superfinanciera.gov.co](http://www.superfinanciera.gov.co).)
- Cryer, D. J., y Chan, K. S. (2008). *Time Series Analysis with Applications in R*. N.Y., USA: Springer-Verlag.
- Dickey, D. A., y Fuller, W. A. (1981). Likelihood Ratio Statistical for Autoregressive Time Series with a Unit Root. *Ecnometrica*, 49(4), 1057-1072.
- Evolución del sistema financiero colombiano*. (2015). Superintendencia Financiera

- 
- de Colombia. (Disponible en: [www.superfinanciera.gov.co](http://www.superfinanciera.gov.co).)
- Fishman, G. S. (1995). *Monte carlo: Concepts, algorithms, and applications*. N.Y., USA: Springer-Verlag.
- Flowers, C., Flowers, R. S., y Trejo, F. R. (2014). Evaluación de criterios de selección de modelos probabilísticos con series de valores máximos simulados. *Tecnología y Ciencias del Agua*, 5(5), 189-197.
- Giraldo, Y. W. (2010). *Determinantes de la morosidad de la cartera en el sistema financiero colombiano*. Universidad Icesi. Santiago de Cali, Colombia. (Disponible en: [hdl.handle.net/10906/5394](http://hdl.handle.net/10906/5394))
- Gómez, A. L. (2010). *Análisis del combustible de un reactor nuclear modular de helio con turbina de gas*. Universidad Nacional Autónoma de México, México, D.F. (Disponible en: <http://132.248.52.100:8080/xmlui/handle/132.248.52.100/720>)
- Granger, W. J., y Engle, R. F. (1987). Co-integration and error correction: representation, estimation and testing. *Econometrica*, 55(2), 251-276.
- Gujarati, D. N. (2010). *Econometría*. México, D.F., México: McGraw-Hill/Interamericana de México, S.A. de C.V.
- Lando, D. (2004). *Credit risk modeling: theory and applications*. Princeton, USA: Princeton University Press.
- Law, A. M., y Kelton, W. D. (2000). *Simulation modeling and analysis*. N.Y., USA: McGraw-Hill, Inc.
- López, G. A., Rodríguez, V. J., y Pulido, S. A. (2004). *Notas de clase: Curso de predicción económica y empresarial*. Universidad Autónoma de Madrid. Madrid, España. (Disponible en: [www.uam.es/predysim](http://www.uam.es/predysim))
- Mackinnon, J. G. (2010). Critical Values for Cointegration Tests. *Queen's Economics Department Working Paper*(1227), 1-17.
- Martinez, S. E. (2014). *Metodología para estimar el deterioro de cartera de crédito de libre destinación sin libranza otorgado por instituciones de crédito a personas naturales en Colombia*. Universidad de Medellín. Medellín, Colombia. (Disponible en: [repository.udem.edu.co/handle/11407/1108](http://repository.udem.edu.co/handle/11407/1108))
- Mata, H. L. (2004). *Notas de clase: Nociones Elementales de Cointegración Enfoque de Engle-Granger*. Universidad de los Andes. Mérida, Venezuela. (Disponible en: [webdelprofesor.ula.ve/economia/hmata](http://webdelprofesor.ula.ve/economia/hmata))
- Montenegro, G. A. (2011). *Análisis de series de tiempo*. Bogotá, Colombia: Editorial Pontificia Universidad Javeriana.
- Nieto, M. S. (2010). *Crédito al consumo: La estadística aplicada a un problema de riesgo crediticio*. Universidad Autónoma Metropolitana, Unidad Iztapalapa, México. (Disponible en: [mat.izt.uam.mx/mcmai/documentos/tesis/Gen.07-](http://mat.izt.uam.mx/mcmai/documentos/tesis/Gen.07-)
-

- 
- O/Nieto-S-Tesis.pdf)
- Pérez, R. F., y Fernández, C. H. (2009). *Econometría conceptos básicos*. Medellín, Colombia: ECOE EDICIONES.
- Pfaff, B. (2008). *Analysis of Integrated and Cointegrated Time Series with R*. N.Y., USA: Springer-Verlag.
- Raychaudhuri, S. (2008). *Introduction to Monte Carlo Simulation. Proceedings of the 2008 Winter Simulation Conference*.
- Reporte de estabilidad financiera*. (2014). Banco de la República. (Disponible en: [www.banrep.gov.co](http://www.banrep.gov.co).)
- Sagner, A. (2011). El índice de cartera vencida como medida del riesgo de crédito: análisis y aplicación al caso de Chile. *Documentos de trabajo Banco Central de Chile*(618).
- Salazar, V. F. (2013). Cuantificación del riesgo de incumplimiento en créditos de libre inversión: un ejercicio econométrico para una entidad bancaria del municipio de Popayán, Colombia. *Estudios Gerenciales*(29), 416-427.
- Shapiro, S. S., y Wilk, M. B. (1965). An analysis of variance test for normality (complete samples). *Biometrika*, 52(3/4), 591-611.
-