

# Modelos TAR en series de tiempo financieras

## TAR models in financial time series

Edna C. Moreno<sup>a</sup>  
ednamoreno@usantotomas.edu.co

Fabio H. Nieto<sup>b</sup>  
fhnietos@unal.edu.co

### Resumen

En este artículo, se evalúa el desempeño de un modelo autorregresivo de umbrales (TAR), en el análisis de series de tiempo financieras. Se utilizan datos del mercado accionario brasilero y norteamericano a fin de ajustar un modelo; además se realiza una comparación con los modelos GARCH vía los momentos condicionales.

**Palabras clave:** heterocedasticidad condicional, modelo TAR, series temporales financieras.

### Abstract

The performance of TAR models to analyse financial time series is evaluated. Empirically, and using data from the Brazilian stock market, the TAR model is compared with GARCH models via conditional moments.

**Keywords:** conditional heteroscedasticity, financial time series, TAR model.

## 1. Introducción

En muchas situaciones prácticas es necesaria la construcción de un modelo que represente el comportamiento dinámico de un proceso estocástico el cual genera una serie de tiempo financiera. Si bien los modelos más utilizados son los ARCH Y GARCH que permiten una especificación de la heterocedasticidad condicional, varios autores han encontrado apropiado utilizar modelos no lineales para reflejar los hechos estilizados característicos de este tipo de series temporales, dentro de los que se encuentran, aunque no son muy usados, los modelos STAR (Smooth Threshold Autorregresivo), SETAR (Self-Exciting Threshold Autorregresivo) de Tong (1990), los de cambio de régimen de Hamilton (1994) y los modelos TAR (Threshold Autoregressive) de Tsay (1998).

<sup>a</sup>Profesor. Universidad Santo Tomás, Colombia.

<sup>b</sup>Profesor. Universidad Nacional de Colombia, Colombia.

Los modelos STAR, SETAR, TAR y de cambio de régimen, se encuentran muy relacionados, pues los cuatro se caracterizan por la presencia de un número determinado de regímenes; sin embargo, todos han puesto en manifiesto dificultades en los procesos de identificación y estimación de los parámetros estructurales (número de regímenes y valores umbrales). Los SETAR y Cambio de Régimen se caracterizan por la presencia de un número discreto de regímenes y en el modelo SETAR la variable de umbrales es la misma variable de interés rezagada  $d$  periodos. Finalmente, los modelos TAR dan cuenta de una variable de umbrales diferente a la variable objetivo y un número de regímenes mayor, permitiendo una mayor generalidad. Recientemente Nieto (2005) desarrolló una metodología con el propósito de analizar series de tiempo, en presencia de datos faltantes, que provienen de un modelo no lineal TAR, usando el enfoque Bayesiano para la identificación y estimación de parámetros. En el análisis de series de tiempo económicas, Hoyos (2006) usa dicha metodología; mientras que Nieto (2005) la utilizó en el análisis de series de tiempo hidrológicas/ meteorológicas.

Teniendo en cuenta las ventajas del modelo TAR frente a otros de su misma naturaleza, y dado que rara vez son considerados en el ámbito de las finanzas resulta entonces apropiado evaluar el comportamiento de dichos modelos en series de tiempo financieras; ya que se ha observado que una trayectoria típica de un proceso estocástico TAR refleja conglomerados de valores grandes.

De esta manera, a continuación se ilustra el desempeño de los modelos TAR y de la metodología de análisis propuesta por Nieto (2005), en series de tiempo financieras, mediante una aplicación al índice de la bolsa de valores de Sao Paulo (BOVESPA) y de la bolsa de valores de norteamericana, analizando el DOWJONES como variable de umbrales. La relación entre estas dos variables tiene soporte empírico dado que el mercado financiero norteamericano tiene influencia sobre el resto de mercados financieros del mundo, tal como está documentado en diversos trabajos. Entre otros Cathy W.S. Chen (2006) y las referencias allí presentadas.

Además, se presenta la comparación del modelo TAR con los modelos GARCH en términos de algunos de los momentos marginales y condicionales del proceso estocástico subyacente. Esta comparación inicial dará luces para una investigación más profunda sobre el tema. Adicionalmente, este ejercicio es importante en el contexto de riesgo financiero, ya que se ilustran diferentes medidas de volatilidad a partir del modelo TAR lo cual proporciona una aproximación simple del valor en riesgo de una posición financiera.

Como se indicará en la sección 2, la metodología se basa fundamentalmente en el supuesto de que el proceso de ruido blanco que genera el proceso de interés es Gaussiano.

El artículo se organiza de la siguiente manera: en la sección 2 se muestran los prerrequisitos teóricos más importantes en el modelo TAR. La sección 3 presenta una aplicación a datos reales, específicamente a la serie de rendimientos del BOVESPA. Las conclusiones se exponen en la sección 4.

## 2. El modelo TAR

### 2.1. Especificación del modelo

Un modelo TAR representa un sistema dinámico con un proceso de entrada  $\{Z_t\}$  y uno de salida  $\{X_t\}$ , tales que

$$X_t = a_0^{(j)} + \sum_{i=1}^{k_j} a_i^{(j)} X_{t-i} + h^{(j)} \varepsilon_t, \quad t \in \mathbb{Z}, \quad (1)$$

si  $Z_t \in R_j = (r_{j-1}, r_j)$  para algún  $j$ ,  $j = 1, \dots, l$ , con  $r_0 = -\infty$  y  $r_l = \infty$ . Además,  $a_i^{(j)}$ ,  $h^{(j)}$  ( $j = 1, \dots, l$ ;  $i = 0, 1, \dots, k_j$ ) son números reales. Los números reales  $r_j$  con  $r_1 < \dots < r_{l-1}$ , que definen  $l$  regímenes para el proceso  $Z_t$ , son llamados los umbrales y los números enteros no negativos  $k_1, \dots, k_l$  representan los órdenes autorregresivos en cada régimen. Adicionalmente, el proceso  $\{\varepsilon_t\}$  es Gaussiano de media cero y varianza 1 y es mutuamente independientes de  $\{Z_t\}$ . Se usa el símbolo  $TAR(l; k_1, \dots, k_l)$  para representar el modelo y  $l; r_1, \dots, r_{l-1}; k_1, \dots, k_l$  se llaman los parámetros estructurales. Los demás parámetros se llaman no estructurales.

En la definición anterior el proceso  $\{Z_t\}$  es exógeno, en el sentido de que no existe retroalimentación de  $\{X_t\}$  hacia él y su comportamiento estocástico está descrito por una cadena de Markov homogénea de orden  $p$ , con distribución inicial  $F_1(z, \theta_z)$  y distribución kernel  $F(z_t/z_{t-1}, \dots, z_{t-p}, \theta_z)$ , donde  $\theta_z$  es un vector de parámetros en un espacio numérico apropiado.

Con base en esta clase de modelos se puede explicar cierto tipo de heterocedasticidad condicional en  $\{X_t\}$ , ya que una realización típica de él puede exhibir conglomerados de datos extremos, como sucede en ciertas variables financieras.

Nieto (2005) desarrolló una metodología estadística para el análisis del modelo en presencia de datos faltantes que consta de las fases de identificación, estimación y validación del modelo. Esta se basa en el enfoque Bayesiano y en los métodos de simulación MCMC.

### 2.2. Estimación del modelo

El problema que se plantea aquí es la estimación de los parámetros no estructurales del modelo,  $a_i^{(j)}$  y  $h^{(j)}$  ( $j = 1, 2, \dots, l$ ); suponiendo que este ya se tiene identificado, es decir que se conocen los parámetros estructurales  $l; r_1, \dots, r_{l-1}$  y  $k_1, \dots, k_l$ . La metodología se fundamenta en el muestreador de Gibbs, para lo cual es necesario disponer de las densidades condicionales completas y de las densidades *a priori*.

Sean  $\theta$  el vector de parámetros desconocidos de  $\{X_t\}$  y  $\{Z_t\}$ ,  $\theta_j = (a_0^{(j)}, a_1^{(j)}, \dots, a_{k_j}^{(j)})$ , ( $j = 1, \dots, l$ ),  $\mathbf{h} = (h^{(1)}, \dots, h^{(l)})$  y  $\theta_x = (\theta_1, \dots, \theta_l, \mathbf{h})$ . Así  $\theta = (\theta_x, \theta_z)$  y la densidad condicional de interés es  $p(\theta|\mathbf{x}, \mathbf{z})$ .

Las densidades *a priori* para las componentes de  $\theta_x$  son las siguientes:

$$\theta_j \sim N(\theta_{o,j}, V_{o,j}^{-1}), \quad j = 1, \dots, l,$$

con  $\theta_{o,j}$  un vector de  $k_j + 1$  componentes y  $V_{o,j}^{-1}$  una matriz diagonal, y

$$[h^{(j)}]^2 \sim IG(\gamma_{o,j}, \beta_{o,j})$$

con  $j = 1, \dots, l$ . «IG» representa la distribución Gamma inversa. Bajo estas consideraciones las densidades *a posteriori* se presentan en las siguientes Proposiciones planteadas por Hoyos (2006). Se asume *a priori* que los parámetros entre regímenes son independientes, que  $\theta_j$  y  $\mathbf{h}^{(j)}$  son también independientes, y que  $\theta_x$  no depende de  $\theta_z$ .

**Proposición 1.** Para cada  $j = 1, \dots, l$ , sea  $\{t : j_t = j\} = \{t_{1,j}, \dots, t_{n_j,j}\}$  donde  $n_j$  denota el número de observaciones de  $X$  en el régimen  $j$ , la distribución condicional completa de  $\theta_j$  dado  $\theta_i; i \neq j; \mathbf{h}; \theta_z, \mathbf{x}$  y  $\mathbf{z}$ , es multinormal con media

$$\theta_j^* = V_j^{-1}[(1/\{h^{(j)}\}^2)\mathbf{W}_j' \mathbf{X}_j + V_{o,j}\theta_{o,j}]$$

y varianza

$$V_j^{-1} = [(1/\{h^{(j)}\}^2)\mathbf{W}_j' \mathbf{W}_j + V_{o,j}]^{-1}$$

con  $\mathbf{W}_j = (w_{t_{1,j}}, \dots, w_{t_{n_j,j}})'$ ,  $\mathbf{w}_t' = (1, x_{t-1}, \dots, x_{t-k_j})$ ,  $\mathbf{X}_j = (x_{t_{1,j}}, \dots, x_{t_{n_j,j}})'$ .

**Proposición 2.** La densidad condicional completa de  $[h^{(j)}]^{-2}$  dado  $\theta_j, \theta_z, \mathbf{x}$  y  $\mathbf{z}; j = 1, \dots, l$ ; corresponde a una distribución Gamma con parámetro de forma  $\gamma_{o,j} + \frac{n_j}{2}$  y parámetros de escala  $\beta_{o,j} + (1/2)(\mathbf{X}_j - \mathbf{W}_j\theta_j)'(\mathbf{X}_j - \mathbf{W}_j\theta_j)$ .

Se recomienda estimar  $\theta_z$  de acuerdo con el problema a la mano. Una vez se tengan las densidades condicionales completas se usa el muestreador de Gibbs a fin de obtener las estimaciones de los parámetros, tomando los promedios muestrales desde las muestras generadas, véase Nieto (2005) para mayores detalles.

### 2.3. Validación del modelo

Para cada  $t = 1, \dots, T$ , sea

$$\hat{e}_t = (X_t - X_{t|t-1})/h^{(j)}$$

si  $Z_t \in B_j$  para algún  $j(j = 1, \dots, n)$ , donde

$$X_{t|t-1} = a_0^{(j)} + \sum_{i=1}^{k_j} a_i^{(j)} X_{t-i|t-1}$$

es el predictor un paso adelante de  $X_t$ . En este documento, se utilizan las gráficas CUSUM y CUSUMSQ y la función de autocorrelación parcial de los residuos para chequear heterocedasticidad en  $\{\varepsilon_t\}$ , la especificación del modelo y la no correlación serial en los residuos, con base en el conjunto  $\{\hat{e} : t = k_m, \dots, T\}$ ,  $k_m = \max\{k_1, \dots, k_l\}$ .

## 2.4. Identificación del modelo

El primer paso es el ajuste del modelo TAR es la identificación de los parámetros estructurales. Se debe (1) identificar el número de regímenes y los valores umbrales y (2) especificar los órdenes autorregresivos.

Como una alternativa a las consideraciones empíricas y al criterio AIC, que son los métodos frecuentemente utilizados en la identificación del modelo TAR, Nieto (2005) propone usar los métodos MCMC y el enfoque Bayesiano para resolver los problemas (1) y (2). Con la perspectiva Bayesiana, es posible establecer un conjunto de distribuciones *a priori* para el número de regímenes y los órdenes autorregresivos en cada uno y proceder a obtener las *a posteriori* y de allí lograr los estimadores óptimos. Para mayor información véase Nieto (2005).

## 2.5. Otras características del modelo TAR

Con la intención de explorar más adelante la relación entre los momentos marginales y condicionales, tanto en el modelo TAR como en el modelo GARCH, se hace necesario especificar algunos de estos momentos teóricos ya utilizados en algunas aplicaciones. Para esto, se debe conocer la distribución marginal y condicional de  $X_t$  con  $t = 1, 2, \dots, \infty$ .

Es posible demostrar que, para cada  $t$ , la función de distribución acumulada marginal de  $X_t$  está dada por

$$\begin{aligned} F_t(x) &= \sum_{j=1}^l P((X_t \leq x)/Z_t \in B_j)P(Z_t \in B_j) \\ &= \sum_{j=1}^l p_j F_{t,j}(x), \end{aligned} \quad (2)$$

donde  $F_{t,j}(x) = P((X_t \leq x)/Z_t \in B_j)$  y  $P(Z_t \in B_j) = p_j$ . Nótese además que  $\sum_{j=1}^l p_j = 1$ , por lo que  $F_t(x)$  es una mezcla de funciones de distribución acumuladas condicionales. Ahora, siempre que las raíces del polinomio  $\phi_j(y) = 1 - \sum_{i=1}^{k_j} a_i^{(j)} y^i$  en el modelo (1), estén por fuera del círculo unitario y  $\varepsilon_t \sim N(0, 1)$  es posible verificar, para cada  $j = 1, \dots, l$ , que

$$X_t/Z_t \in R_j \sim N\left(\psi_j(1) a_0^{(j)}, (h^{(j)} \sigma_j)^2\right), \quad (3)$$

donde  $\psi_j(y) = \frac{1}{\phi_j(y)} = \sum_{i=0}^{\infty} \psi_i^{(j)} y^i$ ,  $\sigma_j^2 = \sum_{i=0}^{\infty} (\psi_i^{(j)})^2$  y  $\Phi_{0,1}(\cdot)$  denota la función de distribución acumulada de una normal estándar.

En segundo lugar, denotando  $\tilde{x}_{t-1} = \{x_{t-1}, \dots, x_1\}$ , la distribución condicional de  $X_t$  dado  $\tilde{x}_{t-1}$  está dada por

$$F_t(x/\tilde{x}_{t-1}) = \sum_{j=1}^l p_j P((X_t \leq x)/Z_t \in B_j, \tilde{x}_{t-1}). \quad (4)$$

Bajo el supuesto de que  $\varepsilon_t \sim N(0, 1)$ ,  $P((X_t \leq x)/Z_t \in B_j, \tilde{x}_{t-1})$  corresponde a la distribución acumulada de la normal con media  $a_0^{(j)} + \sum_{i=1}^{k_j} a_i^{(j)} x_{t-i}$  y varianza  $[h^{(j)}]^2$ , es decir

$$X_t/Z_t \in R_j, \tilde{x}_{t-1} \sim N\left(a_0^{(j)} + \sum_{i=1}^{k_j} a_i^{(j)} x_{t-i}, [h^{(j)}]^2\right). \quad (5)$$

A partir de los resultados anteriores, podemos obtener la función de medias y la función de varianzas del proceso  $\{X_t\}$ . Por definición del valor esperado y de (2), se obtiene que, para todo  $t$ ,

$$E(X_t) = \int_{\mathbb{R}} X_t dF_t(x) = \sum_{j=1}^l p_j \int_{\mathbb{R}} X_t dF_{j,t}(x).$$

Notando  $\int_{\mathbb{R}} X_t dF_{j,t}(x) = \mu_{t,j,1}$ , que se entiende como el primer momento de  $X_t$  condicional al régimen  $j$ , se obtiene que

$$E(X_t) = \sum_{j=1}^l p_j \mu_{t,j,1}.$$

la cual no depende de  $t$ .

De manera similar, con  $\int_{\mathbb{R}} X_t^2 dF_{j,t}(x) = \mu_{t,j,2}$  y teniendo en cuenta que  $\text{var}(X_t) = E(X_t^2) - (E(X_t))^2$ , se deduce la siguiente expresión para la varianza marginal:

$$\text{var}(X_t) = \sum_{j=1}^l p_j \mu_{t,j,2} - \left( \sum_{j=1}^l p_j \mu_{t,j,1} \right)^2$$

para todo  $t$ . De (3) es posible observar que

$$\mu_{t,j,1} = \frac{a_0^{(j)}}{\phi_j(1)}$$

y

$$\mu_{t,j,2} = (h^{(j)}\sigma_j)^2 + \left(\frac{a_0^{(j)}}{\phi_j(1)}\right)^2$$

para todo  $t$ . De estos dos resultados se resalta que la función de medias y la función de varianzas son constantes, cuando las raíces de los polinomios característicos en cada régimen estén por fuera del círculo unitario.

Por último, condicionando únicamente en la información observada para la variable  $X$  hasta el tiempo  $t - 1$  y usando (4) y (5), se puede probar que Moreno (2010)

$$E(X_t/\tilde{x}_{t-1}) = \sum_{j=1}^l p_j a_0^{(j)} + \sum_{j=1}^l p_j a_1^{(j)} x_{t-1} + \cdots + \sum_{j=1}^l p_j a_{k_j}^{(j)} x_{t-k_j}$$

y

$$\begin{aligned} & var(X_t/\tilde{x}_{t-1}) \\ &= \sum_{j=1}^l p_j \left[ (h^{(j)})^2 - \left( a_0^{(j)} + \sum_{i=1}^{k_j} a_i^{(j)} x_{t-i} \right)^2 \right] - \left[ \sum_{j=1}^l p_j \left( a_0^{(j)} + \sum_{i=1}^{k_j} a_i^{(j)} x_{t-i} \right) \right]^2. \end{aligned}$$

Los otros momentos condicionales están dados en las expresiones (3) y (5). De esta manera se distinguen tres tipos de condicionamiento, en los regímenes, en el conjunto de información de  $X$  hasta el tiempo  $t - 1$  y en ambos casos.

### 3. Una aplicación a datos reales

Determinar patrones de comportamiento en mercados financieros ha sido un asunto de mucho interés. Esta sección busca evaluar el desempeño del modelo TAR propuesto por Nieto (2005), en el análisis de los hechos estilizados del mercado accionario. Se utilizará la serie financiera de la bolsa de valores de Sao Paulo. Paralelamente, se realizará una comparación con los modelos GARCH, muy conocidos en el ámbito financiero.

De esta manera, la variable de interés en la aplicación corresponde a los rendimientos diarios del índice BOVESPA y la variable de umbrales es el índice industrial Dow Jones, el principal índice del mercado estadounidense. Entre las acciones que conforman el Dow Jones se encuentran nombres mundialmente reconocidos como Coca Cola, General Motors, McDonalds, Procter and Gamble, IBM y Microsoft.

Los datos fueron obtenidos del sistema de información Bloomberg y corresponden a los rendimientos de los índices en los días hábiles de las bolsas de los tres países. Sin embargo, para efectos del ajuste de los modelos, se completaron las series de tal manera que los valores corresponden a los rendimientos de los días lunes a viernes, para esto se utilizó la metodología propuesta por Nieto & Ruiz (2002).

Es importante decir que no se tiene en cuenta el factor «fin de semana». Además, cabe recordar que por volatilidad se entenderá la desviación estándar condicional a la información pasada de los retornos. Más exactamente, tomamos

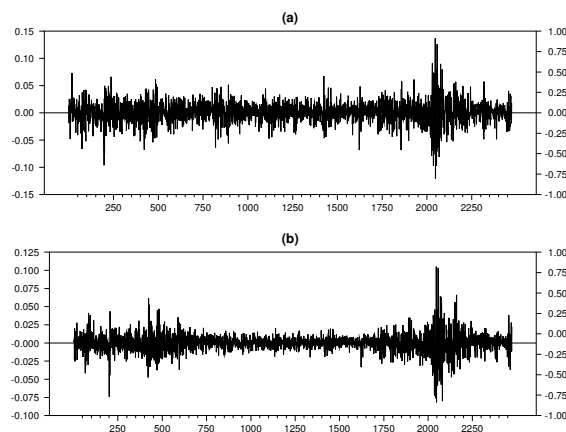
$$X_t = \ln(BOVESPA_t) - \ln(BOVESPA_{t-1})$$

$$Z_t = \ln(DOWJONES_t) - \ln(DOWJONES_{t-1}).$$

Se consideran 2473 datos que corresponden a las observaciones registradas desde el viernes 8 de diciembre de 2000 hasta el miércoles 2 de junio de 2010. En la Figura 1 se aprecia claramente los llamados conglomerados de valores extremos, bien conocidos en el ámbito financiero.

La Figura 2 muestra correlaciones ligeramente significativas en los rezagos 3 y 7 para la serie de retornos del BOVESPA, en tanto que en la Figura 3, se observa la alta dependencia que existe en los retornos al cuadrado de la variable de interés.

Figura 1: *retornos logarítmicos: (a) BOVESPA, (b) DOWJONES. Fuente: elaboración propia.*



### 3.1. Ajuste de un modelo TAR

El primer paso al ajustar un modelo TAR es explorar si el proceso  $X_t$  es no lineal. Con este fin, se utiliza la estadística de prueba propuesta por Tsay (1998) para modelos de umbrales multivariados, en la que se contrasta la hipótesis nula de linealidad contra la alternativa de no linealidad explicada por el modelo TAR. En este ejemplo, la prueba se realiza para diferentes rezagos de  $Z_t$ , esto es, para  $Z_{t-d}$  con  $d = 0, 1, 2, 3, \dots, 7$ . El orden autorregresivo considerado, con base en los criterios AIC y BIC, fue  $\bar{k} = 4$ . Los resultados se presentan en el Tabla 1, en la que se observa que la variable de umbrales apropiada debe ser la misma  $Z_t$  sin rezagar, que es donde se obtiene el mayor valor de la estadística de prueba y menor valor  $p$ .



Figura 2: función de autocorrelación (a) y de autocorrelación parcial (b) de la serie de retornos del BOVESPA. Fuente: elaboración propia.

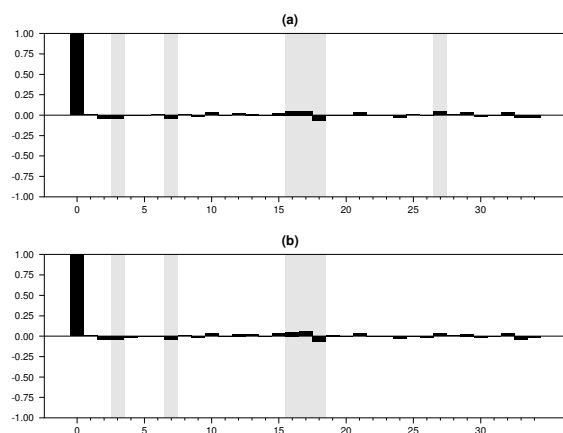


Figura 3: función de autocorrelación (a) y de autocorrelación parcial (b) de la serie de retornos al cuadrado del BOVESPA. Fuente: elaboración propia.

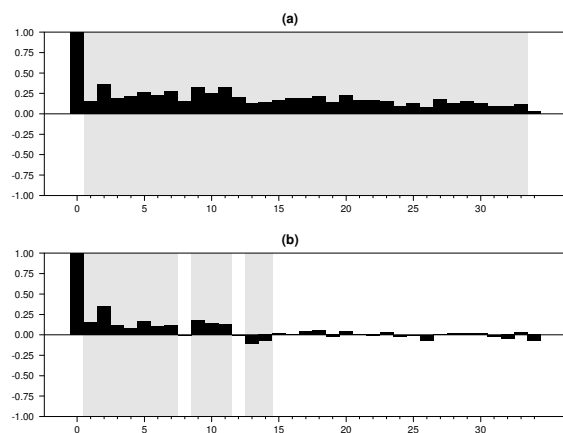


Tabla 1: prueba de no linealidad para BOVESPA. Fuente: elaboración propia.

Test	Rezago de la variable umbral							
	0	1	2	3	4	5	6	7
$\chi^2$	841.75	9.21	2.99	3.02	3.38	23.12	7.06	17.26
p-valor	0.000	0.021	0.000	0.700	0.362	0.000	0.058	0.000

En lo que sigue, será necesario especificar las densidades *a priori* para los parámetros no estructurales, dadas en la sección 2.2. Los parámetros utilizados para estas

densidades *a priori* fueron  $\theta_{o,j} = \mathbf{0}$ ,  $V_{0,j}^{-1} = 0.01\mathbf{I}$  donde  $\mathbf{I}$  es la matriz identidad,  $\gamma_{o,j} = 1.5$ ,  $\beta_{o,j} = \tilde{\sigma}^2/2$ , con  $j = 1, 2, 3$  y  $\tilde{\sigma}^2 = 0.000374$  que corresponde a la varianza de los residuos de ajustar un modelo lineal  $AR(4)$  a la serie de rendimientos del BOVESPA. Además, el orden autorregresivo máximo escogido para todos los regímenes fue el mismo y corresponde a  $\bar{k} = 4$ . Las distribuciones *a priori* para el número de regímenes y para los órdenes autorregresivos utilizadas en la identificación de  $l$ , fueron  $\pi_2 = \pi_3 = 0.5$  y  $p(k_{il}/l) = 0.2$  donde  $i = 1, \dots, l$  y  $k_{il} = 0, 1, 2, 3, 4$ .

Para la identificación de los umbrales se tomaron como posibles candidatos los percentiles de la forma  $5k$ , con  $k = 1, \dots, 19$ , de la distribución empírica de  $Z$ . Para la escogencia del umbral del modelo con dos regímenes ( $M_2$ ), se realizó una búsqueda entre el conjunto de posibles candidatos y para la detección de los umbrales del modelo  $M_3$  la búsqueda se hizo entre todas las combinaciones de las parejas  $r_i < r_j$ ,  $i < j$  de los cuantiles seleccionados. Los umbrales son seleccionados de acuerdo al mínimo NAIC (Tong 1990),

$$NAIC = \frac{\sum_{j=1}^k AIC_j}{\sum_{j=1}^k n_j},$$

donde  $AIC_j$  y  $n_j$  denotan el criterio de información de Akaike (AIC) y el número de observaciones en el  $j$ -ésimo régimen.

Los resultados para los posibles regímenes se presentan en la Tabla 2.

Tabla 2: *identificación de umbrales para el BOVESPA. Fuente: elaboración propia.*

$l$	Umbrales	Órdenes autorregresivos	Mínimo NAIC
2	0.0004	3, 4	-1.65
3	-0.0054, 0.0031	2, 0, 3	-2.17

Con este conjunto de posibles umbrales, se calculó la función de probabilidad *a posteriori* para el número de regímenes, la cual se presenta en la tabla 3. Para esto se realizaron 3.000 iteraciones y se tomó el 10 % de las muestras como punto de calentamiento. La convergencia del muestreador fue monitoreada via el enfoque de estacionariedad. Los resultados muestran que el número de regímenes apropiado para los datos es dos.

Tabla 3: *distribución a posteriori para el número de regímenes en el caso del BOVESPA. Fuente: elaboración propia.*

$l$	
2	3
0.55	0.44

El umbral definitivo para el modelo con dos regímenes corresponde al identificado en la Tabla 2, esto es  $r_1 = 0.0004$ . Sin embargo, el modelo ajustado no fue apropiado para explicar la heterocedasticidad marginal en los datos. Así, con base en el criterio de información NAIC se seleccionó el modelo TAR con tres regímenes y con umbrales  $r_1 = -0.0054$  y  $r_2 = 0.0057$ .

Condicional en  $\hat{l} = 3$ , los órdenes autorregresivos  $k_1, k_2, k_3$  fueron reestimados (Tabla 4). Se realizaron 6.000 iteraciones y se tomó el 10 % de las muestras como punto de calentamiento. La convergencia del muestreador fue monitoreada via el enfoque de estacionariedad.

Se realizó un análisis de sensibilidad modificando los valores a priori de los parámetros no estructurales y las probabilidades a priori de los órdenes autorregresivos y los resultados anteriores se mantienen. Los órdenes autorregresivos estimados finalmente son  $\hat{k}_1 = 2$ ,  $\hat{k}_2 = 0$  y  $\hat{k}_3 = 4$ .

Tabla 4: *distribución a posteriori para los órdenes autorregresivos en el caso del BOVESPA. Fuente: elaboración propia.*

Orden	Régimen		
	1	2	3
0	0.20	0.33	$6.48 \times 10^{-4}$
1	0.22	0.24	$1.00 \times 10^{-4}$
2	0.25	0.17	$2.37 \times 10^{-5}$
3	0.15	0.14	0.09
4	0.17	0.12	0.90

Con lo anterior se concluye la etapa de identificación del modelo, obteniéndose un  $TAR(3; 2, 0, 4)$  con umbrales  $r_1 = -0.0054$  y  $r_2 = 0.0057$ , los cuales corresponden a los percentiles 25 y 75 de los rendimientos logarítmicos del DOWJONES. La estimación de los parámetros de este modelo se presenta en la Tabla 5, junto con sus respectivos errores estándar e intervalos de credibilidad del 90 %. Se verificó la convergencia del muestreador a través de las funciones de autocorrelación para todos los parámetros. Estos resultados pueden ser suministrados por los autores.

Los resultados de la Tabla 5 muestran que todos los coeficientes autorregresivos son estadísticamente significativos al 10 % y que la varianza de los términos de error no varía mucho en los tres regímenes. En cuanto al análisis de sensibilidad de las estimaciones de los parámetros, los parámetros estimados no presentan mayores alteraciones cuando se modifican los valores *a priori*. Esto resultados tampoco cambian cuando se realiza un número mayor de iteraciones o cuando se toma un punto de calentamiento grande. El modelo ajustado es el siguiente:

$$X_t = \begin{cases} -0.0127 + 0.1113X_{t-1} - 0.0685X_{t-2} \\ + 0.0198\varepsilon_t, & \text{si } Z_t < -0.0054, \\ 6.81 \times 10^{-4} + 0.0137\varepsilon_t, & \text{si } -0.0054 \leq Z_t < 0.0057, \\ 0.0135 - 0.0837X_{t-1} - 0.0684X_{t-2} \\ - 0.1687X_{t-3} - 0.0633X_{t-4} + 0.0191\varepsilon_t, & \text{si } Z_t \geq 0.0057 \end{cases} \quad (6)$$

Al examinar los residuales del modelo, se encontró que no hay autocorrelación en los residuos estandarizados (Figura 4 parte (a)). Las gráficas CUSUM y CUSUMSQ (Figura 5) indican que no hay evidencia de una especificación incorrecta del modelo, aunque se observa un poco de heterocedasticidad marginal. Pese a esto, el modelo tiene un ajuste aceptable, a la luz de los criterios de Nieto (2005).

Por otro lado, se observan correlaciones parciales significativas en los residuos al cuadrado y en sus valores absolutos. Las gráficas de la función de autocorrelación de los residuos al cuadrado y de sus valores absolutos se presentan con la intención de evaluar el desempeño del modelo TAR en la explicación de algunos hechos estilizados típicos de las series financieras, pero es importante recordar que su objetivo no es validar el modelo TAR.

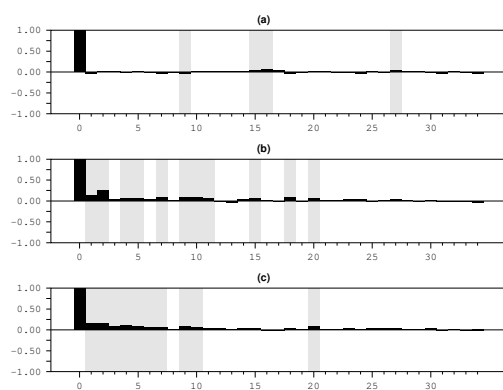


Figura 4: FACP del modelo  $TAR(3;2,0,4)$  en el caso del BOVESPA: (a) Residuos estandarizados (b) Residuos estandarizados al cuadrado y (c) Valores absolutos de los residuos estandarizados. Fuente: elaboración propia.

Tabla 5: *parámetros estimados del modelo TAR(3; 2, 0, 4) para el BOVESPA. Errores estándar en paréntesis e intervalos creíbles en corchetes cuadrados. Fuente: elaboración propia.*

	Régimen 1	Régimen 2	Régimen 3
$a_0^{(j)}$	-0.0127 ( $8.0 \times 10^{-4}$ ) [ -0.0139 , -0.0113 ]	$6.81 \times 10^{-4}$ ( $3.88 \times 10^{-4}$ ) [ $2.85 \times 10^{-5}$ , 0.0013 ]	0.0135 ( $7.67 \times 10^{-4}$ ) [ 0.0122 , 0.0147 ]
$a_1^{(j)}$	0.1112 (0.0363) [0.0508 , 0.1705 ]		-0.0837 (0.0379) [ -0.1458 , -0.0212 ]
$a_2^{(j)}$	-0.0685 (0.0349) [ -0.1274 , -0.0111 ]		-0.0684 (0.0384) [ -0.1325 , -0.0039 ]
$a_3^{(j)}$			-0.1687 (0.0347) [ -0.2261 , -0.1133 ]
$a_4^{(j)}$			-0.0633 (0.0376) [ -0.1263 , -0.0019 ]
$h^{(j)}$	0.0198	0.0137	0.0191

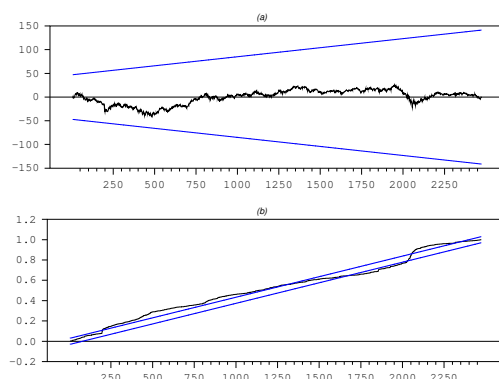


Figura 5: *CUSUM y CUSUMSQ, (a) y (b) respectivamente, para los residuos del modelo TAR(3;2,0,4) en el caso del BOVESPA. Fuente: elaboración propia.*

### 3.2. Ajuste del modelo GARCH a la serie de rendimientos del BOVESPA

Se aplica la metodología convencional para ajustar un modelo  $ARMA(p, q)$  a la media condicional de los datos, posteriormente se ajusta un modelo GARCH para

la varianza condicional y finalmente se hace la estimación conjunta de todos los parámetros del modelo.

Por parsimonia y por dar mejores resultados en la estadística  $Q$  para los residuales, el modelo *ARMA* más adecuado para la serie temporal observada es

$$X_t = 5.73 \times 10^{-4} - 0.04X_{t-3} + a_t$$

Las estadísticas  $t$  y sus niveles de significancia para la media y el coeficiente autorregresivo son 1.53 (0.12) y -2.19 (0.03). Aunque la constante no es significativamente distinta de cero, se decide dejarla en el modelo dado que ella determina en cuanto al retorno esperado no condicional.

Los correlogramas de los retornos al cuadrado (Figura 3) y de los residuos después de ajustar el modelo para la media (Figura 6), muestran claramente la existencia de heterocedasticidad condicional y las pruebas estadísticas no son necesarias. La FACP ilustra que en caso de ajustar un modelo ARCH se necesitarían muchos rezagos. En efecto, un análisis preliminar arroja que un ARCH(11) es un posible modelo. Sin embargo, como veremos a continuación, el modelo más apropiado para la dinámica de la volatilidad de la serie es un GARCH(1,1) con una distribución  $t$  para el proceso  $\{\epsilon_t\}$ . La estimación de los grados de libertad fue de 7.

El modelo estimado es el siguiente:

$$\begin{aligned} X_t &= 0.0014 - 0.0393X_{t-3} + a_t, \\ a_t &= \epsilon_t \sigma_t, \\ \sigma_t^2 &= 7.512 \times 10^{-6} + 0.0736a_{t-1}^2 + 0.9062\sigma_{t-1}^2. \end{aligned} \quad (7)$$

En este modelo todos los parámetros son significativos al 5 %. Los valores de la estadística  $t$  y el valor  $p$  para la constante en la ecuación de la media y el coeficiente autorregresivo son 4.52 ( $6.08 \times 10^{-6}$ ) y -1.92 (0.045), para la ecuación de la varianza estos valores son, respectivamente, de 3.14 (0.0017) 5.97 (0.00) 58.36 (0.00).

La Figura 7 muestra ninguna correlación serial en los residuos, residuos al cuadrado, y en sus valores absolutos, indicando que los modelos para la media y la varianza son adecuados. La estadística de Ljung-Box de los residuales estandarizados es  $Q(40) = 43.81$  con  $p - valor = 0.32$  y de  $\tilde{a}_t^2$  es  $Q(40) = 39.07$  con  $p - valor = 0.51$ . Además, las gráficas CUSUM y CUSUMSQ (Figura 8), muestran que no hay evidencia de una especificación incorrecta del modelo o de heterocedasticidad marginal en  $\{X_t\}$ .

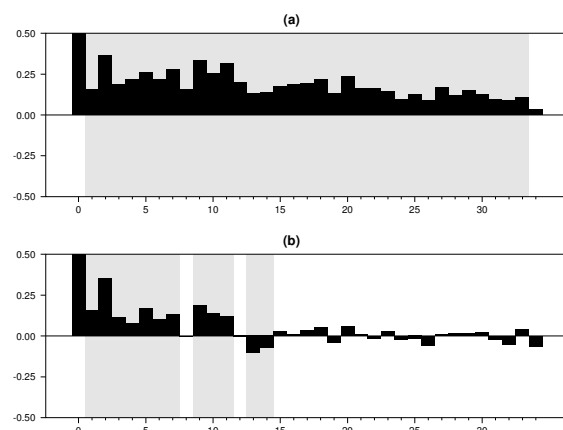


Figura 6: función de autocorrelación (a) y de autocorrelación parcial (b), de los residuos al cuadrado del modelo  $X_t = 5.73 \times 10^{-4} - 0.04X_{t-3} + a_t$  para el caso del BOVESPA. Fuente: elaboración propia.

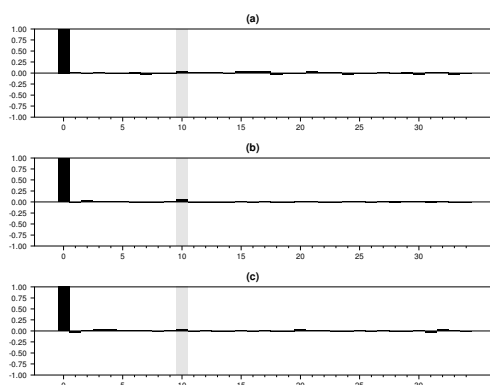


Figura 7: diagnóstico del modelo GARCH para la serie de retornos del BOVESPA. FACP muestral para: (a) los residuales estandarizados (b) los residuos estandarizados al cuadrado y (c) los valores absolutos de los residuos estandarizados. Fuente: elaboración propia.

### 3.3. Momentos condicionales y no condicionales de los modelos TAR y GARCH para la serie de rendimientos del BOVESPA

#### Modelo TAR

El modelo TAR ajustado a la serie de rendimientos logarítmicos del BOVESPA es presentado en la ecuación (6). En este se obtiene que  $p_1 = 0.25$ ,  $p_2 = 0.50$  y

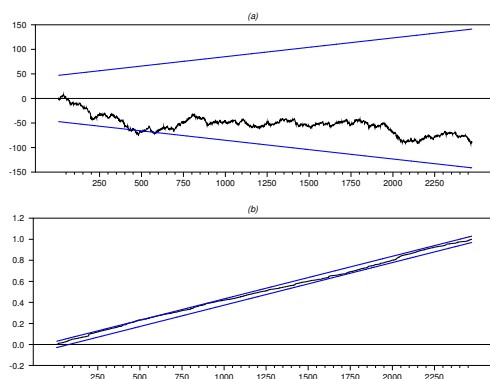


Figura 8: *CUSUM* y *CUSUMSQ* (a) y (b) respectivamente para los residuos estandarizados del modelo GARCH para el BOVESPA. Fuente: elaboración propia.

$p_3 = 0.25$ . Por otro lado, se verificó que las raíces de los polinomios característicos en cada régimen están por fuera del círculo unitario. A partir de esta información se tiene que  $\psi^{(1)}(1) = 1.045$ ,  $\psi^{(2)}(1) = 1$  y  $\psi^{(3)}(1) = 0.722$ . Además,  $\sigma_1^2 = 1.016$ ,  $\sigma_2^2 = 1$  y  $\sigma_3^2 = 1.039$ .

Los valores  $\mu_{t,j,1} = a_0^{(j)}/\phi_j(1)$  y  $\mu_{t,j,2} = (h^{(j)}\sigma_j)^2 + \left(a_0^{(j)}/\phi_j(1)\right)^2$  ( $j = 1, 2, 3$ ), necesarios para el cálculo de los momentos no condicionales se calcularon y los momentos condicionales y no condicionales se presentan en la Tabla 6, recordando que para este modelo existen tres tipos de condicionamiento, en los regímenes, en la información pasada y en ambos conjuntos de información.

Los tres regímenes del modelo TAR deben interpretarse con cuidado. Los umbrales elegidos en el modelo corresponden a los percentiles 25 y 75 de la distribución empírica del DOWJONES, así que los valores presentes en el primer y tercer régimen se pueden interpretar como eventos aislados en el valor de los retornos del índice estadounidense.

En el primer régimen se tienen grandes pérdidas en la bolsa estadounidense, que se reflejan en un retorno esperado del índice brasilero del  $-1.32\%$  ( $100E(X_t/Z_t \in R_1)$ ) y en una desviación estándar diaria del  $2\%$  ( $100\sqrt{\text{var}(X_t/Z_t \in R_1)}$ ); en el segundo régimen, el retorno esperado en el índice de Brasil corresponde al  $0.07\%$  ( $100E(X_t/Z_t \in R_2)$ ) y la desviación es del  $1.41\%$  ( $100\sqrt{\text{var}(X_t/Z_t \in R_2)}$ ); por último, en el tercer régimen de grandes ganancias para el DOWJONES, se observa un retorno esperado en la bolsa brasilera correspondiente al  $0.97\%$  ( $100E(X_t/Z_t \in R_3)$ ) con una desviación diaria del  $2\%$  ( $100\sqrt{\text{var}(X_t/Z_t \in R_3)}$ ).



Tabla 6: *medias y varianzas para el modelo TAR en el caso del BOVESPA. Fuente: elaboración propia.*

Momento	Valor
Media no condicional	$E(X_t) = \sum_{j=1}^l p_j \mu_{t,j,1} = -0.0005$
Varianza no condicional	$var(X_t) = \sum_{j=1}^l p_j \mu_{t,j,2} - \left[ \sum_{j=1}^l p_j \mu_{t,j,1} \right]^2 = 0.0003$
Media condicional a los regímenes	$E(X_t/Z_t \in R_j) = \psi_j(1) a_0^{(j)}$ $E(X_t/Z_t \in R_1) = -0.0132$ $E(X_t/Z_t \in R_2) = 0.0007$ $E(X_t/Z_t \in R_3) = 0.0097$
Varianza condicional a los regímenes	$var(X_t/Z_t \in R_j) = (h^{(j)} \sigma_j)^2$ $Var(X_t/Z_t \in R_1) = 0.0004$ $Var(X_t/Z_t \in R_2) = 0.0002$ $Var(X_t/Z_t \in R_3) = 0.0004$
Media condicional a los regímenes y a la información hasta el tiempo $t - 1$	$E(X_t/Z_t \in R_j, \tilde{x}_{t-1}) = a_0^{(j)} + \sum_{i=1}^{k_j} a_i^{(j)} x_{t-i}$ $E(X_t/Z_t \in R_1, \tilde{x}_{t-1}) = -0.0127 + 0.1113x_{t-1} - 0.0685x_{t-2}$ $E(X_t/Z_t \in R_2, \tilde{x}_{t-1}) = 6.81 \times 10^{-4}$ $E(X_t/Z_t \in R_3, \tilde{x}_{t-1}) = 0.0135 - 0.0837x_{t-1} - 0.0684x_{t-2} - 0.1687x_{t-3} - 0.0633x_{t-4}$
Varianza condicional a los regímenes y a la información hasta el tiempo $t - 1$	$var(X_t/Z_t \in R_j, \tilde{x}_{t-1}) = [h^{(j)}]^2$ $Var(X_t/Z_t \in R_1, \tilde{x}_{t-1}) = 0.0004$ $Var(X_t/Z_t \in R_2, \tilde{x}_{t-1}) = 0.0002$ $Var(X_t/Z_t \in R_3, \tilde{x}_{t-1}) = 0.0004$
Media condicional a la información hasta el tiempo $t - 1$	$E(X_t/\tilde{x}_{t-1}) = \sum_{j=1}^l p_j (a_0^{(j)} + \sum_{i=1}^{k_j} a_i^{(j)} x_{t-i})$ $E(X_t/\tilde{x}_{t-1}) = 0.0005 + 0.0069x_{t-1} - 0.0338x_{t-2} - 0.0416x_{t-3} - 0.0156x_{t-4}$
Varianza condicional a la información hasta el tiempo $t - 1$	$var(X_t/\tilde{x}_{t-1}) = 0.0003 - 0.2475(-0.0127 + 0.1113x_{t-1} - 0.0685x_{t-2})^2 - 0.2467(0.0135 - 0.0837x_{t-1} - 0.0684x_{t-2} - 0.1687x_{t-3} - 0.0633x_{t-4})^2 - (0.0005 + 0.0069x_{t-1} - 0.0338x_{t-2} - 0.0416x_{t-3} - 0.0156x_{t-4})^2 - (0.0005 + 0.0069x_{t-1} - 0.0338x_{t-2} - 0.0416x_{t-3} - 0.0156x_{t-4})^2$
Curtosis no condicional	$\alpha_4 = 3.52$

Se observa que si la rentabilidad de los precios en la bolsa estadounidense tiene rendimientos negativos por debajo del  $-0.54\%$  o rendimientos positivos por encima del  $0.57\%$ , esto representa una mayor variabilidad diaria en el mercado brasilero para el primer y tercer régimen, con una desviación estándar del  $2\%$  diario en ambos casos, por lo que no se observa un efecto asimétrico de esta medida de variabilidad sobre el rendimiento esperado. Por otro lado, el efecto de la varianza condicional del segundo régimen en la rentabilidad es un poco menor, con un valor de  $1.41\%$ .

Nótese también que el retorno esperado condicional en cada uno de los regímenes y en la información pasada ( $E(X_t/Z_t \in R_j, t-1)$ ) es en el segundo régimen una constante equivalente a  $0.068\%$ , en cambio en el primer y último depende de los retornos en los dos periodos anteriores.

Por otro lado, el retorno esperado en la bolsa brasilera corresponde a  $-0.05\%$  ( $100E(X_t)$ ), con una varianza marginal de  $0.0003$  ( $var(X_t)$ ) que equivale a una desviación estándar marginal diaria de  $1.73\%$ .

Por último, la curtosis de  $3.52$  refleja una distribución leptocúrtica.

### Modelo GARCH

En la Tabla 7 se presentan los momentos condicionales y no condicionales del proceso  $\{X_t\}$  que obedece el modelo GARCH de la ecuación (7).

Tabla 7: *medias y varianzas para el modelo GARCH estimado para la serie de rendimientos del BOVESPA. Fuente: elaboración propia.*

Momento	Valor
Media no condicional	$E(X_t) = \frac{\phi_0}{1 - \sum_{i=1}^p \phi_i} = 0.0014$
Varianza no condicional	$var(X_t) = 3.71 \times 10^{-4}$
Media condicional	$E(X_t/\tilde{x}_{t-1}) = 0.0014 - 0.0393x_{t-3}$
Varianza condicional	$var(X_t/\tilde{x}_{t-1}) = 7.512 \times 10^{-6} + 0.0736a_{t-1}^2 + 0.9062\sigma_{t-1}^2 = var(a_t/\tilde{a}_{t-1})$
Curtosis no condicional	$\alpha_4 = 10.91$

Con respecto al modelo GARCH, el valor esperado del retorno diario para el BOVESPA es de  $0.14\%$  ( $100E(X_t)$ ), y la desviación estándar marginal diaria del  $1.92\%$  ( $100\sqrt{var(X_t)}$ ). La rentabilidad condicional depende principalmente de la rentabilidad observada tres días antes.

### Análisis comparativo:

Si bien el modelo TAR no es un modelo diseñado para explicar la evolución a través del tiempo en cuanto a la volatilidad condicional, el modelo tiene un ajuste aceptable a la luz de los criterios de Nieto (2005). Sin embargo, al analizar los residuos al cuadrado de ambos modelos se encuentra que estos son adecuadamente representados por el modelo GARCH únicamente. De igual manera ocurre con los valores absolutos de los residuos en cada modelo, su estructura de correlación es

capturada por el modelo GARCH, pero no por el modelo TAR. Es importante puntualizar que el modelo TAR resulta adecuado para explicar el comportamiento no lineal de la serie de rendimientos del índice brasilero en términos de los regímenes del DOWJONES.

Tabla 8: *momento comunes de los modelos TAR y GARCH. Fuente: elaboración propia.*

Modelo TAR	
$E(X_t)$	-0.0005
$var(X_t)$	$3 \times 10^{-4}$
$E(X_t/\tilde{x}_{t-1})$	$0.0005 + 0.0069x_{t-1} - 0.0338x_{t-2} - 0.0416x_{t-3} - 0.0156x_{t-4}$
$var(X_t/\tilde{x}_{t-1})$	$0.0003 - 0.2475(-0.0127 + 0.1113x_{t-1} - 0.0685x_{t-2})^2 - 0.2467(0.0135 - 0.0837x_{t-1} - 0.0684x_{t-2} - 0.1687x_{t-3} - 0.0633x_{t-4})^2 - (0.0005 + 0.0069x_{t-1} - 0.0338x_{t-2} - 0.0416x_{t-3} - 0.0156x_{t-4})^2 - (0.0005 + 0.0069x_{t-1} - 0.0338x_{t-2} - 0.0416x_{t-3} - 0.0156x_{t-4})^2$
Modelo GARCH	
$E(X_t)$	0.0014
$var(X_t)$	$3.71 \times 10^{-4}$
$E(X_t/\tilde{x}_{t-1})$	$0.0014 - 0.0393x_{t-3}$
$var(X_t/\tilde{x}_{t-1})$	$7.512 \times 10^{-6} + 0.0736a_{t-1}^2 + 0.9062\sigma_{t-1}^2$

La Tabla 8 muestra los momentos que son comparables en los dos modelos y a partir de ellos podemos decir que:

(i) el retorno esperado marginal es un poco diferente, (ii) la varianza marginal es básicamente la misma para ambos modelos, (iii) con respecto al retorno esperado condicional a la información pasada este depende de sus valores en los cuatro días anteriores en el modelo TAR, y en el modelo GARCH dependen del valor que tomó tres días antes, y (iv) la varianza condicional en el modelo GARCH corresponde al modelo de volatilidad de la variable latente  $\sigma_t^2$ , mientras que en el TAR la varianza condicional a  $\tilde{x}_{t-1}$  depende de las variables observables  $X_{t-i}$  ( $i = 1, \dots, 4$ ).

## 4. Conclusiones

Si bien el modelo GARCH representa la dinámica de la volatilidad condicional, a la información pasada en la serie, de un forma más simple, es importante mencionar que si el modelo TAR se puede ajustar, este ofrece un nutrido conjunto de medidas de variabilidad, condicionales a los regímenes, a la información pasada o en ambos, que permiten dar interpretaciones importantes en el fenómeno estudiado.

Una práctica común es incorporar el efecto de la variabilidad en el modelo de los rendimientos, el modelo TAR permite representar el hecho de que la rentabilidad de un activo está relacionada con las expectativas de riesgo sobre este. Además la presencia del término  $h^{(j)}$  en el modelo permitiría observar algunos «efectos de apalancamiento», en cuanto a que rentabilidades negativas en el índice estadounidense podrían reflejar una mayor variabilidad condicional en los retornos de la variable de interés, que las rentabilidades positivas; por otro lado, los modelos GARCH no permiten representar el efecto de apalancamiento financiero y tampoco reflejar la relación de estas medidas de riesgo con los rendimientos.

Se detectó dificultad para ajustar un modelo TAR, pues en algunos casos este no captura en forma adecuada la heterocedasticidad marginal de la serie. Además, no explica la estructura de autocorrelación de los valores absolutos y de los cuadrados de los retornos.

Esto nos lleva a concluir que los conglomerados de datos extremos explicados por un TAR son de naturaleza distinta a los explicados por un GARCH. En el modelo TAR, estos se explican a través de los regímenes de la variable de umbrales, mientras que en el GARCH, estos conglomerados se deben al comportamiento dinámico de la heterocedasticidad condicional.

**Recibido: 08 de agosto de 2014**

**Aceptado: 30 de octubre de 2014**

## Referencias

- Cathy W.S. Chen, M. K. S. (2006), 'On a threshold heteroscedastic model', *International Journal of Forecasting* **22**, 73–89.
- Hamilton, J. (1994), 'Time series analysis', *Journal of the American Statistical Association*.
- Hoyos, N. (2006), Una aplicación del modelo tar en economía, Master's thesis, Universidad Nacional de Colombia, sede Bogotá.
- Moreno, E. (2010), Una aplicación del modelo tar en series financieras, Master's thesis, Universidad Nacional de Colombia, sede Bogotá.
- Nieto, F. (2005), 'Modeling bivariate threshold autoregressive processes in the presence of missing data', *Communications in Statistics: Theory and Methods* **34**, 905–930.
- Nieto, F. & Ruiz, F. (2002), 'About a prompt strategy for estimating missing data in long time series', **26**(100), 411–418.
- Tong, H. (1990), *Non-linear Time Series. A Dynamical System Approach*, Oxford Science Publications.

Tsay, R. (1998), 'Testing and modeling multivariate threshold models', *Sociological Methodology* **93**(1188-1202).