

Teoría de precios de arbitraje. Evidencia empírica para Colombia a través de redes neuronales

Charle Augusto Londoño H.

Mauricio Lopera Castaño*

Sergio Restrepo

Recibido: Noviembre 2009 – Aprobado: Marzo 2010

Resumen. Esta investigación utiliza una red neuronal multicapa para relacionar el Índice General de Bolsa de Valores de Colombia (IGBC) con fundamentales macroeconómicos y variables financieras. Proponemos dos modelos: un modelo APT (fundamentales macroeconómicos) y un modelo APT modificado (fundamentales macroeconómicos + indicador de las bolsas del mundo); de acuerdo a nuestro análisis el APT tradicional se ajusta mejor para predecir el mercado de valores Colombiano. Los resultados confirman que las redes neuronales artificiales (ANN) son más efectivas que los modelos estadísticos tradicionales por su capacidad explicativa y precisión.

Palabras clave: Teoría de precios de arbitraje, variables macroeconómicas, mercado de valores, redes neuronales artificiales.

Clasificación JEL: G12, G15, F31, C45.

Abstract. This research uses a multilayer neural network to relate the General Index from Colombia Stock Exchange (IGBC) with macroeconomic fundamentals and financial variables. We propose two models: an APT (macroeconomic fundamentals) one and a modified APT (macroeconomic fundamentals + international stock markets indicator); according to our analysis the traditional APT predicts more accurately the behavior of the Colombian stock market. The results confirm that the artificial neural network (ANN) approach is more effective than traditional statistical models given its explanatory power and precision.

Keywords: Arbitrage pricing theory, macroeconomic variables, stock market, artificial neural networks.

JEL classification: G12, G15, F31, C45.

*Universidad de Antioquia. Este ensayo se deriva del trabajo de grado de pregrado en economía de Charle Augusto Londoño H. Dicho trabajo obtuvo mención de honor a la mejor tesis del semestre 2009/1 y ganó el segundo puesto en el premio a la investigación estudiantil 2009 en el área de Ciencias Exactas, Naturales, Ingeniería y Economía de la Universidad de Antioquia. Correo electrónico: chlondono@yahoo.es.

Los autores de esta investigación agradecen todas las observaciones y sugerencias hechas por la economista Diana M. Vélez y por el evaluador anónimo.

1. Introducción

En Colombia, a raíz de la fusión de las tres bolsas de valores (Bogotá, Medellín y Occidente) en el año 2001 el mercado accionario se ha vuelto más eficiente, integrado, y rentable; sin embargo, el aumento de la capitalización bursátil y la liquidez hacen menos predecible el comportamiento futuro de los activos e incrementa los niveles de incertidumbre y de riesgo. Por tal motivo, académicos e inversionistas se han interesado en diseñar estrategias para minimizar el riesgo de mercado, que basándose en teorías microfundamentadas (Ver Markowitz, 1952; Sharpe, 1964; Merton, 1973; Ross, 1976; Solnik, 1983, Connor, 1984; entre otros) han elaborado modelos econométricos cuyo objeto es el de capturar todos los factores sistemáticos que puedan afectar la toma de decisiones de inversión.

Estas teorías se encuentran inscritas en las tres versiones de las Hipótesis de los Mercados Eficientes según las cuales se establecen diversas estrategias de valoración de activos de acuerdo a la información que se posee: forma débil, semifuerte y fuerte (véase Fama, 1991, para más detalles).¹ Al momento de realizar esta investigación, en lo que respecta a la valoración de activos financieros colombianos, sólo se han contrastado la forma débil y algunas aproximaciones de la semifuerte.² No obstante, ninguna ha analizado la forma semifuerte empleando la metodología de redes neuronales artificiales (ANN),³ la cual, a través de su estructura no lineal permite capturar de forma conjunta los diferentes choques macroeconómicos y financieros que influyen asimétricamente sobre la valoración de activos, lo que permite visualizar cuál es el comportamiento de los pesos con los que tales variables entran en el retorno de los activos y cómo son utilizadas para la predicción.⁴ Por tanto, este trabajo se convierte en una buena herramienta econométrica para la toma de decisiones de inversión en un ambiente de incertidumbre.

¹La Hipótesis de los Mercados Eficientes establece tres tipos de inversionistas: i) forma débil, que sólo tienen acceso a información de los precios históricos, a través del análisis técnico (estudio de tendencias e indicadores de los precios de los activos) para la toma de decisiones de inversión; ii) forma semifuerte, que disfrutan de la información anterior, y también de información sobre fundamentales microeconómicos y macroeconómicos. Con el fin de para predecir el comportamiento futuro del precio de los activos ésta estrategia utiliza tanto la psicología del mercado, en lo que atañe a los factores que la mueven (anuncios de venta, fusiones, adquisiciones, información macroeconómica, entre otras variables) como modelos econométricos microfundamentados. Por último se tiene iii) la forma fuerte, que ostenta información monopólica del mercado, es decir, aquella información de la economía que pocos agentes poseen, antes de llegar a ser pública, anticipando las decisiones que tomará el mercado.

²En relación a este aspecto se consolidan tres líneas de investigación: Una es la administración de portafolios como Agudelo y Arango (2008), Medina (2003), Reveiz (2008), Reveiz y León (2008) y Reveiz y Rojas (2008). Otra es la construcción de curvas de rendimiento, Melo y Becerra (2006), Santana (2008). Y por último, en hacer pronósticos como Botero y Cano (2008), Castaño et. al (2008), Gallón y Gómez (2007), Jalil y Misas (2007), Rodríguez (2001) Villamil y Delgado (2007) y Villamil (2009).

³Por sus siglas en inglés *Artificial Neural Network*.

⁴La importancia conjunta de esgrimir argumentos tales como los precios históricos de los activos y variables macroeconómicas para predecir el comportamiento de los valores es ilustrada en Chen et. al (1986); Koutoulas y Kryzanowski (1994); Refenes et. al (1995); Ahmaldi (1996); Altay (2003) y Shajarundi y Fung (2009) quienes muestran la relevancia de las variables macroeconómicas en la valoración de activos.

Más aún, la metodología de ANN ha mostrado resultados sobresalientes aplicados a la predicción de variables económicas por encima de modelos lineales. Para Colombia se han elaborado diversas investigaciones para la predicción de la inflación como es el trabajo de Misas et. al (2002) y Aristizábal (2006). Cabe destacar que este último hace una buena revisión de aplicaciones con modelos de ANN a diversas áreas del conocimiento. Del mismo modo, Jalil y Misas (2007), Villamil y Delgado (2007) y Villamil (2009) muestran la importancia de ANN en problemas tan complejos e importantes como la predicción del tipo de cambio, el cual puede ser afectado por una serie de factores que se presentan de forma no lineal. Por último, Santana (2008) muestra que las ANN son una buena estrategia en la captura del comportamiento de las curvas de rendimiento de los bonos, que son de gran relevancia para la asignación óptima de los recursos económicos.

En el caso internacional, hay una extensa literatura en la cual se destacan algunos trabajos como los de Cheng y Titterton (1994) y Kuan y White (1994) los cuales cimientan las ANN desde una perspectiva estadística y econométrica, respectivamente. Por otro lado, Qi (1994) hace una comparación de modelos tradicionales y de ANN, en la cual muestra su interpretación e implementación aplicados a las finanzas. Igualmente, Franses y Dijk (1999) exponen una buena metodología de ANN para pronósticos financieros.

Desde este punto de vista, el objetivo de esta investigación es evaluar los múltiples factores económicos nacionales e internacionales que puedan afectar la valoración de activos en Colombia. Para tal propósito, se utilizarán los principales argumentos teórico - empíricos del esquema *Arbitrage Pricing Theory* (APT) que por medio de un modelo de ANN y su respectivo análisis de sensibilidad identifiquen las variables macroeconómicas más relevantes, clasificándolas según sea su preponderancia.

El presente trabajo se divide en seis partes. La primera es esta introducción. La segunda presenta algunos de los modelos teóricos empleados en finanzas, la formalización del modelo APT y una justificación del modelo de ANN como la mejor herramienta para el contraste de las hipótesis planteadas bajo este esquema teórico. La tercera presenta la metodología de ANN. La cuarta describe los datos utilizados y la construcción de las variables. La quinta reporta los hallazgos obtenidos de la aplicación de dos modelos APT (tradicional y modificado), tanto en lo que respecta a la validación del modelo, como en relación a los resultados ofrecidos por el análisis de sensibilidad. Por último, se presentan las conclusiones en las cuales se vislumbra la superioridad explicativa del modelo APT bajo ANN con respecto a su forma modificada.

2. Teorías financieras microfundamentadas

2.1. Revisión de la literatura

El objetivo de la teoría financiera microfundamentada es encontrar una estrategia óptima para la minimización del riesgo en el mercado de valores. Entre los trabajos realizados en este campo, se resalta el de Von Neumann y Morgenstein

(1944) que ha provisto los cimientos teóricos en la materia, al describir a los individuos como agentes aversos al riesgo que, por medio de la función de utilidad esperada caracterizada, evalúan y determinan la forma como los agentes toman decisiones bajo condiciones de riesgo (Sharpe, 1964). Siguiendo esta línea, Markowitz (1952) se centra en describir el comportamiento diversificador de los inversionistas, estableciendo una regla normativa para la elección de un portafolio, según la cual para minimizar el riesgo se debe elegir un grupo de activos que contenga el retorno esperado máximo sujeto a una varianza dada.

Sharpe (1964) muestra que la regla de Markowitz (1952) lleva a resultados subóptimos, porque no utiliza los activos individuales con mayor retorno esperado. Para resolver tal dificultad, el autor diseña una medida de riesgo, β , inherente a cada activo, que es obtenida de la relación esperada entre su retorno y el de mercado, e instituye el criterio de elegir los activos con los β más altos.

De la anterior teoría se desprenden dos líneas de investigación: una que concede relevancia únicamente a las variables internas del mercado y otra que utiliza fundamentales macroeconómicos como factores explicativos de riesgo. En la primera, Merton (1973) construye un modelo con una estructura intertemporal, que avala una mayor consistencia entre la maximización de la utilidad y la minimización del riesgo. Por su parte, Miller (1977) introduce el supuesto de expectativas heterogéneas, postulando que todos los agentes, al enfrentarse con incertidumbre en la toma de decisiones, llegan a resultados disímiles e inciertos en sus pronósticos debido a la gran variedad de activos a tener en cuenta.⁵

En la segunda, Ross (1976), su principal exponente, acepta el cumplimiento de la ley de los grandes números, estableciendo que el riesgo depende de fundamentales macroeconómicos como las tasas de interés, la inflación, la actividad real, entre otros; lo que ayuda a explicar el comportamiento futuro de los inversionistas y, por tanto, esclareciendo la estrategia a seguir para cubrirse ante la exposición al riesgo. Solnik (1983) encuentra que las decisiones de inversión son tomadas en mercados globalizados, siendo riesgos latentes que no pueden ser cubiertos por el modelo de Sharpe (1964). Así, encuentra que el modelo de Ross (1976) puede llegar a resultados de equilibrio en una estructura internacional, al permitir que todos los factores de riesgo interno y externos entren en la ecuación de valoración de forma no observable.

Connor (1984) realiza un modelo isomorfo al de Ross (1976), el cual admite la elección de un portafolio bien diversificado,⁶ lo que se logra a través de la elección de un grupo de activos altamente dependientes a los factores de riesgo, proveyendo una generalización del β de mercado del modelo de Sharpe (1964).

En conclusión, las teorías anteriores, bajo los supuestos convencionales del modelo neoclásico, inician un proceso de ajuste que es logrado por medio de la crítica a los trabajos de otros autores, caracterizando de una forma más realista

⁵La propuesta de Miller (1977) establece una relación de corto plazo y no una generalización del comportamiento de los inversionistas que elegirán un portafolio que maximice su utilidad de forma homogénea garantizando el equilibrio como lo hacen notar Sharpe (1964) y Merton (1973).

⁶Una de las críticas principales al modelo de Ross (1976) es que no se sabe cuales activos deben ser seleccionados para que se cumpla la ley de los grandes números.

el comportamiento de los inversionistas, y por ende, ofreciendo a los agentes herramientas económicas cada vez más diversas y sofisticadas para la toma de decisiones de inversión bajo condiciones de riesgo.

2.2. Teoría de precios de arbitraje (APT)

El enfoque en el que se sustenta este trabajo es el de modelar a los agentes de acuerdo al esquema APT de Ross (1976). El modelo posee la siguiente estructura: los agentes están caracterizados por una función de utilidad U la cual cumple las propiedades clásicas de preferencias como son la concavidad y monotonidad, tal que

$$U(x) = G[U(x; R)] \quad (1)$$

de ahí, $U(x; R)$ es la función de utilidad, es la canasta elegida de activos y es el riesgo relativo.

Dadas las características generales de los agentes se supone la existencia de un activo i considerado libre de riesgo para un periodo de tenencia, t . Este activo compensa a los inversionistas con una tasa libre de riesgo ρ_0 . Los inversionistas creen (subjettivamente) que los retornos son generados por el modelo en la ecuación (2), además, estos no son insignificantes en el mercado en el sentido de que todos siguen el mismo modelo

$$\begin{aligned} \tilde{x}_i &= E_i + \beta_{i1}\tilde{\delta}_1 + \dots + \beta_{ik}\tilde{\delta}_k + \tilde{\epsilon}_i \\ &= E_i + \beta_{i1}\tilde{\delta} + \tilde{\epsilon}_i \end{aligned} \quad (2)$$

donde $\tilde{x}_i$ es el retorno de un portafolio de activos, $E_i (= \rho_0 + \beta_{i1}\rho_1 + \dots + \beta_{ik}\rho_k)$ es el retorno esperado, con ρ_i la prima de riesgo, β_{ij} son las cargas factoriales (medidas de riesgo) indicando la sensibilidad del i -ésimo activo al j -ésimo factor, $\tilde{\delta}_i$ son factores latentes y $\tilde{\epsilon}_i$'s es el termino de perturbación con las siguientes propiedades $E[\tilde{\epsilon}_i] = 0$, $E[\tilde{\epsilon}_i\tilde{\epsilon}_j] = 0 \ \forall i \neq j$ y $var(\tilde{\epsilon}_i) = \tilde{\sigma}_i^2$.

Adicionalmente, todos los inversionistas tienen expectativas homogéneas, lo cual significa que ningún inversionista espera ganar un retorno superior al esperado E_i dada la elección de un portafolio óptimo, lo que garantiza el cumplimiento de la ley de los grandes números, que satisface la condición según la cual al aumentar el número de activos, el retorno esperado puede tender a igualarse a la tasa libre de riesgo. Por tanto, la riqueza ω de un inversionista estará invertida en un grupo de activos fijo que conforma su portafolio $\tilde{x}_i \geq 0$.

En resumen, el modelo busca caracterizar a los inversionistas y ofrecer una herramienta predictiva acerca de como éstos toman sus decisiones, y de tal manera instituir un portafolio óptimo para agentes aversos al riesgo, del cual, dependiendo de sus preferencias diversificadoras, obtendrán al menos la tasa libre de riesgo ρ_0 como rendimiento mínimo.

2.3. *Justificación de la metodología de redes neuronales para el contraste empírico del modelo APT*

Existen diversos trabajos que han tratado el problema de la diversificación del riesgo a través del modelo APT, en varias economías.⁷ Sin embargo, hay diferencias en la utilización de variables y en las metodologías econométricas. En primer lugar, no hay un consenso claro sobre las variables a emplear,⁸ dado que el mercado de valores es multidimensional y disímil en su estructura económica, implicando que no se cuenta con las variables deseadas en los distintos mercados, debido, entre otras razones, a no encontrarse medidas en todas las economías, no tener los mismos efectos en cada mercado, ni hallarse con la periodicidad necesaria para el contrastaste empírico. En segundo lugar, las metodologías econométricas utilizadas han sido distintas en la extracción de los factores de riesgo, debido a que la aplicación empírica del APT requiere usar variables latentes para explicar el retorno esperado de los activos.

Tanto Chen et. al (1986), como Koutoulas y Kryzanowski (1994) construyen variables independientes esperadas, y empleando modelos estadísticos tradicionales llegan a resultados satisfactorios de contraste empírico. Otros, como Engle et. al (1990); Altay (2003) y Shajarundi y Fung (2009) utilizan análisis factorial para extraer las variables explicativas, y mediante modelos estadísticos tradicionales encuentran una adecuada caracterización del retorno promedio de los activos. Por último, Refenes et. al (1995) y Ahmadi (1996), utilizando modelos de ANN, esbozan la superioridad en los resultados de estimación en relación a los modelos estadísticos tradicionales.

No obstante, hay algunas objeciones para las primeras dos metodologías en lo que se refiere a la extracción de los factores de riesgo. Chen et. al (1986) encuentran que es difícil obtener variables esperadas con una buena periodicidad, lo cual hace inviable la realización del contraste empírico; además, la selección de las variables puede llevar a problemas de multicolinealidad de las variables.⁹ Con respecto a la utilización de análisis factorial, en ocasiones los factores no existen y, en caso de ser hallados, se presentan dos limitantes: el primero es que su carácter estático hace que los resultados sean diferentes al utilizar distintas muestras;¹⁰ el segundo, y más importante, ya sea que tengan o no interpretabilidad económica, estos no determinan la forma en que las fuerzas económicas los modifican¹¹ (Fama, 1991; Refenes et. al, 1995 y Ahmaldi, 1996),

⁷Por ejemplo, para el mercado estadounidense están Chen et al (1986); Engle et. al (1990) y Ahmadi (1996). Para Canadá, Koutoulas y Kryzanowski (1994); para el Reino Unido, Refenes et. al (1995); para Alemania y Turquía, Altay (2003) y para Malasia, Shajarundi y Fung (2009). Para una revisión más amplia del modelo APT y sus aplicaciones ver a Fama (1991); Altay (2003) y García y García (2006).

⁸Ver Altay (2003) y García y García (2006) para una revisión del estado del arte.

⁹En la actualidad, la construcción de variables esperadas ya no es una desventaja, debido a los avances en el software econométrico.

¹⁰Claro está que el uso de distintos periodos muestrales hace que difieran los valores estimados de los parámetros en cualquier modelo estadístico, pero esto no implica cambios tan abruptos o la inexistencia de los factores (coeficientes) como los que se presentan en el análisis factorial.

¹¹Fama (1991) halla que, si bien hay una alta relación entre los factores comunes y el re-

lo cual implica encontrar problemas de especificación en la ecuación estimada.

En contraposición, la metodología de ANN¹² se presenta como una buena opción en el contraste del APT. Esto gracias a su estructura no lineal, que admite, entre otras cosas, la entrada de las variables explicativas (variables de entrada) de forma latente; avalando un ajuste que potencia una mejor descripción de los fundamentales macroeconómicos sobre el retorno de los activos al capturar los cambios recónditos en la economía (ruido), el cómo son valoradas la información pública y privada (ineficiencias del mercado), cómo reaccionan los agentes ante caídas (subidas) de los fundamentales y variables financieras (asimetrías), asimismo, esclarece como es evaluada por los inversionistas (regularidades), lo que posibilita una mayor generalización de comportamiento del mercado de valores en comparación a los modelos estadísticos tradicionales como lo muestran Refenes et. al (1995), Ahmaldi (1996), Karyl, et. al (2005), Aristizábal (2006), Jalil y Misas (2007), Santana (2008), Londoño (2009) y Villamil (2009); entre otros.

Más aún, en los mercados emergentes, en comparación a los países desarrollados, el retorno de los valores se presenta con una mayor volatilidad y, en consecuencia, una alta incertidumbre para predecir el comportamiento futuro de estos. Por tal motivo, es relevante contar con un modelo que recoja todos los patrones que puedan afectar a los inversionistas en el momento de la toma de decisiones de inversión, esto con el propósito de diseñar estrategias de cobertura cada vez más precisas ante la exposición al riesgo (Karyl et. al, 2005).

3. Metodología

Antes de proceder cabe efectuar un contraste terminológico entre modelos estadísticos tradicionales y modelos de ANN. En la actual investigación se usaran indistintamente tales significados terminológicos

Tabla 1. Comparación terminológica entre modelos estadísticos tradicionales y ANN

Terminología estadística	Terminología de redes neuronales
Observación	Patrón
Muestra	Datos de entrenamiento
Muestra de validación	Datos de validación, test
Variables independientes	Variables de entrada (input)
Variables dependientes	Variables de salida (output)

Continúa

torno esperado con dicha técnica, esta provoca una pérdida de importantes características de la economía en lo que se refiere a cómo se relacionan estos con el consumo y las oportunidades de inversión.

¹²En la siguiente sección se describe esta metodología.

Modelo	Arquitectura
Residual	Error
Estimación	Entrenamiento, aprendizaje
Coeficientes	Pesos de conexión

Fuente: Tomado de Haykin (1994)

3.1. Red neuronal artificial¹³

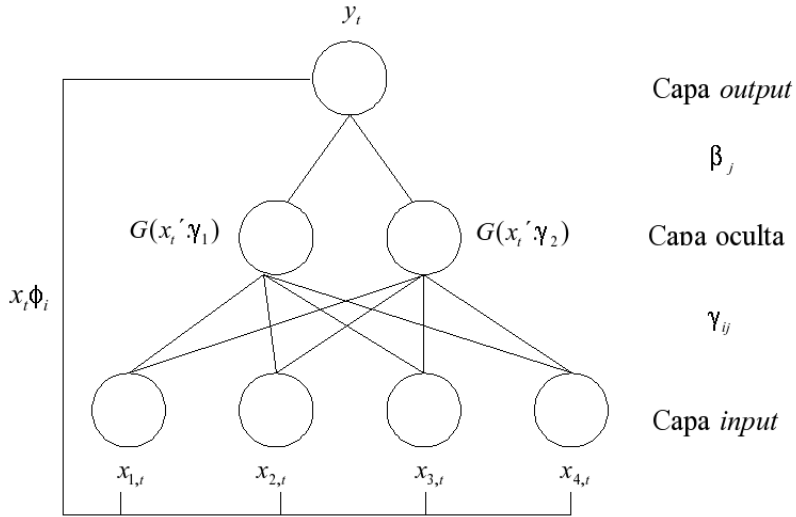
Un ANN, como un mapa universal de funciones no lineales, es un modelo interconectado y retroalimentado, que incorpora dos características: reacción dinámica entre sus unidades, lo cual permite una conexión bidireccional entre todos sus componentes, y un nivel de generalidad, que por medio de un proceso de prueba y error busca obtener el nivel óptimo de activación (Anderson, 2007; Kuan y White, 1994; Londoño, 2009).

Su nomenclatura es ANN(k, q), con k representando el número de variables de entrada utilizadas tanto en la parte lineal como en la no lineal y q el número de neuronas en el modelo. En la Figura 1 se muestra su estructura con $k = 4$ y $q = 2$: en la parte inferior se encuentra la capa de entrada que está compuesta de variables de entrada x_t . Estas, a través de un proceso paralelo son multiplicadas por los parámetros ϕ_i en la parte lineal y por conectores de ponderación (*connection strengths*) γ_{ij} en la parte no lineal. Estos últimos, por medio de su función de activación excitan o inhiben la información que reciben de las variables de entrada para ingresarla de forma asimétrica y óptima en la capa oculta (*hidden layer*).¹⁴

En la capa oculta la combinación lineal $x'_t \gamma_j$ recibe una transformación sigmoideal por medio de su función de activación $G(\bullet)$. Luego multiplicada por las ponderaciones β_j produciéndose el valor de la variable de salida y_t . Este modelo es conocido como red neuronal multicapa de una capa o superficie oculta “aumentado” (“augmented” *hidden layer feedforward network model*).

¹³Para una introducción sobre aspectos básicos sobre modelos de ANN remítase a Villamil y Delgado (2007) y para una literatura más avanzada véase Villamil (2009).

¹⁴El nombre de capa oculta surge porque este factor no es directamente observable ni interpretable. Por tanto, Franses y Dijk (1999) muestran que esta característica es tanto una ventaja como una desventaja, lo primero en lo que respecta a los connection strengths. Lo segundo se debe a que no se sabe con qué fuerza entra cada variable de entrada en la arquitectura. Para solventar tales deficiencias estos autores proponen realizar análisis de sensibilidad que permiten observar como cada variable independiente entra en la ANN.

Figura 1. Red neuronal artificial

Concretamente, el modelo de ANN es expresado como:

$$y_t = F(x_t; \Theta) + \varepsilon_t \quad (3)$$

donde

$$F(x_t; \Theta) = x'_t \phi_i + \sum_{j=1}^q \beta_j G(x'_t \gamma_{ij})$$

siendo y_t la variable de salida, $x_t = (1, \tilde{x}'_t)'$ el vector de variables de entrada, al interior del cual se encuentran p rezagos de y_{t-j} y r variables exógenas $x_{j,t}$ con sus m rezagos. Así, $\phi_i = (\phi_0, \phi_1, \dots, \phi_p, \phi_{11}, \dots, \phi_{1m}, \phi_{21}, \dots, \phi_{2m}, \dots, \phi_{r1}, \dots, \phi_{rm})$ es el vector de parámetros a estimar en la parte lineal, β_j son las ponderaciones de conexión de la capa oculta, $\gamma_{ij} = (\gamma_{0j}, \gamma_{1j}, \dots, \gamma_{pj}, \gamma_{11j}, \dots, \gamma_{1mj}, \gamma_{21j}, \dots, \gamma_{2mj}, \dots, \gamma_{r1j}, \dots, \gamma_{rmj})$ es el vector de pesos con los que se incorpora cada variable de entrada en la capa oculta j ; y $G(\bullet)$ es la función de activación, en este caso es la función logística operando en el intervalo $[0,1]$, cuya expresión es la siguiente

$$G(z) = \frac{1}{1 + \exp(-z)} \quad (4)$$

La ecuación (3) puede ser descrita como una arquitectura del perceptrón multicapa (MLP) siendo una extensión al modelo de regresión clásico, ya que, además de tener un componente lineal, su función de activación utiliza las variables de entrada x_t de forma no lineal en la variable de salida y_t (Kuan y White, 1994). Todo esto sin la necesidad de cumplir supuestos restrictivos como la normalidad, la estacionaridad y la continuidad de las series, además,

no requieren un modelo paramétrico específico (Darrat y Zhong, 2000). Lo que admite capturar las asimetrías, el ruido y en suma toda la dinámica contenida en las variables de entrada.

3.2. Método de estimación

En la literatura de redes neuronales se proponen diversas metodologías de estimación, unas que son “optimizadas” derivando y, otras, libres de derivación.

¹⁵ Ahora bien, estos métodos le permiten a la red adquirir entrenamiento, experiencia y, consecuentemente, la capacidad de predicción de la variable de interés. Así pues, la función objetivo es la siguiente

$$E_n(\Theta) = \sum_{t=1}^n [y_t - F(x_t; \Theta)]^2 \quad (5)$$

siendo $\Theta = \{\phi_i, \gamma_{ij}, \beta_j\}_{k+1+q(k+2)}$ el vector de parámetros. Para evitar que la ecuación (5) tenga altos valores en algunos de sus parámetros estimados, Franses y Dijk (1999) sugieren introducir una medida de penalización conocida como decaimiento de los pesos (*weight decay*), lo que hace que la función objetivo sea

$$E_n(\Theta) = \sum_{t=1}^n [y_t - F(x_t; \Theta)]^2 + r_\phi \sum_{i=0}^k \phi_i^2 + r_\beta \sum_{j=1}^q \beta_j^2 + r_\gamma \sum_{j=1}^q \sum_{i=0}^k \gamma_{ij}^2 \quad (6)$$

siendo r_ϕ, r_β y r_γ los parámetros de decaimiento de los pesos, los cuales tienen una propiedad de gran relevancia en las estimación del modelo de ANN, la cual es permitir que el modelo no presente un sobreajuste en su proceso de estimación, lo que posibilita una minimización más óptima de los errores. Para que lo anterior tenga sentido las variables de entrada deben ser comparables en magnitud, por lo tanto se requiere la transformación de escala al intervalo $[0,1]$ (Franses y Dijk, 1999). Ahora, para optimizar la ecuación (6) se utilizará el algoritmo quasi - Newton,¹⁶ que está definido como la segunda derivada de la función objetivo $E_n(\Theta)$ ¹⁷

$$E(\Theta) \approx E(\Theta_{now}) + g(\Theta)'(\Theta - \Theta_{now}) + \frac{1}{2}(\Theta - \Theta_{now})'H(\Theta - \Theta_{now}) \quad (7)$$

¹⁵En cuanto a las primeras están el *steepest descent*, *backpropagation*, el gradiente conjugado, delta bar delta, *quickprop* y el método Newton (dividido en cuatro: Clásico, modificado, quasi - Newton y Gauss - Newton). Con respecto a las segundas, se encuentran algoritmos genéticos, *simulate annealing*, *random search* y *downhill simples search* (ver Haykin, 1994; Jang et. al, 1997; y Martin Del Brio y Sanz, 2001, para más detalles).

¹⁶De acuerdo con Franses y Dijk (1999) el algoritmo de estimación *backpropagation*, ampliamente usado en la literatura, ha mostrado ser inferior a otros algoritmos no lineales como puede ser el quasi-Newton. Véase el anexo 1 para una explicación de cómo funciona este primero.

¹⁷Para que se cumpla (6) es necesario suponer que Θ_{now} está suficientemente cercano a un mínimo local. Además, se omiten los términos de alto orden, debido al supuesto de que $\|\Theta - \Theta_{now}\|$ es considerablemente pequeño.

con

$$g(\Theta) = \nabla E(\Theta) = \left[\frac{\partial E(\Theta)}{\partial \Theta_1}, \frac{\partial E(\Theta)}{\partial \Theta_2}, \dots, \frac{\partial E(\Theta)}{\partial \Theta_n} \right]$$

aquí, los subíndices (*now*) y (*next*) representan respectivamente el valor inicial y final de los pesos, $g(\Theta)$ es el gradiente de la función objetivo en primeras diferencias; y H es la matriz Hessiana, la cual actualiza la información procesada iterativamente por el método, en este caso, el propuesto por Broyden–Fletcher–Goldfarb–Shanno en 1970, que es conocido como el método BFGS (véase a Jang et al., 1997, para más detalles).

4. Datos

Las series estadísticas utilizadas se obtuvieron de la página Web del Grupo Aval. El periodo considerado va desde el 6 de mayo de 2004 al 19 de mayo de 2009, para un total de 1176 observaciones. Debido a que los mercados de activos son tan cambiantes en el corto plazo, dada su alta volatilidad, se utilizarán únicamente 60 observaciones para evaluación, esto con el objetivo de mostrar cómo se podría emplear este esquema de valoración de la forma más veraz. Esto es, cuando se realiza un modelo estadístico lo que se propone es usar el 70 % para estimación y el 30 % para evaluación, no obstante, dado que son tantos datos el modelo especificado puede cambiar su grupo de información (variables explicativas), lo cual puede hacer a este modelo poco práctico en su aplicación real en el momento de hacer pronósticos por fuera del periodo de entrenamiento. Ahora bien, estas variables fueron transformadas por medio de la fórmula $r_t = \ln(x_{jt}) - \ln(x_{j,t-1})$. Luego fueron estandarizadas en el intervalo $[0,1]$, esto último con el objetivo de “mejorar las propiedades del método numérico de estimación no lineal” (Franses y Dijk, 1999).

En la tabla 2 se encuentran resumidas todas las variables. Con respecto a la variable objetivo se seleccionó el Índice General de la Bolsa de Valores de Colombia (*IGBC*). Su importancia como variable dependiente reside en que incluye acciones que poseen características deseables para los inversionistas, como son una alta capitalización bursátil (esto es un valor adecuado de las empresas a precio de mercado) y una buena liquidez, indicando que tales acciones se transan fácilmente en el mercado de valores. Igualmente, es una herramienta que sirve para gestionar profesionalmente los portafolios, debido a que ofrece referencias transparentes de desempeño de las empresas, y avala una representación óptima del universo de posibilidades de inversión, lo que permite la minimización del riesgo de una forma más eficiente (fuente: www.bvc.com.co).¹⁸

En cuanto a los factores de riesgo se eligieron 20, los cuales están justificados principalmente por tres aplicaciones empíricas: la primera es la de Chen et. al (1986) quienes muestran que los fundamentales macroeconómicos son variables significativas para explicar el comportamiento del mercado de valores estadounidense. La segunda es el trabajo de Koutoulas y Kryzanowsk (1994),

¹⁸Otros contrastes del APT que han usado índices de bolsas como portafolios son Burmeister y McElroy (1987) y Koutoulas y Kryzanowsk (1994).

quienes hallan que el mercado de valores canadiense se ve afectado por una serie de factores económicos, entre ellos los fundamentales macroeconómicos de los Estados Unidos, encontrando que hay una integración entre ambos mercados. Por último, Shaharudin y Fung (2009) analizan el mercado de valores de Malasia, y vislumbran la importancia de las variables microeconómicas y macroeconómicas en los mercados emergentes.

Tabla 2. Variables utilizadas para el contraste empírico del APT para Colombia

Variable	Símbolo	Método de agregación	Variable transformada
Índice General de la Bolsa de Valores de Colombia	$IGBC_t$		
Letras del tesoro de EE.UU.			
6 meses	$T6_t$		
2 años	$T24_t$	Componentes	
5 años	$T60_t$	Principales	$Bonos_t$
10 años	$T120_t$	Asintóticos	
30 años	$T360_t$		
Índices de Bolsas del Mundo			
Brasil	$Bovespa_t$		
EE.UU.	$DowJones_t$		
EE.UU.	$Nasdaq_t$		
EE.UU.	$S\&P500_t$	Componentes	
Frankfurt	DAX_t	Principales	$Indice1_t$
Hong Kong	$HangSeng_t$	Asintóticos	
Londres	$Ptse100_t$		

Continúa

Madrid	IBEX _t		
Paris	CAC40 _t		
Tokio	Nikkei _t		
Monedas		Componentes	
Tasa representativa del mercado	TRM _t	Principales	Moneda _t
Euro/peso	TRMUE _t	Asintóticos	
Tasas de interés			
Tasa interbancaria (Colombia)	TIB _t		
Proxy de la inflación en Colombia			
Unidad de Valor Real	UVR _t		
Prima de Riesgo			
Á Emergent Market Bond Index	EMBI _t		

Con las anteriores variables se construyeron tres componentes principales. El primero está compuesto por los rendimientos de las letras del tesoro de los Estados Unidos. Esta funciona como una *proxy* de la tasa de interés de referencia de la Reserva Federal. Arango et. al (2007) realizan un estudio sobre curvas de rendimientos de bonos colombianos y encuentran que para explicar el comportamiento de tales curvas no sólo se requiere emplear argumentos tales como las tasas de interés de Colombia, sino también las tasas de interés de los Estados Unidos. De este modo se encuentra una importante conexión entre las decisiones tomadas por los inversionistas en el mercado de valores y las decisiones tomadas por los encargados de la política monetaria. Así pues, se extrajo el primer componente, que explica el 68.94 % de la varianza total, y que está dado por

$$\begin{aligned} \text{Bonos}_t = & 0.3457 * T6_t + 0.4680 * T24_t + 0.4937 * T60_t + 0.4674 * T120_t \\ & + 0.4462 * T360_t \end{aligned}$$

Aquí todos los pesos son positivos y cercanamente ponderados. Denominamos a esta variable Componente del Movimiento General del Retorno de las Letras del Tesoro.

El segundo componente contiene los índices de las bolsas de valores más importantes del mundo, y muestra el comportamiento del mercado mundial

de valores. Dada la integración del mercado accionario colombiano a los principales mercados financieros del mundo, se espera una relación positiva entre éste componente y el IGBC. Además, este refleja el comportamiento de la producción industrial mundial a través del desempeño de las empresas inscritas en cada bolsa (Chen et. al, 1986). Para este caso el primer componente explica una variabilidad del 60.87 % y se denomina Retorno del Índice General de las Bolsas del Mundo.

$$\begin{aligned}\text{Indice1}_t = & 0.3185 * \text{Bovespa}_t + 0.3309 * \text{DowJones}_t + 0.3164 * \text{Nasdaq}_t \\ & + 0.3549 * \text{Dax}_t + 0.2162 * \text{HangSeng}_t + 0.3460 * \text{Ptse100}_t \\ & + 0.3441 * \text{IBEX}_t + 0.3476 * \text{CAC40}_t + 0.2117 * \text{Nikkei}_t\end{aligned}$$

El tercer componente es un componente de monedas, cuya varianza total es del 83.84 %. Lo consideramos un índice general, denominado Movimiento General del Mercado de divisas en Dólares y Euros

$$\text{Moneda}_t = 0.7071 * \text{TRM}_t + 0.7071 * \text{TRMUE}_t$$

La relevancia de este factor como predictor del comportamiento del retorno de los activos es establecida por Jalil y Misas (2007) quienes exponen que el tipo de cambio es determinado por las preferencias (grado de aversión) de los diferentes agentes que participan en la economía global. Por tanto, las decisiones de inversión estarán influenciadas por los cambios en el precio de la divisa. En el caso de que esta presente un comportamiento demasiado volátil,¹⁹ el riesgo y la incertidumbre de los agentes aumentará, produciendo secuelas negativas en el retorno de los activos. Por otro lado, la tendencia de la moneda puede condicionar las decisiones de inversión en los mercados de valores. Dejando todo lo demás constante, si estas tienen una tendencia esperada a apreciarse (depreciarse), esto influirá en la compra (venta) de activos financieros de un país.

La *TIB*, como proxy de las decisiones de política monetaria, es un indicador de la liquidez del mercado que está relacionada negativamente con el precio de las acciones debido a tres razones: primera, altas tasas de interés reducen los beneficios de las empresas al aumentar los costos de los préstamos. Segunda, desplaza la inversión del sector productivo a otras actividades más rentables. Tercera, reducen el retorno sobre la inversión, puesto que la compra de acciones se hace más costosa (Shajarundi y Fung, 2009).²⁰

La *UVR* como una *proxy* del nivel de precios agregado, tal y como la *TIB*, tiene una relación negativa con respecto al retorno de las acciones (Chen et. al,

¹⁹Reconociéndose que los mercados de divisas normalmente tienen un componente volátil, pero en distintas proporciones.

²⁰Si bien, la política monetaria tiene efectos sobre el flujo de caja esperado de las firmas, estos efectos son asimétricos según sea su tamaño. Por ejemplo, dado un endurecimiento de las tasas de interés una firma grande no tendrá tantas restricciones de crédito al momento de adquirir o pagar sus préstamos como lo tendrá una empresa pequeña (Shajarundi y Fung (2009) y Thorbecke (1997). Igualmente, las empresas grandes tienen información más veraz en el momento de realizar un préstamo, lo que les permitirá tomar decisiones más acertadas sobre cuando y cuanto invertir (Ehrmann y Fratzscher (2004).

1986), ya que influye sobre el retorno esperado de las empresas y, por tanto, sobre las posibilidades de inversión. Si los inversionistas, tanto en el sector real como en el financiero, esperan una inflación alta, esto puede ocasionar volatilidad e incertidumbre en el consumo y la actividad real futura (Edwing, 2002), a causa de que la confianza en el manejo macroeconómico por parte de la autoridad monetaria no es creíble, lo cual provoca un desplazamiento de la inversión a otro tipo de instrumentos financieros y economías que tengan condiciones más propicias para invertir.

El *EMBI* es una medida de riesgo país, determinada por la sobretasa que pagan los bonos de un país emergente en relación a los bonos de los Estados Unidos (probabilidad de incumplimiento de los pagos de deuda externa). Esta permite establecer la rentabilidad y, de este modo, el grado de aversión al riesgo de los inversionistas a invertir en dicho país. La medida tiene una relación incierta con el retorno de los activos de tales países, puesto que este es afectado por dos vías diferentes. Si esta es excesivamente alta implica condiciones económicas, sociales, políticas y geográficas inestables, lo cual produce una salida de capitales (o una imposibilidad a la entrada de estos), presentándose un bajo crecimiento económico (fuente: BBC). Si el anterior efecto no se cumple, tendrá secuelas positivas, debido a que entre mayor sea el riesgo esperado por el inversionista en un país específico, mayor será su rentabilidad.

5. Resultados

En esta sección se elaboraron dos modelos diferentes, uno en el que las variables explicativas sólo son los fundamentales macroeconómicos (*APT*) y otro que incluye el indicador de las bolsas del mundo (*APT* modificado), esto con el propósito de investigar la posible superioridad predictiva del modelo *APT* en su forma tradicional con respecto al *APT* modificado (Chen et. al, 1986). En otras palabras, si efectivamente se satisface el predominio empírico de tal modelo para el mercado de valores colombiano, o lo contrario, se cumple otro tipo de hipótesis como es la del modelo de Sharpe (1964).

Como variables independientes de la red, se consideraron variables contemporáneas y sus seis primeros rezagos, este conjunto de información fue seleccionado buscando capturar la tendencia de corto plazo del mercado que se supone relativamente eficiente en el descuento de la información. Así pues, se eligió el mejor modelo de cada versión mediante el procedimiento *stepwise* (PSSVS) de selección secuencial de variables.²¹ Por tanto, esta sección a su

²¹En el caso en que no existan variables exógenas en el modelo, Frances y Dijk (1999) proponen elegir k y q mediante la siguiente estrategia: estimar todos los posibles ANN y elegir el óptimo según los criterios de selección de modelos dentro y fuera de la muestra. Aquí, el número de variables ocultas (o neuronas) y el número de variables de entrada se eligen paralelamente. En este caso, debido a la inclusión de variables exógenas se usan todas estas tanto en la parte lineal y no lineal, dado que la selección recursiva de k y q podría traer grandes costos en el proceso de estimación. Igualmente, se quiso tener un grupo de información que pudiera capturar los diferentes choques en el mercado de valores, es decir, se buscó que el modelo procurase ser estable en sus variables en sus distintos horizontes para de esta forma ser una herramienta empírica consistente con las necesidades que se tienen de

vez se divide en dos, una que muestra todos los resultados del modelo APT y otra en la cual se comparan estos resultados con el modelo APT modificado.

5.1. Modelo APT

Tras aplicar el PSSVS se llega al siguiente modelo lineal

$$\text{IGBC}_t = x'_t \phi_i + \varepsilon_t \quad (8)$$

con

$$x'_t = [1\text{Bonos}_t \text{Bonos}_{t-5} \text{Moneda}_t \text{Moneda}_{t-2} \text{Moneda}_{t-5} \text{TIB}_{t-5} \text{UVR}_{t-5} \\ \text{UVR}_{t-6} \text{EMBI}_t \text{EMBI}_{t-2} \text{EMBI}_{t-3}]$$

donde, x'_t representa un vector de factores de riesgo sistémico. Si bien el modelo de ANN puede tener una capacidad predictiva superior a los modelos lineales, esto puede ser falso si las variables de entrada no tienen un comportamiento no lineal con respecto a la variable de salida (Franses y Dijk, 1999). Por consiguiente, se realizó la prueba de no linealidad de Lee, White y Granger (1993), cuyos resultados se resumen en la tabla 3, y se rechaza la hipótesis nula de no linealidad

Tabla 3. Resultados de la prueba de no linealidad para el modelo APT

Replicaciones	Periodo de entrenamiento	Modelo seleccionado	P-Valor
10	06/05/04-05/02/09	APT	0.000257

Fuente: Cálculos Propios

Por consiguiente, el modelo APT a estimar, bajo el esquema de ANN es el siguiente

$$\text{IGCB}_t = x'_t \phi_i + \sum_{j=1}^q \beta_j G(x'_t \gamma_{ij}) + \varepsilon_t \quad (9)$$

Para establecer que el vector de riesgos sistémicos x'_t está capturando cabalmente el comportamiento del IGBC_t , se recurre a la metodología de pronósticos “rolling” que muestra la capacidad predictiva del modelo a lo largo del tiempo, y también permite observar como dichas variables recogen el comportamiento del IGBC_t en sus distintos tramos. En otros términos, si se mantiene una buena tendencia explicativa, o por el contrario, las variables conjuntamente van perdiendo importancia en su predicción.

Con ese propósito se debe elegir el mejor modelo de ANN, que es aquel en el cual el número q de neuronas permite que la red aprenda, en lugar de predecir el comportamiento futuro de los activos.

memorizar. Por tanto, para entrenar la ecuación (9) se usará $q = 2$ hasta $q = 13$.²² Dado que los últimos 60 datos son para evaluación del modelo, con cada horizonte de pronósticos de 5 observaciones, se necesitó estimar 12 arquitecturas distintas por neurona. No obstante, debido a que la minimización de la función $E_n(\Theta)$ en (6) puede arrojar muchos mínimos locales, independientemente de la convergencia de los estimadores, fue necesario, como lo proponen Franses y Dijk (1999), estimar cada arquitectura 30 veces, en cada horizonte de pronóstico, seleccionando aquella que arrojara el error cuadrático medio más bajo.

Por lo tanto, se evaluó cada arquitectura dentro y fuera del periodo de entrenamiento con los estadísticos habituales de selección de modelos (ver Aristizábal, 2006 y Londoño, 2009, para más detalles). Esto con el fin de seleccionar el mejor modelo, ya que si hay un exceso de neuronas, el modelo estimado puede llegar a presentar un sobreajuste, provocando la memorización de los patrones e imposibilitando la predicción del comportamiento futuro de la variable de salida. Inversamente, pocas neuronas pueden traer como consecuencia un proceso incompleto de aprendizaje, presentándose una mala especificación. Por tal razón, es necesario que haya un equilibrio entre la evaluación de la arquitectura dentro y fuera del periodo de entrenamiento. Los estadísticos utilizados para evaluar el desempeño dentro de la muestra, tabulados en el Apéndice 1, muestran la superioridad del modelo ANN para $q = 2.6$ y 9 en comparación al modelo lineal.

Siguiendo a Aristizábal (2006) y Jalil y Misas (2007) se realizan los pronósticos por tramos. Como se observa en la Figura 2, la primera estimación utiliza los datos 1 a 1116, y pronostica de 1117 a 1121, la segunda usa los datos 5 a 1121 y pronostica de 1122 a 1126 y así sucesivamente hasta completar el periodo de evaluación.

De esta forma se evalúa la consistencia del modelo en los h horizontes de tiempo, con $h = 1, 2, \dots, 12$, por arquitectura $j = 1, 2, \dots, 12$. Esta estrategia produce 12 modelos por cada neurona seleccionada con su respectivo vector de pronósticos, las cuales son evaluadas con medidas simétricas y asimétricas de selección de modelos fuera del periodo de entrenamiento.

Como lo señalan Aristizábal (2006) y Jalil y Misas (2007) hay varias medidas de error de pronóstico ε_{t+h} . Por un lado, se dan las medidas simétricas de pronósticos, la más habitual es la diferencia entre el valor predicho $\hat{y}_{t+h}$ y el valor observado y_{t+h}

$$\varepsilon_{t+h} = \hat{y}_{t+h} - y_{t+h} \quad (10)$$

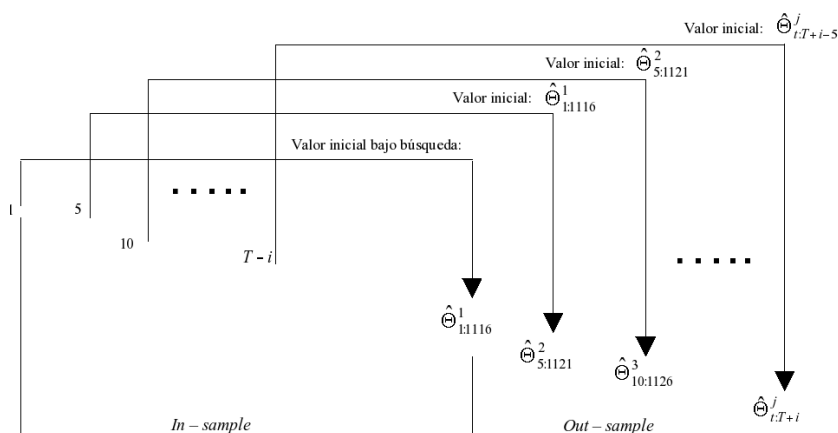
Otras medidas que resultan de la ecuación (10), son el error absoluto,

$$AE_{t+h} = |\hat{y}_{t+h} - y_{t+h}| \quad (11)$$

y el cuadrático

$$SE_{t+h} = (\hat{y}_{t+h} - y_{t+h})^2 \quad (12)$$

²²Una regla que se sigue en modelos ANN aplicados a pronósticos financieros es la de estimar varias arquitecturas cambiando el número de neuronas, con $q = 1$ hasta $q = q^*$, siendo q^* el número de variables de entrada + 1, y se selecciona el mejor modelo por medio de medidas de desempeño (Londoño, 2009).

Figura 2. Esquema *Rolling* de pronósticos

Estas presentan las siguientes características: son simétricas, en la medida que sancionan con igual peso cada error de pronóstico sin importar lo alejado que esté del origen de pronóstico. Asimismo, el grado de penalización es mayor cuando se dan grandes desviaciones del valor del pronóstico con respecto al valor observado. No obstante, tales medidas desconocen los costos que acarrearía tener un valor predicho por encima o por debajo del valor observado. Por ejemplo, en el mercado de valores diariamente se toman decisiones de inversión de compra y venta de activos financieros, estas decisiones no se realizan de forma determinística, de manera que si no coincide lo predicho con lo observado, se generarán costos que son asimétricos.

Por ende, los inversionistas, como agentes aversos al riesgo y optimizadores de su riqueza, buscarán seleccionar la opción que les ofrece la mayor utilidad esperada dada la información disponible. De acuerdo a esto, será más costoso que un activo sea sobrevalorado que subvalorado. Por tanto, cada resultado se penalizará de forma asimétrica. Con tal propósito, se evaluará el modelo con una función de costos asimétrica. Siguiendo a Aristizábal (2006) y Jalil y Misas (2007) se utilizó la función LINLIN, cuya forma es lineal a ambos lados del valor objetivo, pero de manera asimétrica, determinada por la razón a/b , su fórmula es

$$LLC \{ \hat{y}_{t+h}, y_{t+h} \} = \begin{cases} a|\hat{y}_{t+h} - y_{t+h}| & \text{para } \hat{y}_{t+h} < y_{t+h} \\ 0 & \text{para } \hat{y}_{t+h} = y_{t+h}, b > a \\ b|\hat{y}_{t+h} - y_{t+h}| & \text{para } \hat{y}_{t+h} > y_{t+h} \end{cases} \quad (13)$$

En este caso a representa los costos en los que se incurre por la subvaloración y b representa el costo de la sobrevaloración. Además, a y b son parámetros suavizadores que penalizan en menor medida los valores pronosticados más alejados del origen de pronóstico, obteniéndose una mejor evaluación. Así, se valoró cada red fuera de la muestra con medidas simétricas y asimétricas para cada horizonte de pronóstico. En el Apéndice 2 y 3 se resumen los resultados

para el modelo APT aplicando OLS y ANN, respectivamente. Allí, se puede observar que en general el modelo de ANN tiene una capacidad explicativa superior al modelo lineal. Asimismo, se puede ver que tales medidas son estables de forma conjunta en los distintos horizontes.

Ahora bien, para probar la significancia de las variables al igual que la dirección de su signo, Franses y Dijk (1999) proponen efectuar análisis de sensibilidad que determine la influencia de cada variable de entrada sobre la variable de salida. Para esto se toman diferentes valores de $x_{i,t}$ y se fija el resto de variables de entrada $x_{1,t}, x_{2,t}, \dots, x_{i-1,t}, \dots, x_{k,t}$ en su valor medio en la siguiente derivada

$$\frac{\partial y_t}{\partial x_{i,t}} = \phi_i - \sum_{j=1}^q \beta_j \gamma_{ij} [G(x'_t \gamma_j)] [1 - G(x'_t \gamma_j)] \quad (14)$$

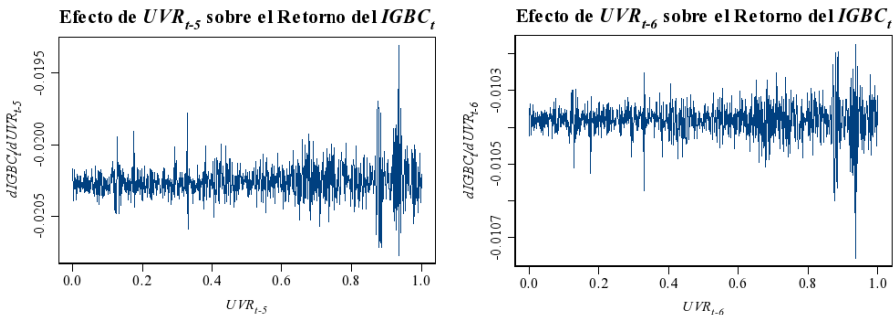
con

$$\frac{\partial G(x'_t \gamma_j)}{\partial x_{i,t}} = -\gamma_{ij} [G(x'_t \gamma_j)] [1 - G(x'_t \gamma_j)]$$

Su interpretación es que entre mayor sea $\frac{\partial y_t}{\partial x_{i,t}}$ en términos absolutos, $x_{i,t}$ tendrá una mayor efecto sobre y_t . En este trabajo se realizará el análisis de sensibilidad considerando el modelo con 2 neuronas dados sus buenos resultados (Apéndice 3).

En la figura 3, sus dos paneles muestran el efecto que tiene la variable UVR rezagada 5 (UVR_{t-5}) y 6 periodos (UVR_{t-6}) sobre el retorno del $IGBC_t$. En este caso, al efecto de la UVR presenta el signo esperado. Esto se explica por la aversión de los agentes a tener activos que sean más sensibles a los cambios nominales (Chen et. al, 1986), lo cual influye las perspectivas de inversión real y, en consecuencia, el valor futuro de las acciones. También se vislumbra en esta figura que, como es de esperarse, los inversionistas le dan mayor peso a la información más inmediata (UVR_{t-5}) que a la información pasada (UVR_{t-6}).

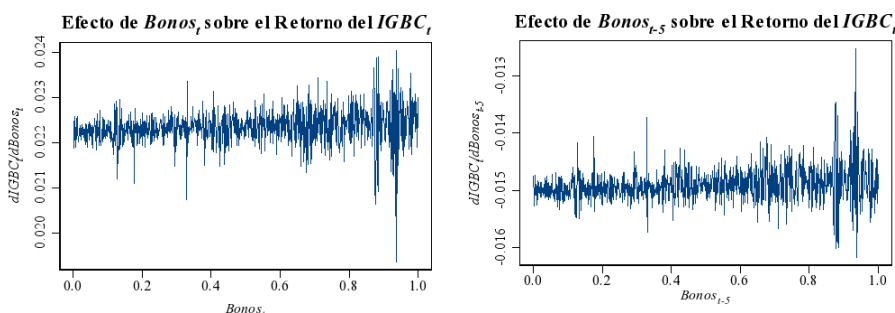
Figura 3. Efecto de UVR en sus distintos rezagos



En el caso de $Bonos_t$ en la figura 4 se presenta una relación positiva que no coincide con lo esperado, según la cual ésta debería estar afectando negativamente el retorno del $IGBC_t$, dado que ante un aumento de la tasa de interés

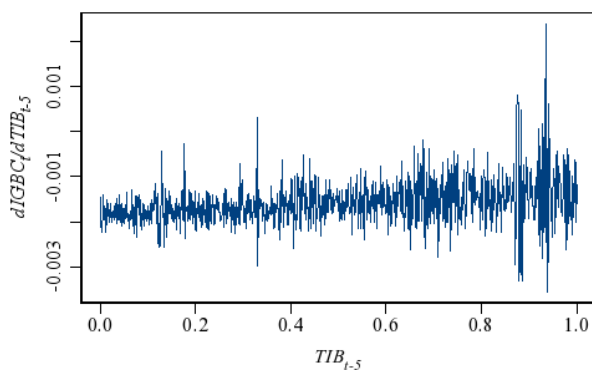
externa se produce un desincentivo a la inversión en empresas nacionales, reduciéndose el retorno del portafolio de acciones mantenido. En lo concerniente a $Bonos_{t-5}$ se da el efecto esperado dado que un aumento de la tasa de interés extranjera provoca una recomposición del portafolio hacia activos que generen una mayor rentabilidad a más largo plazo (bonos). Asimismo, del panel derecho se observa que el efecto se mantiene muy estable, alrededor de -0.014, aunque a medida que el valor de $Bonos_{t-5}$ aumenta se da un aumento de la volatilidad en su efecto.

Figura 4. Efecto de $Bonos$ en sus distintos rezagos



Siguiendo con la misma línea, el signo de TIB_{t-5} coincide con lo esperado, pero su efecto es muy bajo. Esto se explica en parte por las pocas intervenciones hechas en ese periodo, a la alta credibilidad que tiene el Banco Central después de la implementación de la inflación *targeting* que no provoca muchas sorpresas en las expectativas de los agentes que anticipan la información, entre otros factores (ver Arango et. al, 2007, para más detalles) que determinan las perspectivas de inversión de los agentes.

Figura 5. Efecto de TIB_{t-5} sobre el Retorno del $IGBC_t$

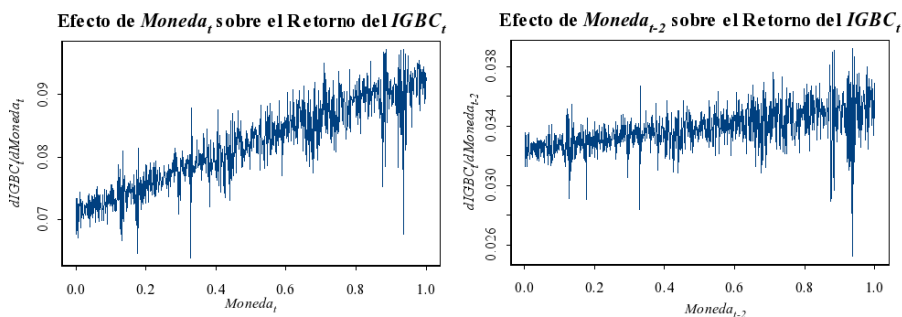


Por otra parte, los efectos de las variables $Moneda_t$ y $Moneda_{t-2}$ de la figura 6 tienen el signo esperado. Esto muestra la gran importancia que tiene el

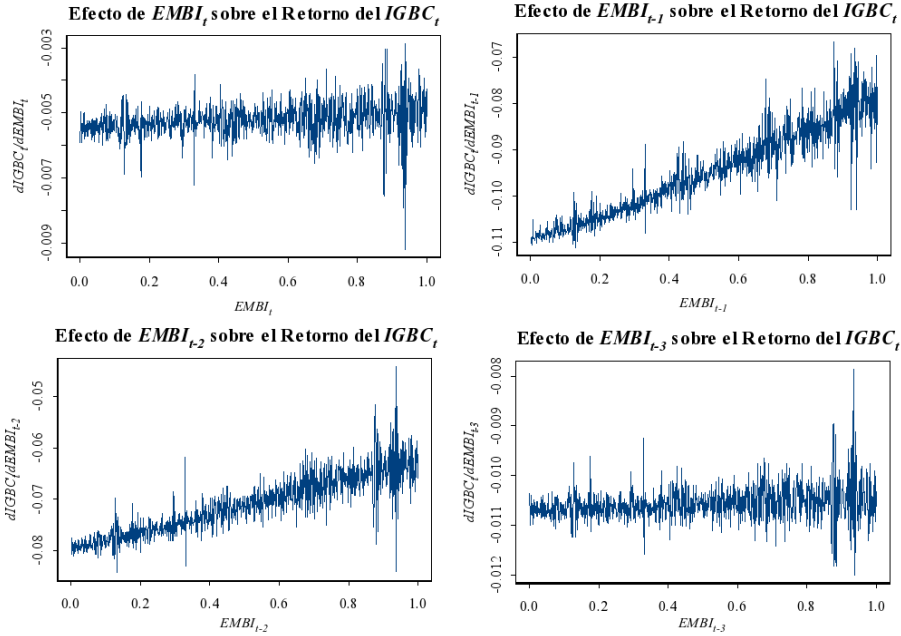
comportamiento del mercado de divisas en la toma de decisiones de inversión por parte de inversionistas nacionales y extranjeros en el mercado accionario nacional, en otras palabras, la actividad de inversión financiera está altamente influenciada por el comportamiento del mercado de divisas, que muestra en última instancia el ambiente de confianza en los fundamentales macroeconómicos, determinando la entrada y salida de capitales.

En cuanto al efecto del $EMBI$ sobre $IGBC_t$, este tiene el signo esperado, es decir, ante un aumento del riesgo país se espera una disminución en el retorno del portafolio de mercado, ya que habrá una salida de capitales hacia otras economías más confiables. Con relación a la figura 7, se puede decir que los inversionistas valoraran en mayor proporción el efecto que tienen $EMBI_{t-1}$ y $EMBI_{t-2}$ sobre el $IGBC_t$, con respecto a $EMBI_t$ y $EMBI_{t-3}$, lo cual podría explicarse por las expectativas de los agentes, quienes esperarían cierto tiempo para esclarecer los efectos que tiene la subida de la prima de riesgo y así tomar una nueva posición financiera.

Figura 6. Efecto de *Moneda* en sus distintos regazos



En resumen, el análisis de sensibilidad mostró que las variables más relevantes para explicar los retornos del $IGBC$ son *Monedas* y $EMBI$, las cuales tienen los mayores efectos en términos absolutos, seguidos por *Bonos*, UBR y TIB . Por tanto, tales variables son adecuadas para explicar el comportamiento del $IGBC$. Sin embargo, faltaría observar que efectos puede tener la introducción de variables financieras como el *Indice1* sobre la capacidad predictiva del modelo y la significancia individual de las variables. El propósito de hacer esto es el de poder determinar cual de los dos modelos (APT y APT modificado) predomina en la descripción del comportamiento del mercado accionario colombiano.

Figura 7. Efecto de $EMBI$ en sus distintos regazos

5.2. Comparación del modelo APT y APT modificado

Según Chen et. al (1986) los índices de los mercados financieros pueden sesgar los resultados, ya que su movimiento puede tener efectos más inmediatos en las decisiones de los inversionistas que las variables macroeconómicas. En la ecuación (15) se puede observar los resultados del PSSV introduciendo el *Indice1*. Por lo que el modelo queda como sigue

$$IGBC_t = z'_t \phi_i + \varepsilon_t \quad (15)$$

con

$$z'_t = [1 \text{ Bonos}_{t-3} \text{ Bonos}_{t-6} \text{ Indice1}_t \text{ Indice1}_{t-1} \text{ Indice1}_{t-3} \text{ Moneda}_{t-2} \text{ UVR}_{t-5} \text{ UVR}_{t-6}]$$

donde z'_t representa un vector con factores de riesgo sistemáticos. Por tanto, el modelo APT bajo el esquema de ANN a estimar es el siguiente

$$IGBC_t = z'_t \phi_i + \sum_{j=1}^q \beta_j G(z'_t \gamma_{ij}) + \varepsilon_t \quad (16)$$

Comparando los dos modelos APT modificado (lineal y ANN) en lo que respecta a sus medidas dentro del periodo de entrenamiento (Apéndice 4) y

evaluación de pronósticos del modelo lineal (Apéndice 5) y ANN (Apéndice 6) muestran la misma tendencia observada en el modelo APT, en otras palabras, se sigue dando la superioridad de este último sobre el primero. En cuanto al análisis de sensibilidad para el modelo de la ecuación (16) (véase Apéndice 7), se muestra que es la variable más importante, dándose lo esperado por Chen et. al (1986) en lo que respecta al sesgo de las demás variables y, por consiguiente, mostrando la importancia que tiene el mercado de valores mundial sobre el mercado accionario colombiano. Así, globalmente, se encontró que la variable *Indice1* tiene el mayor poder explicativo. En consecuencia, como es documentado por Chen et. al (1986), las variables macroeconómicas, aunque sean importantes para explicar el retorno de los activos, pierden dinámica cuando se consideran conjuntamente con variables financieras, debido a que estas últimas capturan los movimientos del mercado de valores más rápidamente.

Ahora bien, comparando los resultados de los modelos APT (Apéndice 1 y 3) y APT modificado (Apéndice 4 y 6) bajo el esquema de ANN se puede observar que el primero presenta en general mejores estadísticos y, por tanto, la superioridad de este modelo sobre el APT modificado. Además, según el análisis de sensibilidad los fundamentales macroeconómicos del APT tienen mayores efectos con respecto a los del modelo APT modificado. Encontrándose en definitiva que el modelo APT en su forma tradicional se ajusta mejor para predecir el comportamiento del mercado de valores colombiano.

6. Conclusiones

Este trabajo ha estudiado un grupo de variables macroeconómicas y financieras como una tentativa para explicar el retorno del mercado accionario colombiano, haciendo énfasis principalmente en los trabajos de Chen et al (1986), Koutoulas y Kryzanowski (1994) y Shaharudin y Fung (2009). Se encontró que aunque las variables financieras tienen un alto poder explicativo en relación a los fundamentales macroeconómicos, se da una dominancia del modelo APT en relación al APT modificado, tales resultados son vistos a través del análisis de sensibilidad y el esquema “*Rolling*” de pronósticos. El modelo de ANN permitió llegar a resultados más sobresalientes a los que se podría llegar con modelos estadísticos tradicionales, teniendo una mejor capacidad explicativa y precisión, como lo muestran sus medidas simétrica y asimétrica de evaluación.

De igual manera, el análisis de sensibilidad hizo posible un estudio más adecuado para explicar el comportamiento del retorno del *IGBC* ante los distintos choques no lineales que pueden ocasionar las variables macroeconómicas, estableciendo el nivel de activación, influencia y preponderancia, según sean las variables. Sin embargo, estos efectos deben de tener un estudio más detallado para una mejor comprensión del comportamiento del mercado de valores, con la intención de desarrollar nuevas hipótesis. Con tal objeto, sería también conveniente diseñar estadísticos que no sólo capturen el efecto individual que las variables de entrada tienen sobre la variable de salida, sino también estadísticos que tengan en cuenta un conjunto de variables homogéneas (cada variable explicativa en forma rezagada) para determinar el efecto total de tal variable

sobre la de salida y conocer de este modo su impacto global.

Por otro lado, los contrastes empíricos del APT, citados previamente en este trabajo utilizan series de tiempo mensuales y encuentran una baja e incluso nula significancia estadística en la variable índice de mercado (según sea su construcción). En nuestra investigación se presenta el efecto contrario, pues al utilizar datos diarios muestra que ésta variable presenta una amplia capacidad explicativa, sesgando los resultados en contra de los fundamentales macroeconómicos, al ser la primera un determinante clave de las decisiones de inversión en el mercado de valores (Chen et. al, 1986).

Referencias

- Agudelo, D. A., Arango, M. A. (2008). “La curva de rendimientos a plazo y las expectativas de tasas de interés en los mercados colombianos de renta fija, 2002-2007” . *Lecturas de Economía* 68, 39–66.
- Ahmadi, H. (1996). “Testability of the arbitrage pricing theory by neural networks” . En Trippi, R. R., Turban, E. (Editores). *Neural Networks in Finance and Investing: Using artificial intelligence to improve real – world performance*. McGraw Hill Interamericana; 543–554.
- Altay, E. (2003). “The effect of macroeconomic factors on asset returns: A comparative analysis of the German and the Turkish stock markets in an APT framework” . *Betriebswirtschaftliche Diskussionsbeiträge* 48, 38.
- Anderson, J. A. (2007). *Redes Neuronales*. Alfaomega Grupo Editor: México.
- Arango, L. E., González, A., León, J. J., Melo, L. F. (2007). “Cambios en la tasa de intervención y su efecto en la estructura a plazo de Colombia” . Banco de la República. Gerencia Técnica. Unidad de Estudios Económicos 424.
- Aristizábal, M. C. (2006). “Evaluación asimétrica de una red neuronal artificial: Aplicación al caso de la inflación en Colombia” . *Lecturas de Economía* 65, 73 – 116.
- Botero, S., Cano, J. A. (2008). “Análisis de series de tiempo para la predicción de los precios de la energía en la bolsa de Colombia” . *Cuadernos de Economía* 27 (48), 173 – 208.
- Castaño, E., Gómez, K., Gallón, S. (2008). “Pronóstico y estructuras de volatilidad multiperiodo de la tasa de cambio del peso colombiano” . *Cuadernos de Economía* 27 (48), 2 – 26.
- Chen, N. F., Roll, R., Ross, S. A. (1986). “Economic forces and the stock market” . *The Journal of Business* 59 (4), 383 – 403.
- Cheng, N., Titterton, D. M. (1994). “Neural network: A review from a statistical perspective” . *Statistical Science* 9 (1), 2 – 54.
- Connor, G. (1984). “A unified beta pricing theory” . *Journal of Economic Theory* 34 (1), 13 – 31.

- Darrat, A. F., Zhong, M. (2000). "On testing the random – walk hypothesis: A model comparison approach" . *The Financial Review* 35, 105 – 124.
- Edwing, B. T. (2002). "Macroeconomic news and the returns of financial companies" . *Managerial and Decision Economics* 23, 439 – 446.
- Ehrmann, M., Fratzscher, M. (2004). "Taking stock: Monetary policy transmission to equity markets" . *Journal of Money, Credit and Banking*. 36 (4), 719 – 737.
- Engle, R. F., NG, V., Rothschild, M. (1990). "Asset pricing with a factor-ARCH covariance structure: Empirical estimates for treasury bills" . *Journal of Econometrics* 45, 213 – 237.
- Fama, E. F. (1991). "Efficient capital markets: II" . *The Journal of Finance* 46 (5), 1575 – 1617.
- Frances, H. P., Dijk, D. (1999). *Nonlinear time series models in empirical finance*. Cambridge University Press.
- Gallón, S., Gómez, K. (2007). "Distribución condicional de los retornos de la tasa de cambio colombiana: Un ejercicio empírico a partir de modelos GARCH multivariados" . *Revista de Economía del Rosario* 10 (2), 127 – 152.
- García, Y., García, J. (2006). "Revisión bibliográfica de la evidencia empírica de los modelos multifactoriales de valoración de activos financieros" . *Cuadernos de Economía* 25 (44), 197 – 224.
- Haykin, S. (1994) *Neural Networks. A Comprehensive Foundation*. Macmillan, New York, NY.
- Jalil, M. A., Misas, M. (2007). "Evaluación de pronósticos del tipo de cambio utilizando redes neuronales y funciones de pérdida asimétrica" . *Revista Colombiana de Estadística* 30 (1), 143 – 161.
- Jang, S. J., Sun, C. T., Mizutani, E. (1997). *Neuro-fuzzy and soft computing: A computational approach to learning and machine intelligent*. Pritice Hall: Upper Saddle River.
- Karil, Q. C., Leggio, K. B., Schniederjans, M. J. (2005). "A comparison between Fama and French's model and artificial neural networks in predicting the Chinese Stock Market" . *Computers & Operations Research* 32, 2499 – 2512.
- Koutoulas, G., Kryzanowski, L. (1994). "Integration or segmentation of the Canadian stock market: Evidence based on the APT" . *The Canadian Journal of Economics / Revue Canadienne d'Economie* 27 (2), 329 – 351.
- Kuan, C. M., White, H. (1994). "Artificial neural networks: An econometric perspective" . *Econometric Reviews* 13 (1), 1 – 91.
- Lee, T. H., White, H., Granger, C. W. (1993). "Testing for neglecting nonlinearity in time series models" . *Journal of Econometrics* 56, 269-290.
- Londoño, C. A. (2009). "Teoría de precios de arbitraje. Evidencia empírica para Colombia a través de redes neuronales" . Tesis no publicada. Universidad de Antioquia.

- Markowitz, H. (1952). "Portfolio selection". *The Journal of Finance* 7, 77–91.
- Medina, L. Á. (2003). "Aplicación de la teoría del portafolio en el mercado accionario colombiano". *Cuadernos de Economía* XXII (39), 129 – 168. Bogotá
- Melo, L. F., Becerra, O. R. (2006). "Una aproximación a la dinámica de las tasas de interés de corto plazo en Colombia a través de modelos GARCH multivariados". Banco de la República, Gerencia Técnica 366.
- Merton, R. C. (1973). "An intertemporal capital asset pricing model". *Econometrica* 41 (5), 867 – 887.
- Miller, E. M. (1977). "Risk, uncertainty, and divergence of option". *The Journal of Finance* 32 (4), 1151 – 1168.
- Misas, M., López, E., Querubín, P. (2002). "La inflación en Colombia una aproximación desde redes neuronales". *Ensayos Sobre Política Económica*, 42, 143-214
- Qi, M. (1994). "Financial application of artificial neural networks". Manddala, G.S., Rao, C. R (eds). *Handbook of statistiscs 14: Statistical methods in finance*. Elsevier Science B.V.
- Refenes, A. P., Zapranis, A.D., Francis, G. (1995). "Modeling stock returns in the framework of APT: A comparative study with regression model". En Refenes, A. P (Editores). *Neural networks in the capital markets*. John Wiley; 101 – 125.
- Reveiz, A. (2008), "The Factor-Portfolios Approach to Asset Management Using Genetic Algorithms". Borradores de Economía 511.
- Reveiz, A., León, C. E. (2008). "Efficient portfolio optimization in the wealth creation and maximum drawdown space". Borradores de Economía 520.
- Reveiz, A., Rojas, S. (2008). "The case for active management from the perspective of complexity theory". Borradores de Economía 495.
- Rodriguez, N. (2001). "Bayesian estimation and model selection for the weekly Colombian exchange rate". *Revista de Economía del Rosario* 4, 143 – 172.
- Ross, S. (1976). "The arbitrage theory of capital asset pricing". *Journal of Economic Theory* 13, 341–353.
- Santana, J. C. (2008). "La curva de rendimientos: Una revisión metodológica y nuevas aproximaciones de estimación". *Cuadernos de Economía* 27 (48), 71 – 113.
- Shaharudin, R. S., Fung, H. S. (2009). "Does size really matter? a study of size effect and macroeconomic factors in Malaysian stock returns". *International Research Journal of Finance and Economics* 24, 101–116.
- Sharpe, W. (1964). "Capital asset prices: A theory of market equilibrium under conditions of risk". *The Journal of Finance* 19 (3), 425 – 442.
- Solnik, B. (1983). "International arbitrage pricing theory". *The Journal of Finance* XXXVIII (2), 450 – 457.

- Thorbecke, W. (1997). “On stock market returns and monetary policy” . *The Journal of Finance* LII (2), 635 – 654.
- Villamil, J., Delgado, A. (2007). “Entrenamiento de una red neuronal multicapa para la tasa de cambio euro-dólar” . *Ingeniería e Investigación* 27 (3), 106 – 117.
- Villamil, J. (2009). “Aproximación no lineal al modelo de overshooting usando redes neuronales multicapa para el tipo de cambio dólar-peso” . *Cuadernos de Economía* 28 (50), 117 – 156.
- Von Neumann, J., Morgenstern, O. (1944). *Theory of games and economic behavior*. Princeton: Princeton University Press.

APÉNDICES

Apéndice 1: Medidas de evaluación del modelo APT dentro del periodo de entrenamiento

Tipo	Q	AIC	BIC	RMSE	RMSPE	MAE	MAPE
APT	L	-9.690100	-9.442821	0.007489	85.538170	0.005057	5.050346
	2	-9.743252	-9.558917	0.007384	88.13439	0.004981	4.84227
	6	-9.652997	-9.216887	0.007347	95.305880	0.004969	5.045205
	9	-9.567779	-8.942839	0.007384	87.848260	0.004980	4.833468

Nota: con $q = L$ es el modelo APT lineal y $q = 2, 6, 9$ el modelo

APT de redes neuronales.

Apéndice 2: Medidas de evaluación del pronóstico medidas simétricas y asimétricas para el modelo APT aplicando OLS

Horizonte	RMSE	RMSPE	MAE	MAPE	U-THEIL	LINLIN
1	0.006305	2.207788	0.005528	1.740760	0.014097	0.003149
2	0.005082	1.210528	0.004344	1.075389	0.011364	0.003624
3	0.005311	0.850209	0.004727	0.831503	0.011873	0.004489
4	0.004891	1.929023	0.004354	1.636267	0.010937	0.004786

5	0.004706	2.861266	0.003384	1.793046	0.010522	0.004368
6	0.004742	3.102444	0.004197	2.598428	0.010604	0.003104
7	0.002746	1.783929	0.002275	1.577734	0.006140	0.003228
8	0.003019	1.738435	0.002563	1.570543	0.006749	0.002430
9	0.005549	1.373863	0.003868	1.127102	0.012407	0.002447
10	0.001814	1.716402	0.001397	1.071505	0.004056	0.001039
11	0.010087	1.463411	0.008936	1.291188	0.022547	0.009356
12	0.004325	1.318441	0.003707	1.028159	0.009671	0.003074

Apéndice 3: Medidas de evaluación del pronóstico medidas simétricas y asimétricas para el modelo APT aplicando ANN

Horizonte	Q	RMSE	RMSPE	MAE	MAPE	U-THEIL	LINLIN
1	2	0.006611	2.40571	0.005822	1.872729	0.014782	0.003331
2	9	0.004907	1.068322	0.004104	0.963568	0.010973	0.003428
3	9	0.005309	0.840427	0.004700	0.819776	0.011868	0.004453
4	6	0.004579	1.867320	0.004074	1.557179	0.010239	0.004472
5	2	0.004423	2.593765	0.003323	1.668506	0.009891	0.004196
6	2	0.004401	3.219067	0.003996	2.631095	0.009840	0.003034
	6	0.004956	2.778603	0.004270	2.417942	0.011082	0.002979
7	2	0.002727	1.926058	0.002388	1.754154	0.006098	0.003340
	6	0.002820	1.676221	0.002141	1.387818	0.006305	0.003231
8	2	0.003372	1.794891	0.002798	1.624197	0.007540	0.002696
9	2	0.005486	1.514407	0.003824	1.244920	0.012268	0.002421
	6	0.005444	1.333668	0.003795	1.013849	0.012172	0.002400

10	2	0.002795	1.844388	0.002137	1.259593	0.006248	0.001755
11	2	0.010132	1.480592	0.008963	1.299240	0.022647	0.009311
12	2	0.003922	1.155934	0.003339	0.913305	0.008768	0.002766

Apéndice 4: Medidas de evaluación del modelo por dentro del periodo de entrenamiento

Tipo	Q	AIC	BIC	RMSE	RMSPE	MAE	MAPE
APTM	L	-9.823175	-9.387065	0.006748	33.80204	0.004583	3.674423
	2	-9.925196	-9.740861	0.006742	34.53489	0.004580	3.686562

Nota: con $q = L$ es el modelo APT modificado lineal y $q = 2$, el modelo

APT modificado de redes neuronales.

Apéndice 5: Medidas de evaluación del pronóstico medidas simétricas y asimétricas para el modelo APT modificado aplicando OLS

Horizonte	RMSE	RMSPE	MAE	MAPE	U-THEIL	LINLIN
1	0.001431	0.428090	0.001178	0.291788	0.003200	0.001224
2	0.005360	1.599355	0.005283	1.440111	0.011984	0.005456
3	0.005905	2.221207	0.005037	1.567317	0.013200	0.005759
4	0.004226	2.965028	0.003137	1.728507	0.009450	0.002141
5	0.005704	2.332270	0.004744	2.062996	0.012754	0.003331
6	0.004954	4.512094	0.004549	3.589287	0.011078	0.004010
7	0.004786	3.710147	0.004235	3.248124	0.010700	0.002795
8	0.006750	10.827390	0.004577	5.795485	0.015092	0.003277

9	0.004958	3.929121	0.004368	2.856746	0.011087	0.004493
10	0.003231	2.717220	0.003004	1.989896	0.007224	0.002539
11	0.008521	1.608260	0.007360	1.177966	0.019047	0.007334
12	0.003121	1.047732	0.002721	0.828504	0.006979	0.002479

Apéndice 6: Medidas de evaluación del pronóstico medidas simétricas y asimétricas para el modelo APT modificado aplicando ANN

Horizonte	Q	RMSE	RMSPE	MAE	MAPE	U-THEIL	LINLIN
1	2	0.001428	0.421737	0.00117	0.281492	0.003194	0.001496
2	2	0.005353	1.594511	0.005277	1.437158	0.011969	0.005447
3	2	0.005887	2.217283	0.005012	1.562646	0.013161	0.005726
4	2	0.004206	2.952141	0.003116	1.718333	0.009404	0.002123
5	2	0.005690	2.319344	0.004732	2.053303	0.012723	0.003325
6	2	0.004946	4.512327	0.004539	3.587248	0.011059	0.004003
7	2	0.004766	3.687041	0.004219	3.231812	0.010656	0.002787
8	2	0.006741	10.765410	0.004572	5.771229	0.015073	0.003279
9	2	0.004953	3.903398	0.004363	2.843044	0.011074	0.004482
10	2	0.003226	2.700203	0.002998	1.980363	0.007213	0.002534
11	2	0.008528	1.607694	0.007369	1.178840	0.019062	0.007348
12	2	0.003120	1.045430	0.002719	0.827022	0.006975	0.002479

Apéndice 7: Análisis de sensibilidad APT modificado

Figura 8. Efecto de *Bonos* en sus distintos regazos

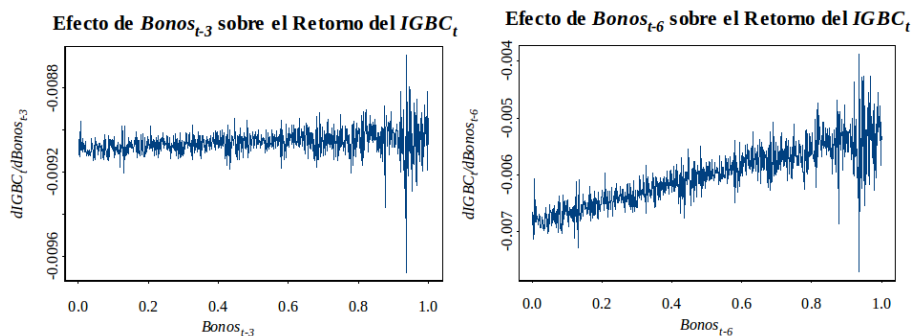


Figura 9. Efecto de *UVR* en sus distintos regazos

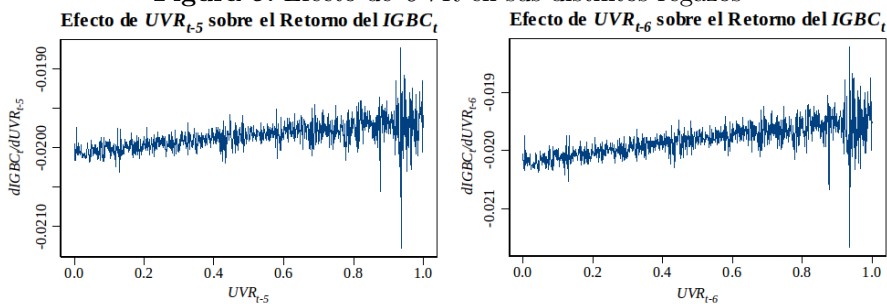


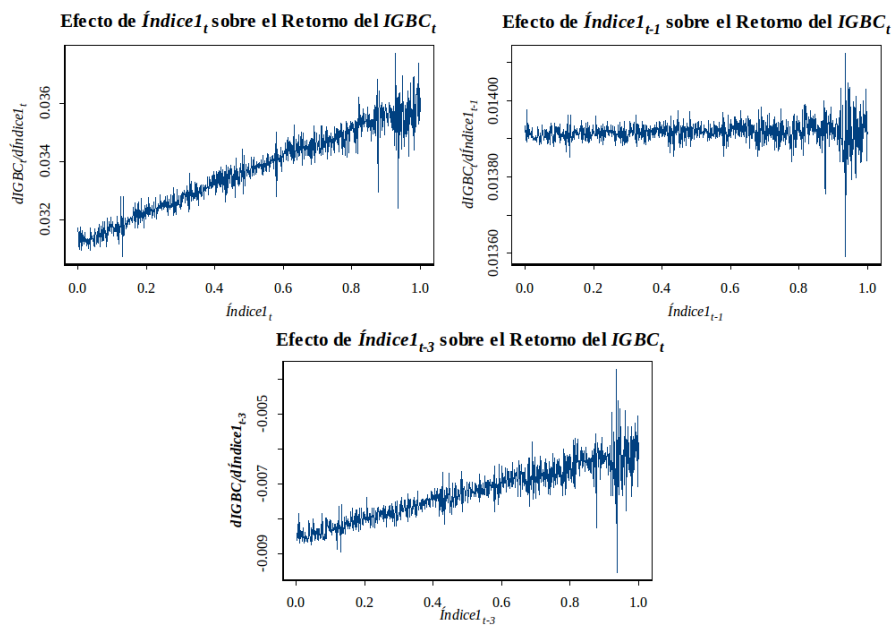
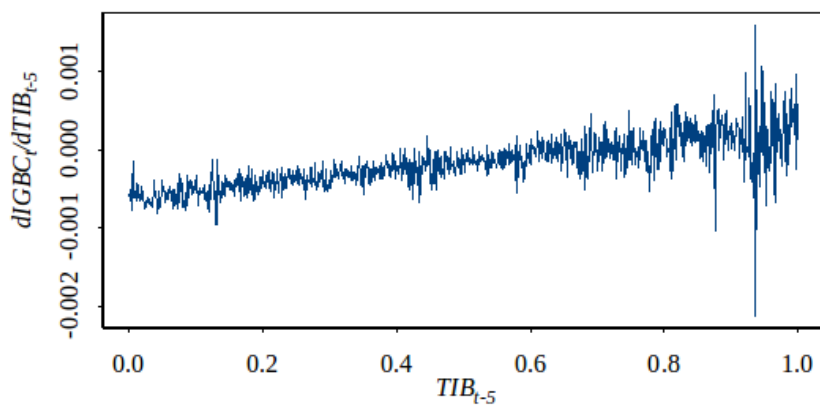
Figura 10. Efecto de Índice1 en sus distintos regazos**Figura 11.** Efecto de TIB_{t-5} sobre el Retorno del IGBC_t 

Figura 12. Efecto de *Moneda* sobre sus distintos regazos