UN MODELO ESTOCÁSTICO SOBRE LA PREDICTIBILIDAD DEL SIGNO DEL RETORNO Y SU RELACIÓN CON LA NO LINEALIDAD EN MEDIA*

*Javier Humberto Ospina Holguín***
*Edinson Caicedo Cerezo****

* Los autores agradecen los comentarios de los dos árbitros asignados por *Cuadernos de Administración*, al igual que las gestiones realizadas por el profesor Ignacio Vélez-Pareja y por los editores de la revista durante el proceso de edición. Así mismo, a la Facultad de Administración de la Universidad del Valle, Cali, Colombia, por su apoyo al grupo de investigación de Solvencia y Riesgo Financiero. El artículo se recibió el 03-10-2007 y se aprobó el 18-03-2008.

** Magíster en Ciencias de la Organización, Universidad del Valle, Cali, Colombia (2007); Físico, Universidad del Valle, (2001); profesor adjunto, Universidad del Valle. Miembro del grupo de investigación de Solvencia y Riesgo Financiero.
Correo electrónico: javierospina@gmail.com

*** Magíster en Ciencias de la Organización, Universidad del Valle, Cali, Colombia (1999); Estadístico, Universidad del Valle, (1992); jefe del Departamento de Contabilidad y Finanzas, Universidad del Valle. Miembro del grupo de investigación de Solvencia y Riesgo Financiero.
Correo electrónico: edccaiced@hotmail.com.

Un modelo estocástico sobre la predictibilidad del signo del retorno y su relación con la no linealidad en media

Resumen

Se examina cómo es la relación entre la predictibilidad del signo del retorno y los momentos condicionales. Para este propósito se utiliza un modelo no lineal de series de tiempo que luego se restringe a una expansión de Taylor de orden uno en las innovaciones. Como resultado se muestra por qué la predictibilidad del signo, cuando existe, depende sólo de los momentos condicionales de orden impar (de la media condicional en el modelo restringido), lo que hace interesante evaluar la linealidad o no linealidad en media antes de intentar una predicción direccional. A modo de aplicación, se explora la existencia de no linealidad en media en el índice de la bolsa de Colombia IGBC mediante el test BDS bajo un filtro ARMA-GARCH y el test de White basado en redes neuronales bajo un filtro AR. Ambos tests muestran que la serie del IGBC exhibe no linealidad en media.

Palabras clave: predictibilidad, econometría no lineal, no linealidad en media, test BDS, test de White.

A Stochastic Model for Return Sign Predictability and How it Relates to Mean Nonlinearity

Abstract

This work examines the relationship between predictability of return sign and conditional moments. A nonlinear time series model, later restricted to a first-order Taylor expansion in the innovations, is used for this purpose. As a result, it is shown why sign predictability depends only on odd-numbered order conditional moments (on the conditional mean in the restricted model), which makes testing for linearity or non linearity in the mean interesting before attempting directional forecasting. As an application, the presence of nonlinearity in the mean in the Colombian stock exchange index IGBC is examined applying the BDS test to the residuals of an ARMA-GARCH filter and the White Neural Network Test for Nonlinearity to the residuals of an AR filter. Both tests show that the IGBC exhibits nonlinearity in the mean.

Key Words: predictability, nonlinear econometrics, nonlinearity in the mean, BDS test, White Neural Network Test for Nonlinearity.

Introducción

Una buena parte de la literatura financiera asume la impredictibilidad de los retornos financieros debido a la hipótesis de la eficiencia débil de los mercados. Sin embargo, hay evidencia contundente de la presencia de patrones predecibles en muchas de las series examinadas de forma empírica (Lo & MacKinlay, 1988). La gran mayoría de la literatura de predictibilidad que existe se ha centrado en modelos de predicción lineales en media, como los modelos MA (*moving average*), AR (*autorregresive*) y ARMA (*autoregressive moving average*) (Hamilton, 1994, caps. 3-4; Tsay, 2001, cap. 2).

A pesar de esto, dado que la linealidad es un supuesto fuerte (Kuchta, 2004), en principio esta debería rechazarse con facilidad. Por una parte, el mercado accionario está sometido a las perturbaciones no lineales del ambiente; por otra, como Pesaran & Potter (1992) lo han enfatizado, el predominio de los modelos lineales en economía no se debe a que las teorías económicas sean intrínsecamente lineales, sino a que los modelos lineales son más sencillos. Ahora bien, si miramos a los practicantes financieros, la relevancia de la no linealidad es aparente: en términos teóricos, el análisis técnico consiste precisamente en la extracción de patrones no lineales de datos con ruido (Lo, Mamaysky & Wang, 2000; De Wachter, 2004, p. 15). El consenso general respecto a la no linealidad en las series de tiempo puede ponerse en palabras de Ramsey: "All these authors testify to one general result. There is abundant evidence that economic and financial data provide many varied indications of widespread stochastic nonlinearity, even though the main effects seem to be exhibited in the variances of the respective distributions" (1996).

Una interesante pregunta de investigación relacionada con lo anterior se refiere a cuál sería el papel de la linealidad y de la no linealidad (tanto en media como en varianza) en la predictibilidad del signo del retorno. Al elaborar un modelo de predicción de los log-retornos de una serie financiera es muy importante hacer la distinción entre las categorías de linealidad y de no linealidad, ya que los modelos no lineales tienden a sobreajustarse con muestras pequeñas gracias a su gran flexibilidad y a la amplia gama de alternativas disponible. Así, si hay un sobreajuste cuando el verdadero proceso generador de datos (DGP) es lineal con una varianza residual dada, el modelo no lineal tenderá a encontrar espuriamente una varianza estimada menor (Teräsvirta, Tjøstheim & Granger, 1994, p. 2919) y producirá predicciones erróneas en el periodo fuera de la muestra.

El objetivo de este artículo es ofrecer un modelo general que relacione la predictibilidad del signo del retorno financiero con los momentos condicionales de diferentes órdenes de la serie de los retornos. De esta manera, podremos determinar el papel que desempeña la linealidad y la no linealidad en media y en varianza en la construcción de algoritmos de predicción. El resultado al que llegamos establece que si el signo de los log-retornos es predecible, sólo los patrones en los términos que dependen de los momentos condicionales de orden impar son relevantes para la predicción del signo. En caso de que, por simplicidad, el modelo se restrinja a una aproximación de orden dos en los momentos condicionales (de orden uno en las innovaciones), el resultado obtenido equivale a afirmar que únicamente los patrones en el término que depende de la media condicional son relevantes para la predicción del signo. Por ende, si existe no linealidad en la serie de retornos, en este modelo la única no linealidad que será necesario examinar para la predicción en una aproximación hasta el segundo momento condicional es la no linealidad en media.

El modelo desarrollado es, hasta donde sabemos, una novedad, aunque está inspirado directamente en los comentarios de Kanzler (1998), de acuerdo con los cuales quienes descartan la hipótesis del camino aleatorio por correlación en el segundo momento condicional cometen un grave error (en nuestro modelo queda claro cómo la correlación en el segundo momento condicional no permite la predictibilidad ni siquiera del signo). La construcción de nuestro modelo se basa en definiciones de linealidad y de no linealidad hoy estándares. Es muy posible que para el econometrista este modelo sea tan claro como obvio; sin embargo, no son pocos los trabajos que han llegado a resultados erróneos por ignorar estas "ideas obvias".

Si el signo de los retornos depende principalmente de los patrones en el primer momento condicional cuando los retornos son predecibles, cabe preguntarse ¿tales patrones son lineales o no lineales? Como ejemplo de la aplicabilidad del modelo, en este artículo estudiamos la no linealidad en media en la serie de los retornos del Índice Accionario de la Bolsa de Colombia.

En lo que concierne a nuestros antecedentes metodológicos, desde el punto de vista teórico, el capítulo de modelos no lineales del *Handbook of Econometrics* por Teräsvirta, Tjøstheim & Granger (1994), al igual que Campbell, Lo & MacKinlay (1997, cap. 12) y Tsay (2001), los últimos con énfasis en series financieras, describen algunas de las principales teorías y metodologías para establecer y para modelar la no linealidad en una serie de tiempo determinada. Por su parte, Ramsey (1996) y Markellos (2002) son excelentes introducciones a la no linealidad en finanzas y econometría al tiempo que proveen referencias a una buena parte de los principales artículos que constituyen el tema, comentando los resultados empíricos pero sin profundizar en los detalles teóricos.

Cabe anotar que existe una amplia gama de test de no linealidad. Una introducción relativamente detallada a los test de no linealidad fue hecha por Barnett *et al*. (1997) y Brock & de Lima (1995), quienes ofrecen referencias a algunos de los principales artículos teóricos y trabajos empíricos sobre no linealidad en series de datos económicos. Aquí utilizamos dos test para detectar no linealidad en media: el test BDS, inspirado en ideas de sistemas dinámicos (toma su nombre de Brock, Dechert y Scheinkman), y el test de White, basado en redes neuronales. Estos test se describen, junto con sus respectivas referencias especializadas, en los numerales 2.1.2 y 2.2.1 respectivamente. Ambos han sido utilizados en diversos contextos y por diferentes autores, por ejemplo: Abhyankar, Copeland & Wong (1997) utilizan tanto el test BDS como el test de White en cuatro de los principales mercados internacionales y ofrecen otras referencias sobre no linealidad en los mercados accionarios; Abhyankar, Copeland & Wong (1995) y Opong *et al.* (1999) emplean el test BDS en el mercado británico; Barkoulas (1998) aplica el test BDS en el mercado griego; Kosfeld & Robé (2001) usan el test BDS en acciones bancarias alemanas; Reitgruber & Sterlina (1995) utilizan un filtro ARMA-GARCH y criterios de información como nuestro trabajo, pero adoptan el test BDS para probar la adecuación de los modelos de predicción y no para buscar no linealidad en media; Kyrtsoua, Labys & Terraza (2001) se valen tanto del test BDS como del test de White en un mercado de metales, y Kyrtsoua & Serletis (2006) utilizan, entre otros, el test de White en la tasa de cambio canadiense. Cabe anotar que Reitgruber & Sterlina (1995) concluyen que el mercado austriaco es aún "inmaduro" y logran un modelo de predicción simple para uno de sus índices, modelo que resulta rentable incluso teniendo en cuenta los costos de transacción.

En suma, la sección 1 del artículo introduce el modelo teórico que relaciona la predictibilidad del signo de los log-retornos con los momentos condicionales de distintos órdenes. La sección 2, como ejemplo de la aplicabilidad del modelo, establece dos metodologías para estudiar la no linealidad en media en la serie de los retornos del índice accionario de la Bolsa de Colombia (el Índice General de la Bolsa de Colombia o IGBC) con base en el marco teórico construido. Las secciones 3 y 4 introducen los datos del IGBC y los resultados de los test obtenidos. Por último, se exponen las conclusiones.

1. Modelo teórico

1.1 *Linealidad y no linealidad de una serie de tiempo*

Existen diversas definiciones de linealidad y no linealidad en la literatura de series de tiempo, no todas compatibles unas con otras. Siguiendo a Campbell, Lo & MacKinlay (1997, p. 468-469), usaremos un modelo típico en el que una serie de tiempo r_t es función de una sucesión de choques (incrementos o innovaciones) ε_t. En el análisis lineal de series de tiempo, r_t es una función lineal de las innovaciones, que se asumen incorrelacionadas, pero no necesariamente independientes e idénticamente distribuidas (IID) (Campbell, Lo & MacKinlay, 1997, p. 468) (de acuerdo con esto, los modelos de camino aleatorio RW1, RW2 y RW3 son modelos lineales, por ejemplo).

Sin embargo, en el análisis no lineal de series de tiempo se asume que las innovaciones ε_t son IID, pero que r_t depende de una función no lineal de las innovaciones. Seguiremos la convención de que las innovaciones ε_t son IID (incluso cuando se trate de funciones lineales).

1.2 *Un modelo de no linealidad para los retornos y su relación con la predictibilidad*

Imaginemos que r_t es una función de una sucesión de innovaciones IID:

$$r_t = f(\varepsilon_t, \varepsilon_{t-1}, \varepsilon_{t-2}, \ldots) \quad (1)$$

Obtengamos la serie de Taylor alrededor de $\varepsilon_t = 0$, para $\varepsilon_{t-1}, \varepsilon_{t-2}, \ldots$ dados:[1]

$$r_t = f(0,\varepsilon_{t-1},\ldots) + \varepsilon_t \frac{d}{dt} f(0,\varepsilon_{t-1},\ldots) + \frac{1}{2}\varepsilon_t^2 \frac{d^2}{dt^2} f(0,\varepsilon_{t-1},\ldots) + \ldots \qquad (2)$$

Supongamos ahora que r_t es un retorno (logarítmico) y que queremos llevar a cabo una predicción de este retorno con base en la expansión de Taylor (2) para realizar una compra o una venta de un activo financiero. Vamos a asumir que lo primero que debemos saber para tomar tal decisión es el signo del retorno, puesto que compraremos si el precio va a subir y venderemos si el precio va a bajar. Una de las primeras observaciones que debemos hacer sobre la serie de Taylor (2) es que el signo de todos los términos de orden mayor o igual a uno depende del signo de la innovación actual ε_t, signo que por definición (i.e., por su naturaleza estocástica) es impredecible. A pesar de esto, el signo de todos los términos de orden par es predecible y corresponde al signo de cada derivada, ya que sin importar el signo de ε_t todas sus potencias pares son positivas. Esto nos lleva a concluir que sólo los términos de orden impar tienen un signo impredecible. Así, si queremos predecir el signo del retorno, este signo estará determinado por términos que son predecibles (los términos de orden par, incluyendo el de orden cero) y por términos que son impredecibles. La predictibilidad del signo del retorno sólo será posible en cuanto que el aporte de todos los términos de orden impar sea mucho menor (una cantidad insignificante) en comparación al aporte de todos los términos de orden par.

En conclusión, si el signo del retorno es predecible, el aporte de los términos de orden impar debe ser una cantidad insignificante, y la predictibilidad del signo sólo dependerá de los términos de orden par. En lo que viene asumiremos que el signo del retorno es predecible, por ende, nos concentraremos exclusivamente en la predicción de los términos de orden par.[2]

1.3 *Interpretación del modelo en términos de momentos condicionales*

Regresando a Campbell, Lo & MacKinlay (1997, p. 469), si la magnitud de la innovación actual ε_t es pequeña, los términos de orden dos en adelante podrían llegar a ser muy pequeños, pues son proporcionales a una potencia de ε_t. Despreciemos entonces los términos de la expansión de Taylor (2) de orden dos en adelante y renombremos los términos restantes, así:

$$r_t = g(0,\varepsilon_{t-1},\ldots) + \varepsilon_t h(0,\varepsilon_{t-1},\ldots) \qquad (3)$$

donde

$$g(0,\varepsilon_{t-1},\ldots) \equiv f(0,\varepsilon_{t-1},\ldots)$$

$$\text{y } h(0,\varepsilon_{t-1},\ldots) \equiv \frac{d}{dt} f(0,\varepsilon_{t-1},\ldots) \qquad (4)$$

[1] Esta no es la única forma en que se puede descomponer la serie, pero es una de las más sencillas. Para apreciar otras posibles representaciones, véase Teräsvirta, Tjøstheim & Granger (1994) y Tsay (2001).

[2] Aunque la predicción de los términos de orden impar no proporciona información directa sobre el signo del retorno, entre mayor sea la contribución de estos términos, mayor será el "nivel de ruido" que tendrá cualquier predicción. Por esta razón, tal contribución sólo ayuda a determinar el riesgo de la inversión.

La función g representa la media de r_t condicionada a la información pasada (Teräsvirta, Tjøstheim & Granger, 1994, p. 2925), puesto que

$$E[r_t \mid r_{t-1}, r_{t-2}, \ldots] = g(\varepsilon_{t-1}, \varepsilon_{t-2}, \ldots) \quad (5)$$

Por su parte, la función h^2 representa la varianza de r_t condicionada a la información pasada (Teräsvirta, Tjøstheim & Granger, 1994, p. 2926), puesto que

$$\begin{aligned} &E[(r_t - g(\varepsilon_{t-1}, \varepsilon_{t-2}, \ldots))^2 \mid r_{t-1}, r_{t-2}, \ldots] \\ &= h(\varepsilon_{t-1}, \varepsilon_{t-2}, \ldots)^2 \end{aligned} \quad (6)$$

Vale la pena notar que al suponer que la serie se expande sólo hasta el segundo momento, la variación temporal de los momentos condicionales de orden más alto queda ligada a la variación temporal del segundo momento, pues

$$\begin{aligned} &E[(r_t - E[r_t \mid r_{t-1}, r_{t-2}, \ldots])^p \mid r_{t-1}, r_{t-2}, \ldots] \\ &= h(\varepsilon_{t-1}, \varepsilon_{t-2}, \ldots)^p \, E[\varepsilon_t^{\,p}] \end{aligned}$$

(Campbell, Lo & MacKinlay 1997, p. 469). Esta restricción es el precio a pagar por la mayor sencillez del modelo.

De acuerdo con Campbell, Lo & MacKinlay (1997, p. 469), los modelos que usan una función g no lineal son llamados modelos no lineales en media, mientras que los modelos que usan una función h^2 no lineal son llamados modelos no lineales en varianza (cf. Tsay, 2001, p. 127).[3] Los términos no linealidad en media, o respectivamente en varianza, condicional también son usados en la literatura (véase por ejemplo Peel & Speight, 1998).

1.4 *Predictibilidad del signo en términos de los momentos condicionales*

De manera más general, podríamos decir que el término de orden cero de la expansión de Taylor (2) depende del primer momento condicional de r_t; el término de primer orden, del segundo momento condicional de r_t, y así sucesivamente. De la anterior discusión de la predictibilidad del signo del retorno, habíamos concluido que si el signo del retorno era predecible, la predictibilidad del signo sólo dependería de los términos pares. Si nos restringimos a la expansión de Taylor de orden dos (ecuación (3)), se puede concluir que si asumimos que el (signo del) retorno es predecible, para predecirlo en términos prácticos deberemos restringirnos a estudiar específicamente los patrones en la media condicional g, puesto que los patrones en la varianza condicional h^2 no nos permitirán saber nada sobre el signo de los incrementos.

1.5 *Descomposición de la media condicional*

Después de concluir que sólo los patrones en la media condicional son de nuestro interés, conviene averiguar qué tipo de patrones pueden ser estos. Una manera de representar la función g es mediante un modelo polinomial (paramétrico) que expanda la función en una serie de Taylor alrededor de $\varepsilon_{t-1} = \varepsilon_{t-2} = \ldots = 0$ (Campbell, Lo & MacKinlay, 1997, p. 471) (lo que equivale a una versión discreta de la serie de Volterra de g):

[3] A menudo, cuando se usa el término no lineal sin hacer referencia a ningún momento (como en "correlación no lineal", "dependencia no lineal" o "no linealidad"), tal no linealidad incluye la dependencia lineal en momentos condicionales altos (por ejemplo, iguales o mayores a dos) (Kanzler, 1999).

$$\begin{aligned} &g(\varepsilon_{t-1},\varepsilon_{t-2},\ldots) = \\ &\sum_{i=1}^{\infty} a_i \varepsilon_{t-i} + \sum_{i=1}^{\infty}\sum_{j=i}^{\infty} b_{ij}\varepsilon_{t-i}\varepsilon_{t-j} \\ &+\sum_{i=1}^{\infty}\sum_{j=i}^{\infty}\sum_{k=j}^{\infty} c_{ijk}\varepsilon_{t-i}\varepsilon_{t-j}\varepsilon_{t-k} + \ldots \end{aligned} \quad (7)$$

donde el índice *j* empieza en *i* y el índice *k* empieza en *j* para no contar dos veces los productos cruzados de las innovaciones.

El primer término de la suma es un término estándar de media móvil ($MA(\infty)$) lineal, mientras que los demás términos son no lineales. Esto nos motiva a escribir *g* como una suma de un término lineal y de un término no lineal, así:

$$\begin{aligned} &g(\varepsilon_{t-1},\varepsilon_{t-2},\ldots) \\ &= g_l(\varepsilon_{t-1},\varepsilon_{t-2},\ldots) + g_{nl}(\varepsilon_{t-1},\varepsilon_{t-2},\ldots) \end{aligned} \quad (8)$$

con

$$g_l(\varepsilon_{t-1},\varepsilon_{t-2},\ldots) = \sum_{i=1}^{\infty} a_i \varepsilon_{t-i}$$

$$\begin{aligned} &g_{nl}(\varepsilon_{t-1},\varepsilon_{t-2},\ldots) \\ &= \sum_{i=1}^{\infty}\sum_{j=i}^{\infty} b_{ij}\varepsilon_{t-i}\varepsilon_{t-j} + \\ &\sum_{i=1}^{\infty}\sum_{j=i}^{\infty}\sum_{k=j}^{\infty} c_{ijk}\varepsilon_{t-i}\varepsilon_{t-j}\varepsilon_{t-k} + \ldots \end{aligned} \quad (9)$$

Nótese que la definición que estamos usando de linealidad es análoga a la definición de no linealidad: una función es lineal si es una función lineal de una sucesión de innovaciones IID. Esta definición se puede comparar con la definición de linealidad IID de Brock y Potter en Brock & de Lima (1995, p. 5):

> Un proceso estocástico estrictamente estacionario de media cero Y_t con suficiente regularidad para que posea una representación de Wold (causal) unilateral es IID (MDS) lineal si tiene una representación $Y_t = \sum \alpha_i \varepsilon_t$ con $\{\varepsilon_t\}$ una sucesión IID (una sucesión de diferencias de martingala (MDS)) (traducción nuestra).[4]

1.6 *Resumen del modelo*

En resumen, la serie r_t de los retornos se puede expresar como

$$\begin{aligned} r_t &= g(0,\varepsilon_{t-1},\ldots) + \varepsilon_t h(0,\varepsilon_{t-1},\ldots) \\ &= g_l(\varepsilon_{t-1},\varepsilon_{t-2},\ldots) + g_{nl}(\varepsilon_{t-1},\varepsilon_{t-2},\ldots) \\ &+\varepsilon_t h(0,\varepsilon_{t-1},\ldots) \end{aligned} \quad (10)$$

es decir, como una media condicional que tiene una parte lineal y una parte no lineal, más un término que depende de la raíz de la varianza condicional. Cuando hablemos de hallar patrones de no linealidad, nos interesará encontrar exclusivamente el término no lineal de la media condicional. Muchos de los test (Portmanteau) con poder contra varias alternativas no lineales no tienen un poder específico contra no linealidad en la media condicional: por esta razón, en las dos metodologías que se utilizarán para establecer la no linealidad en media será necesario un pre-filtrado lineal, y en una de ellas habrá que

[4] Otra definición convencional de linealidad que se corresponde con la de esta sección es aquella según la cual un proceso estocástico es lineal si es un filtro lineal de entradas IID (Barnett *et al.,* 1997). Un proceso ARIMA es un filtro lineal de orden finito, mientras que una expansión de Volterra de primer orden (con ningún kernel de orden más alto) es infinito-dimensional y cubre el espacio de los filtros lineales (Barnett *et al.,* 1997).

descartar además la no-linealidad debida a la varianza condicional.

2. Metodología

En el modelo anterior se concluyó que, despreciando los términos que dependen del tercer momento condicional en adelante, siempre que el signo del retorno sea predecible, tal signo dependerá exclusivamente de la media condicional g. El propósito de esta sección es exponer con base en el modelo dos posibles metodologías que permiten evaluar si los patrones en la media condicional son no lineales o no, es decir, si el término g_{nl} es no nulo o no.

La primera metodología consta de un filtrado lineal de la media condicional y de un filtrado de la varianza condicional, después de lo cual se aplica un test de IID. A continuación, cada uno de los dos filtros se justifica teóricamente, se explican sus detalles técnicos y se expresa la serie de los retornos en términos de estos filtros. Posteriormente, se presenta el test de IID y se interpreta y explica su funcionamiento técnico.

La segunda metodología consta de un filtrado lineal y de un test de no linealidad en media basado en redes neuronales. Para explicar esta metodología se ofrece una interpretación simple de lo que hace el test.

2.1 *Primera metodología (filtro lineal de la media condicional - filtro de la varianza condicional - test de IID)*

La primera metodología consiste en cuatro pasos:

- Adoptar un modelo para la parte lineal de la media condicional.
- Adoptar un modelo para la varianza condicional.
- Filtrar los términos que dependen de la parte lineal de la media condicional y de la varianza condicional estimando los parámetros del filtro con base en los modelos adoptados.
- Establecer mediante un test apropiado si los residuos del filtro (que en teoría equivalen al término no lineal de la media condicional g_l) son ruido independiente o si contienen patrones que indiquen que el término no lineal no es nulo.

2.1.1 *Modelos para la parte lineal de la media condicional y para la varianza condicional*

2.1.1.1 *Un modelo para la parte lineal de la media condicional*

El Teorema de Representación de Wold establece que todo proceso estacionario de tiempo discreto puede expresarse como la suma de dos procesos: uno puramente determinista y uno puramente indeterminista (Chatfield, 2003, p. 50). Por puramente determinista se quiere decir que una serie x_t se puede escribir como una combinación lineal (posiblemente infinita) de $x_{t-1}, x_{t-2}, \dots$ sin error (cf. Bierens, 1996, p. 136). Nótese que esta definición no se corresponde con el sentido tradicional en que se usa la palabra determinista (Chatfield, 2003, p. 51).

El teorema también establece que el componente puramente indeterminista puede escribirse como la suma lineal de una sucesión de variables aleatorias (Chatfield, 2003, p. 51). Si los elementos de la sucesión se distribuyen normalmente, la incorrelación implica independencia, y podemos expresar la sucesión como un proceso de media móvil $\mathrm{MA}(\infty)$ (Chatfield, 2003, p. 51) definido como:

$$\mathrm{MA}(\infty)=\sum_{i=1}^{\infty} a_i \varepsilon_{t-i} \quad (11)$$

De acuerdo con esto, el término lineal de la media condicional g_l puede interpretarse precisamente como el proceso $\mathrm{MA}(\infty)$ que representa la componente puramente indeterminista de g_l en la descomposición de Wold (la componente puramente determinista estaría ausente).

Sin embargo, desde el punto de vista computacional, y en la práctica, no es muy útil expresar un proceso lineal puramente indeterminista como g_l en la forma de un proceso $\mathrm{MA}(\infty)$ (Chatfield, 2003, p. 51), ya que habría que estimar demasiados parámetros. Usualmente, para obtener representaciones más parsimoniosas de la representación de Wold se busca una representación mediante un modelo ARMA (Rothman, 2001). Un conocido teorema (Schmidt, 2005, p. 235, teorema 12.3) establece que todo proceso ARMA $(p,\ q)$ puede representarse por un proceso MA con un número suficientemente largo de rezagos (posiblemente infinito). De acuerdo con esto, podemos reescribir el término lineal de la media condicional g_l en términos de un proceso ARMA (p,q)

$$\begin{aligned} & g_l(\varepsilon_{t-1}, x_{t-1}, \varepsilon_{t-2}, x_{t-2}, \ldots) \\ &= \sum_{i=1}^{\infty} a_i \varepsilon_{t-i} \\ &= \sum_{i=1}^{\infty} a_i^{*} \varepsilon_{t-i} + \sum_{i=1}^{\infty} a_i^{**} \varepsilon_{t-i} \qquad (12) \\ &= \sum_{i=1}^{q} \theta_i \varepsilon_{t-i} + \sum_{i=1}^{p} \rho_i x_{t-i} + \sum_{i=1}^{\infty} a_i^{**} \varepsilon_{t-i} \end{aligned}$$

donde los dos primeros términos de la última línea corresponden a un proceso ARMA que captura parsimoniosamente la mayor parte de los patrones lineales del proceso generador de datos (DGP), mientras que el tercer término es comparativamente despreciable.[5] La constante del modelo ARMA se omite aquí sin pérdida de generalidad. En la práctica, la distinción entre un filtro lineal de orden muy alto y un proceso no lineal es de poco uso, por esto nos restringimos a un filtro lineal ARMA (cf. Barnett *et al.*, 1997). Así, la detección o el filtrado de la linealidad de la serie implicará simplemente el uso de un filtro lineal de orden bajo. La discusión anterior justifica la validez teórica de aplicar un filtro ARMA para extraer la linealidad en la media condicional.

[5] Compárese con uno de los modelos paramétricos no lineales, el modelo bilineal:

$$g(\varepsilon_{t-1}, x_{t-1}, \varepsilon_{t-2}, x_{t-2}, \ldots) = \sum_{i=1}^{\infty} \rho_i \varepsilon_{t-i} + \sum_{i=1}^{\infty} \theta_i x_{t-i} + \sum_{i=1}^{\infty}\sum_{j=1}^{\infty} \gamma_{ij} x_{t-i} \varepsilon_{t-i}$$

(Campbell, Lo & MacKinlay, 1997, p. 471; véase Tsay, 2001, p.128, para una versión con sumas finitas). Los dos primeros términos corresponderían a la expresión lineal que hemos llamado g_l.

2.1.1.2 *Un modelo para la varianza condicional*

Existen múltiples modelos para la varianza condicional: ARCH (*Autorregresive Conditional Heteroskedasticity*); GARCH (*Generalized ARCH*, una suerte de ARMA para varianzas condicionales); IGARCH (*Integrated GARCH*, una suerte de ARIMA para varianzas condicionales); FIGARCH (*Fractionally Integrated GARCH*, una suerte de ARFIMA para varianzas condicionales); ARCH-M (*ARCH in mean*), EGARCH (*Exponential GARCH*); HARCH (*Hysteresis-GARCH*) (cf. Hamilton, 1994) y los modelos A-PARCH (*Asymmetric Power ARCH*), entre otros.[6] En finanzas, son comunes los modelos GARCH (especialmente GARCH(1,1)), GARCH-M y EGARCH. En este artículo utilizaremos un modelo GARCH, con el supuesto de que el apoyo empírico que ha recibido confirma que puede ser suficientemente general para captar los patrones de la varianza condicional (cf. Hansen & Lunde, 2005).

Los modelos GARCH (*Generalized* ARCH) fueron desarrollados por Bollerslev (1986) y Taylor (1986) como una generalización de los modelos ARCH introducidos originalmente por Engle (1982). La representación de la varianza GARCH (r,s) es:

$$\sigma_t^2 = \omega + \sum_{i=1}^{s} \alpha_i \varepsilon_{t-i}^2 + \sum_{j=1}^{r} \beta_j \sigma_{t-j}^2 \quad (13)$$

donde *r* es el orden de los términos GARCH autorregresivos y *s* es el orden de los términos ARCH de media móvil (cf. Hamilton, 1994, cap. 21).

Este modelo permite predecir el valor de la varianza en el período actual con base en el promedio de largo plazo (la constante), las varianzas predichas en los períodos anteriores (términos GARCH) y la información sobre la volatilidad observada en los períodos anteriores (términos ARCH). El modelo es consistente con el agrupamiento de la volatilidad observado a menudo en los datos financieros, de acuerdo con el cual los cambios grandes en los retornos suelen ser seguidos de cambios grandes y los cambios pequeños de cambios pequeños. Puede consultarse a Bollerslev, Chou, & Kroner (1992), Bollerslev, Engle & Nelson (1994) o a Hamilton (1994, cap. 21) para más detalles.

2.1.1.3 *Representación de los retornos con base en los modelos para la parte lineal de la media condicional y para la varianza condicional*

De acuerdo con los modelos que hemos escogido, los retornos se pueden expresar como:

$$\begin{aligned} r_t &= g_{nl}(\varepsilon_{t-1}, \varepsilon_{t-2}, \ldots) + g_l(\varepsilon_{t-1}, \varepsilon_{t-2}, \ldots) + \varepsilon_t h(0, \varepsilon_{t-1}, \ldots) \\ &= g_{nl}(\varepsilon_{t-1}, \varepsilon_{t-2}, \ldots) + \left(\lambda + \sum_{i=1}^{p} \rho_i x_{t-i} + \sum_{i=1}^{q} \theta_i \varepsilon_{t-i} \right) \\ &\quad + \varepsilon_t \left(\omega + \sum_{i=1}^{s} \alpha_i \varepsilon_{t-1}^2 + \sum_{j=1}^{r} \beta_j \sigma_{t-j}^2 \right)^{1/2} \end{aligned} \quad (14)$$

[6] Los modelos A-PARCH cubren a los modelos particulares ARCH, GARCH, TS-GARCH, GJR-GARCH, T-GARCH, N-GARCH y Log-GARCH (McKenzie *et al.*, 2001, Wurtz, Chalabi & Luksan, 2006).

donde el segundo término representa un modelo ARMA (p,q) y el último término es una innovación por la raíz de un modelo GARCH (r,s). (La constante λ se introduce en caso de que la media del proceso no sea cero).

Después de estimar los parámetros del modelo ARMA-GARCH es posible filtrar los retornos para evaluar si los residuos del filtro no son independientes o si están hechos de ruido IID. Si los residuos no son independientes, el término teórico no lineal de la media condicional será no nulo (pues, en términos más laxos, los residuos tendrán un "patrón" que no ha sido capturado por el modelo). En cambio, si los residuos del filtro ARMA-GARCH son IID, concluiremos que el modelo ARMA-GARCH describe con suficiencia los datos y que, por ende, no existe no linealidad en media condicional en la serie de los retornos. Con el fin de evaluar si los residuos del modelo ARMA-GARCH son o no IID, tendremos que aplicar un test de IID (véase el numeral 2.1.3 sobre los detalles específicos de cómo son teóricamente estos residuos).

2.1.2 *Un test de IID: el test BDS*

El test de IID que se utilizará con el fin de evaluar si los residuos del modelo ARMA-GARCH son o no IID es el test BDS. El test BDS fue desarrollado por William Brock, Davis Dechert y Jose Scheinkman en 1986, aunque su trabajo original no fue publicado hasta 1996 en compañía de Blake LeBaron, que ayudó a escribir un programa en Fortran y luego en C para el test (Brock *et al*., 1996). Otros artículos originales de interés incluyen LeBaron (1997) y Scheinkman & LeBaron (1989), y el libro Brock, Hsieh & LeBaron (1991). El test BDS es un test de independencia con poder contra virtualmente todos los tipos de desviaciones lineales o no lineales (aunque se han hallado excepciones teóricas sin relevancia práctica, v.g. Dechert (1988)). El test no hace asunciones sobre la distribución de los datos (fuera de que son IID en la hipótesis nula) ni asume la existencia de momentos superiores.

En realidad, el test BDS no es un test de no linealidad (ni de caos), aunque algunos autores lo describan (erróneamente) así (cf. Casti, 1992, cap. 4.7). Sin embargo, en conjunción con determinados filtros, el test puede detectar no linealidad. Así, por ejemplo, si se aplica a datos filtrados linealmente, el test tendría poder contra prácticamente cualquier tipo de desviaciones estocásticas no lineales (en media, en varianza, etc.) o deterministas no lineales (caos). A pesar de lo anterior, es importante aclarar que el test no permite averiguar, sin una estrategia metodológica apropiada que podría incluir a otros test o filtros, qué tipo de no linealidad está presente.

De acuerdo con la estructura de la primera metodología que se ha propuesto (numeral 2.1.1), para establecer la presencia de no linealidad en media utilizaremos un filtro ARMA-GARCH. Una de las ventajas del test BDS es que es teóricamente robusto si se aplica a residuos de filtros ARMA (Brock & de Lima, 1995) y GARCH (Kanzler, 1999). Sin embargo, por razones aún desconocidas, y a pesar de su robustez teórica, los residuos aditivos GARCH parecen sesgar el estadístico del test BDS (Kanzler, 1999), aunque este problema se puede contrarrestar utilizando *bootstrapping* para estimar la significancia del test (Kanzler, 1999), tal como se hizo en este trabajo.

2.1.2.1 *El funcionamiento del test BDS*

El test BDS está descrito en detalle en Brock *et al.* (1996), pero vale la pena ofrecer una idea general sobre su funcionamiento. Supongamos, bajo la hipótesis nula, que una serie de tiempo x_t es IID. Si calculáramos la probabilidad de que una historia de dos observaciones cumpliera cada una con la condición de cercanía de estar en una bola cerrada de radio ε, tendríamos para cualquier $i \neq j$

$$\begin{aligned} & c_2(\varepsilon) \\ & \equiv P\left(\left|x_i - x_j\right| \leq \varepsilon, \left|x_{i-1} - x_{j-1}\right| \leq \varepsilon\right) \\ & = P\left(\left|x_i - x_j\right| \leq \varepsilon\right) P\left(\left|x_{i-1} - x_{j-1}\right| \leq \varepsilon\right) \quad (15) \\ & = P\left(\left|x_i - x_j\right| \leq \varepsilon\right)^2 \\ & = c_1^{\,2}(\varepsilon) \end{aligned}$$

con

$$c_1(\varepsilon) \equiv P\left(\left|x_i - x_j\right| \leq \varepsilon\right) \quad (16)$$

donde la segunda y la tercera igualdad en la ecuación (15) se deben al carácter IID de la serie.

En general, este resultado se puede generalizar para *m*-historias (*m*-dimensiones), y las hipótesis nula y alternativa del test BDS son precisamente que

$$H_o{:}c_m(\varepsilon) = c_1^{\,m}(\varepsilon)$$

$$H_1{:}c_m(\varepsilon) \neq c_1^{\,m}(\varepsilon) \quad (17)$$

lo cual es prácticamente equivalente a probar IID frente a casi cualquier otra alternativa. Como al trabajar con datos no es posible calcular las probabilidades $c_m(\varepsilon)$ y $c_1(\varepsilon)$ de manera directa, estas se estiman recorriendo todas las posibles *m*-historias mientras que se cuenta el número de *m*-historias que satisfacen la condición de cercanía. En términos matemáticos, para una serie con *n* observaciones, tenemos:

$$c_{m,n}(\varepsilon) = \frac{2}{(n-m+1)(n-m)} \sum_{s=1}^{n-m+1} \sum_{t=s+1}^{n-m+1} \prod_{j=0}^{m-1} u\left(\varepsilon - \left|x_{s+j} - x_{t+j}\right|\right) \quad (18)$$

donde

$$u(x) = \begin{cases} 0, x < 0 \\ 1, x > 0 \end{cases} \quad (19)$$

es la función de Heaviside, que en este caso ayuda a enumerar cada *m*-historia que cumpla con la condición de cercanía. Precisamente, la entidad $c_{m,n}(\varepsilon)$ es la integral de correlación hecha popular por Grassberger & Procaccia (1983), una medida de la probabilidad media de que los estados de dos instantes diferentes estén cercanos entre sí.

Estas estimaciones muestrales de las probabilidades pueden usarse para construir el estadístico

$$b_{m,n}(\varepsilon) = c_{m,n}(\varepsilon) - c_{1,n-m+1}(\varepsilon)^m \quad (20)$$

que en tanto se cumpla la asunción IID debería estar cerca de cero (el segundo término descarta las últimas *m–1* observaciones para que contenga el mismo número de términos que el primero).

De hecho, Brock *et al.* (1996) prueban que

$$\lim_{n\to\infty}\sqrt{n-m+1}\,\frac{b_{m,n}(\varepsilon)}{\sigma_{m,n}(\varepsilon)}\sim N(0,1) \quad (21)$$

con

$$\sigma^2_{m-n}(\varepsilon)=4k^{m}+8\sum_{j=1}^{m-1}k^{m-j}c_1^{\ 2j}+4(m-1)^2c_1^{\ 2m}-4m^2kc_1^{\ 2m-2} \quad (22)$$

donde el parámetro c_1 puede estimarse por medio de $c_{1,n}$, y el parámetro k es la probabilidad de que una tripla de puntos estén a una cercanía ε los unos de los otros y puede estimarse contando el número de triplas que satisfacen esa condición

$$k_n(\varepsilon)=\frac{2}{n(n-1)(n-2)}\sum_{t=1}^{n}\sum_{s=t+1}^{n}\sum_{r=s+1}^{n}(I_\varepsilon(x_t,x_s)I_\varepsilon(x_s,x_r)+I_\varepsilon(x_t,x_r)I_\varepsilon(x_r,x_s)+I_\varepsilon(x_s,x_t)I_\varepsilon(x_t,x_r)) \quad (23)$$

donde

$$I_\varepsilon(x,y)\equiv u(\varepsilon-|x-y|) \quad (24)$$

2.1.3 *Resumen de la primera metodología*

De acuerdo con los modelos que se escogieron para la parte lineal de la media condicional y para la varianza condicional, los log-retornos se pueden describir, con base en la ecuación (14), así:

$$\begin{aligned} r_t &= g_{nl}(\varepsilon_{t-1},\varepsilon_{t-2},\ldots) + g_l(\varepsilon_{t-1},\varepsilon_{t-2},\ldots)+\varepsilon_t h(0,\varepsilon_{t-1},\ldots)\\ &= g_{nl}(\varepsilon_{t-1},\varepsilon_{t-2},\ldots) + \left(\lambda+\sum_{i=1}^{p}\rho_i x_{t-i}+\sum_{i=1}^{q}\theta_i\varepsilon_{t-i}\right)+\ldots \\ &\quad \varepsilon_t\left(\omega+\sum_{i=1}^{s}\alpha_i\varepsilon^2_{t-i}+\sum_{j=1}^{r}\beta_j\sigma^2_{t-j}\right)^{1/2} \end{aligned} \quad (25)$$

donde el primer término de la última línea es el término no lineal en media, el segundo término corresponde al modelo ARMA del término lineal en media y el último término al residuo escrito en términos del modelo GARCH.

Si el filtro ARMA-GARCH está correctamente especificado, después de aplicar la parte ARMA del filtro el término ARMA desaparecerá de los retornos, y después de estandarizar los residuos del filtro ARMA con base en el modelo GARCH para la varianza (es decir, después de dividir los residuos del filtro ARMA entre la raíz de la varianza GARCH), los residuos estandarizados (i.e. los retornos filtrados y estandarizados) quedarán como:

$$\frac{\varepsilon^*_t}{\sigma_t}=\frac{g_{nl}(\varepsilon_{t-1},\varepsilon_{t-2},\ldots)}{\left(\omega+\sum_{i=1}^{s}\alpha_i\varepsilon^2_{t-i}+\sum_{j=1}^{r}\beta_j\sigma^2_{t-j}\right)^{1/2}}+\varepsilon_t \quad (26)$$

donde se utiliza la expresión ε_t^* para denotar los residuos de la ecuación en media del filtro ARMA y la división entre σ_t para denotar la estandarización GARCH de estos residuos.

En otras palabras, en términos teóricos, los residuos estandarizados del filtro ARMA-GARCH resultan de eliminar de los retornos de la ecuación (25) el término ARMA $\left(\lambda + \sum_{i=1}^{p}\rho_i\varepsilon_{t-i} + \sum_{i=1}^{q}\theta_i x_{t-i}\right)$ y de dividir el resultado entre la raíz de la varianza GARCH $\left(\omega + \sum_{i=1}^{s}\alpha_i\varepsilon_{t-i}^2 + \sum_{j=1}^{r}\beta_j\sigma_{t-j}^2\right)^{1/2}$. Si no existe no linealidad en media (i.e. si $g_{nl} = 0$), el primer término de la expresión (26) será cero. Por ende, de acuerdo con el modelo teórico, los residuos estandarizados serán IID:

$$\frac{\varepsilon_t^*}{\sigma_t} = \varepsilon_t \text{ cuando } g_{nl} = 0 \quad (27)$$

En cambio, si existe no linealidad (i.e. si $g_{nl} \neq 0$), no hay ninguna razón *a priori* para que los residuos estandarizados sean IID. Precisamente, como paso final de esta primera metodología para detectar no linealidad en media, se aplica el test BDS de IID a los residuos estandarizados del filtro ARMA-GARCH.

2.2 *Segunda metodología (filtro lineal de la media condicional – test basado en redes neuronales)*

La segunda metodología que se utilizará para detectar no linealidad en media en la serie del IGBC consiste en evaluar si una función mensurable de la historia del proceso se correlaciona con los residuos de un filtro lineal. Esta metodología es la que se ha desarrollado en el test de White de no linealidad (White, 1989).[7] En este test, la serie de tiempo se ajusta por medio de una red neuronal con conexiones hacia delante (*feed-forward*) con una única capa oculta, la cual se usa para establecer si los residuos de un filtro (lineal) AR exhiben estructura no lineal (en media). Esta metodología tiene la ventaja de que es innecesario un prefiltrado de la varianza condicional.

2.2.1 *Test de White de no linealidad basado en redes neuronales*

Los detalles técnicos del test de White pueden verse en White (1989) o en Lee (2001) para una exposición resumida. En este apartado sólo describiremos en términos generales la lógica del test, como es expuesta por Barnett *et al*. (1997) y por Lee (2001). Para comenzar, el test de White ajusta la serie de tiempo mediante una red neuronal con conexiones hacia delante (*feed-forward*) con una única capa oculta (para una introducción accesible a este tipo de redes neuronales se puede ver Heaton, 2005, cap. 5). La salida de la red viene dada por:

$$y_t = x_t\beta + \sum_{j=1}^{q}\delta_j\psi(x_t\gamma_j) + \varepsilon_t \quad (28)$$

donde x_t es la entrada, β es un vector columna que representa la fuerza de las conexiones entre la capa de entrada y la capa de salida, γ_j es un vector columna que representa la fuerza de las conexiones entre la capa de entrada y cada una de las q unidades ocultas, δ_j

7 Algunos autores lo denotan como el test de Lee-White-Granger (v.g. McNelis, 2005, p. 91) o test de White basado en redes neuronales. Este test no debe confundirse con el conocido test de heteroscedasticidad de White.

es un escalar que representa la fuerza de la conexión entre la *j*-ésima unidad oculta y la unidad de salida, y ψ es una función de base radial o una función compresora.

La red funciona enviando señales desde las unidades de entrada hasta las unidades ocultas intermedias; luego cada unidad oculta produce una activación ψ que envía una señal a la unidad de salida.

Mediante la red se determina si hay cualquier estructura no lineal en los residuos de un filtro AR sobre la misma serie de tiempo. La hipótesis nula del test es la linealidad en media. Si la hipótesis de linealidad en media se cumple, los residuos de un filtro lineal (como el filtro AR) no se deberían correlacionar con ninguna función mensurable de la historia del proceso. El test de White utiliza la red neuronal ajustada para generar la función mensurable de la historia del proceso y, de esta manera, puede examinar la hipótesis de que la función ajustada no se correlaciona con los residuos del proceso AR.

Específicamente, el test de White se basa en una función $h(x_t)$ escogida como las activaciones de las unidades ocultas $\psi(x_t\Gamma_j)$, donde Γ_j son vectores aleatorios independientes de x_t. Según lo anterior, bajo la hipótesis nula de linealidad en media se cumple que $E[\psi(x_t\Gamma_j)\varepsilon_t^*|\Gamma_j]=E[\psi(x_t\Gamma_j)\varepsilon_t^*]=0$, donde ε_t^* son los residuos del mejor filtro lineal. Así, evidencia de correlación en ε_t^* con el vector de las activaciones $\psi(x_t\Gamma_j)$ es evidencia en contra de la hipótesis nula. Si existe correlación, al aumentar más unidades ocultas con activaciones $\psi(x_t\Gamma_j)$ la red debería aumentar su eficiencia.

Con base en lo anterior se puede conseguir un test basado en la correlación muestral de los errores de la red afín con las unidades de activación ocultas. En efecto, dadas ciertas condiciones de regularidad, el siguiente estadístico de prueba tiene una distribución Chi cuadrado asintótica bajo la hipótesis de no linealidad en media (Lee, 2001):

$$\left(n^{-1}\sum_{t=1}^{n}\begin{pmatrix}\psi(x_t\Gamma_1)\\ \vdots \\ \psi(x_t\Gamma_n)\end{pmatrix}\hat{\varepsilon}_t\right)^T \hat{W}_n^{-1}\left(n^{-1}\sum_{t=1}^{n}\begin{pmatrix}\psi(x_t\Gamma_1)\\ \vdots \\ \psi(x_t\Gamma_n)\end{pmatrix}\hat{\varepsilon}_t\right)\xrightarrow{d}\chi^2 \tag{29}$$

donde $\hat{W}_n$ es un estimador consistente de la matriz de covarianza asintótica. White ha sugerido que si se satisfacen ciertas asunciones los parámetros de la red no tienen que ser estimados, sino que basta con realizar la regresión y la extracción de los componentes principales para llegar a un test asintóticamente equivalente (cf. Barnett *et al*., 1997).

El orden del filtro lineal AR se escoge por medio del criterio de información de Schwartz (BIC), siguiendo a Jungeilges (1993) que muestra que utilizar este criterio mejora el poder de la prueba. Como el test es condicional a la escogencia de la dirección, esta dirección se escoge al azar (véase Barnett *et al*., 1997). Múltiples simulaciones y comparaciones (v.g. Lee *et al*., 1993 y Teräsvirta, 1996 con una interpretación del artículo anterior o Barnett *et al*., 1997) han mostrado que este test tiene poder contra una gran variedad de procesos no lineales en media (sin verse afectado por otros procesos no lineales, pero lineales en media, como los procesos GARCH).

3. Datos

En esta sección especificamos cuáles fueron los datos utilizados en este trabajo. El estudio se enfoca en el Índice General de la Bolsa de Colombia (IGBC) del cual examinamos 1095 datos diarios de cierre[8] (publicados por la Superintendencia Financiera de Colombia) durante los días transados desde su formación en 03/07/2001 hasta 21/12/2005, inclusive. Tal enfoque tiene sus limitaciones,[9] pero la investigación del comportamiento de datos diarios del índice de un mercado corresponde a una práctica usual de la econometría financiera. Una de las ventajas de utilizar datos diarios es que los errores muestrales son menores debido a que existen más muestras que si se consideraran datos entre períodos más largos (como semanas).

4. Resultados

4.1 *Resultados de la primera metodología (filtro lineal de la media condicional - filtro de la varianza condicional - test de IID)*

A continuación se presentan los resultados obtenidos al examinar la no linealidad en media de los log-retornos del IGBC mediante la primera metodología. De acuerdo con esta metodología, para establecer no linealidad en media, primero se estimó el mejor filtro ARMA; con base en este mejor filtro se estimó el mejor filtro ARMA-GARCH; y finalmente a los residuos estandarizados de este mejor filtro ARMA-GARCH se les aplicó el test BDS de IID. En ausencia de no linealidad en media, se esperaría que los residuos estandarizados del mejor filtro ARMA-GARCH fueran IID. Después de aplicar los diagnósticos comunes, el criterio de información de Hannan-Quinn indicó que el mejor modelo ARMA-GARCH es un modelo ARMA(1,0)-GARCH(0,5), es decir, un modelo AR(1)-ARCH(5) (omitimos los extensos detalles, pero están disponibles para quienes los soliciten).

Los resultados del test BDS (cuadro 1A del apéndice) muestran que la hipótesis de IID de los residuos estandarizados del mejor filtro ARMA-GARCH de los log-retornos del IGBC se puede rechazar (por lo menos a un nivel de significancia del 10%) para todas las dimensiones examinadas. De acuerdo con el modelo teórico, la implicación directa de este resultado es que la hipótesis de que no existe no linealidad en media en la serie ha quedado rechazada.

4.2 *Resultados de la segunda metodología (filtro lineal de la media condicional – test basado en redes neuronales)*

En el cuadro 2A del apéndice se presentan los resultados empíricos obtenidos al caracterizar la no linealidad en media de la serie del IGBC mediante la segunda metodología que

8 En este tipo de estudio, los datos diarios de cierre deben preferirse a los datos diarios promedio debido a que estos últimos suelen introducir sesgos no deseados (véase Dimson & Mussavian, 1998).

9 Cf. Darrat (1990) y Lo & MacKinlay (1988). Sin embargo, hacemos notar que en el caso de un índice de muchas acciones, como el IGBC, las limitaciones más comunes se atenúan.

utiliza el test de White basado en redes neuronales. Los resultados del test de White sobre los residuos del mejor filtro AR (en este caso un filtro AR(1)) muestran que la hipótesis de que no existe no linealidad en media en los log-retornos del IGBC puede rechazarse.

Conclusiones

Este artículo presentó un modelo general que relaciona la predictibilidad del signo del log-retorno financiero con los momentos condicionales de diferentes órdenes de la serie de los log-retornos. El modelo permite establecer qué papel desempeña la linealidad y la no linealidad en media y en varianza en la construcción de algoritmos de predicción. La conclusión principal es que si el signo del retorno se supone predecible, cuando se consideran términos hasta el segundo momento condicional, la única no linealidad que es de interés para la predicción del signo es la no linealidad en media.

Hemos comentado previamente que es posible que para el econometrista este modelo sea tan claro como obvio (podría tratarse de un "*folk theorem*" para algunos expertos); sin embargo, no son pocos los autores que parecen ignorar estas "ideas obvias" cayendo en serios errores de interpretación. Así, por ejemplo, si nos interesa la predictibilidad de los log-retornos, no se debería rechazar la hipótesis del camino aleatorio y concluir predictibilidad sobre la base de que existe correlación en el segundo momento condicional, ya que esta correlación no se puede explotar ni siquiera para predecir el signo del retorno (al menos no a partir de un modelo de nolinealidad basado en innovaciones IID). Pues bien, no falta en la literatura quienes hayan intentado rechazar o explicar el rechazo de la hipótesis del camino aleatorio precisamente sobre estas bases erróneas (cf. Milevsky, 1995; Barnett & Serletis, 2000; Dahl & Nielsen, 2001). De hecho, lo que se puede concluir de una correlación en el segundo momento condicional es que si esta correlación implica alta volatilidad, el practicante no debería fiarse mucho –en el momento de construir modelos de predicción– de los datos en el período que la exhibe (pues la alta volatilidad podría haber dejado inservibles los patrones aprovechables en el primer momento condicional). Algorítmicamente, esto podría sugerir que tal vez deberíamos darles un menor peso o incluso ignorar estos datos en la construcción de los modelos de predicción que se diseñen.

El modelo que ofrecemos esclarece estos conceptos, pero también brinda un marco teórico para detectar aquellos patrones no lineales en los datos que sirven de base para la predicción. Así, por ejemplo, el modelo permite derivar estrategias metodológicas para establecer si los patrones en el primer momento condicional son lineales o no, lo que resulta de interés al haberse establecido que el signo de los retornos depende sólo de estos patrones cuanto los retornos son predecibles. Precisamente abordamos como un ejemplo práctico el estudio de la no linealidad en media condicional para el caso del IGBC por medio de dos metodologías diferentes, pero complementarias, una que usa ideas de sistemas dinámicos (por medio del test BDS) y otra que utiliza técnicas de computación blanda (específicamente redes neuronales a través del test de White).

Era de esperarse que se detectaría con facilidad la presencia de no linealidad, dado que la linealidad es una suposición fuerte, pero quedaba como una pregunta abierta si tal no linealidad comprendería también los patrones en la media condicional. Los resultados obtenidos presentaron evidencia sólida y complementaria de que sí existe no linealidad en la media condicional en el caso del IGBC. De acuerdo con estos resultados, si se pretendiese construir un modelo bien determinado para la predicción del IGBC se esperaría, *a priori*, que este tendría que modelar la no linealidad en media. Aunque examinamos sólo el caso de un índice accionario, no puede descartarse que otros índices tengan comportamientos similares, y el modelo ofrecido puede servir de guía para establecer o refutar tal hipótesis.

Una de las enseñanzas que nos queda del ejemplo particular estudiado es que cuando se desea examinar la presencia de no linealidad en media es mejor utilizar una estrategia metodológica compuesta para subsanar las limitaciones de las diferentes metodologías particulares. Así, por ejemplo, las metodologías que usan técnicas de computación blanda, como el test de White bajo un filtro AR, y que tienen por ende una alta dependencia de los pesos particulares de la red neuronal se ven completadas por metodologías más robustas, pero con más asunciones, como la del test BDS bajo un filtro ARMA-GARCH cuyo éxito depende de la suposición de que la volatilidad no exhibe un efecto de apalancamiento. Por supuesto, es posible concebir metodologías mucho más sofisticadas como un test de White bajo filtros ARFIMA o un test BDS bajo filtros A-PARCH, pero aún queda mucho por explorarse en cuanto a la robustez y la conveniencia de estas aproximaciones particulares.

Una línea de investigación empírica que puede ser de interés futuro concierne al examen de datos de alta frecuencia, ya de índices ya de activos financieros aislados, pues en tales datos se podría esperar que los diversos términos no lineales, incluso de órdenes superiores, puedan ser aún más relevantes. Por último, cabe también la posibilidad de extender nuestro modelo hacia el desarrollo de nuevos test. Una posible contribución a este respecto sería el perfeccionamiento de un test que en una aproximación hasta el segundo momento condicional de los logretornos pudiera establecer en qué secciones de la serie de tiempo el aporte de los momentos condicionales de orden par es mayor al aporte de los momentos de orden impar. La importancia de un test como este sería fundamental para la predicción, pues permitiría comprobar cuantitativamente en qué instantes (en qué secciones de la serie) es más plausible el éxito de la predicción, ya que de acuerdo con nuestro modelo la viabilidad de la predicción mejora cuando el aporte de la varianza condicional es insignificante respecto al aporte de la media condicional.

Lista de referencias

Abhyankar, A., Copeland, L. S., and Wong, W. (1995). Nonlinear Dynamics in Real-Time Equity Market Indices: Evidence from the United Kingdom. *The Economic Journal*, 105 (431), 864-880.

Abhyankar, A., Copeland, L. S., and Wong, W. (1997). Uncovering Nonlinear Structure in Real-Time Stock-Market Indexes: The S&P 500, the DAX, the Nikkei 225, and the FTSE-100. *Journal of Business & Economic Statistics*, 15 (1), 1-14.

Barkoulas, J. (1998). Chaos in an emerging capital market? The case of the Athens Stock Exchange. *Applied Financial Economics*, 8 (3), 231-243.

Barnett, W. A. and Serletis, A. (2000). Martingales, nonlinearity, and chaos. *Journal of Economic Dynamics and Control*, 24 (5), 703-724.

Barnett, W. A., Gallant, A. R., Hinich, M. J., Jungeilges, J. A., Kaplan, D. T. and Jensen, M. J. (1997). A single-blind controlled competition among tests for nonlinearity and chaos. *Journal of Econometrics*, 82 (1), 157-192.

Bierens, H. J. (1996). *Topics in Advanced Econometrics: Estimation, Testing, and Specification of Cross-Section and Time Series Models*. Cambridge: Cambridge University Press.

Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, 31 (3), 307-327.

Bollerslev, T., Chou, R. Y. and Kroner, K. F. (1992). ARCH Modeling in Finance: A Review of the Theory and Empirical Evidence. *Journal of Econometrics*, 52 (1), 5-59.

Bollerslev, T., Engle, R. F., and Nelson, D. B. (1994). ARCH Models (Cap. 49). En: R. Engle and D. McFadden (eds.), *Handbook of Econometrics*, 4 (pp. 2959-3038). Amsterdam: North-Holland.

Brock, W. A. and de Lima, P. J. F. (1995). *Nonlinear Time Series, Complexity Theory, and Finance*. Working paper. Social Systems Research Institute, University of Wisconsin.

Brock, W. A., Dechert, W. D., and Scheinkman, J. A. and LeBaron, BD. (1996). A test for independence based on the correlation dimension. *Econometric Reviews*, 15 (3), 197-235.

Brock, W. A., Hsieh, D. A. LeBaron, B. D. (1991). *Nonlinear Dynamics, Chaos, and Instability: Statistical Theory and Economic Evidence*. Cambridge: MIT Press.

Campbell, J., Lo, A. and MacKinlay, C. (1997). *The Econometrics of Financial Markets*. Princeton: Princeton University Press.

Casti, J. L. (1992). *Reality Rules*. London: Wiley-Interscience.

Chatfield, C. (2003). *The Analysis of Time Series*. Boca Raton: Chapman and Hall/CRC.

Dahl, C. M. and Nielsen, S. (2001). *The random walk of stock prices: implications of recent nonparametric tests*. Working paper. Department of Economics, Copenhagen Business School.

Darrat, A. F. (1990). Stock Returns, Money, and Fiscal Deficits. *The Journal of Financial and Quantitative Analysis*, 25 (3), 387-398.

De Wachter, S. (2004). Course notes for Financial Econometrics. Department of Economics, Nuffield College, Oxford University.

Dechert, W. A. (1988). Characterization of Independence for a Gaussian Process in Terms of the Correlation Dimension. SSRI Working Paper 8812. University of Wisconsin at Madison.

Dimson, E. and Mussavian, M. (1998). A brief history of market efficiency. *European Financial Management*, 4 (1), 91-103.

Engle, R. F. (1982). Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation. *Econometrica*, 50 (4), 987-1008.

Grassberger, P. and Procaccia, I. (1983). Measuring the strangeness of strange attractors. *Physica D: Nonlinear Phenomena*, 9 (1-2), 189-208.

Hamilton, J. D. (1994). *Time Series Analysis*. Princeton: Princeton University Press.

Hansen, P. R. and Lunde, A. (2005). A Forecast Comparison of Volatility Models: Does Anything Beat a GARCH (1, 1). *Journal of Applied Econometrics,* 20 (7), 873-889.

Heaton, J. T. (2005). *Introduction to Neural Networks with Java*. Chesterfield: Heaton Research, Inc.

Jungeilges, J. A. (1993). Operational characteristics of White's test for neglected nonlinearities. En: W. Barnett, A. Kirman, & M. Salmon (Eds.), *Nonlinear dynamics in economics*. Cambridge: Cambridge University Press.

Kanzler, L. (1998). A Study of the Efficiency of the Foreign Exchange Market Through Analysis of Ultra-high Frequency Data. D. Phil. Thesis. Sub-Faculty of Economics, University of Oxford.

Kanzler, L. (1999). Very Fast and Correctly Sized Estimation of the BDS Statistic. Manuscript. Department of Economics, University of Oxford.

Kosfeld, R. and Robe, S. (2001). Testing for nonlinearities in German bank stock returns. *Empirical Economics*, 26 (3), 581-597.

Kuchta, S. (2004). Nonlinearity and Chaos in Macroeconomics and Financial Markets. Essay. Department of Economics, University of Connecticut.

Kyrtsoua, C. and Serletis, A. (2006). "Univariate tests for nonlinear structure", *Journal of Macroeconomics,* 28 (1), 154-168.

Kyrtsoua, C., Labys, W. and Terraza, M. (2001). Heterogeneity and Chaotic Dynamics in Commodity Markets. Working Paper. LAMETA, University of Montpellier I.

LeBaron, B. (1997). A Fast Algorithm for the BDS Statistic. *Studies in Nonlinear Dynamics and Econometrics*, 2 (2), 53-59.

Lee, T-H. (2001). Neural Network Test and Nonparametric Kernel Test for Neglected Nonlinearity in Regression Models. *Studies in Nonlinear Dynamics and Econometrics*, 4 (4), 1063-1063.

Lee, T-H., White, H. and Granger, C.W.J. (1993). Testing for Neglected Nonlinearity in Time Series Models: A Comparison of Neural Network Methods and Alternative Tests. *Journal of Econometrics,* 56 (3), 269-290.

Lo, A. W. and MacKinlay, A. C. (1988). Stock market prices do not follow random walks: evidence from a simple specification test. *Review of Financial Studies*, 1 (1), 41-66.

Lo, A. W., Mamaysky, H., and Wang, J. (2000). Foundations of Technical Analysis: Computational Algorithms, Statistical Inference, and Empirical Implementation. *Journal of Finance*, 55 (4), 1705-1765.

Markellos, R.N. (2002). Nonlinear Dynamics in Economics and Finance. Working Paper 8/02. Dept. of Management Science and Technology, Athens Uinversity of Economics and Business.

McKenzie, M. D., Mitchell, H., Brooks, R. D. and Faff, R. W. (2001). Power ARCH modelling of commodity futures data on the London Metal Exchange. *The European Journal of Finance*, 7 (1), 22-38.

McNelis, P. D. (2005). *Neural Networks in Finance: Gaining Predictive Edge in the Market*. Amsterdam: Elsevier Academic Press.

Milevsky, M. A. (1995). An Empirical Examination of the Geometric Brownian Motion Hypothesis Via the Space-Time Duality of a Stochastic Process. Working Paper. Faculty of Administrative Studies, York University.

Opong, K. K., Mulholland, G., Fox, A. F., and Farahmand, K. (1999). The behaviour of some UK equity indices: An application of Hurst and BDS tests. *Journal of Empirical Finance*, 6 (3), 267-282.

Peel, D. A. and Speight, A. E. H. (1998). Modelling Business Cycle Nonlinearity in Conditional Mean and Conditional Variance: Some International and Sectoral Evidence. *Economica*, 65 (5), 211-229.

Pesaran, M. H. and Potter, S. M. (1992). Nonlinear Dynamics and Econometrics: An Introduction. *Journal of Applied Econometrics,* 7 (1), 1-7.

Ramsey, J. B. (1996). If nonlinear models cannot forecast, what use are they? *Studies in Nonlinear Dynamics and Econometrics*, 1 (2), 65-86.

Reitgruber, W. and Sterlina, I. (1995). On the forecastability of share prices on the Viennese stock exchange. *Empirical Economics*, 20 (3), 415-433.

Rothman, P. (2001). MA, AR, and ARMA Models (Cap. 7). Economics 5000: Time Series Analysis Course Notes. Department of Economics, East Carolina University.

Scheinkman, J. A. and LeBaron, B. (1989). Nonlinear Dynamics and Stock Returns. *The Journal of Business*, 62 (3), 311-337.

Schmidt, S. (2005). *Econometrics*. New York: McGraw-Hill Irwin.

Taylor, S. (1986). *Modelling financial time series*. New York: Wiley.

Teräsvirta, T. (1996). Power Properties of Linearity Tests for Time Series, *Studies in Nonlinear Dynamics and Econometrics,* 1 (1), 3-10.

Teräsvirta, T., Tjøstheim, D. and Granger, C. W. J. (1994). Aspects of modelling nonlinear time series. En: R. Engle & D. McFadden (Eds.), *Handbook of Econometrics,* 4 (pp. 2917-2957). Amsterdam: North-Holland.

Tsay, R. (2001). *Analysis of Financial Time Series*. New York: Wiley.

White, H. (1989). Some Asymptotic Results for Learning in Single Hidden-Layer Feedforward Network Models. *Journal of the American Statistical Association*, 84 (408), 1003-1013.

Wurtz, D., Chalabi, Y. and Luksan, L. (2006). Parameter Estimation of ARMA Models with GARCH/APARCH Errors An R and SPlus Software Implementation. A aparecer en *Journal of Statistical Software*.

Apéndice

Cuadro 1A

Test BDS (Brock *et al*., 1996) para los residuos estandarizados del mejor modelo ARMA-GARCH de los log-retornos del IGBC

Dimensión	Estadístico BDS	Error Estándar	Estadístico z	Valor p	Valor p Bootstrapping
2	0.004347	0.002442	1.780304	0.0750	0.0828
3	0.010614	0.003938	2.695451	0.0070	0.0110
4	0.012882	0.004757	2.707823	0.0068	0.0110
5	0.014306	0.005031	2.843721	0.0045	0.0088
6	0.014584	0.004922	2.963168	0.0030	0.0066
7	0.015060	0.004575	3.291609	0.0010	0.0034
8	0.014688	0.004102	3.580670	0.0003	0.0016
9	0.014623	0.003580	4.084923	0.0000	0.0010

Fuente de la serie de tiempo: Superintendencia Financiera de Colombia.

El cuadro 1A muestra los resultados del test BDS cuya hipótesis nula para este caso es que los residuos estandarizados del mejor modelo ARMA-GARCH son IID. De acuerdo con el modelo teórico, en ausencia de no linealidad en media, si el modelo ARMA-GARCH está bien especificado, los residuos estandarizados deberían ser IID. El test se evaluó para todas las dimensiones m (con que se definen las m-historias) entre 2 y 9. Para cada dimensión, el tamaño de la cercanía ε se escogió de tal forma que la integral de correlación de dimensión 1 $c_{1,n}(\varepsilon)$ fuera igual a 0.71, siguiendo la recomendación de Kanzler (1999, p. 53). Las columnas de la derecha del cuadro muestran los valores p del estadístico BDS basados en la distribución normal, y los valores p basados en *bootstrapping* con 10 000 simulaciones. Los resultados del test revelan que la hipótesis de IID se rechazó (al 2%) para todas las dimensiones, (excepto para $m = 2$ que se rechazó al 10%). De acuerdo con el modelo teórico, la implicación directa de este resultado es que la hipótesis de que no existe no linealidad en media en los log-retornos del IGBC fue rechazada.

Cuadro 2A

Test de White (1989) de no linealidad basado en redes neuronales para los log-retornos del IGBC

Estadístico (chi^2)	Valor p
88.2464	$< 2.2 \cdot 10^{-16}$
72.1035	$2.22 \cdot 10^{-16}$
57.8758	$2.71 \cdot 10^{-13}$
76.1923	$2.20 \cdot 10^{-16}$
63.5619	$1.58 \cdot 10^{-14}$

Fuente de la serie de tiempo: Superintendencia Financiera de Colombia.

Resultados del test de White repetido cinco veces con diferentes pesos iniciales para la red neuronal. La hipótesis nula del test es la linealidad en media de la serie de los log-retornos, y el estadístico de prueba que resulta tiene una distribución Chi cuadrado asintótica (en este caso con dos grados de libertad) bajo la hipótesis de no linealidad en media. El test evalúa si una función mensurable de la historia del proceso se correlaciona con los residuos de un filtro lineal AR(p) ajustando la serie de tiempo por medio de una red neuronal con conexiones hacia delante (*feed-forward*) con una única capa oculta. Esta metodología tiene la ventaja de que es innecesario un prefiltrado de la varianza condicional. El orden del filtro AR se escoge por medio del criterio de información de Schwartz (BIC), siguiendo a Jungeilges (1993), y dado que no existían términos MA en el mejor filtro ARMA de la serie se tomó el orden AR de este mejor filtro (i.e. el orden fue 1). Los resultados del test rechazan la hipótesis nula de no linealidad en media en los log-retornos del IGBC.