

ESTIMACIÓN BAYESIANA DE UN MODELO GARCH-M BIVARIADO

A BAYESIAN ESTIMATION OF BIVARIATE GARCH-M MODELS

CRISTIAN CRUZ-TORRES¹

Received: 27/Jan/2023; Accepted: 08/Jan/2024

Revista de Matemática: Teoría y Aplicaciones is licensed under a Creative Commons
Reconocimiento-NoComercial-Compartigual 4.0 International License.
Creado a partir de la obra en <http://www.revistas.ucr.ac.cr/index.php/matematica>



¹ Universidad Nacional Autónoma de Honduras, Departamento de Estadística Matemática, Tegucigalpa, Honduras. E-Mail: cristian.cruz@unah.edu.hn

Resumen

El modelo GARCH (modelo autorregresivo condicional heterocedástico generalizado) es un modelo estadístico para series de tiempo usado para describir la varianza del error como una función de los errores al cuadrado pasados y de las varianzas. Estos modelos GARCH son usados para modelar la volatilidad variando en el tiempo y los clusters de volatilidad. Si además el efecto de la varianza es incluido en las observaciones para predecir la media, se tiene un GARCH-M (GARCH en media). En este artículo, se analizan estos modelos en un contexto bayesiano para series de tiempo bivariadas, donde las observaciones son asumidas de comportarse como un modelo VAR-GARCH-M. Una aplicación del modelo bivariado es ajustado para medir el efecto de la variabilidad de la inflación y crecimiento del producto en la media de la inflación y crecimiento del producto.

Palabras clave: modelos bivariados GARCH-M; inferencia bayesiana; Monte Carlo Hamiltoniano; inflación y crecimiento del producto.

Abstract

The generalized autoregressive conditional heteroskedasticity (GARCH) model is a statistical model for time series used to describes the variance of the current error as a function of past squared errors terms and previous variances. These GARCH models are commonly used in modeling time varying volatility and volatility clustering. If, in addition, the effect of the variance is included in the observations to predict the mean, we have the GARCH-M (GARCH in mean) models. In this paper, the above issues are analyzed in a bayesian approach to modeling a bivariate time series, where the observations is assumed to behave as a VAR-GARCH-M model. An application of a bivariate model is fitted to measure the effects of inflation variability and uncertainty growth on inflation and output growth mean.

Keywords: bivariate GARCH-M models; bayesian inference; Hamiltonian Monte Carlo; inflation and output growth.

Mathematics Subject Classification: Primarios: 62H12, 62M10. Secundarios: 62F15.

1 Introducción

La volatilidad es un indicador de riesgo que es definido como la desviación estándar condicional del retorno de un activo. La volatilidad se refiere al nivel de incerteza o el riesgo asociado con el tamaño de las variaciones de un valor y muestra cuán grandes fluctúan los valores de un activo alrededor del precio medio. Un valor alto en la volatilidad indica que el valor de una inversión potencialmente puede distribuirse en un rango más amplio de valores. Por tanto, se puede concluir que para tomar una decisión de inversión es importante conocer la volatilidad y comparar frente a una cartera de referencia, favoreciendo aquellas de menor volatilidad.

Modelar y predecir correctamente la volatilidad es por tanto crítico. Además, el modelado de la volatilidad de una serie temporal puede aumentar

la eficiencia de la estimación de los parámetros y la precisión de los intervalos de predicción. La premisa de que la volatilidad no es directamente observable es una de sus características distintivas. Las observaciones en series de retornos semanales y de mayor frecuencia de activos económicos y financieros no son independientes. A pesar del hecho de que las observaciones en estas series casi no están correlacionadas, existe una dependencia de orden superior en la serie.

Algunas características que muestran las series con volatilidad es que suelen presentarse en cluster (es decir, volatilidad que puede ser alta en ciertos periodos y baja en otros). En segundo lugar, la volatilidad evoluciona continuamente; los picos de volatilidad son poco comunes. Tercero, la volatilidad no diverge hacia el infinito; fluctúa dentro de un rango dado, lo que implica que la volatilidad es frecuentemente estacionaria. En cuarto lugar, el efecto de apalancamiento describe cómo reacciona la volatilidad ante un aumento o disminución importante de la volatilidad. Estas características son esenciales en el desarrollo de modelos de volatilidad [37].

En la literatura se encuentran diferentes tipos de modelos de volatilidad, el primer artículo que se puede mencionar es el modelo ARCH (“*autoregressive conditional heteroskedasticity*”) de [9]. También tenemos el modelo GARCH (“*generalized autoregressive conditional heteroskedasticity*”) [5] y sus extensiones, como ser el IGARCH (“*integrated GARCH*”) [29], EGARCH (“*exponential GARCH*”) [30], GARCH - M (“*GARCH in mean*”) [19] y TGARCH (“*Threshold GARCH*”) [43] por mencionar algunos. Por otro lado, existe otra familia de modelos donde es asumido que la ecuación de la volatilidad posee un choque con componente aleatoria, estos modelos son conocidos como modelos de volatilidad estocástica (SV) [22].

En una literatura mas reciente, [10] introduce el GARCH extendido (GARCH - X), el cual incluye a la varianza como una variable exógena en la ecuación de volatilidad y [40] demuestra el poder predictivo de un modelo GARCH(1,1) con los índices de retornos de Standard & Poor’s (S& P500). Para estimar la volatilidad de los precios de las opciones del índice bursátil de Taiwán, [38] integra un modelo EGARCH y una red neuronal, mientras [41] estudia una regresión cuantílica para un modelo GARCH lineal utilizando una representación ARCH(∞) de las volatilidades. Además, [36] implementó un *wavelet* junto con un modelo autorregresivo con medias móviles integrado (ARIMA) y el modelo GARCH para predecir precios futuros.

Asimismo, [26] caracteriza el comportamiento asintótico de las volatilidades en un modelo GARCH(1,1) no estacionario; [35] utilizó una aproximación racional para la función de volatilidad en el modelo GARCH como una técnica de aproximación más prometedora en comparación con la aproximación polinomial utilizada en otros modelos de tipo ARCH y [18] investigó los efectos indirectos de la crisis financiera mundial en los mercados bursátiles asiáticos emergentes. Los modelos GARCH BEKK bivariados confirmaron la presencia de efectos indirectos de NYSE en las economías emergentes en los tres casos (es decir, antes, durante y después de la crisis financiera). Sumado a lo anterior, [42] propusieron el modelo GARCH-

MIDAS basado en cuantiles para examinar la influencia de la incertidumbre de la política económica mensual en el valor en riesgo diario en los mercados del petróleo crudo de West Texas Intermediate. En particular, se descubrió que un aumento en la incertidumbre de la política económica genera un mayor riesgo de mercado del petróleo crudo. Finalmente, [39] hace un análisis de volatilidad basado en modelos tipo GARCH para encontrar evidencia del mercado bursátil chino.

El modelo ARCH es un modelo de series de tiempo que expresa la varianza del término de error actual o innovación en función del valor actual del error y de los términos de error de periodos anteriores; la varianza está frecuentemente relacionada con los cuadrados de las innovaciones pasadas. El modelo ARCH es apropiado cuando la varianza del error sigue un modelo autorregresivo (AR). Si se asume un modelo autorregresivo con medias móviles (ARMA) para la varianza del error, tenemos un modelo GARCH.

Los modelos GARCH se emplean comúnmente en el modelado de series de tiempo que exhiben volatilidad variable en el tiempo y cluster de volatilidad. Los modelos tipo GARCH son frecuentemente clasificados incorrectamente como modelos de volatilidad estocástica (SV), esto es erróneo porque la volatilidad en el tiempo t está totalmente predeterminada (determinista) con base a valores previos. Debido a su eficacia para simular la rentabilidad de los activos y la inflación, los modelos GARCH se emplean ampliamente en economía y finanzas. Intentan reducir los errores de pronóstico teniendo en cuenta los anteriores y mejorando la calidad de las predicciones en curso.

El objetivo de este trabajo es proponer una metodología para ajustar un modelo VAR-GARCH-M bivariado en un contexto bayesiano, donde se implementa el Monte Carlo hamiltoniano como método de estimación. Esta implementación se hace suponiendo choques Normales y t-Student, donde en esta última se estima asumiendo que las observaciones poseen distintos grados de libertad. Finalmente, se aplica este modelo para analizar el comportamiento de dos variables económicas como ser inflación y el crecimiento del producto para dos economías.

El artículo es organizado de la siguiente manera: en la sección 2 se revisa las técnicas de estimación de estadística Bayesiana. La sección 3 explica la metodología econométrica para modelos GARCH univariados y multivariados. En la sección 4, se presenta un ejemplo de simulación para validar la metodología explicada. La sección 5, presenta una aplicación para datos reales con resultados de estimación de ajuste para un modelo GARCH-M bivariado. Finalmente, la sección 6 ofrece las principales conclusiones.

2 Estadística bayesiana

La estadística bayesiana se basa en la probabilidad subjetiva, trabaja con la actualización de la evidencia considerando el conocimiento adquirido previo a una investigación, más la evidencia obtenida de los datos. El teorema de Bayes se utiliza para calcular y actualizar las probabilidades después de obtener nuevos datos.

El teorema de Bayes define la probabilidad condicional de una ocurrencia basada en datos e información previa o creencias sobre el evento o las condiciones asociadas.

2.1 Regla de Bayes

El paradigma bayesiano se construye en función del teorema de Bayes, el cual es un mecanismo que permite conocer sobre los parámetros a ser actualizados cuando nuevos datos son presentados. En forma general, sea θ un vector que contiene todos los parámetros de interés, sea y todos los datos hasta el tiempo presente. Si una distribución a priori $p(\theta)$ es dada y una función de verosimilitud $L(\theta; y)$ puede ser propuesta, la actualización de los parámetros puede ser hecha a través de la siguiente ecuación:

$$p(\theta|y) = \frac{p_Y(y|\theta)p(\theta)}{p_Y(y)} = \frac{p_Y(y|\theta)p(\theta)}{\int p_Y(y|\theta)p(\theta)d\theta} = \frac{L(\theta; y)p(\theta)}{\int L(\theta; y)p(\theta)d\theta} \propto L(\theta; y)p(\theta). \quad (2.1)$$

Dado que y es conocido, el denominador de (2.1) es una constante de proporcionalidad, el cual es útil para asegurar que la distribución a posteriori es propia. En muchos casos, $p_Y(y)$ no puede ser calculada de forma explícita, justificando el uso de algoritmos de Monte Carlo vía cadenas de Markov (MCMC) [16].

2.2 Monte Carlo hamiltoniano

El método de Monte Carlo hamiltoniano (HMC) [28] es un algoritmo de Metrópolis-Hastings, con una evolución dinámica hamiltoniana simulada utilizando un integrador numérico reversible en el tiempo y que conserva el volumen (generalmente el integrador de salto) para proponer un movimiento a un nuevo punto en el espacio de estado. En comparación con el algoritmo Metrópolis-Hastings, que suele utilizar un paseo aleatorio gaussiano o una distribución de propuesta independiente, el Monte Carlo hamiltoniano reduce la correlación entre estados muestreados sucesivos al proponer movimientos a estados distantes que tienen una alta probabilidad de aceptación debido a la energía aproximada, conservando las propiedades de la dinámica hamiltoniana simulada cuando se usa un integrador simpléctico.

Suponga que la distribución objetivo es $f(x)$ y una cadena de muestras X_0, X_1, X_2, \dots es requerida. Las ecuaciones de Hamilton son

$$\frac{dx_i}{dt} = \frac{\partial H}{\partial p_i} \quad y \quad \frac{dp_i}{dt} = -\frac{\partial H}{\partial x_i},$$

donde x_i y p_i son la i -ésima componente de la posición y el vector de momento respectivamente y H es el hamiltoniano. Sea M la matriz de masa, que es simétrica y definida positiva, entonces el hamiltoniano es

$$H(x, p) = U(x) + \frac{1}{2}p^T M^{-1}p,$$

donde $U(x)$ es la energía potencial, la cual esta dada por $U(x) = -\ln f(x)$.

El algoritmo requiere un número entero positivo para el número de pasos de salto de rana L y un número positivo para el tamaño del salto Δt . Suponga la cadena en $X_n = x_n$, sea $x_n(0) = x_n$. Un momento aleatorio Gaussiano $p_n(0)$ es generado de la $N(0, M)$. Luego, la partícula será corrida bajo la dinámica hamiltoniana para el tiempo $L\Delta t$, esto será hecho resolviendo la ecuación de Hamilton numéricamente usando el algoritmo de salto de rana. La posición y vector de momento después del tiempo Δt usando el algoritmo de salto de rana es

$$p_n(t + \frac{\Delta t}{2}) = p_n(t) - \frac{\Delta t}{2} \nabla U(x)|_{x=x_n(t)}, \quad (2.2)$$

$$x_n(t + \Delta t) = x_n(t) + \Delta t M^{-1} p_n(t + \frac{\Delta t}{2}), \quad (2.3)$$

$$p_n(t + \Delta t) = p_n(t + \frac{\Delta t}{2}) - \frac{\Delta t}{2} \nabla U(x)|_{x=x_n(t+\Delta t)} \quad (2.4)$$

Estas ecuaciones son aplicadas a $x_n(0)$ y $p_n(0)$ L veces para obtener $x_n(L\Delta t)$ y $p_n(L\Delta t)$. Debido a que el algoritmo de salto de rana es un método numérico y no resuelve exactamente las ecuaciones de Hamilton, se utiliza un paso de Metrópolis-Hastings. La transición de $X_n = x_n$ a X_{n+1} es

$$X_{n+1}|X_n = x_n = \begin{cases} x_n(L\Delta t), & \text{con probabilidad } \alpha(x_n(0), x_n(L\Delta t)) \\ x_n(0), & \text{otro caso} \end{cases},$$

donde

$$\alpha(x_n(0), x_n(L\Delta t)) = \min\left(1, \frac{\exp[-H(x_n(L\Delta t), p_n(L\Delta t))]}{\exp[-H(x_n(0), p_n(0))]} \right) \quad (2.5)$$

Esto es repetido para obtener X_{n+1}, X_{n+2}, \dots ([4]).

La dinámica hamiltoniana tiene una interpretación física que puede proporcionar intuiciones útiles. En dos dimensiones, podemos visualizar la dinámica como la de un disco sin fricción que se desliza sobre un superficie de altura variable. El estado del sistema consiste de la *posición* del disco, dado por un vector q , y el momento del disco (su masa por su velocidad), es dado por el vector p . La energía potencial, $U(q)$, del disco es proporcional a la altura de la superficie en su posición actual, y su energía cinética, $K(p)$, es igual a $|p|^2/(2m)$, donde m es la masa del disco. En la parte nivelada de la superficie, el disco se mueve a una velocidad constante, igual a p/m . Si encuentra una pendiente ascendente, el impulso del disco le permite continuar, con una disminución de su energía cinética y un aumento de su energía potencial, hasta que la energía cinética (y, por tanto, p) sea cero, en cuyo punto volverá a deslizarse hacia abajo (con la energía cinética aumentando y la energía potencial disminuyendo).

En aplicaciones no físicas de la dinámica hamiltoniana, la posición corresponderá a las variables de interés. La energía potencial será el negativo del logaritmo de la densidad de probabilidad para estas variables. Las variables de momento, una para cada variable de posición, se introducirán artificialmente [28].

Por ejemplo, una intuición económica podría ser que el conjunto de parámetros a estimar representan diferentes activos en un mercado financiero. Cada combinación de valores de parámetros corresponde a una cartera de inversiones en ese mercado, los momentos en el HMC se pueden relacionar con la velocidad de cambio en el mercado. La dinámica hamiltoniana se asemeja a cómo los precios de los activos y el valor de las carteras cambian con el tiempo debido a factores económicos y financieros, esto puede incluir noticias, datos económicos, tasas de interés, etc.

Varias propiedades de la dinámica hamiltoniana son cruciales para su uso en la construcción de las actualizaciones del MCMC. En primer lugar, la dinámica hamiltoniana es reversible, la reversibilidad de la esta es importante para mostrar que las actualizaciones de MCMC dejan invariable la distribución deseada. Una segunda propiedad de la dinámica es que mantiene el invariante hamiltoniano (es decir, conservado), para las actualizaciones de Metrópolis utilizando una propuesta encontrada por la dinámica hamiltoniana, la probabilidad de aceptación es 1 si H se mantiene invariable. Una tercera propiedad fundamental es que conserva el volumen en el espacio (x, p) (un resultado conocido como el teorema de Liouville), la importancia de la conservación del volumen para MCMC es que no necesitamos tener en cuenta ningún cambio en el volumen en la probabilidad de aceptación de las actualizaciones de Metrópolis.

La principal ventaja del HMC es muestrear mucho más eficientemente que los métodos simples como el paseo aleatorio de Metrópolis, esto se debe a que podemos evitar el paseo aleatorio. Un beneficio adicional del HMC es su mejor escalado con la dimensionalidad que los métodos de Metrópolis.

Un impedimento práctico para el uso del HMC es la necesidad de seleccionar valores adecuados para el tamaño de paso de salto, Δt , y el número de pasos de salto de rana, L , que juntos determinan la longitud de la trayectoria en un tiempo ficticio, $L\Delta t$. Otra limitación es que solo se aplica a partir de variables continuas con densidades diferenciables. Si algunas de las variables son discretas o tienen densidades no diferenciables, entonces podemos usar movimientos HMC solo en un subconjunto de las variables.

El “*No U-Turn Sampler*” (NUTS) [21] es una extensión para controlar L automáticamente. Sintonizar L es crítico, si L es muy grande, la partícula oscilará y esto incrementará el costo computacional, para L muy pequeño, la partícula se comportará como una caminata aleatoria. NUTS ejecuta la dinámica hamiltoniana tanto hacia delante como hacia atrás en el tiempo de forma aleatoria hasta que se satisface la condición de giro “*U-Turn*”. Cuando eso sucede, se elige un punto aleatorio de la ruta para la muestra de MCMC y el proceso se repite desde ese

nuevo punto. En detalle, se construye un árbol binario para rastrear la ruta de los pasos de salto de rana. Para producir una muestra de MCMC, se lleva a cabo un procedimiento iterativo. Una variable $U_n \sim U(0, \exp(-H[\mathbf{x}_n(0), \mathbf{p}_n(0)]))$ es generada. Sea \mathbf{x}_n^+ y \mathbf{p}_n^+ la posición y el momento de la partícula delantera respectivamente. Similarmente, \mathbf{x}_n^- y \mathbf{p}_n^- para la partícula trasera. En cada iteración, el árbol binario selecciona aleatoriamente de manera uniforme para mover la partícula delantera hacia adelante o la partícula trasera hacia atrás en el tiempo. Además, para cada iteración, el número de pasos de salto de rana aumenta en un factor de 2.

El proceso iterativo continua hasta que la condición “*U-Turn*” es satisfecha, esto es

$$(\mathbf{x}_n^+ - \mathbf{x}_n^-) \cdot \mathbf{p}_n^- < 0 \quad \text{o} \quad (\mathbf{x}_n^+ - \mathbf{x}_n^-) \cdot \mathbf{p}_n^+ < 0$$

o cuando el hamiltoniano se vuelve inexacto

$$\exp[-H(\mathbf{x}_n^+, \mathbf{p}_n^+) + \delta] < U_n \quad \text{o} \quad \exp[-H(\mathbf{x}_n^-, \mathbf{p}_n^-) + \delta] < U_n.$$

Una vez la condición “*U-Turn*” es satisfecha, la siguiente muestra del MCMC, \mathbf{x}_{n+1} , se obtiene muestreando uniformemente la ruta de salto trazada por el árbol binario $\{\mathbf{x}_n^-, \dots, \mathbf{x}_n(-\Delta t), \mathbf{x}_n(0), \mathbf{x}_n(\Delta t), \dots, \mathbf{x}_n^+\}$ el cual satisface

$$U_n < \exp[-H(\mathbf{x}_{n+1}, \mathbf{p}_{n+1})].$$

3 Metodología econométrica

Sea y_t el valor observado de una variable considerada estacionaria. La suposición principal detrás de un estudio de la volatilidad es que mientras el momento de primer orden de la serie no está serialmente correlacionado o tiene una correlación serial de orden inferior, el momento de segundo orden de la variable exhibe algún tipo de correlación serial.

Para poner los modelos de volatilidad en contexto, considere la media condicional y la varianza de y_t dado el conjunto de información, D_{t-1} . Los primeros dos momentos son definidos como, $\mu_t = E(y_t | D_{t-1})$ y $\sigma_t^2 = \text{var}(y_t | D_{t-1})$. D_{t-1} consiste típicamente de todas las funciones lineales de los valores observados pasados de esta variable. Generalmente, μ_t es asumido de ser un modelo estacionario con una estructura *ARMA*(p, q).

3.1 Modelos univariados ARMA-GARCH-M

El modelo autorregresivo de orden p y media móvil de orden q y heterocedasticidad condicional autorregresiva generalizada de orden k es dada por

$$y_t = \mu + \sum_{j=1}^p \phi_j y_{t-j} + \sum_{j=1}^q \theta_j u_{t-j} + \sum_{j=1}^k \psi_j h_{t-j} + u_t, \quad (3.1)$$

donde $u_t \sim N(0, h_t)$, $t = p + 1, \dots, N$ y h_t es definida como

$$h_t = \omega + \sum_{j=1}^r \beta_j h_{t-j} + \sum_{j=1}^s \alpha_j u_{t-j}^2 \quad (3.2)$$

con la restricción $\sum_{j=1}^s \alpha_j + \sum_{j=1}^r \beta_j < 1$ para garantizar estacionariedad.

Sin pérdida de generalidad considere $\mu = 0$ y $\omega = 0$. Sean las observaciones $\mathbf{y}_0 = (y_1, \dots, y_p)'$ conocidas y sea $u_p = u_{p-1} = \dots = u_{p-q} = u_{p-q+1} = 0$, con $q \leq p$ y $k \leq p$. La ecuación (3.1) puede ser reescrita como

$$\mathbf{y} = \mathbf{X}\phi + \mathbf{A}\theta + \tilde{\mathbf{H}}\psi + \mathbf{u}, \quad (3.3)$$

donde $\mathbf{y} = (y_{p+1}, \dots, y_N)'$, $\phi = (\phi_1, \dots, \phi_p)'$, $\mathbf{u} = (u_{p+1}, \dots, u_N)'$, $\theta = (\theta_1, \dots, \theta_q)'$, $\psi = (\psi_1, \dots, \psi_k)'$ y

$$\mathbf{X} = \begin{pmatrix} y_p & y_{p-1} & \dots & y_1 \\ y_{p+1} & y_p & \dots & y_2 \\ \vdots & \vdots & \ddots & \vdots \\ y_{N-1} & y_{N-2} & \dots & y_{N-p} \end{pmatrix}, \quad \mathbf{A} = \begin{pmatrix} u_p & \dots & u_{p-q+1} \\ u_{p+1} & \dots & u_{p-q+2} \\ \vdots & \ddots & \vdots \\ u_{N-1} & \dots & u_{N-q} \end{pmatrix}$$

y

$$\tilde{\mathbf{H}} = \begin{pmatrix} h_p & \dots & h_{p-k+1} \\ h_{p+1} & \dots & h_{p-k+2} \\ \vdots & \ddots & \vdots \\ h_{N-1} & \dots & h_{N-k} \end{pmatrix}.$$

De la ecuación (3.1) podemos definir u_t , para $t = p+1, \dots, N$ recursivamente por

$$u_t = y_t - \sum_{j=1}^p \phi_j y_{t-j} - \sum_{j=1}^q \theta_j u_{t-j} - \sum_{j=1}^k \psi_j h_{t-j}.$$

Por tanto, la función de verosimilitud es dada por

$$L(\phi, \theta, \psi \mid \mathbf{X}, \mathbf{A}, \mathbf{y}, \mathbf{y}_0) \propto |H|^{-1/2} \exp \left\{ -\frac{1}{2} (\mathbf{y} - \mathbf{X}\phi - \mathbf{A}\theta - \tilde{\mathbf{H}}\psi)' H^{-1} (\mathbf{y} - \mathbf{X}\phi - \mathbf{A}\theta - \tilde{\mathbf{H}}\psi) \right\}, \quad (3.4)$$

donde $H = \text{diag}(h_{p+1}, \dots, h_N)$. Esta función de verosimilitud L es proporcional con la distribución normal.

3.2 Modelo GARCH-M bivariado

La volatilidad en las variables pueden estar relacionada en muchas circunstancias, cuando los impactos de los choques contemporáneos están correlacionados entre sí. Los modelos GARCH multivariados intentan estimar las volatilidades condicionales de las variables en el mismo instante de tiempo. Además, estos modelos pueden ser especialmente relevantes en la investigación que involucra decisiones de asignación de cartera que están influenciadas por el grado de covarianza de los precios de las acciones o la correlación en las volatilidades de un valor que sigue a un choque.

Considere la serie bivariada y_t indexada en tiempo discreto $t \in \mathbb{N}_+$ de dimensión 2, un modelo estacionario VARMA(p, q)-GARCH(r, s)-M(k) es dado por [33]

$$y_t^l = \mu^l + \sum_{m=1}^2 \sum_{j=1}^p \phi_j^{lm} y_{t-j}^m + \sum_{m=1}^2 \sum_{j=1}^q \theta_j^{lm} u_{t-j}^m + \sum_{m=1}^2 \sum_{j=0}^k \psi_j^{lm} h_{t-j}^m + u_t^l, \quad (3.5)$$

donde $u_t^l \sim N(0, h_t^l)$, para $t = p + 1, \dots, N$, $l = 1, 2$ y

$$h_t^l = \omega^l + \sum_{m=1}^2 \left(\sum_{j=1}^r \beta_j^{lm} h_{t-j}^m + \sum_{j=1}^s \alpha_j^{lm} u_{m,t-j}^2 \right) \quad (3.6)$$

con la restricción $\sum_{m=1}^2 (\sum_{j=1}^s \alpha_j^{lm} + \sum_{j=1}^r \beta_j^{lm}) < 1$, donde todos los coeficientes son positivos.

Una estructura especial del modelo GARCH bivariado en notación matricial, se denomina como representación BEKK, es dada en [11]:

$$u_t \sim N(0, H_t); \quad H_t = B_0' B_0 + \sum_{k=1}^s B_k' u_{t-k} u_{t-k}' B_k + \sum_{l=1}^r C_l' H_{t-l} C_l. \quad (3.7)$$

Este es un modelo paramétrico con restricción de precisión positiva, proporcionando así un modelo para modelar la volatilidad. El modelo (3.7) asegura que la matriz H_t es simétrica y definida positiva, mientras que también usa los valores pasados de u_t y H_t , imitando la estructura del modelo GARCH univariado. Dado que B_0 no multiplica ningún otro factor que no sea el mismo, debe ser triangular para que los parámetros sean identificables.

Engle y Kroner demostraron que esta representación es suficientemente general ya que incluye todas las representaciones diagonales definidas positivas y las representaciones *vech*. Si bien este modelo supera una gran debilidad de la representación *vech*, aún involucra la cantidad de parámetros $(1 + r + s)m^2 = O(m^2)$, para nuestro caso bivariado, $m = 2$ [20]. Esto limita severamente la aplicabilidad del modelo BEKK a escenarios con un pequeño número de variables. Las versiones escalar y diagonal del modelo BEKK también se propusieron como formulaciones más manejables del modelo BEKK. Una desventaja de la formulación BEKK es que los parámetros no pueden interpretarse fácilmente y se pierde la intuición de los efectos sobre los parámetros [15].

Proponer un modelo de regresión que incluya tanto la influencia de los datos rezagados como la matriz de covarianza en el vector de medias sería una extensión natural de esto. Como resultado, la nueva estructura de un VAR(p)-M(1) será

$$y_t = A_0 + \sum_{j=1}^p A_j y_{t-j} + G \text{vech}(H_t) + u_t, \quad (3.8)$$

donde u_t tiene la estructura de (3.7). De esta forma todos los parámetros están identificados.

Este enfoque se usará para modelar dos variables observables y_t^1 y y_t^2 , $y_t = (y_t^1, y_t^2)$. Se adopta el enfoque VAR-GARCH-M para tener en cuenta la posible influencia de la incertidumbre de las dos variables en la media de ambas. Los efectos de la incertidumbre sobre las dos variables observables o premio del

riesgo son captados por los elementos de la matriz G . En este caso, los elementos de la matriz G representan el premio de riesgo que la volatilidad de cada variable ejerce sobre la media de cada variable, adicionalmente se mide el impacto de la covarianza entre las volatilidades sobre la media de cada variable.

En esta formulación en orden de establecer el enlace entre el modelo (3.5)-(3.6) y el modelo (3.7)-(3.8). Note que

$$A_0 = \begin{bmatrix} \mu^1 \\ \mu^2 \end{bmatrix}, \quad A_j = \begin{bmatrix} \phi_j^{11} & \phi_j^{12} \\ \phi_j^{21} & \phi_j^{22} \end{bmatrix}, \quad j = 1, \dots, p; \quad G = \begin{bmatrix} \psi_0^{11} & \psi_0^{(1,2)} & \psi_0^{12} \\ \psi_0^{21} & \psi_0^{(2,1)} & \psi_0^{22} \end{bmatrix}.$$

En esta formulación, no estamos considerando el bloque del efecto de la media móvil. El modelo propuesto tiene en cuenta la posible no diagonalidad en las estructuras de covarianza, que es una distinción esencial entre los enfoques utilizados aquí. $\psi_0^{(1,2)}$ y $\psi_0^{(2,1)}$ son el coeficiente de la covarianza $h_t^{(1,2)}$ para cada observación. Para los demás parámetros

$$H_t = \begin{bmatrix} h_t^1 & h_t^{(1,2)} \\ h_t^{(1,2)} & h_t^2 \end{bmatrix}, \quad \text{vech}(H_t) = \begin{bmatrix} h_t^1 \\ h_t^{(1,2)} \\ h_t^2 \end{bmatrix}, \quad B_0' B_0 = \begin{bmatrix} \omega^1 & \omega^{(1,2)} \\ \omega^{(1,2)} & \omega^2 \end{bmatrix},$$

donde B_k con $k = 1, \dots, s$ y C_l con $l = 1, \dots, r$ son matrices cuadradas de dimensión 2. Note que, el modelo (3.7)-(3.8) es mas general que (3.5)-(3.6), porque toma en cuenta el efecto de la covarianza de la matriz H_t en la media de las observaciones (3.8) y en las volatilidades (3.7).

El modelo se estimará utilizando la media posteriori, siendo H_t una matriz de covarianza condicional definida positiva.

3.3 Modelo VAR-GARCH-M t-Student con múltiples grados de libertad

En vista de la formulación anterior, se propone el modelo GARCH t-Student condicional generalizada, reemplazando el supuesto de normalidad condicional del error por la distribución t-Student condicional generalizada. La forma asumida por este modelo es [8]

$$\begin{aligned} y_t &= A_0 + \sum_{j=1}^p A_j y_{t-j} + G \text{vech}(H_t) + u_t; \quad u_t | H_t, \Lambda_t \sim N(0, \Lambda_t^{1/2} H_t \Lambda_t^{1/2}); \\ \Lambda_t^{1/2} &= \text{diag}(\lambda_{1,t}^{1/2}, \lambda_{2,t}^{1/2}); \quad \lambda_{j,t} | v_j \sim \text{IG}(\frac{\nu_j}{2}, \frac{\nu_j}{2}), \quad j = 1, 2; \\ H_t &= B_0' B_0 + \sum_{k=1}^s B_k' u_{t-k} u_{t-k}' B_k + \sum_{l=1}^r C_l' H_{t-l} C_l. \end{aligned} \quad (3.9)$$

Téngase en cuenta que, en lugar de usar una sola variable de mezcla en el modelo, se propone usar más de una variable de mezcla, una para cada elemento diagonal en la matriz de covarianza H_t . La matriz $\Lambda_t^{1/2} H_t \Lambda_t^{1/2}$ y H_t tienen la misma matriz de correlación, pero la mezcla de variables por componente puede ser útil para detectar cambios estructurales en la ecuación de las observaciones.

Además, nótese que cuando H_t es diagonal, el modelo general se reduce a distribuciones t-Student independientes univariadas con diferentes grados de libertad, que cuando se integra el parámetro λ_{jt} en la formulación normal-gamma en el caso de la diagonal, se obtiene la distribución t-Student con ν_j grados de libertad para u_t^l con función de densidad [2], [6]

$$f(u_t^l) = \frac{\Gamma((\nu_j + 1)/2)}{\Gamma(\nu_j/2)} ((\nu_j - 2)h_t^l)^{-1/2} \left(1 + \frac{(u_t^l)^2}{(\nu_j - 2)h_t^l}\right)^{-1/2(\nu_j+1)}, \quad l = 1, 2. \quad (3.10)$$

Una versión manejable del GARCH bivariado es aplicable cuando se usa una distribución que permite múltiples grados de libertad y es condicionalmente normal. La matriz H_t ya no representa la covarianza en su totalidad, pero puede tomarse como una matriz de escala. La matriz $\Lambda_t^{1/2}$ es de menor importancia ya que solo esta presente para permitir una distribución de probabilidad condicionalmente normal.

3.4 Distribuciones a priori

Siguiendo [25], para A_j con $j = 1, \dots, p$, la distribución a priori escogida para todos los parámetros en la matriz es $N(0, 1)$ independiente en cada uno. Dado que esta priori se centra en 0, se tiene la suposición de que la influencia de los valores rezagados es inexistente, favoreciendo un modelo más simple. Además, el hecho de que sea ideal dejar que los datos determinen la estructura de dependencia sugiere que los prioris sean independientes. Finalmente, una varianza de 1 debería permitir dar más peso a los datos en la distribución posterior, pero, por otro lado, asegura la condición de estacionariedad.

La matriz de parámetros G es uno de los factores más esenciales en un contexto inferencial ya que determina el impacto de la volatilidad de cada factor en el vector de medias. También se observó que era uno de los más difíciles de estimar. Asumiendo que el impacto de la volatilidad en la media, mientras exista, no debe ser excesivo, la priori seleccionada para cada elemento de G es $N(0, 1)$ independiente.

Los parámetros restantes residen en H_t , deben de ser no muy informativa ya que están sumergidos en la estructura del modelo. Entonces B_k con $k = 1, \dots, s$ y C_l con $l = 1, \dots, r$ son prioris $N(0, 1)$ para todos los parámetros de la matriz independientes entre sí.

La distribución a priori de ν_j , para el modelo t-Student generalizado, $\nu_j, j = 1, 2$ es dada por [14], el cual son prioris independientes de Jeffrey, que se ha demostrado que conducen a posteriores propias y tienen buenas propiedades frecuentistas

$$p(\nu_j) \propto \left(\frac{\nu_j}{\nu_j + 3}\right)^{1/2} \left\{ \varphi'(\nu_j/2) - \varphi'((\nu_j + 1)/2) - \frac{2(\nu_j + 3)}{\nu_j(\nu_j + 1)} \right\}^{1/2},$$

donde $\varphi(a) = d \log \Gamma(a)/da$ y $\varphi'(a) = d\varphi(a)/da$ son las funciones digamma y trigamma respectivamente.

3.5 Estimación

El algoritmo para hacer la estimación del modelo VAR-GARCH-M por el HMC puede ser descrito a grandes rasgos de la siguiente manera:

1. Inicializar los parámetros del modelo VAR-GARCH-M. Definir la función objetivo como la distribución a posteriori de los parámetros.
2. Implementar el HMC para muestrear de la distribución a posteriori, es decir, calcular el gradiente de la función objetivo con respecto a los parámetros.
3. Generar una propuesta definiendo el número de salto de rana L y el tamaño de salto Δt . Aceptar la propuesta en un paso de Metropolis con probabilidad de aceptación dada en (2.5). O generar automáticamente por medio del NUTS.
4. Actualizar los parámetros del modelo utilizando las muestras obtenidas mediante el HMC.
5. Repetir los pasos 2-4 hasta que se cumpla algún criterio de convergencia.

3.6 Función de Impulso Respuesta

La respuesta a un impulso de un sistema es la que se da en la salida cuando se introduce un impulso a la entrada. Las funciones de impulso respuesta se utilizan para representar cómo reacciona la economía a lo largo del tiempo ante impulsos exógenos, a menudo conocidos como choques, y se modelan con frecuencia en el contexto de un vector autorregresivo. Los cambios en el gasto público, las tasas impositivas y otros parámetros de la política fiscal con frecuencia se tratan como exógenos desde un punto de vista macroeconómico; cambios en la base monetaria u otros parámetros de política monetaria; cambios en la productividad u otros parámetros tecnológicos. La reacción de las variables macroeconómicas endógenas como la producción, el consumo, la inversión y el empleo en el momento del choque y en períodos futuros se describe mediante funciones de impulso respuesta.

Las funciones de impulso respuesta rastrean el impacto dinámico de un “choque” o cambio en una entrada en un sistema, [27]. Esperamos que los choques sean transitorios en los sistemas estacionarios y que converjan con el tiempo. Dependiendo de las restricciones impuestas al modelo VAR estructural, el sistema puede o no converger al estado original.

La función de impulso respuesta se calcula para el modelo VAR-GARCH-M t-Student generalizado, en [32] proponen un método para calcularla cuando los choques tienen distribución Normal. Siguiendo [7], adaptamos el método para calcular la función impulso respuesta del modelo cuando los choques tienen la distribución t-Student generalizada, esta viene dada por

$$y_t = A_0 + A_1 y_{t-1} + G h_t + u_t, \quad u_t \sim N(0, \Lambda_t^{1/2} H_t \Lambda_t^{1/2})$$

$$h_t = \text{vech}(H_t) = \text{vech}(B_0' B_0 + B_1' u_{t-1} u_{t-1}' B_1 + C_1' H_{t-1} C_1).$$

Para calcular la función impulso respuesta, suponga que en el tiempo 0, $u_0 = [1, 0]'$, nos interesa estudiar el efecto sobre el sistema durante los periodos $t = 1, 2, \dots$, si no ocurren más choques, es decir, $u_1 = 0, u_2 = 0, \dots$. Nótese que, cuando se coloca el impulso $u_0 = [1, 0]'$, se aplica un impulso sobre la primera variable, se quiere ver este efecto en la primera y segunda variable a lo largo del tiempo. Un choque unitario en la primera variable en $t = 0$ en este sistema se obtiene como:

$$h_0 = \text{vech}(H_0) = \text{vech}(B'_0 B_0)$$

$$y_0 = A_0 + G h_0 + u_0,$$

$$h_1 = \text{vech}(B'_0 B_0 + B'_1 u_0 u'_0 B_1 + C'_1 H_0 C_1).$$

$$y_1 = A_0 + A_1 y_0 + G h_1,$$

$$h_2 = \text{vech}(B'_0 B_0 + C'_1 H_1 C_1).$$

$$y_2 = A_0 + A_1 y_1 + G h_2$$

$$\vdots$$

$$h_n = \text{vech}(B'_0 B_0 + C'_1 H_{n-1} C_1).$$

$$y_n = A_0 + A_1 y_{n-1} + G h_n$$

Un impulso en el primer choque en el tiempo 0 da una respuesta sobre las variables observables para cada tiempo t . Entonces y_t^i representa la respuesta de la variable i en el momento t del impulso desde el primer choque, $i = 1, 2$ y $t = 1, 2, \dots$. De igual manera se hace con la respuesta de las variables observables ante un impulso para un choque de la segunda variable, es decir, cuando $u_0 = [0, 1]'$.

4 Ejemplo

En esta sección, se realiza un experimento de simulación para investigar la consistencia de los estimadores y evaluar el desempeño de los diferentes métodos para estimar los parámetros. El método de HMC en el paradigma bayesiano se utiliza como algoritmo para la estimación de los parámetros. El algoritmo se implementa en R usando la librería de varstan [1].

4.1 t-VAR(1)-GARCH(1,1)-M(1)

Ahora presentaremos un ejemplo simulado para datos bivariados, queremos estudiar la efectividad de estimar los parámetros en el modelo propuesto. Considere 100 simulaciones de series de tiempo bidimensional con 300 observaciones cada una, gobernadas por la ecuación

$$y_t = \begin{pmatrix} 0,1 & 0 \\ 0 & -0,6 \end{pmatrix} y_{t-1} + \begin{pmatrix} 0,05 & 0,15 & 0,15 \\ 0,1 & 0,1 & 0,2 \end{pmatrix} \text{vech}(H_t) + u_t, \quad u_t \sim N(0, H_t). \quad (4.1)$$

Para la ecuación del GARCH considere la matriz de coeficientes dada por:

$$B_0 = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}, \quad B_1 = \begin{pmatrix} 0,1 & 0 \\ 0 & 0,1 \end{pmatrix} \text{ y } C_1 = \begin{pmatrix} 0,2 & 0 \\ 0 & 0,3 \end{pmatrix}.$$

En estas simulaciones, los coeficientes de la matriz G representan el impacto de las varianzas h_{1t} , h_{2t} y la covarianza entre las observaciones $cov(h_{1t}, h_{2t})$ sobre la media de las observaciones. Un valor distinto de cero en los coeficientes de la matriz G representa que esta varianza (o covarianza) tiene un impacto directo en esa variable observable, lo que nos lleva a concluir que el modelo, al tener presencia de volatilidad, afecta directamente la media de las observaciones.

Para este ejemplo, téngase en cuenta que nuestros datos se generan a partir de un modelo Normal, pero lo estimaremos con un modelo t-Student generalizado definido en (3.9). Nuestra expectativa es que los grados de libertad del modelo t-Student generalizado tengan en cuenta este comportamiento Normal, es decir, el valor estimado de los grados de libertad sea grande. Usando el algoritmo HMC con la extensión NUTS para optimizar el salto de rana L , tomando 50000 iteraciones de la cadena, descartando 5000 y manteniendo cada décima iteración para reducir la dependencia en cada simulación. En la tabla 1 tenemos el promedio de las medias a posteriori y los intervalos de credibilidad del 95 % de los parámetros de las simulaciones efectuadas.

Podemos notar que en general, las medias a posteriori capturan el comportamiento de los parámetros con sus respectivos signos. Además, los dos grados de libertad obtienen el comportamiento de la distribución Normal, es decir, los valores estimados son grandes, dado que en las distribuciones t-Student cuando los grados de libertad aumentan, estas distribuciones convergen a la distribución Normal.

5 Datos y resultados

En esta sección presentamos una aplicación del modelo propuesto para estudiar el comportamiento de la inflación y el crecimiento del producto real para dos países del MERCOSUR. El objetivo es investigar si la volatilidad en la inflación y el crecimiento del producto afectan directamente el comportamiento del valor medio de estas dos variables en ambos países.

Cabe mencionar que en la literatura existen algunos trabajos donde se aplican los modelos GARCH para estudiar variables económicas como la inflación, precios, tipo de cambio, tasa de interés, etc. Entre los cuales podemos mencionar [17], [31] y [12].

5.1 Datos

Se utiliza datos mensuales sobre el Índice de Precios al Consumidor (IPC) y el Índice de Producción Industrial (IPI). Los datos se refieren a Paraguay (1994-2016) y Uruguay (2003-2015). Los datos del IPC fueron tomados del sitio web del

Tabla 1: Estimaciones del t-VAR(1)-GARCH(1,1)-M(1).

Matriz	Parámetros	Valor verdadero	Media	Intervalo de credibilidad del 95 %
μ	μ_1	0	0.01	[-0.14,0.11]
	μ_2	0	0.00	[-0.13,0.10]
A_1	$A_{1(1,1)}$	0.1	0.11	[0.01,0.23]
	$A_{1(1,2)}$	0	0.01	[-0.08,0.09]
	$A_{1(2,1)}$	0	0.00	[-0.08,0.11]
	$A_{1(2,2)}$	-0.6	-0.60	[-0.67,-0.46]
B_0	$B_{0(1,1)}$	0	0.03	[-0.11,0.16]
	$B_{0(1,2)}$	0	0.01	[-0.06,0.09]
	$B_{0(2,2)}$	0	0.06	[-0.03,0.14]
B_1	$B_{1(1,1)}$	0.1	0.14	[0.02,0.23]
	$B_{1(1,2)}$	0	-0.01	[-0.09, 0.06]
	$B_{1(2,1)}$	0	0.02	[-0.08,0.12]
	$B_{1(2,2)}$	0.1	0.09	[-0.01,0.18]
C_1	$C_{1(1,1)}$	0.2	0.25	[0.17,0.32]
	$C_{1(1,2)}$	0	0.03	[-0.12,0.16]
	$C_{1(2,1)}$	0	0.00	[-0.20,0.19]
	$C_{1(2,2)}$	0.3	0.27	[0.13,0.40]
G	$G_{1,1}$	0.05	0.03	[-0.05,0.12]
	$G_{1,2}$	0.15	0.12	[0.03,0.20]
	$G_{1,3}$	0.15	0.18	[0.09,0.29]
	$G_{2,1}$	0.1	0.11	[0.02,0.19]
	$G_{2,2}$	0.1	0.07	[-0.01,0.16]
	$G_{2,3}$	0.2	0.24	[0.13,0.34]
ν	ν_1	-	35.67	[22.89,63.04]
	ν_2	-	64.41	[33.96,198.37]

Fondo Monetario Internacional, International Financial Statistics (FMI/IFS) y los datos del IPI del sitio web de Bloomberg.

La inflación se mide por la diferencia mensual anualizada del índice de precios al consumidor IPC ($y_{1t} = CPI_t/CPI_{t-1} - 1$) y el crecimiento del producto real por la diferencia mensual anualizada del IPI ($y_{2t} = IPI_t/IPI_{t-1} - 1$). Note que en los períodos estudiados para ambos países incluye la crisis financiera global del 2008-2009, esta crisis representa valores atípicos que producen un aumento en la volatilidad, lo cual puede ser capturado por el modelo propuesto y como esta volatilidad afecta la media de las observaciones.

Las estadísticas descriptivas sobre la inflación y el crecimiento de la producción se presentan en la tabla 2. Estas estadísticas incluyen resultados sobre la media, la asimetría y la curtosis. Cuando los datos son asimétricos y con colas pesadas, esto puede ser un indicador de que un modelo GARCH con asimetría en la volatilidad es el más apropiado [13]. Incluso, si los datos no tuvieran un coeficiente de asimetría muy alto pero la curtosis es persistente, esto sería evidencia de que un modelo GARCH con choques t-Student es el más pertinente [23].

Tabla 2: Estadísticas descriptivas: inflación y crecimiento del producto en Paraguay y Uruguay.

País	Variable	Media	Asimetría	Curtosis	KS
Paraguay	Inflación	0.0061	0.6261	4.619	0.491
	Crecimiento del producto	0.0038	0.1962	6.698	0.463
Uruguay	Inflación	0.0062	0.0091	2.800	0.498
	Crecimiento del producto	0.0088	0.1585	3.355	0.440

En nuestros datos podemos ver que existe evidencia de colas pesadas para ambas variables en los dos países, esto se puede observar por el coeficiente de curtosis, en la mayoría de los casos es superior a 3, lo que indica que la distribución de estos datos tiene colas mas pesadas que la distribución normal. Este es un indicador de que quizás existen valores extremos en los datos, lo que sugiere que una distribución t-Student es más adecuada para capturar este comportamiento, además de provocar un aumento en la variabilidad de los datos, lo que lleva a la presencia de la volatilidad en la media de las observaciones [3].

Para dar mayor evidencia a la hipótesis de no normalidad, se realizó la prueba de Kolmogorov-Smirnov (KS) de normalidad, donde el valor crítico varía según el tamaño de la muestra. En nuestros datos los valores críticos son de 0,117 para los datos de Paraguay y 0,157 para los datos de Uruguay, note que el valor de KS en la tabla supera estos valores críticos, por tanto se rechaza la hipótesis de normalidad. Es importante mencionar que la curtosis en los datos de Paraguay es mayor que en los datos de Uruguay, es decir que los datos de Paraguay poseen colas mas pesadas, lo cual se evidencia cuando se estiman los grados de libertad

en la tabla 3, valores más bajos de grados de libertad indican colas más pesadas en los datos.

Tabla 3: Grados de libertad para la distribución t-Student generalizada.

País	ν_1	ν_2
Paraguay	4.79 [3.92,6.23]	5.76 [4.71,7.00]
Uruguay	15.75 [9.59,29.25]	8.91 [5.91,16.85]

5.2 Resultados

Dado que existe un indicio de posible volatilidad con colas pesadas, procedemos a estimar la media posteriori utilizando el método NUTS del HMC, el cual se implementó en R por medio de la librería `varstan` [1]. Al igual que el ejemplo presentado en la sección 4, realizaremos la estimación de la media a posteriori de los parámetros y sus respectivos intervalos de credibilidad para estudiar la significancia para los parámetros autorregresivos y de la volatilidad.

Para ajustar el modelo es necesario decidir qué orden del modelo autorregresivo (VAR) de los datos es más apropiado, el criterio para decidir el orden es a través de la verosimilitud marginal, la cual se calcula con el estimador gamma desplazado propuesto por [34]. En la tabla 4 se reporta la verosimilitud marginal para el modelo VAR de orden 1, 2 y 3 en ambos países.

Tabla 4: Verosimilitud marginal en los modelos VAR.

País	VAR(1)	VAR(2)	VAR(3)
Paraguay	1976.60	1983.24*	1978.07
Uruguay	1126.15	1128.01*	1122.04

El criterio para elegir el mejor modelo por la verosimilitud marginal es aquel que tenga el mayor valor. Esto es particularmente útil, ya que si queremos comparar entre modelos, la razón de verosimilitudes marginales nos da el factor de Bayes, lo que implica que al tener un valor más alto, el factor de Bayes será mayor a 1 en comparación con el resto de modelos, lo que se puede concluir que los datos favorecen a este modelo en relación con los demás [24].

Nótese en la tabla 4, que para ambos países la verosimilitud marginal es mayor para un VAR de orden 2, lo que nos lleva a concluir que el modelo VAR(2) ajusta mejor. Pero hay que tener en cuenta que la diferencia entre la verosimilitud marginal del VAR(1) y VAR(2) no es tan significativa. En términos del factor de Bayes, el VAR(2) es mejor que el VAR(1) por una fracción de 1,0034 en Paraguay y de 1,0016 en Uruguay, es decir, el VAR(2) es apenas superior al VAR(1) en ambos países. Por ello en lo que sigue de este trabajo se tomará un VAR(1) para la

estimación en ambos países, otra razón para preferir un VAR(1) es la parsimonia, tomando un VAR(1) el número de parámetros a ser estimado se reduce en comparación con otros VAR(p), $p > 1$.

5.3 Modelo VAR-GARCH-M

Para tener en cuenta posibles relaciones de contemporaneidad y complementar la estimación, se procede a mostrar el ajuste de los parámetros de GARCH-M bivariado. En esta aplicación, ajustamos el modelo vectorial autorregresivo bivariado GARCH-en-media (VAR(1)-GARCH(1,1)-M(1)). Cuando los choques tienen una distribución normal, el modelo ahora se puede escribir como

$$H_t = B_0' B_0 + B_1' u_{t-1} u_{t-1}' B_1 + C_1' H_{t-1} C_1,$$

donde

$$B_0 = \begin{pmatrix} b_{11} & b_{12} \\ 0 & b_{22} \end{pmatrix}, B_1 = \begin{pmatrix} \beta_{11} & \beta_{12} \\ \beta_{21} & \beta_{22} \end{pmatrix}, C_1 = \begin{pmatrix} \gamma_{11} & \gamma_{12} \\ \gamma_{21} & \gamma_{22} \end{pmatrix} \text{ y } G = \begin{pmatrix} g_{11} & g_{12} & g_{13} \\ g_{21} & g_{22} & g_{23} \end{pmatrix}.$$

En el caso de que los choques tengan una distribución t-Student generalizada, considerando la expresión dada en la ecuación (3.9) con la definición de las matrices B_0 , B_1 , C_1 y G de la de la misma manera y u_t como

$$u_t | H_t, \Lambda_t \sim N(0, \Lambda_t^{1/2} H_t \Lambda_t^{1/2}),$$

$$\Lambda_t^{1/2} = \text{diag}(\lambda_{1,t}^{1/2}, \lambda_{2,t}^{1/2}), \quad \lambda_{j,t} | \nu_j \sim IG\left(\frac{\nu_j}{2}, \frac{\nu_j}{2}\right), \quad j = 1, 2.$$

Estas ecuaciones muestran la matriz de varianza-covarianza condicional para choques de inflación y crecimiento del producto. Note que la matriz G es de 2×3 porque el vector $\text{vech}(H_t)$ es la vectorización de la triangularización inferior de H_t . Además, g_{11} representa el impacto de la varianza de la inflación en la media de la inflación, g_{12} representa el impacto de la covarianza entre la inflación y el crecimiento del producto en la media de la inflación, g_{13} el impacto de la varianza del crecimiento del producto sobre la inflación media. Por otro lado, g_{21} representa el impacto de la varianza de la inflación sobre la media del crecimiento del producto, g_{22} representa el impacto de la covarianza entre la inflación y el crecimiento sobre la media del crecimiento del producto y g_{23} representa el impacto de la varianza del crecimiento del producto en la media del crecimiento de la producción.

Usando el algoritmo HMC con la extensión NUTS para optimizar el salto de rana L , se corren 50000 iteraciones, descartando 5000 y manteniendo cada décima iteración para reducir la dependencia, los resultados de la estimación a la matriz G para ambos países del modelo VAR-GARCH-M se reportan en la tabla 5 y en la tabla 6. Las estimaciones de la media de los coeficientes (elementos de la matriz G) e intervalos de credibilidad del 95 % para ambos modelos, modelos normal y t-Student generalizado para ambos países.

Tabla 5: Estimación de la matriz G basado en el modelo VAR-GARCH-M normal y t-Student generalizado para los datos de Paraguay.

Modelo	g_{11}	g_{12}	g_{13}	g_{21}	g_{22}	g_{23}
Normal	-6.67 [-8.23, -4.97]	0.47 [-4.94, 6.02]	6.82 [5.21, 8.26]	-1.49 [-5.31, 2.34]	0.26 [-5.06, 5.57]	1.51 [-2.31, 5.30]
t-Student	-3.60 [-6.61, -0.38]	0.74 [-4.75, 6.11]	3.69 [0.35, 6.70]	-1.34 [-5.05, 2.36]	0.03 [-6.39, 6.26]	1.39 [-2.28, 5.23]

Tabla 6: Estimación de la matriz G basado en el modelo VAR-GARCH-M normal y t-Student generalizado para los datos de Uruguay.

Modelo	g_{11}	g_{12}	g_{13}	g_{21}	g_{22}	g_{23}
Normal	-0.67 [-1.41, 0.17]	-2.14 [-6.42, 2.07]	0.69 [0.23, 0.92]	-2.33 [-4.45, -0.15]	0.16 [-6.22, 6.63]	2.41 [0.27, 4.45]
t-Student	-1.07 [-1.83, -0.08]	-2.21 [-6.74, 2.34]	1.08 [0.41, 1.39]	-2.54 [-5.14, 0.09]	-0.10 [-6.94, 6.58]	2.61 [0.02, 5.19]

Existe evidencia del efecto de la varianza o covarianza en la inflación y el crecimiento del producto. En Paraguay, hay dos impactos significativos, uno negativo de la varianza de la inflación (g_{11}) y uno positivo de la varianza del crecimiento del producto (g_{13}) para la media de la inflación. Nótese que en ambos modelos (normal y t-Student generalizado) g_{11} y g_{13} son significativos, podemos concluir que la volatilidad en la inflación y el crecimiento del producto tienen efectos directos sobre la media de la inflación.

Por otro lado, en Uruguay la situación es un poco diferente, en el modelo de choques normales, hay tres impactos significativos, uno negativo y dos positivos; uno negativo de la varianza de la inflación (g_{21}) sobre la media del crecimiento del producto, dos positivos de la varianza del crecimiento del producto (g_{13} y g_{23}) sobre la media de la inflación y del crecimiento del producto, respectivamente. En el caso del modelo con choques t-Student generalizado, la conclusión es un poco diferente, hay tres impactos significativos, uno negativo de la varianza de la inflación (g_{11}) sobre la media de la inflación, y dos impactos positivos de la varianza del crecimiento del producto (g_{13} y g_{23}) sobre la media de la inflación y el crecimiento del producto, respectivamente. Nótese que para ambos países, los signos de los parámetros estimados de la matriz G coinciden tanto para el modelo de choques normal como para el modelo t-Student generalizado.

Una ventaja de utilizar el modelo t-Student generalizado es que nos permite estimar los grados de libertad ν , ya que la formulación de nuestro modelo permite una mezcla de tal manera que se puede estimar un grado de libertad diferente para cada serie de tiempo univariada, esto será útil para determinar qué tan pesada es la distribución de la cola para cada serie de datos. En la tabla 3, se presentan las estimaciones de los grados de libertad $\nu = (\nu_1, \nu_2)$ de la distribución t-Student generalizada para los datos de ambos países. Hay alguna evidencia de colas pesadas, lo que indica que estas series de inflación y crecimiento del producto no siguen la distribución normal para ambos países. Nótese que en Paraguay, la estimación de los grados de libertad es menor que en Uruguay, indicando que la distribución de sus datos tiene colas más pesadas, y que esto va de acuerdo a los resultados que se presentaron en la tabla 2, donde la la curtosis fue mayor para los datos de Paraguay, es decir, estos datos tienen más valores extremos que en los datos de Uruguay. Por esta razón se obtiene más evidencia a preferir el modelo t-Student generalizado y por tanto a sus conclusiones. En las figuras 1 y 2 se presenta la estimación de las cadenas MCMC, histograma y función de autocorrelación de los grados de libertad $\nu = (\nu_1, \nu_2)$ de la distribución t-Student generalizada para las dos naciones, donde se observa que en ambas se tiene una buena convergencia.

5.4 Función de impulso respuesta

Se calculó la función de impulso respuesta para el modelo VAR-GARCH-M t-Student generalizado utilizando las muestras a posteriori de los parámetros. Cabe mencionar que para cada muestra posterior θ_j , con $j = 1, \dots, M$, se ha calculado una función impulso respuesta para cada choque, así, y_{jt}^i , representa la

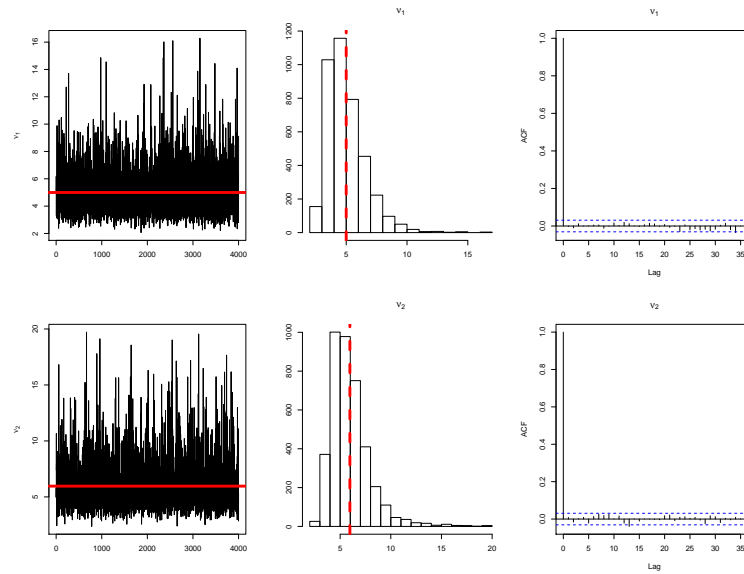


Figura 1: Cadenas MCMC, histograma y ACF para los grados de libertad de la distribución t-Student generalizada para los datos de Paraguay.

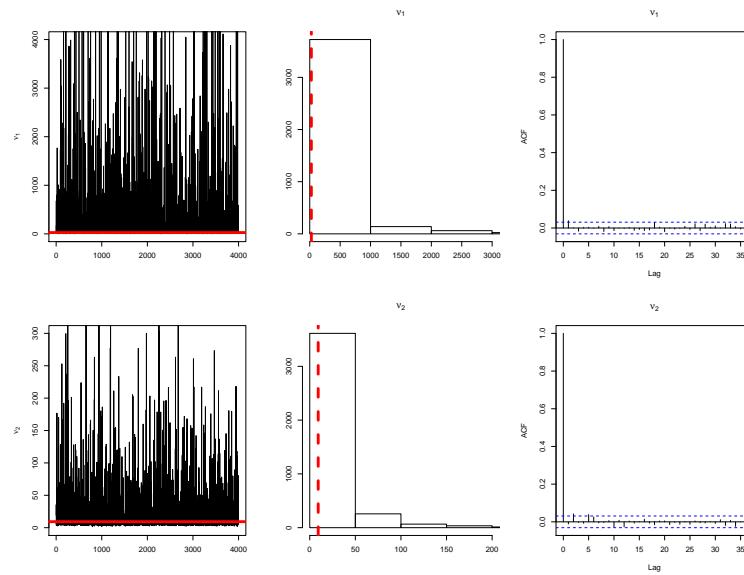


Figura 2: Cadenas MCMC, histograma y ACF para los grados de libertad de la distribución t-Student generalizada para los datos de Uruguay.

respuesta de la variable i en el momento t del impulso del choque, para la muestra θ_j . Por lo tanto, es posible obtener una muestra de funciones de respuesta de impulso para cada choque.

Las figuras 3 y 4 muestran la respuesta media de la inflación y crecimiento del producto para los dos impulsos en los choques (inflación y crecimiento del producto) para los países de Paraguay y Uruguay.

Podemos concluir que en Paraguay, en la figura 3, un impulso en la inflación y el crecimiento del producto no provocan un efecto significativo en estas variables, sus valores están alrededor de 0,10, estos efectos se desvanecen a los 5 periodos después de haber ocurrido los choques por ambas variables. En Uruguay la situación es diferente como podemos ver en la figura 4, un choque en la inflación provoca un gran efecto en la inflación y el crecimiento del producto, su magnitud es de 3 unidades y estos efectos se mantienen durante 5 periodos después de que se producen los choques; tenga en cuenta que un choque en el crecimiento del producto no tiene un efecto significativo en la inflación y el crecimiento del producto. Por tanto, podemos concluir que la función impulso respuesta refuerza la conclusión obtenida anteriormente de que la volatilidad de las variables tiene un efecto sobre sus valores medios.

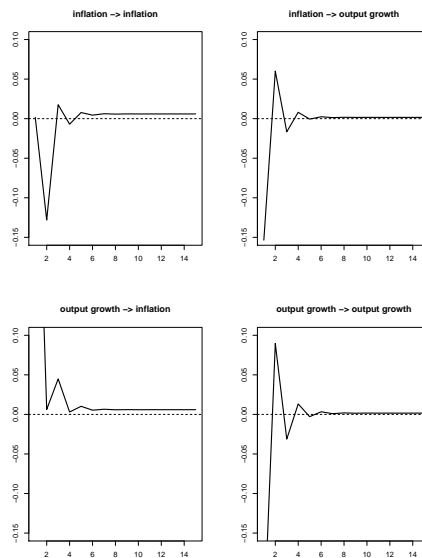


Figura 3: Función de impulso respuesta de la inflación y crecimiento del producto en Paraguay.

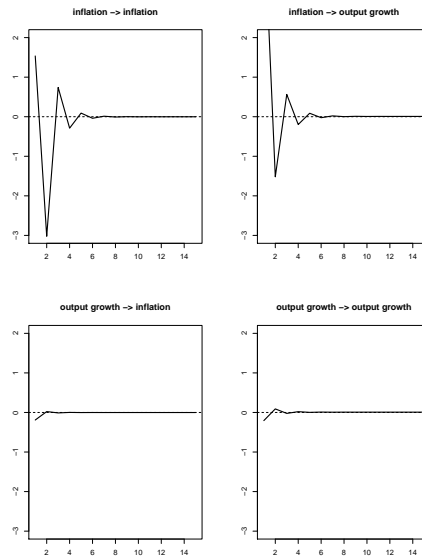


Figura 4: Función de impulso respuesta de la inflación y crecimiento del producto en Uruguay.

6 Conclusiones

Este trabajo demuestra que los modelos VAR-GARCH-M sugeridos son bastante prometedores en términos de ofrecer modelos parsimoniosos para la covarianza condicional en el tiempo, garantizando que sus estimaciones sean matrices definidas positivas. Este trabajo también asegura que el modelo VAR-GARCH-M con choques t-Student generalizado permite estimar diferentes niveles de pesadez en las colas para datos bivariados, lo cual es una ventaja sobre el modelo tradicional de choques t-Student que asume un único grado de libertad para datos bivariados.

Calculamos la función de impulso respuesta como consecuencia de que el modelo tiene una estructura VAR y lo adaptamos a un modelo con volatilidad, lo que nos permitió ver los efectos del choque de una variable sobre la media de las otras variables, lo cual es particularmente útil cuando se trabaja con datos económicos dependientes en el tiempo.

Este trabajo también demuestra que bajo un enfoque bayesiano se considera una fuente adicional de incertidumbre, la cual tiene implicaciones para medir la volatilidad y utilizando el algoritmo HMC obtenemos una convergencia rápida en los parámetros.

Esta metodología propuesta puede ser utilizada para estimar modelos de tres dimensiones o más haciendo sus debidas adaptaciones.

Agradecimientos

El autor agradece a Asael Alonzo por su discusión y ayuda en el código del proyecto.

Soporte financiero

Este trabajo fue parcialmente apoyado por Naciones Unidas. Fondo: 32BTS TFPProjectsECLAC.

Referencias

- [1] I. A. Alonzo Matamoros y C. A. Cruz Torres, *varstan: An R package for bayesian analysis of structured time series models with Stan*. Arxiv. 2020. DOI: [10.48550/arXiv.2005.10361](https://doi.org/10.48550/arXiv.2005.10361)
- [2] D. Ardia y L. Hoogerheide, *Bayesian Estimation of the GARCH(1,1) Model with Student-t Innovations*. The R Journal **2**(2010), no. 2, 41-47. DOI: [10.32614/RJ-2010-014](https://doi.org/10.32614/RJ-2010-014)
- [3] T. Bali, *An Extreme Value Approach to Estimating Volatility and Value at Risk*. The Journal of Business **76**(2003), no. 1, 83-108. DOI: [10.1086/344669](https://doi.org/10.1086/344669)
- [4] M. Betancourt, *A Conceptual Introduction to Hamiltonian Monte Carlo*. 2018. arXiv: [1701.02434](https://arxiv.org/abs/1701.02434) [[stat.ME](https://arxiv.org/abs/1701.02434)].
- [5] T. Bollerslev, *A conditionally heteroskedastic time series model for speculative prices and rates of return*. The Review of Economics and Statistics **69**(1987), no. 3, 542-547. DOI: [10.2307/1925546](https://doi.org/10.2307/1925546)
- [6] S. Chib y S. Ramamurthy, *DSGE models with Student-t errors*. Econometrics Reviews **33**(2014), no. 1-4, 152-171. DOI: [10.1080/07474938.2013.807152](https://doi.org/10.1080/07474938.2013.807152)
- [7] C. Chua, S. Suardi y S. Tsiaplias, *An impulse-response function for a VAR with multivariate GARCH-in-Mean that incorporates direct and indirect transmission of shocks*. Economics Letters **117**(2012), no. 2, 452-454. DOI: [10.1016/j.econlet.2012.06.031](https://doi.org/10.1016/j.econlet.2012.06.031)
- [8] C. Cruz, *Modelos Dinâmicos Estocásticos de Equilíbrio Geral com Choques Heterocedásticos*. Tese de Doutorado, UFRJ, 2015.
- [9] R. Engle, *Autoregressive Conditional Heteroscedasticity with Estimates of Variance of United Kingdom Inflation*. Econometrica **50**(1982), no. 4, 987-1008. DOI: [10.2307/1912773](https://doi.org/10.2307/1912773)

- [10] R. Engle, *New frontiers for ARCH models*. Journal of Applied Econometrics **17**(2002), 425-446. DOI: [10.1002/jae.683](https://doi.org/10.1002/jae.683)
- [11] R. Engle y K. Kroner, *Multivariate simultaneous generalized ARCH*. Econometric Theory **11**(1995), no. 1, 122-150. DOI: [10.1017/S0266466600009063](https://doi.org/10.1017/S0266466600009063)
- [12] A. Ferreira Silva, A. Fagundes Carrara y N. Rennó Castro, *Inflation persistence for product groups in Brazil using the ARFIMA-GARCH model*. Macroeconomics and Finance in Emerging Market Economies (2022), 1-20. DOI: [10.1080/17520843.2022.2080345](https://doi.org/10.1080/17520843.2022.2080345)
- [13] T. Fonseca, V. Cerqueira, H. Migon y C. Cruz-Torres, *Evaluating the performance of degrees of freedom estimation in Asymmetric GARCH models with Student-t innovations*. Brazilian Review of Econometrics **40**(2020), no. 2, 347-373. DOI: [10.12660/bre.v40n22020.80292](https://doi.org/10.12660/bre.v40n22020.80292)
- [14] T. Fonseca, M. Ferreira y H. Migon, *Objective Bayesian analysis for the Student-t regression model*. Biometrika **95**(2008), no. 2, 325-333. DOI: [10.1093/biomet/asn001](https://doi.org/10.1093/biomet/asn001)
- [15] C. Francq y J. Zakoian, *GARCH Models, Structure, Statistical Inference and Financial Applications*. John Wiley & Sons Ltd, 2019. DOI: [10.1002/9781119313472](https://doi.org/10.1002/9781119313472)
- [16] D. Gamerman y H. Lopes, *Markov Chain Monte Carlo: Stochastic Simulation for Bayesian Inference*. Chapman & Hall/CRC, 2006. DOI: [10.1201/9781482296426](https://doi.org/10.1201/9781482296426)
- [17] K. Grier y M. Perry, *Inflation, inflation uncertainty, and relative price dispersion: Evidence from bivariate GARCH-M models*. Journal of Monetary Economics **38**(1996), no. 2, 391-405. DOI: [10.1016/S0304-3932\(96\)01280-9](https://doi.org/10.1016/S0304-3932(96)01280-9)
- [18] S. Gulzar et al., *Financial cointegration and spillover effect of global financial crisis: A study of emerging Asian financial markets*. Economic Research-Ekonomska Istraživanja **32**(2019), no. 1, 187-218. DOI: [10.1080/1331677X.2018.1550001](https://doi.org/10.1080/1331677X.2018.1550001)
- [19] J. Hamilton, *Time Series Analysis*. Princeton University Press, 1994. DOI: [10.2307/j.ctv14jx6sm](https://doi.org/10.2307/j.ctv14jx6sm)
- [20] M. Heracleous, *Volatility Modeling Using the Student's t Distribution*. Doctoral Thesis, Virginia Polytechnic Institute and State University, 2003.
- [21] M. Hoffman y A. Gelman, *The No-U-Turn Sampler: adaptively setting path lengths in Hamiltonian Monte Carlo*. Journal of Machine Learning Research **15**(2014), 1593-1623.
- [22] E. Jackier, N. Polson y P. Rossi, *Bayesian Analysis of Stochastic Volatility Models*. Journal of Business & Economic Statistics **12**(1994), no. 4, 371-389. DOI: [10.2307/1392199](https://doi.org/10.2307/1392199)
- [23] E. Jondeau y M. Rockinger, *Conditional volatility, skewness, and kurtosis: existence, persistence, and comovements*. Journal of Economic Dynamics and Control **27**(2003), no. 10, 1699-1737. DOI: [10.1016/S0165-1889\(02\)00079-9](https://doi.org/10.1016/S0165-1889(02)00079-9)

- [24] R. Kass y A. Raftery, *Bayes factor*. Journal of the American statistical association **90**(1995), no. 430, 773-795. DOI: [10.1080/01621459.1995.10476572](https://doi.org/10.1080/01621459.1995.10476572)
- [25] R. Lassance, V. Cerqueira y T. Fonseca, *VARMA-GARCH-M with Multiple Degrees of Freedom*. 2018.
- [26] D. Li, M. Li y W. Wu, *On dynamics of volatilities in nonstationary GARCH models*. Statistics & Probability Letters **94**(2014), 86-90. DOI: [10.1016/j.spl.2014.07.003](https://doi.org/10.1016/j.spl.2014.07.003)
- [27] H. Lutkepohl, *Introduction to Multiple Time Series Analysis*. Springer-Verlag, 2005. DOI: [10.1007/978-3-662-02691-5](https://doi.org/10.1007/978-3-662-02691-5)
- [28] R. Neal, *MCMC Using Hamiltonian Dynamics*. Handbook of Markov Chain Monte Carlo, 2011. DOI: [10.1201/b10905](https://doi.org/10.1201/b10905)
- [29] D. Nelson, *Stationarity and Persistence in the GARCH(1,1) Model*. Econometric Theory **6**(1990), no. 3, 318-334. DOI: [10.1017/S0266466600005296](https://doi.org/10.1017/S0266466600005296)
- [30] D. Nelson, *Conditional Heteroskedasticity in Asset Returns: A New Approach*. Econometrica **59**(1991), no. 2, 347-370. DOI: [10.2307/2938260](https://doi.org/10.2307/2938260)
- [31] E. Nortey, D. Ngoh, K. Doku-Amponsah y K. Ofori-Boateng, *Modeling inflation rates and exchange rates in Ghana: application of multivariate GARCH models*. SpringerPlus **4**(2015), 1-10. DOI: [10.1186/s40064-015-0837-6](https://doi.org/10.1186/s40064-015-0837-6)
- [32] W. Polasek y R. Lei, *Generalized Impulse Response Functions for VAR-GARCH-M Models*. Gaul W., Opitz O., Schader M. (eds) Data Analysis. Studies in Classification, Data Analysis, and Knowledge Organization (2000), 299-311. DOI: [10.1007/978-3-642-58250-9_24](https://doi.org/10.1007/978-3-642-58250-9_24)
- [33] W. Polasek y L. Ren, *A multivariate GARCH-M model for exchange rates in the US, Germany and Japan*. Classification and Information Processing at the Turn of the Millennium (2000), 355-363. DOI: [10.1007/978-3-642-57280-7_39](https://doi.org/10.1007/978-3-642-57280-7_39)
- [34] A. Raftery, M. Newton, J. Satagopan y P. Krivitsky, *Estimating the Integrated Likelihood via Posterior Simulation Using the Harmonic Mean Identity*. Bayesian Statistics **8**(2007), no. 1, 1-45. DOI: [10.1093/oso/9780199214655.003.0015](https://doi.org/10.1093/oso/9780199214655.003.0015)
- [35] T. Takaishi, *Rational GARCH model: An empirical test for stock returns*. Physica A: Statistical Mechanics and Its Applications **473**(2017), 451-460. DOI: [10.1016/j.physa.2017.01.011](https://doi.org/10.1016/j.physa.2017.01.011)
- [36] Z. F. Tan, J. Zhang, J. H. Wang y J. Xu, *Day-ahead electricity price forecasting using wavelet transform combined with Arima and GARCH models*. Applied Energy **87**(2010), no. 11, 3606-3610. DOI: [10.1016/j.apenergy.2010.05.012](https://doi.org/10.1016/j.apenergy.2010.05.012)
- [37] R. Tsay, *Analysis of Financial Time Series*. Wiley Series in Probability and Statistics, 2005. DOI: [10.1002/0471746193](https://doi.org/10.1002/0471746193)

- [38] C. H. Tseng, S. T. Cheng, Y. H. Wang y J. T. Peng, *Artificial neural network model of the hybrid EGARCH volatility of the Taiwan stock index option prices*. *Physica A: Statistical Mechanics and Its Applications* **387**(2008), no. 13, 3192-3200. DOI: [10.1016/j.physa.2008.01.074](https://doi.org/10.1016/j.physa.2008.01.074)
- [39] Y. Wang, Y. Xiang, X. Lei e Y. Zhou, *Volatility analysis based on GARCH-type models: Evidence from the Chinese stock market*. *Economic Research-Ekonomska Istraživanja* **35**(2022), no. 1, 2530-2554. DOI: [10.1080/1331677X.2021.1967771](https://doi.org/10.1080/1331677X.2021.1967771)
- [40] A. Wilhelmsson, *GARCH forecasting performance under different distribution assumptions*. *Journal of Forecasting* **25**(2006), 561-578. DOI: [10.1002/for.1009](https://doi.org/10.1002/for.1009)
- [41] Z. J. Xiao y R. Koenker, *Conditional Quantile Estimation for Generalized Autoregressive Conditional Heteroscedasticity Models*. *Journal of the American Statistical Association* **104**(2009), no. 488, 1696-1712. DOI: [10.1198/jasa.2009.tm09170](https://doi.org/10.1198/jasa.2009.tm09170)
- [42] Y. Xu, X. Wang y H. Liu, *Quantile-based GARCH-MIDAS: Estimating value-at-risk using mixed-frequency information*. *Finance Research Letters* **43**(2021). DOI: [10.1016/j.frl.2021.101965](https://doi.org/10.1016/j.frl.2021.101965)
- [43] J.-M. Zakoian, *Threshold heteroskedastic models*. *Journal of Economic Dynamics and Control* **18**(1994), no. 5, 931-955. DOI: [10.1016/0165-1889\(94\)90039-6](https://doi.org/10.1016/0165-1889(94)90039-6)