

MODELO VAR INTEGRADO CON VOLATILIDAD
ESTOCÁSTICA MULTIVARIADA Y ERRORES DE
COLA PESADA

INTEGRATED VAR MODEL WITH
MULTIVARIATE STOCHASTIC VOLATILITY AND
HEAVY-TAILED ERRORS

CRISTIAN ANDRÉS CRUZ TORRES¹

MARVIN LEVI VILAFRANCA RIVERA²

Received: 19/Jul/2024; Accepted: 28/Nov/2025

Revista de Matemática: Teoría y Aplicaciones is licensed under a Creative Commons
Reconocimiento-NoComercial-CompartirIgual 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

² Universidad Nacional Autónoma de Honduras, Departamento de Estadística Matemática, Tegucigalpa, Honduras. E-Mail: marvin.villafranca@unah.hn

Resumen

Los modelos autorregresivos vectoriales (VAR) son usados para capturar las relaciones dinámicas de series de tiempo multivariadas. Por otro lado, los modelos de volatilidad estocástica multivariada (MSV) permiten modelar la varianza cuando cambia en el tiempo. La distribución t de Student es usada para modelar valores en las series de tiempo que a menudo son de magnitud extrema. Por lo anterior, en este artículo se propone la integración de un modelo VAR, un modelo MSV y una distribución t de Student (VAR-MSV-t). La elección del orden VAR-MSV-t más adecuado se lleva a cabo por medio del Criterio de Información de Desviación (DIC). Se presentan fórmulas para estimar la asimetría de Mardia y la curtosis de Koziol del modelo. Se hizo una aplicación a tres variables macroeconómicas clave para los Estados Unidos. Agregamos el índice del mercado de valores S&P 500 y se interpretaron los resultados. Para estimar los parámetros se usan métodos de Monte Carlo vía Cadenas de Markov (MCMC). Los resultados indican que el modelo captura las relaciones dinámicas, así como la varianza cambiando en el tiempo y los valores de magnitud extrema de manera eficaz.

Palabras clave: volatilidad estocástica multivariada; asimetría multivariada; curtosis multivariada; error de colas pesadas.

Abstract

Vector autoregressive (VAR) models are used to capture the dynamic relationships among multivariate time series. On the other hand, multivariate stochastic volatility (MSV) models allow modeling the variance as it changes over time. Student's t-distribution is used to model extreme values in time series. Therefore, this article proposes the integration of a VAR model, an MSV model, and a Student's t-distribution (VAR-MSV-t). The selection of the most appropriate VAR-MSV-t order is carried out using the Deviance Information Criterion (DIC). Formulas are presented to estimate the Mardia skewness and the Koziol kurtosis of the model. The model is applied to three key macroeconomic variables for the United States. The S&P 500 stock market index is added, and the results are interpreted. Parameter estimation is carried out using Markov chain Monte Carlo (MCMC) methods. The results indicate that the model effectively captures the dynamic relationships, time-varying variance, and extreme magnitude values.

Keywords: multivariate stochastic volatility; multivariate skewness; multivariate kurtosis; heavy tails error.

Mathematics Subject Classification: Primary: 62H12; secondary: 62F15, 62M10, 62P20.

1. INTRODUCCIÓN

Las variables macroeconómicas permiten a los macroeconomistas estudiar el crecimiento y fluctuaciones de un país desde una perspectiva amplia. Para modelar la dinámica y relaciones causales entre un conjunto de variables macroeconómicas,

Christopher A. Sims [56, 57, 58, 59] propuso el uso de los modelos autorregresivos vectoriales (VAR). La configuración de los modelos VAR es tal que los valores actuales de un conjunto de variables se explican en parte por los valores pasados de las variables involucradas. Debido a que los modelos VAR involucran valores actuales y pasados de múltiples series de tiempo, capturan comovimientos que no pueden capturarse en modelos univariados. Estadísticas estándar de resumen VAR como las pruebas de causalidad Granger, las funciones de respuesta impulso y las descomposiciones de varianza son métodos bien aceptados y ampliamente utilizados para representar estos comovimientos.

Sin embargo, los modelos VAR mantienen la suposición de choques con distribución normal y matriz de covarianza constante en el tiempo. Hay muchas razones para pensar que tal suposición puede ser demasiado restrictiva en muchas aplicaciones macroeconómicas. Por ejemplo, antes y después del periodo conocido como la Gran Moderación (1984–2006), las variables macroeconómicas son más volátiles (movimientos de baja frecuencia) y con choques raros (movimientos de alta frecuencia) (ver, por ejemplo, [3, 12, 20, 49, 64]). En las últimas décadas, un número creciente de estudios ha proporcionado evidencia que cuestiona la práctica del supuesto de normalidad. Algunos argumentan que los modelos VAR con estos supuestos difícilmente podrían explicar o predecir los eventos con movimientos de baja y alta frecuencia, principalmente porque estos modelos les asignan probabilidad casi cero (ver, por ejemplo, [18, 46]). Lawrence J. Christiano [8] comenta que la distribución de los residuos de los análisis VAR estándar presenta un marcado exceso de curtosis, y concluye que la evidencia en contra de la suposición de normalidad es sustancial. Frederic Mishkin [40] advierte contra los estudios cuantitativos de política monetaria óptima que suponen que los choques que afectan la economía tienen distribución normal invariable en el tiempo, y argumenta que en realidad “la distribución de los choques que afectan a la economía es más compleja” y “pueden exhibir exceso de curtosis, es decir, colas pesadas”.

En finanzas se suelen usar modelos de volatilidad estocástica multivariada (MSV) para capturar la volatilidad en múltiples retornos financieros. Los modelos MSV postulan que la volatilidad es conducida por su propio proceso estocástico. Para poder capturar los movimientos de baja frecuencia, los investigadores macroeconomistas han integrado los modelos VAR con modelos MSV (VAR-MSV). Por ejemplo, Harald Uhlig [69] introduce la volatilidad estocástica multivariada sin restricciones en el contexto de los modelos VAR, escogiendo la distribución beta multivariante para explotar una conjugación entre esa distribución y la distribución Wishart, de modo que la integración sobre el choque no observado en la matriz de precisión se pueda realizar de forma cerrada. Giorgio Primiceri [49] estimó un modelo VAR-MSV con parámetros que varían en el tiempo (TVP-VAR-MSV) estructural con el fin de investigar las causas potenciales del pobre desempeño económico de los Estados Unidos en los años 70 y principios de los 80 y en qué medida la política monetaria desempeñó un papel importante en esos episodios de alto desempleo e inflación. La literatura de los modelos VAR-MSV y TVP-VAR-MSV es extensa (ver, por ejemplo, [2, 9, 10, 11, 12, 14, 15, 20, 42, 43, 67]).

Cúrdia, Del Negro y Greenwald [13] en el contexto de los modelos de equilibrio general dinámico estocástico (DSGE) advierten que la evaluación de los movimientos de baja frecuencia puede potencialmente verse distorsionada por la presencia de movimientos de alta frecuencia. Argumentan que estimar un modelo que solo permite movimientos de baja frecuencia, sin tomar en cuenta los movimientos de alta frecuencia, resultará en que la volatilidad estocástica intentará ajustarse a los residuos, por lo que los movimientos de alta frecuencia pueden interpretarse como cambios persistentes en la volatilidad cuando, en realidad, pueden ser realizaciones raras de un proceso con distribución invariable en el tiempo. Para superar esta dificultad, algunos autores han integrado a los modelos VAR-MSV distribuciones de colas pesadas (VAR-MSV-t) para poder capturar los movimientos de alta frecuencia. Por ejemplo, Chiu, Mumtaz y Pintér [7] proponen un modelo VAR-MSV con errores t de Student para investigar el ajuste dentro de la muestra y el rendimiento de pronóstico fuera de la muestra.

En macroeconomía, normalmente se supone que los choques de las ecuaciones de la volatilidad estocástica son independientes de los choques de las variables endógenas. Además, la evolución de la volatilidad se modela como una caminata aleatoria con barreras reflectantes (ver, por ejemplo, [2, 7, 9, 11, 12, 14, 15, 20, 49]). Uhlig y Triantafyllopoulos [67, 69] configuran la volatilidad estocástica modelando la matriz de precisión del error como un componente no observado con choques extraídos de una distribución beta multivariante, pero los choques de las ecuaciones de volatilidad estocástica son independientes de los choques de las variables endógenas. Clark y Ravazzolo [10] comparan modelos VAR alternativos con volatilidad que varía en el tiempo en el que los choques de las ecuaciones de volatilidad estocástica son independientes de los choques de las variables endógenas, pero uno de los modelos que proponen trata la evolución de la volatilidad siguiendo un proceso autorregresivo de orden 1. Mumtaz [42, 43] configura los choques de las variables endógenas y los choques de la volatilidad siguiendo una distribución normal multivariada, permitiendo la correlación entre los choques en el tiempo t . Cruz y Villafranca [50, 65] proponen un modelo VAR-MSV en el que los choques de las variables endógenas en el tiempo t están correlacionados con los choques de la volatilidad en el tiempo $t+1$. Sin embargo, el modelo no captura los movimientos de alta frecuencia.

El primer objetivo de este artículo es proponer un modelo VAR-MSV-t en el que los choques de las variables endógenas en el tiempo t están correlacionados con los choques de la volatilidad en el tiempo $t+1$ y, de esta forma, medir el efecto de los choques de las variables endógenas en el tiempo t en los choques de la volatilidad en el tiempo $t+1$. La evolución de la volatilidad sigue un proceso VAR de orden 1, lo que permite medir la persistencia de la volatilidad. Asimismo, la volatilidad estocástica está diseñada para capturar los movimientos de baja frecuencia. La introducción de una variable aleatoria de cola pesada en la estructura de choque capturarán los movimientos de alta frecuencia que a menudo son de magnitud extrema. El número de grados de libertad, que determina la probabilidad de observar choques grandes raros, se estima a partir de los datos. El modelo VAR-

MSV-t determinará si la variación temporal en la estructura del modelo se deriva de eventos raros o de cambios persistentes en la volatilidad. El segundo objetivo es proporcionar algunas propiedades del modelo VAR-MSV-t. El tercer objetivo es desarrollar una metodología con el fin de ajustar el modelo VAR-MSV-t para estimar los parámetros. Para lograr este fin se usan métodos de Monte Carlo vía Cadenas de Markov (MCMC). El cuarto objetivo es aplicar el modelo para analizar el comportamiento de series macroeconómicas clave de los Estados Unidos: producción industrial, tasa de desempleo e índice de precios al consumidor. Además, debido a la reciente evidencia sobre la influencia de las variables financieras en las variables macroeconómicas (ver, por ejemplo, [7, 63]), agregamos el índice de rendimiento del mercado de valores S&P 500 con el fin de medir su influencia en las variables macroeconómicas.

El resto de este artículo está organizado de la siguiente manera. En la Sección 2 presentamos el modelo VAR-MSV-t y algunas de sus propiedades. En la Sección 3 presentamos el método de estimación. En la Sección 4 exponemos el método para escoger el orden VAR-MSV-t. En la Sección 5 presentamos un ejemplo con datos simulados. En la Sección 6 aplicamos el modelo VAR-MSV-t a datos reales. Por último, en la Sección 7 exponemos las conclusiones.

2. MODELO VAR INTEGRADO CON VOLATILIDAD ESTOCÁSTICA MULTIVARIADA

El modelo presentado en este artículo es un modelo autorregresivo vectorial con volatilidad estocástica multivariada y errores t de Student. La volatilidad estocástica y los errores t de Student modelados son los propuestos por Ishihara y Omori [24]. En este modelo, los choques de las variables endógenas están correlacionados; el modelo se diseñó así para capturar las posibles relaciones lineales entre ellos. Lo mismo ocurre con los choques de la volatilidad. Los choques de las variables endógenas en el tiempo t y los choques de la volatilidad en el tiempo $t + 1$ están correlacionados, lo que permite medir el efecto de los choques de las variables endógenas en el tiempo t en los choques de la volatilidad en el tiempo $t + 1$. Esto significa que un choque de ε_t afecta a α_{t+1} en lugar de α_t [70]. Esto se puede observar, ya que $\text{cov}(\varepsilon_t, \alpha_{t+1}) = \text{cov}(\varepsilon_t, \Phi\alpha_t + \eta_t) = \text{cov}(\varepsilon_t, \eta_t)$ (ver la Observación 3 del Apéndice A). De esta manera se puede medir el efecto de los choques en la varianza condicional de las variables macroeconómicas [42]. Los errores de cola pesada se modelan como un proceso con distribución invariable en el tiempo, diseñado para capturar pesos latentes que ajustan la intensidad de los choques, permitiendo representar eventos raros y extremos. En condiciones normales, estos pesos son aproximadamente 1; sin embargo, cuando toman valores mayores, amplifican la magnitud de los choques, lo que refleja situaciones atípicas con alta variabilidad. Las matrices de coeficientes están diseñadas para medir la dependencia lineal de las observaciones pasadas sobre las observaciones actuales, en otras palabras, miden la fuerza con la que las observaciones pasadas afectan a las actuales. El modelo es

de la siguiente manera

$$\mathbf{y}_t = \mathbf{v} + \mathbf{A}_1 \mathbf{y}_{t-1} + \dots + \mathbf{A}_k \mathbf{y}_{t-k} + \mathbf{w}_t, \quad (2.1)$$

$$\mathbf{w}_t = \lambda_t^{-1/2} \mathbf{V}_t^{1/2} \boldsymbol{\varepsilon}_t, \quad t = 1, \dots, n, \quad (2.2)$$

$$\boldsymbol{\alpha}_{t+1} = \boldsymbol{\Phi} \boldsymbol{\alpha}_t + \boldsymbol{\eta}_t, \quad t = 1, \dots, n-1, \quad (2.3)$$

$$\boldsymbol{\alpha}_1 \sim N_p(\mathbf{0}, \boldsymbol{\Sigma}_0), \quad (2.3)$$

$$\mathbf{V}_t^{1/2} = \text{diag}\left(\exp\left(\frac{\alpha_{1t}}{2}\right), \dots, \exp\left(\frac{\alpha_{pt}}{2}\right)\right), \quad (2.4)$$

$$\boldsymbol{\Phi} = \text{diag}(\phi_1, \dots, \phi_p), \quad (2.5)$$

$$\begin{pmatrix} \boldsymbol{\varepsilon}_t \\ \boldsymbol{\eta}_t \end{pmatrix} \sim N_{2p}(\mathbf{0}, \boldsymbol{\Sigma}), \quad \boldsymbol{\Sigma} = \begin{pmatrix} \boldsymbol{\Sigma}_{\varepsilon\varepsilon} & \boldsymbol{\Sigma}_{\varepsilon\eta} \\ \boldsymbol{\Sigma}_{\eta\varepsilon} & \boldsymbol{\Sigma}_{\eta\eta} \end{pmatrix}, \quad (2.6)$$

$$\text{vec}(\boldsymbol{\Sigma}_0) = (\mathbf{I}_{p^2} - \boldsymbol{\Phi} \otimes \boldsymbol{\Phi})^{-1} \text{vec}(\boldsymbol{\Sigma}_{\eta\eta}), \quad (2.7)$$

$$\lambda_t \sim \text{Gamma}\left(\frac{\nu}{2}, \frac{\nu}{2}\right), \quad t = 1, \dots, n, \quad (2.8)$$

$$\nu \sim \text{Gamma}(m_0^\nu, s_0^\nu), \quad (2.9)$$

donde \mathbf{y}_t , $t = -k + 1, \dots, 0, 1, \dots, n$, son variables disponibles de dimensión $p \times 1$. $\boldsymbol{\varepsilon}_t$, $t = 1, \dots, n$, son choques gaussianos de dimensión $p \times 1$. El vector \mathbf{v} es un término de intercepción de dimensión $p \times 1$, mientras que \mathbf{A}_i , $i = 1, \dots, k$, son matrices de coeficientes de dimensión $p \times p$. La variable λ_t es un escalar con distribución gamma común. Usamos $\text{Gamma}(a, b)$ para denotar una distribución gamma con media $\frac{a}{b}$ y varianza $\frac{a}{b^2}$. El parámetro ν representa los grados de libertad, mientras que m_0^ν y s_0^ν son hiperparámetros. El vector $\boldsymbol{\alpha}_t = \mathbf{h}_t - \boldsymbol{\mu}_h$ es de dimensión $p \times 1$, donde \mathbf{h}_t es el correspondiente vector de log volatilidad y $\boldsymbol{\mu}_h$ es el vector de media. El operador vec convierte una matriz $\mathbf{A} = [\mathbf{a}_1, \dots, \mathbf{a}_p]$, donde \mathbf{a}_i es de dimensión $p \times 1$, $i = 1, \dots, p$, en un vector $(\mathbf{a}_1^\top, \dots, \mathbf{a}_p^\top)^\top$ de dimensión $p^2 \times 1$. El operador \otimes es el producto de Kronecker.

Los elementos de las matrices de coeficientes A_{ij}^l , $j = 1, \dots, p$, $l = 1, \dots, k$, denotan la dependencia lineal de y_{it} (valor actual de la serie i) en $y_{j,t-l}$, $j \neq i$, $l = 1, \dots, k$ (valores pasados de las otras series) en la presencia de $y_{i,t-l}$, $l = 1, \dots, k$ (valores pasados de la serie i). Por lo que, A_{ij}^l , $j = 1, \dots, p$, $l = 1, \dots, k$, es el efecto condicional de $y_{j,t-l}$, $j \neq i$, $l = 1, \dots, k$, sobre y_{it} en presencia de $y_{i,t-l}$, $l = 1, \dots, k$. Si $A_{ij}^l = 0$ para todo l y $j \neq i$, entonces y_{it} no depende de $y_{j,t-l}$, $j \neq i$, $l = 1, \dots, k$ pero si de $y_{i,t-l}$, $l = 1, \dots, k$. Por otro lado, si $A_{ij}^l = 0$ para todo l y $j = i$, entonces y_{it} no depende de $y_{i,t-l}$, $l = 1, \dots, k$ pero si de $y_{j,t-l}$, $j \neq i$, $l = 1, \dots, k$. Un ejemplo más detallado de dos dimensiones se puede encontrar en [68].

Los elementos de la matriz $\boldsymbol{\Phi}$ en la ecuación (2.2) están relacionados con la persistencia de los choques a la volatilidad y en el modelo propuesto $-1 < \phi_i < 1$, $i = 1, \dots, p$. La persistencia de los choques a la volatilidad es el efecto del choque actual sobre el pronóstico de la volatilidad y eventualmente desaparece. La vida media de un choque viene dada por $-\log(2)/\log|\phi_i|$, que en series de tiempo mensuales, es el número de meses transcurridos para que el impacto del choque

se reduzca a la mitad [61]. Cuando ϕ_i es cercano a 1 y $\sigma_{ii,\eta\eta}$ es cercano a cero, la evolución de la volatilidad de una serie de tiempo es muy suave [48].

La volatilidad estocástica pretende capturar la posible heterocedasticidad de los choques y las posibles no linealidades en las relaciones simultáneas entre las variables del modelo [49]. En series de tiempo mensuales, un mes en el que $\alpha_t = \mathbf{0}$ puede ser visto como un mes normal. Un mes normal es uno en el que la velocidad de evolución de la volatilidad no es ni demasiado rápida ni demasiado lenta y asume un valor promedio. En un mes normal (condicionado a $\alpha_t = \mathbf{0}$), w_{it} tiene varianza $\sigma_{ii,\varepsilon\varepsilon}$. Luego, $\sigma_{ii,\varepsilon\varepsilon}$ puede ser interpretado como la varianza condicional en un mes normal. La varianza general de w_{it} se denota por $\sigma_{ii,ww}$ y $100(1 - \sigma_{ii,\varepsilon\varepsilon}/\sigma_{ii,ww})$ es el porcentaje de la varianza atribuible a la presencia de heterocedasticidad en la serie temporal i . El flujo de la volatilidad de w_{it} es dado por $\exp\{0,5\sigma_{ii,\eta\eta}/(1 - \phi_i^2)\}$ [60].

La estacionariedad de (2.2) está asegurada por la ecuación (2.7). Las submatrices $\Sigma_{\varepsilon\varepsilon}$ y $\Sigma_{\eta\eta}$ se usan para capturar la posible correlación entre los choques de las variables endógenas y los choques de la volatilidad, respectivamente. Además, la submatriz de covarianza $\Sigma_{\varepsilon\eta}$ se usa para calcular la posible correlación entre los choques de las variables endógenas en el mes actual y los choques de la volatilidad del siguiente mes.

Para el modelo de las ecuaciones (2.1)–(2.9) suponemos que $\mathbf{y}_{-k+1}, \dots, \mathbf{y}_0$ son valores premuestrales disponibles, $\alpha_t = \mathbf{0}$ para $t = 0, -1, -2, \dots$, y $\mathbf{y}_t = \mathbf{0}$ para $t = -k, -k - 1, \dots$, respectivamente. Además, $\text{cov}(\varepsilon_t, \eta_s) = \mathbf{0}$ para todo $t \neq s$ y $\text{cov}(\lambda_t, \varepsilon_{s,i}) = \text{cov}(\lambda_t, \eta_{s,i}) = 0$ para todo $t, s = 1, \dots, n, i = 1, \dots, p$.

Para encontrar la función verosimilitud del modelo de las ecuaciones (2.1)–(2.9) se hace una leve modificación a la verosimilitud de Ishihara y Omori [24] definiendo $\mathbf{Y}_t = [1, \mathbf{y}_{t-1}^\top, \dots, \mathbf{y}_{t-k}^\top]^\top$ y $\boldsymbol{\beta} = \text{vec}(\boldsymbol{\nu}, \mathbf{A}_1, \dots, \mathbf{A}_k)$, de dimensión $(kp + 1) \times 1$ y $(kp^2 + p) \times 1$, respectivamente. Luego reescribimos el modelo de la ecuación (2.1) de la siguiente manera:

$$\mathbf{y}_t = (\mathbf{Y}_t^\top \otimes \mathbf{I}_p)\boldsymbol{\beta} + \mathbf{w}_t. \tag{2.10}$$

Luego se define $\boldsymbol{\theta} = (\boldsymbol{\phi}, \boldsymbol{\Sigma}, \boldsymbol{\beta}, \nu)$, $\boldsymbol{\phi} = (\phi_1, \dots, \phi_p)^\top$, $\boldsymbol{\alpha} = (\alpha_1^\top, \dots, \alpha_n^\top)^\top$, $\boldsymbol{\lambda} = (\lambda_1, \dots, \lambda_n)$, $\mathbf{Y}^n = (\mathbf{y}_1, \dots, \mathbf{y}_n)$, $\mathbf{Y}^k = (\mathbf{y}_{-k+1}, \dots, \mathbf{y}_0)$ y $\mathbf{1}_p = [1, \dots, 1]^\top$, y se obtiene

$$\begin{aligned} f(\boldsymbol{\lambda}, \boldsymbol{\alpha}, \mathbf{Y}^n \mid \boldsymbol{\theta}, \mathbf{Y}^k) &= f(\mathbf{Y}^n, \boldsymbol{\alpha} \mid \boldsymbol{\lambda}, \boldsymbol{\theta}, \mathbf{Y}^k) f(\boldsymbol{\lambda} \mid \boldsymbol{\theta}, \mathbf{Y}^k) \\ &\propto \exp\left\{ \sum_{t=1}^n \ell_t - \frac{1}{2} \boldsymbol{\alpha}_1^\top \boldsymbol{\Sigma}_0^{-1} \boldsymbol{\alpha}_1 \right. \\ &\quad \left. - \frac{1}{2} \sum_{t=1}^{n-1} (\boldsymbol{\alpha}_{t+1} - \boldsymbol{\Phi} \boldsymbol{\alpha}_t)^\top \boldsymbol{\Sigma}_{\eta\eta}^{-1} (\boldsymbol{\alpha}_{t+1} - \boldsymbol{\Phi} \boldsymbol{\alpha}_t) \right\} \\ &\quad \times \left(\prod_{t=1}^n \lambda_t^{\frac{p+\nu}{2}-1} \right) |\boldsymbol{\Sigma}_0|^{-1/2} |\boldsymbol{\Sigma}|^{-\frac{n-1}{2}} |\boldsymbol{\Sigma}_{\varepsilon\varepsilon}|^{-1/2}, \end{aligned} \tag{2.11}$$

donde

$$\begin{aligned} \ell_t = & -\frac{1}{2}(\mathbf{y}_t - ((\mathbf{Y}_t^\top \otimes \mathbf{I}_p)\boldsymbol{\beta} + \boldsymbol{\mu}_t))^\top \boldsymbol{\Sigma}_t^{-1}(\mathbf{y}_t - ((\mathbf{Y}_t^\top \otimes \mathbf{I}_p)\boldsymbol{\beta} + \boldsymbol{\mu}_t)) \\ & - \frac{1}{2}v\lambda_t - \frac{1}{2}\mathbf{1}_p^\top \boldsymbol{\alpha}_t + \text{const}, \end{aligned} \quad (2.12)$$

$$\boldsymbol{\mu}_t = \left(\lambda_t^{-\frac{1}{2}} \mathbf{V}_t^{\frac{1}{2}} \right) \mathbf{m}_t, \quad (2.13)$$

$$\boldsymbol{\Sigma}_t = \left(\lambda_t^{-\frac{1}{2}} \mathbf{V}_t^{\frac{1}{2}} \right) \mathbf{S}_t \left(\lambda_t^{-\frac{1}{2}} \mathbf{V}_t^{\frac{1}{2}} \right). \quad (2.14)$$

$$\mathbf{m}_t = \begin{cases} \boldsymbol{\Sigma}_{\varepsilon\eta} \boldsymbol{\Sigma}_{\eta\eta}^{-1} (\boldsymbol{\alpha}_{t+1} - \boldsymbol{\Phi} \boldsymbol{\alpha}_t), & t < n, \\ \mathbf{0}, & t = n, \end{cases} \quad (2.15)$$

$$\mathbf{S}_t = \begin{cases} \boldsymbol{\Sigma}_{\varepsilon\varepsilon} - \boldsymbol{\Sigma}_{\varepsilon\eta} \boldsymbol{\Sigma}_{\eta\eta}^{-1} \boldsymbol{\Sigma}_{\eta\varepsilon}, & t < n, \\ \boldsymbol{\Sigma}_{\varepsilon\varepsilon}, & t = n. \end{cases} \quad (2.16)$$

2.1. Propiedades del modelo de las ecuaciones (2.1)–(2.9).

Antes de comenzar a describir las propiedades del modelo, comentaremos brevemente la literatura relacionada con los modelos MSV. Es bien conocido que Harvey et al. [22] propusieron el primer modelo MSV multivariado, dado por $y_{it} = \exp(h_{it}/2)\varepsilon_{it}$, donde $h_{it} = \gamma_i + \phi_i h_{it-1} + \eta_{it}$, $i = 1, \dots, p$. En este modelo, $\boldsymbol{\varepsilon}_t$ tiene vector de media cero y matriz de covarianza $\boldsymbol{\Sigma}_{\varepsilon}$, cuyos elementos de la diagonal principal son la unidad y los elementos fuera de la diagonal se indican como ρ_{ij} . Por su parte, $\boldsymbol{\eta}_t$ tiene vector de media cero y matriz de covarianza $\boldsymbol{\Sigma}_{\eta}$. Harvey et al. [22] se concentran en el caso particular donde $\gamma_i = 0$ y $\phi_i = 1$. Demuestran que la matriz de covarianza de $\boldsymbol{\xi}_t = (\xi_{1t}, \dots, \xi_{pt})^\top$, donde $\xi_{it} = \log(\varepsilon_{it}^2) + 1,27$, es dada por $(\pi^2/2)\rho_{ij}^*$, donde ρ_{ij}^* es igual a 1 si $i = j$, y es igual a $\frac{2}{\pi^2} \sum_{l=1}^{\infty} \frac{(l-1)!}{(l/2)!} \rho_{ij}^{2l}$, si $i \neq j$. La mayoría de la literatura hace referencia a este resultado cuando se trata de la matriz de covarianza de $\boldsymbol{\xi}_t$, por ejemplo, Asai y McAleer [1]. Para la elaboración de este artículo se realizó una revisión exhaustiva sobre los segundos, terceros y cuartos momentos de \mathbf{w}_t y no se encontró nada referente a estas propiedades (al menos en la literatura que analizamos, que fue abundante). Tampoco se hallaron referencias sobre la asimetría ni la curtosis de \mathbf{w}_t . Por lo que, a continuación se presentan la matriz de covarianza general, la asimetría de Mardia y la curtosis de Koziol de \mathbf{w}_t . En el Apéndice A se presentan las demostraciones de los resultados principales.

La matriz de covarianza general de \mathbf{w}_t es de la siguiente manera

$$\boldsymbol{\Sigma}_w = \begin{cases} \frac{v\sigma_{ii,\varepsilon\varepsilon}}{v-2} \exp\left(\frac{1}{2} \frac{\sigma_{ii,\eta\eta}}{1-\phi_i^2}\right), & \text{si } i = j, \\ \frac{v\sigma_{ij,\varepsilon\varepsilon}}{v-2} \exp\left(\frac{1}{8} \left(\frac{\sigma_{ii,\eta\eta}}{1-\phi_i^2} + \frac{\sigma_{jj,\eta\eta}}{1-\phi_j^2} + \frac{2\sigma_{ij,\eta\eta}}{1-\phi_i\phi_j} \right)\right), & \text{si } i \neq j. \end{cases} \quad (2.17)$$

La asimetría de Mardia [37] es una de las medidas más utilizadas, basada en los momentos de tercer orden, calculados a partir de un vector aleatorio estandariza-

do. De manera similar, la curtosis de Koziol [31] se define a partir de los momentos de cuarto orden del mismo vector. Encontrar expresiones cerradas para las medidas de asimetría y curtosis en distribuciones multivariadas, como las propuestas por Mardia y Koziol, es un tema activo de investigación [25], especialmente en distribuciones no normales [33].

En el caso de los errores \mathbf{w}_t del modelo VAR-MSV-t definido por las ecuaciones (2.1)–(2.9), las expresiones cerradas de la asimetría de Mardia y la curtosis de Koziol se obtienen de la siguiente manera

$$\|E(Y_{\mathbf{w}_t} \otimes Y_{\mathbf{w}_t} \otimes Y_{\mathbf{w}_t})\|^2 = \sum_{i=1}^{p^3} (L_w^{1i} E(\mathbf{w}_t \otimes \mathbf{w}_t \otimes \mathbf{w}_t))^2 = 0, \tag{2.18}$$

$$\|E(Y_{\mathbf{w}_t} \otimes Y_{\mathbf{w}_t} \otimes Y_{\mathbf{w}_t} \otimes Y_{\mathbf{w}_t})\|^2 = \sum_{i=1}^{p^4} (L_w^{2i} E(\mathbf{w}_t \otimes \mathbf{w}_t \otimes \mathbf{w}_t \otimes \mathbf{w}_t))^2, \tag{2.19}$$

donde $Y_{\mathbf{w}_t} = L_w^{-1}(\mathbf{w}_t - \boldsymbol{\mu}_w) = L_w^{-1}\mathbf{w}_t$, L_w^{1i} y L_w^{2i} corresponden a la i -ésima fila de $L_w^{-1} \otimes L_w^{-1} \otimes L_w^{-1}$ y $L_w^{-1} \otimes L_w^{-1} \otimes L_w^{-1} \otimes L_w^{-1}$, respectivamente. $E(\mathbf{w}_t \otimes \mathbf{w}_t \otimes \mathbf{w}_t)$ y $E(\mathbf{w}_t \otimes \mathbf{w}_t \otimes \mathbf{w}_t \otimes \mathbf{w}_t)$ son los vectores compuestos de todos los terceros y cuartos momentos de \mathbf{w}_t y son definidos en el Apéndice A.

Es bien sabido que, bajo condiciones de estacionariedad, la media y la matriz de covarianza de un modelo autorregresivo vectorial existen (ver, por ejemplo, Lütkepohl [36], quien lo demuestra suponiendo que el proceso se inició en un pasado infinito). Sin embargo, en el modelo propuesto en este trabajo, la ecuación (2.3) tiene el supuesto de que $\boldsymbol{\alpha}_1 \sim N_p(\mathbf{0}, \boldsymbol{\Sigma}_0)$, por lo que no se puede suponer que el proceso de las ecuaciones (2.1)–(2.9) comienza desde el pasado infinito. Aun así, es posible derivar la media y la matriz de covarianza de \mathbf{y}_t cuando el tamaño de la muestra es grande, a saber, cuando $t \rightarrow \infty$. Por lo cual, la media y la matriz de covarianza de \mathbf{y}_t son de la siguiente manera

$$\boldsymbol{\mu}_{y_t} = \lim_{t \rightarrow \infty} E(\mathbf{y}_t) = \mathbf{J}(\mathbf{I}_{pk} - \mathbf{A})^{-1}\mathbf{v}_c, \tag{2.20}$$

$$\text{vec}(\boldsymbol{\Sigma}_{y_t}) = \lim_{t \rightarrow \infty} \text{vec}(\text{cov}(\mathbf{y}_t)) = (\mathbf{J} \otimes \mathbf{J})(\mathbf{I}_{(kp)^2} - \mathbf{A} \otimes \mathbf{A})^{-1} \text{vec}(\boldsymbol{\Sigma}_w), \tag{2.21}$$

donde \mathbf{J} , \mathbf{v}_c , \mathbf{A} y $\boldsymbol{\Sigma}_w$ se definen en el Apéndice A. Una demostración análoga de las ecuaciones (2.20)–(2.21) se puede encontrar en Lütkepohl [36].

La ecuación (2.17) es la matriz de covarianza de \mathbf{w}_t y sus elementos quedan definidos en función de los parámetros $\boldsymbol{\Phi}$, $\boldsymbol{\Sigma}_{\varepsilon\varepsilon}$, $\boldsymbol{\Sigma}_{\eta\eta}$ y v . La ecuación (2.18) es la asimetría de Mardia y es igual a 0, mientras que la ecuación (2.19) es la curtosis de Koziol. Cabe resaltar que las ecuaciones (2.20)–(2.21) aseguran que el modelo propuesto no diverja cuando el tamaño de la muestra tiende a infinito.

3. MÉTODO DE ESTIMACIÓN

Para estimar los parámetros se usa inferencia bayesiana calculando las distribuciones a posteriori por medio del algoritmo MCMC de ocho bloques que es dado por

1. Inicializar $\alpha, \phi, \Sigma, \beta, \lambda$ y ν .
2. Generar $\beta \mid \lambda, \nu, \alpha, \phi, \Sigma, Y^n, Y^k$.
3. Generar $\nu \mid \alpha, \phi, \Sigma, \beta, \lambda, Y^n, Y^k$.
4. Generar $\lambda \mid \nu, \alpha, \phi, \Sigma, \beta, Y^n, Y^k$.
5. Generar $\alpha \mid \phi, \Sigma, \beta, \lambda, \nu, Y^n, Y^k$.
6. Generar $\Sigma \mid \beta, \lambda, \nu, \alpha, \phi, Y^n, Y^k$.
7. Generar $\phi \mid \Sigma, \beta, \lambda, \nu, \alpha, Y^n, Y^k$.
8. Ir a 2.

3.1. Generación de β .

La distribución a priori $f(\beta)$ de β es la de Litterman [16, 32] que corresponde a una distribución normal multivariante con media a priori μ_β y matriz de covarianza a priori Σ_β . dicha distribución a priori de Litterman en este artículo configura $\mu_\beta = \mathbf{0}$ y $\Sigma_\beta \neq \mathbf{0}$. Esto significa que se cree que la dependencia intertemporal de las variables es débil. Nosotros configuramos la matriz de covarianza a priori Σ_β como una matriz diagonal de la siguiente manera

$$\Sigma_{\beta,ij,l} = \begin{cases} \left(\frac{\bar{\lambda}}{l}\right)^2 & \text{si } i = j, \\ \left(\frac{\bar{\lambda}\bar{\theta}\sigma_i}{l\sigma_j}\right)^2 & \text{si } i \neq j, \end{cases} \quad (3.1)$$

donde $\Sigma_{\beta,ij,l}$ es la varianza a priori de A_{ij} del rezago l , $\bar{\lambda}$ es la desviación estándar a priori de los coeficientes $A_{\ell,1}$, $\ell = 1, 2, \dots, k$, $0 < \bar{\theta} < 1$ y, σ_i^2 es el elemento de la i -ésima diagonal de Σ_w o $\Sigma_{\varepsilon\varepsilon}$ actualizado en cada iteración del algoritmo. La función de densidad de probabilidad posterior de β es dada por

$$f(\beta \mid \alpha, \phi, \Sigma, \nu, \lambda, Y^n, Y^k) \propto \exp\left\{-\frac{1}{2}(\beta - \mu_\beta^*)^\top \Sigma_\beta^{*-1} (\beta - \mu_\beta^*)\right\}, \quad (3.2)$$

donde

$$\Sigma_\beta^* = \left[\sum_{t=1}^n (Y_t \otimes I_p)^\top \Sigma_t^{-1} (Y_t^\top \otimes I_p) + \Sigma_\beta^{-1} \right]^{-1},$$

$$\mu_\beta^* = \Sigma_\beta^* \left[\sum_{t=1}^n (Y_t \otimes I_p) \Sigma_t^{-1} (y_t - \mu_t) \right].$$

Se usa el algoritmo de muestreador de Gibbs para generar una muestra de $\beta^* \sim N(\mu_\beta^*, \Sigma_\beta^*)$.

3.2. Generación de ν .

Para generar ν aplicamos el algoritmo MH de Ishihara y Omori [24] sustituyendo \mathbf{y}_t por $\mathbf{y}_t^* = \mathbf{y}_t - (\mathbf{Y}_t^\top \otimes \mathbf{I}_p)\hat{\boldsymbol{\beta}}$. La distribución a priori es $\nu \sim \text{Gamma}(m_0^\nu, s_0^\nu)$ y la función de densidad de probabilidad posterior condicional completa es

$$\pi(\nu | \boldsymbol{\phi}, \boldsymbol{\Sigma}, \boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\lambda}, \mathbf{Y}^n, \mathbf{Y}^k) \propto \pi(\nu) \left\{ \frac{\left(\frac{\nu}{2}\right)^{\frac{\nu}{2}}}{\Gamma\left(\frac{\nu}{2}\right)} \right\}^n \left(\prod_{t=1}^n \lambda_t \exp\left\{-\sum_{t=1}^n \lambda_t\right\} \right)^{\frac{\nu}{2}}. \tag{3.3}$$

Para muestrear esta distribución posterior condicional se usa la transformación $\theta_\nu = \log(\nu)$. Se hace la expansión de Taylor de segundo orden a la distribución posterior condicional $\pi(\theta_\nu | \boldsymbol{\phi}, \boldsymbol{\Sigma}, \boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\lambda}, \mathbf{Y}^n, \mathbf{Y}^k)$ alrededor de $\widehat{\theta}_\nu$ y se aproxima a una densidad normal. Nosotros usamos el algoritmo de Newton-Raphson para encontrar la moda usando la función de puntaje y la matriz de información observada [38]. Se propone un candidato de la distribución normal $\theta_\nu^* \sim N(\mu_\nu, \sigma_\nu^2)$, donde

$$\begin{aligned} \mu_\nu &= \widehat{\theta}_\nu + \sigma_\nu^2 \left[\frac{\partial \log \pi(\theta_\nu | \boldsymbol{\phi}, \boldsymbol{\Sigma}, \boldsymbol{\lambda}, \boldsymbol{\alpha}, \boldsymbol{\beta}, \mathbf{Y}^n, \mathbf{Y}^k)}{\partial \theta_\nu} \Big|_{\theta_\nu = \widehat{\theta}_\nu} \right], \\ \sigma_\nu^2 &= \left[\frac{\partial^2 \log \pi(\theta_\nu | \boldsymbol{\phi}, \boldsymbol{\Sigma}, \boldsymbol{\lambda}, \boldsymbol{\alpha}, \boldsymbol{\beta}, \mathbf{Y}^n, \mathbf{Y}^k)}{\partial \theta_\nu^2} \Big|_{\theta_\nu = \widehat{\theta}_\nu} \right]^{-1}. \end{aligned}$$

Se acepta el candidato con probabilidad $\min \left[\frac{\pi(\theta_\nu^* | \boldsymbol{\phi}, \boldsymbol{\Sigma}, \boldsymbol{\lambda}, \boldsymbol{\alpha}, \boldsymbol{\beta}, \mathbf{Y}^n, \mathbf{Y}^k) f(\theta_\nu | \mu_\nu, \sigma_\nu^2)}{\pi(\theta_\nu | \boldsymbol{\phi}, \boldsymbol{\Sigma}, \boldsymbol{\lambda}, \boldsymbol{\alpha}, \boldsymbol{\beta}, \mathbf{Y}^n, \mathbf{Y}^k) f(\theta_\nu^* | \mu_\nu, \sigma_\nu^2)}, 1 \right]$, donde θ_ν es la muestra actual, y $f(x | \mu, \sigma^2)$ denota una función de densidad de probabilidad normal.

3.3. Generación de λ

Para generar λ aplicamos el algoritmo MH de Ishihara y Omori [24] sustituyendo \mathbf{y}_t por $\mathbf{y}_t^* = \mathbf{y}_t - (\mathbf{Y}_t^\top \otimes \mathbf{I}_p)\hat{\boldsymbol{\beta}}$. La distribución a priori es $\lambda_t \sim \text{Gamma}\left(\frac{\nu}{2}, \frac{\nu}{2}\right)$ y la función de densidad de probabilidad posterior condicional completa es

$$\begin{aligned} \pi(\lambda_t | \boldsymbol{\phi}, \boldsymbol{\Sigma}, \boldsymbol{\alpha}, \boldsymbol{\beta}, \nu, \mathbf{Y}^n, \mathbf{Y}^k) &\propto \lambda_t^{\frac{\nu+p}{2}-1} \exp\left\{-\frac{c_t}{2}\lambda_t + d_t \sqrt{\lambda_t}\right\}, \tag{3.4} \\ c_t &= \nu + \mathbf{y}_t^\top \boldsymbol{\Sigma}_t^{-1} \mathbf{y}_t, \\ d_t &= \mathbf{y}_t^\top \boldsymbol{\Sigma}_t^{-1} \boldsymbol{\mu}_t. \end{aligned}$$

Se propone un candidato $\lambda_t^* \sim \text{Gamma}\left(\frac{\nu+p}{2}, \frac{c_t}{2}\right)$ y se acepta con probabilidad $\min[\exp\{d_t(\sqrt{\lambda_t^*} - \sqrt{\lambda_t})\}, 1]$, donde λ_t es la muestra actual.

3.4. Generación de $\boldsymbol{\alpha}$.

Para generar $\boldsymbol{\alpha}$ se aplica el método de muestra de múltiples movimientos de Ishihara y Omori [24] sustituyendo \mathbf{y}_t por $\mathbf{y}_t^* = \hat{\lambda}_t^{\frac{1}{2}}(\mathbf{y}_t - (\mathbf{Y}_t^\top \otimes \mathbf{I}_p)\hat{\boldsymbol{\beta}})$. Se divide

$\boldsymbol{\alpha} = (\boldsymbol{\alpha}_1^\top, \dots, \boldsymbol{\alpha}_n^\top)^\top$ en $K + 1$ bloques usando el algoritmo de Shephard y Pitt [53]. Se encuentra la distribución completa de densidad conjunta condicional de las perturbaciones del i -ésimo bloque y usando la expansión de Taylor de segundo orden alrededor de la moda y se aproxima a una densidad normal que se usa para el algoritmo de Aceptación-Rechazo (AR). Como la dimensión de la matriz de covarianza crece cuando el tamaño de los bloques crece, se convierte la densidad normal aproximada en un modelo de espacios de estado auxiliar. Se aplica el suavizador de perturbaciones de Koopman [30] repetidas veces al modelo de espacios de estado auxiliar para encontrar la moda y se obtiene un modelo de espacios de estado gaussiano lineal aproximado. Por último, se aplica un algoritmo de Metropolis-Hastings de Aceptación-Rechazo (AR-MH) en el que se utiliza un simulador de perturbaciones [17, 26] al modelo de espacios de estado gaussiano lineal aproximado para generar un candidato. Para más detalles sobre el algoritmo de muestra de múltiples movimientos, ver [24].

3.5. Generación de $\boldsymbol{\phi}$

Para generar $\boldsymbol{\phi}$ aplicamos el algoritmo MH de Ishihara y Omori [24] sustituyendo \mathbf{y}_t por $\mathbf{y}_t^* = \hat{\lambda}_t^{\frac{1}{2}}(\mathbf{y}_t - (\mathbf{Y}_t^\top \otimes \mathbf{I}_p)\hat{\boldsymbol{\beta}})$. La distribución a priori es $\frac{\phi_i+1}{2} \sim \text{Beta}(a_i, b_i)$, $i = 1, \dots, p$, donde $\text{Beta}(a_i, b_i)$ denota la distribución Beta. La función de densidad de probabilidad posterior condicional completa es

$$\pi(\boldsymbol{\phi} \mid \boldsymbol{\Sigma}, \boldsymbol{\alpha}, \boldsymbol{\beta}, \nu, \boldsymbol{\lambda}, \mathbf{Y}^n, \mathbf{Y}^k) \propto h(\boldsymbol{\phi}) \exp \left\{ -\frac{1}{2}(\boldsymbol{\phi} - \boldsymbol{\mu}_\phi)^\top \boldsymbol{\Sigma}_\phi^{-1}(\boldsymbol{\phi} - \boldsymbol{\mu}_\phi) \right\}, \quad (3.5)$$

donde

$$\begin{aligned} h(\boldsymbol{\phi}) &= |\boldsymbol{\Sigma}_0|^{-1/2} \prod_{j=1}^p (1 + \phi_j)^{a_{j-1}} (1 - \phi_j)^{b_{j-1}} \times \exp \left\{ -\frac{1}{2} \boldsymbol{\alpha}_1^\top \boldsymbol{\Sigma}_0 \boldsymbol{\alpha}_1 \right\}, \\ \boldsymbol{\mu}_\phi &= \boldsymbol{\Sigma}_\phi \boldsymbol{b}, \quad \boldsymbol{\Sigma}_\phi^{-1} = \boldsymbol{\Sigma}^{22} \odot \mathbf{A}, \quad \mathbf{A} = \sum_{t=1}^{n-1} \boldsymbol{\alpha}_t \boldsymbol{\alpha}_t^\top, \quad \odot \text{ es el producto de Hadamard,} \\ \boldsymbol{b} &= \text{diag}(B_{11}, \dots, B_{pp}), \quad B_{ii} \text{ es el } (i, i)\text{-ésimo elemento de } \mathbf{B}, \\ \mathbf{B} &= \sum_{t=1}^{n-1} \left\{ \boldsymbol{\alpha}_t \mathbf{y}_t^\top \mathbf{V}_t^{-\frac{1}{2}} \boldsymbol{\Sigma}^{12} + \boldsymbol{\alpha}_t \boldsymbol{\alpha}_{t+1}^\top \boldsymbol{\Sigma}^{12} \right\}, \quad \boldsymbol{\Sigma}^{-1} = \begin{bmatrix} \boldsymbol{\Sigma}^{11} & \boldsymbol{\Sigma}^{12} \\ \boldsymbol{\Sigma}^{21} & \boldsymbol{\Sigma}^{22} \end{bmatrix}. \end{aligned}$$

Muestreamos de (3.5) usando un algoritmo MH generando un candidato de una distribución normal truncada sobre la región R , $\boldsymbol{\phi}^* \sim \text{TN}_R(\boldsymbol{\mu}_\phi, \boldsymbol{\Sigma}_\phi)$ y $R = \{\boldsymbol{\phi} : |\phi_j| < 1, j = 1, \dots, p\}$ y aceptando con probabilidad $\min\left\{\frac{h(\boldsymbol{\phi}^*)}{h(\boldsymbol{\phi})}, 1\right\}$, donde $\boldsymbol{\phi}$ es la muestra actual.

3.6. Generación de $\boldsymbol{\Sigma}$

Para generar $\boldsymbol{\Sigma}$ aplicamos el algoritmo MH de Ishihara y Omori [24] sustituyendo \mathbf{y}_t por $\mathbf{y}_t^* = \hat{\lambda}_t^{\frac{1}{2}}(\mathbf{y}_t - (\mathbf{Y}_t^\top \otimes \mathbf{I}_p)\hat{\boldsymbol{\beta}})$. La distribución a priori es $\boldsymbol{\Sigma} \sim \text{IW}(n_0, \mathbf{R}_0)$,

donde $IW(n_0, \mathbf{R}_0)$ denota la distribución Inversa Wishart. La función de densidad de probabilidad posterior condicional completa es

$$\pi(\boldsymbol{\Sigma} \mid \boldsymbol{\phi}, \boldsymbol{\alpha}, \boldsymbol{\beta}, \nu, \boldsymbol{\lambda}, \mathbf{Y}_n) \propto |\boldsymbol{\Sigma}|^{\frac{n_1+2p+1}{2}} \exp\left\{-\frac{1}{2} \text{tr}(\mathbf{R}_1^{-1} \boldsymbol{\Sigma}^{-1})\right\} \times g(\boldsymbol{\Sigma}), \tag{3.6}$$

donde

$$g(\boldsymbol{\Sigma}) = |\boldsymbol{\Sigma}_0|^{-\frac{1}{2}} |\boldsymbol{\Sigma}_{\varepsilon\varepsilon}|^{-\frac{1}{2}} \exp\left\{-\frac{1}{2} \left(\boldsymbol{\alpha}_1^\top \boldsymbol{\Sigma}_0 \boldsymbol{\alpha}_1 + \mathbf{y}_n^\top \mathbf{V}_n^{-\frac{1}{2}} \boldsymbol{\Sigma}^{-1} \mathbf{V}_n^{-\frac{1}{2}} \mathbf{y}_n\right)\right\},$$

$$n_1 = n_0 + n - 1, \quad \mathbf{R}_1^{-1} = \mathbf{R}_0^{-1} + \sum_{t=1}^{n-1} \mathbf{v}_t \mathbf{v}_t^\top, \quad \mathbf{v}_t = \begin{pmatrix} \mathbf{V}_t^{-\frac{1}{2}} \mathbf{y}_t \\ \boldsymbol{\alpha}_{t-1} - \boldsymbol{\Phi} \boldsymbol{\alpha}_t \end{pmatrix}.$$

Muestreamos de (3.6) usando un algoritmo MH generando un candidato $\boldsymbol{\Sigma}^* \sim IW(n_1, \mathbf{R}_1)$ y aceptando con probabilidad $\min\left\{\frac{g(\boldsymbol{\Sigma}^*)}{g(\boldsymbol{\Sigma})}, 1\right\}$, donde $\boldsymbol{\Sigma}$ es la muestra actual.

4. SELECCIÓN DEL ORDEN VAR-MSV

Para escoger el mejor modelo VAR-MSV-t usamos la metodología propuesta por Ishihara y Omori [24]. Para cada modelo estimado, se calcula el Criterio de Información de Desviación (DIC) de Spiegelhalter, Best, Carlin y Van Der Linde [62]. La medida DIC es definida por

$$DIC = E_{\theta|Y^n}[D(\boldsymbol{\theta})] + P_D, \tag{4.1}$$

donde

$$P_D = E_{\theta|Y^n}[D(\boldsymbol{\theta})] - D(E_{\theta|Y^n}[\boldsymbol{\theta}]), \quad D(\boldsymbol{\theta}) = -2 \log f(\mathbf{Y}^n \mid \boldsymbol{\theta}). \tag{4.2}$$

Para calcular $E_{\theta|Y^n}[D(\boldsymbol{\theta})]$ se puede aproximar mediante $\frac{1}{M} \sum_{m=1}^M D(\boldsymbol{\theta}^{(m)})$, donde $\boldsymbol{\theta}^{(m)}$ son remuestreados a partir de la distribución posterior. El error estándar del estimador se obtiene al estimar repetidamente $E_{\theta|Y^n}[D(\boldsymbol{\theta})]$. Luego, $D(E_{\theta|Y^n}[\boldsymbol{\theta}])$ es igual a $D(\boldsymbol{\theta})$ evaluado en la media posterior. Ishihara y Omori [24] configuraron $M = 100$, $I = 10000$ y repitieron 10 veces $E_{\theta|Y^n}[D(\boldsymbol{\theta})]$ para obtener el error estándar. Utilizan el filtro de partículas auxiliar propuesto por Shephard y Pitt [54] para calcular la función verosimilitud ordinaria dada los parámetros $\log f(\mathbf{Y}^n \mid \boldsymbol{\theta})$ (para una comprensión detallada del filtro de partículas ver [24]). Para escoger el mejor orden VAR-MSV-t se aplican los siguientes pasos:

1. Suponiendo que se sabe que el orden VAR-MSV-t no puede exceder un entero k_1 , se procede a estimar los modelos VAR-MSV-t comenzando desde 0 hasta k_1 y se almacenan sus parámetros estimados $\boldsymbol{\theta}_0, \boldsymbol{\theta}_1, \dots, \boldsymbol{\theta}_{k_1}$, donde $\boldsymbol{\theta}_i$ son los parámetros estimados del modelo i .

2. Para cada modelo se sustituye \mathbf{y}_t por $\mathbf{y}_t^* = \hat{\lambda}_t^{\frac{1}{2}(i)} (\mathbf{y}_t - (\mathbf{Y}_t^\top \otimes \mathbf{I}_p) \hat{\boldsymbol{\beta}}^{(i)})$, donde $\hat{\lambda}_t^{\frac{1}{2}(i)}$ y $\hat{\boldsymbol{\beta}}^{(i)}$ son las variables no observables y matrices de coeficientes estimadas del modelo i , respectivamente. Luego se procede a calcular la correspondiente función verosimilitud ordinaria dados los parámetros $\log f(\mathbf{Y}^n | \boldsymbol{\theta}_i)$ y se procede a calcular la medida DIC para cada modelo.
3. Se escoge el modelo que tenga la menor medida DIC.

5. EJEMPLO ILUSTRATIVO USANDO DATOS SIMULADOS

En esta sección se aplica el método propuesto usando datos simulados para observar que tan bien se estiman los parámetros. Los métodos propuestos en este trabajo requieren el cálculo de algoritmos recursivos de operaciones con matrices, por lo que el costo computacional aumenta a medida que el tamaño de la muestra n y la cantidad de las variables involucradas p aumenta. Debido al gran costo computacional, nuestros resultados computacionales se generaron vinculando C++ con el software libre R por medio de las librerías de álgebra lineal Rcpp y RcppArmadillo. La implementación del proceso de generación de datos simulados, $\boldsymbol{\Sigma}_w$, $\boldsymbol{\Sigma}_{v_t}$, la curtosis de Koziol, así como los cuartos momentos (propiedades 4–8) y la distribución condicional completa de $\boldsymbol{\beta}$, está disponible en el repositorio [66].

La cantidad de parámetros del modelo VAR-MSV-t propuesto en este trabajo crece a medida que la dimensión p y el orden k crecen. En modelos donde la cantidad de parámetros es grande se requieren muestras grandes para obtener estimaciones precisas [29]. Sin embargo, en la práctica, es raro encontrar conjuntos de datos macroeconómicos grandes. Debido a la importancia de medir el desempeño de la metodología propuesta en muestras grandes y pequeñas, se divide esta sección en dos partes. La primera parte consiste en un análisis donde se aplica la metodología a cuatro conjuntos de datos simulados de tamaño $n = 1000, 2000, 3000, 4000$ observaciones. Para analizar el desempeño de la metodología en muestras de tamaño menores que 1000, en la segunda parte se presenta un breve análisis donde se aplica la metodología simulando conjuntos de datos de tamaño $n = 500, 625, 750, 850$, cada tamaño con diferentes órdenes $k = 1, 2, 3, 4, 5$, siendo un total de 20 conjuntos de datos simulados. Sin embargo, como se verá más adelante, no es la intención de esta parte de la sección realizar un análisis exhaustivo sobre muestras pequeñas.

5.1. Muestras grandes.

En esta parte los datos son simulados usando el modelo (2.1)-(2.9). El orden del modelo MSV-VAR-t que se usa para simular datos es $k = 5$, la dimensión es $p = 4$ y $\nu = 12$. La matriz de covarianza $\boldsymbol{\Sigma}$ es tomada de [24]. Configuramos las matrices de coeficientes de tal manera que refleje la creencia de que los coeficientes de los rezagos de orden alto probablemente sean cercanos a cero y, la creencia de que la mayor parte de la variación en cada una de las variables se explica por los rezagos propios. El término de intercepción así como las matrices de coeficientes son dados por $v_i = 0, i = 1, \dots, 4$. $A_{ii}^l = 0,15 - 0,02l, i = 1, \dots, 4, l = 1, \dots, 5$. $A_{ij}^l =$

$0,11 - 0,02l$, $i \neq j$, $l = 1, 2$, y $A_{ij}^l = 0$, $i \neq j$, $l = 3, 4, 5$. Los parámetros Φ y Σ vienen dados por $\phi_i = 0,96$, $i = 1, \dots, 4$. $\sigma_{ii,\varepsilon\varepsilon} = 1,44$, $i = 1, \dots, 4$, y $\sigma_{ij,\varepsilon\varepsilon} = 0,864$, $i \neq j$. $\sigma_{ii,\eta\eta} = 0,04$, $i = 1, \dots, 4$, y $\sigma_{ij,\eta\eta} = 0,028$, $i \neq j$. $\sigma_{ii,\varepsilon\eta} = -0,096$, $i = 1, \dots, 4$, y $\sigma_{ij,\varepsilon\eta} = -0,072$, $i \neq j$.

Los hiperparámetros de la distribución a priori ν y ϕ son los usuales de la literatura, donde $m_0^\nu = 1$, $s_0^\nu = 0,05$, $a_i = 20$, $i = 1, \dots, 4$, $b_i = 1,5$, $i = 1, \dots, 4$. Para los hiperparámetros de la distribución a priori Σ se configuró $n_0 = 2p$ y $R_0 = (n_0\Sigma^*)^{-1}$, donde Σ^* es la matriz de covarianza verdadera. Los hiperparámetros de la a priori de Litterman son los usuales de la literatura, $\mu_\beta = \mathbf{0}$, $\bar{\theta} = 0,2$ y $\bar{\lambda} = 0,5$, mientras que los elementos σ^2_i , $i = 1, \dots, 4$, son tomados de la diagonal de la varianza de $\Sigma_w^{(j)}$ en la iteración j , por lo que son actualizados en cada iteración. Para el parámetro de ajuste se usa $K = 1 \times 10^{-5}n^2$ y para evitar que el método se atasque firmemente debido al rechazo excesivo se sigue la sugerencia de [53] incrementando K unas pocas iteraciones en intervalos regulares.

Para ilustrar el rendimiento de la metodología propuesta en las secciones anteriores generamos una muestra al azar de 4000 observaciones y se estimaron los parámetros usando cada una de las siguientes submuestras: (a) las 4000 observaciones, (b) las primeras 3000 observaciones, (c) las primeras 2000 observaciones y (d) las primeras 1000 observaciones.

A cada una de las submuestras se aplica el método de la Sección 3 a un VAR-MSV-t de orden 5, se generan 120000 muestras y se descartan las primeras 40000. Las restantes 80000 muestras se dividen en 5 submuestras de la siguiente manera: (a) las primeras 10000 muestras, (b) las primeras 20000 muestras, (c) las primeras 40000 muestras, (d) las primeras 60000 muestras y (e) las 80000 muestras. La Tabla 1 muestra los errores de estimación relativos $\|\hat{\theta} - \theta\|_F / \|\theta\|_F$ y el error estándar de $\|\hat{\theta}\|_F$ entre paréntesis. Lo primero que observamos es que, a excepción del parámetro ν , para cada muestra de las observaciones, los parámetros del modelo VAR-MSV-t se estimaron con aproximadamente la misma precisión, independientemente del tamaño de las muestras generadas por las distribuciones condicionales completas. Por otro lado, se observa que el menor error relativo para el parámetro Σ fue para la muestra de 2000 observaciones, siendo aproximadamente 1.064 veces más pequeño que el que se estimó con la muestra completa, y alrededor de 3.6 veces más pequeño que el que se estimó con la muestra de 1000 observaciones. Para el parámetro Φ , el menor error relativo fue para la muestra de 3000 observaciones y es aproximadamente 1.31 veces más pequeño que el de la muestra completa, y alrededor de 2.81 veces más pequeño que el de la muestra de 1000 observaciones. En el caso del parámetro β , el menor error relativo es el que se estimó con la muestra completa, y es aproximadamente 1.18 veces más pequeño que el de la muestra de 3000 observaciones, y aproximadamente 1.92 veces más pequeño que el de la muestra de 1000 observaciones. En general, para el parámetro ν , el menor error relativo fue para la muestra de 2000 observaciones. Específicamente, el valor mínimo fue 0.0004, aproximadamente 133 veces más pequeño que el de la muestra

completa y alrededor de 61 veces más pequeño que el estimado con la muestra de 1000 observaciones.

Por lo mencionado en el párrafo anterior y los resultados de la Tabla 1, se observa que, para el parámetro β , el tamaño de la muestra de las observaciones es muy influyente en la precisión de su estimación. Sin embargo, como se observa en las Tablas 4–6, con una muestra de 2000 observaciones, la estimación de β con la metodología propuesta en la sección 3 es muy cercana a los parámetros verdaderos. Para los parámetros Σ , Φ y ν , con una muestra de 2000 observaciones, existe la posibilidad de obtener estimaciones precisas y esto se refleja en las Tablas 3 y 6. Esto puede ser debido al azar de la muestra de 2000 observaciones, que preserva adecuadamente la estructura del proceso de generación de datos.

Adicionalmente, para demostrar el rendimiento del método de la Sección 4, a cada una de las muestras de tamaño 4000, 3000, 2000 y 1000 observaciones se estimaron los parámetros del modelo VAR-MSV-t con diferentes órdenes, comenzando desde el 1 y llegando hasta el 5. Luego, a todos los modelos estimados se les calcula la medida DIC. La Tabla 2 muestra las medidas DIC de todos los modelos VAR-MSV-t estimados. Se puede observar que, en todas las muestras, la medida DIC más grande es para el VAR(1)-MSV-t y, a medida que el orden crece, la medida DIC va disminuyendo. La menor medida DIC en todas las muestras es para el VAR(5)-MSV-t. Por lo que la metodología propuesta en la Sección 4 es eficiente para capturar el mejor orden VAR-MSV-t.

Las Tablas 3–6 muestran la media posterior e intervalos de credibilidad del 95 % de ν , los elementos de la matriz Φ , la matriz de covarianza Σ , el término de intercepción \mathbf{v} y las matrices de coeficientes \mathbf{A}_1 , \mathbf{A}_2 , \mathbf{A}_3 , \mathbf{A}_4 y \mathbf{A}_5 . Cabe resaltar que estos parámetros fueron estimados con la muestra de 2000 observaciones para demostrar la robustez de la metodología propuesta. Todos los intervalos de credibilidad con un 95 % de confianza contienen los parámetros verdaderos (excepto A_{12}^1 , A_{22}^1 y A_{33}^2). Una observación que se debe tomar en cuenta es que los intervalos de credibilidad de los parámetros estimados A_{14}^2 , A_{11}^5 , A_{21}^1 , A_{24}^2 , A_{22}^5 , A_{41}^2 , A_{44}^4 y A_{44}^5 contienen el cero cuando los parámetros verdaderos son distintos de cero (aunque pequeños, variando desde 0.05 a 0.09); sin embargo, contienen el cero cerca del límite inferior, donde sus límites inferiores varían entre $-0,008$ y 0 . Si el intervalo de credibilidad de un parámetro estimado contiene el cero y está muy cerca de alguno de sus límites cuando el parámetro verdadero es distinto de cero, se puede llegar a creer que es un indicador de que, aumentando el tamaño de la muestra, el intervalo de credibilidad no contendría el cero; sin embargo, el intervalo de credibilidad del parámetro estimado ν_1 contiene el cero cerca de su límite inferior cuando el parámetro verdadero es cero (ver Tabla 6).

En la muestra de 1000 observaciones, siete de los parámetros verdaderos no fueron contenidos por los intervalos de credibilidad del 95 % de sus parámetros estimados. Además, trece intervalos de credibilidad del 95 % de los parámetros

Tabla 1: Errores de estimación relativos $\|\hat{\theta} - \theta\|_F / \|\theta\|_F$ y errores estándar de $\|\hat{\theta}\|_F$ (entre paréntesis) para las submuestras de tamaño 1000, 2000, 3000 y 4000.

	n	Muestras				
		10000	20000	40000	60000	80000
Σ	4000	0.0728 (0.2045)	0.0719 (0.2028)	0.0734 (0.2006)	0.0729 (0.1986)	0.073 (0.197)
	3000	0.1009 (0.2623)	0.0962 (0.2599)	0.0935 (0.2527)	0.0950 (0.2527)	0.0948 (0.2513)
	2000	0.0695 (0.3176)	0.0685 (0.3215)	0.0671 (0.3143)	0.0692 (0.3101)	0.0678 (0.3125)
	1000	0.2372 (0.4999)	0.2493 (0.5090)	0.2471 (0.5156)	0.2461 (0.5254)	0.2509 (0.5410)
Φ	4000	0.0056 (0.0075)	0.0058 (0.0075)	0.0061 (0.0076)	0.006 (0.0076)	0.006 (0.0077)
	3000	0.0044 (0.0083)	0.0047 (0.0084)	0.0045 (0.0083)	0.0045 (0.0084)	0.0044 (0.0084)
	2000	0.0090 (0.0100)	0.0090 (0.0102)	0.0091 (0.0102)	0.0091 (0.0103)	0.0091 (0.0104)
	1000	0.0142 (0.0158)	0.0120 (0.0157)	0.0123 (0.0156)	0.0124 (0.0162)	0.0123 (0.0161)
β	4000	0.2051 (0.0218)	0.2044 (0.0218)	0.2039 (0.0218)	0.2042 (0.0218)	0.2042 (0.0217)
	3000	0.2426 (0.0265)	0.2429 (0.0265)	0.2420 (0.0264)	0.2419 (0.0264)	0.2421 (0.0264)
	2000	0.3197 (0.0343)	0.3195 (0.0342)	0.3199 (0.0342)	0.3204 (0.0343)	0.3207 (0.0343)
	1000	0.3920 (0.0536)	0.3903 (0.0535)	0.3905 (0.0534)	0.3898 (0.0535)	0.3902 (0.0536)
ν	4000	0.0554 (1.1755)	0.0565 (1.1563)	0.0560 (1.1620)	0.0540 (1.1540)	0.0532 (1.1517)
	3000	0.0214 (1.1660)	0.0212 (1.1770)	0.0275 (1.1887)	0.0267 (1.2023)	0.0287 (1.2094)
	2000	0.0131 (1.6257)	0.0033 (1.5161)	0.0004 (1.4848)	0.0036 (1.4865)	0.0050 (1.4890)
	1000	0.0130 (1.9124)	0.0260 (1.9561)	0.0245 (1.9713)	0.0282 (1.9742)	0.0275 (1.9875)

estimados de las matrices de coeficientes contienen el cero cuando sus parámetros verdaderos son distintos de cero (variando entre 0,05 y 0,09).

Debido al espacio limitado, solo presentamos los gráficos de convergencia, histogramas y los gráficos de las funciones de autocorrelación de las muestras generadas por las distribuciones condicionales completas de ν_1 y A_{44}^5 . Sin embargo,

Tabla 2: Estimadores, desviación estándar, valores más grandes y valores más pequeños DIC de los modelos VAR(1)-MSV-t, VAR(2)-MSV-t, VAR(3)-MSV-t, VAR(4)-MSV-t y VAR(5)-MSV-t para las submuestras 1000, 2000, 3000 y 4000.

	n	5	4	3	2	1
DIC	4000	39267.4	39328.49	39573.55	39892.82	41401.83
	3000	29064.25	29129.99	29313.37	29604.39	30728.95
	2000	18957.98	19008.64	19104.84	19299.26	20020.73
	1000	10124.25	10151.52	10209.21	10327.26	10738.93
DIC _{mín}	4000	39266.77	39327.42	39573.22	39892.20	41401.29
	3000	29063.84	29129.45	29312.96	29603.79	30728.08
	2000	18957.53	19008.18	19104.39	19298.77	20020.24
	1000	10124.07	10151.25	10208.79	10327.01	10738.88
DIC _{máx}	4000	39268.04	39328.99	39574.19	39893.39	41402.58
	3000	29064.64	29130.68	29313.9	29605.22	30729.39
	2000	18958.42	19009.02	19105.22	19299.82	20021.17
	1000	10124.45	10151.75	10209.5	10327.56	10738.97
Desviación	4000	0.421641	0.478483	0.315336	0.347745	0.374204
	3000	0.254719	0.364507	0.340827	0.418848	0.347939
	2000	0.278370	0.359721	0.274035	0.353985	0.335569
	1000	0.113382	0.178808	0.226657	0.184318	0.030657

las muestras de los demás elementos del término de intercepción y las matrices de coeficientes presentan los mismos comportamientos. Por otro lado, diagnósticos de las muestras generadas por las distribuciones condicionales completas de los elementos de Σ , Φ y ν se pueden encontrar en [24].

La Figura 1 muestra los gráficos de convergencia, histogramas y las funciones de autocorrelación de las muestras de los parámetros estimados v_1 y A_{44}^5 . En el caso de v_1 , el gráfico de convergencia varía en un rango fijo, lo que es un indicador fuerte de convergencia. La función de autocorrelación alcanza su máximo en 0,333 y decrece lentamente, mostrando una alta dependencia de valores pasados. Esto puede ser debido a la priori de Litterman, ya que las varianzas priori son infinitas para v_i , $i = 1, \dots, p$, reflejando que no se tiene ninguna conjetura a priori para estos coeficientes. En el caso de A_{44}^5 , el gráfico de convergencia varía en un rango fijo, indicando fuerte convergencia. La función de autocorrelación muestra poca correlación, alcanzando su máximo en 0,286 y decreciendo rápidamente, mostrando la poca dependencia de valores pasados. Como era de esperarse, debido a que la distribución condicional completa del término de intercepción y las matrices de coeficientes es normal multivariada, las muestras de sus respectivos elementos están distribuidas normalmente y esto se refleja en los histogramas de las muestras de v_1 y A_{44}^5 . El diagnóstico de convergencia muestra que el enfoque bayesiano para estimar los parámetros es eficiente.

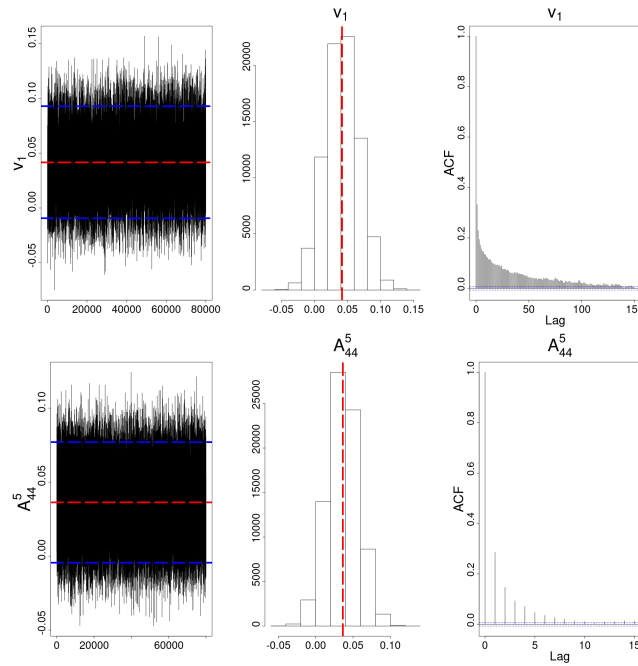


Figura 1: Gráfico de convergencia, histograma y función de autocorrelación del elemento v_1 y A_{44}^5 .

Por la ecuación (2.18), la asimetría de Mardia del modelo definido por las ecuaciones (2.1)-(2.9) es cero, por lo que uno esperaría que los residuales $y_t - (Y_t^T \otimes I_p)\hat{\beta}$, $t = k + 1, \dots, n$, tengan asimetría de Mardia cercana a cero. En la primera fila de la Tabla 7 se observa el estimador de asimetría muestral de Mardia (calculado por medio de la función `mardiaSkew` de la librería `semTools` de R) aplicado a los residuales. A diferencia de la expresión teórica de la asimetría de Mardia del modelo (ecuación (2.19)), que requiere estimar los parámetros Σ , Φ y ν , y sustituirlos en la fórmula, el estimador de asimetría muestral de Mardia solo necesita los residuales. El valor más alto es de 2,264 y corresponde a la muestra de tamaño 1000. A medida que el tamaño de la muestra aumenta, el valor del estimador decrece. El valor más bajo corresponde a la muestra de tamaño 4000 y es de 0,365, que es muy cercano a cero. De esta manera, uno puede usar el estimador de asimetría muestral de Mardia aplicado a los residuales como un indicador práctico de qué tan bien se estimaron los parámetros del modelo.

La curtosis de Koziol calculada con los parámetros verdaderos (Σ , Φ , ν) mediante la ecuación (2.19) es 288,95 (valor teórico), mientras que los valores obtenidos con los parámetros estimados ($\hat{\Sigma}$, $\hat{\Phi}$, $\hat{\nu}$) para los diferentes tamaños de muestra se presentan en la segunda fila de la Tabla 7, variando entre 238,15 y 258,19. El valor más cercano al valor teórico corresponde a la muestra de tamaño 3000 y es de

258,19, lo que implica una diferencia relativa del 10,6% respecto al valor teórico, indicando que los parámetros estimados a partir de esta muestra capturan adecuadamente la heterocedasticidad del modelo (seguido de cerca por la muestra de 2000 con 255,86). El valor más lejano corresponde a los residuales de la muestra de tamaño 1000 y es 238,15, sugiriendo que los parámetros estimados con esta muestra subestiman el nivel de heterocedasticidad presente en el modelo. Estos resultados concuerdan con los errores relativos de la Tabla 1. Por lo cual se aporta mayor evidencia de que las muestras de tamaño 2000 y 3000 preservan mejor la estructura del proceso de generación de datos.

En la tercera fila de la Tabla 7 se observa el estimador de curtosis muestral de Koziol (calculado por medio de la función Kurt de la librería mnt de R) aplicado a los residuales $\mathbf{y}_t - (\mathbf{Y}_t^\top \otimes \mathbf{I}_p)\hat{\boldsymbol{\beta}}$, $t = k + 1, \dots, n$. A diferencia de la expresión teórica de la curtosis de Koziol del modelo (ecuación (2.19)), que requiere estimar los parámetros $\boldsymbol{\Sigma}$, $\boldsymbol{\Phi}$ y ν , y sustituirlos en la fórmula, el estimador muestral de Koziol solo necesita los residuales. Se puede observar que el estimador de curtosis muestral de Koziol converge al valor teórico a medida que el tamaño de la muestra crece. Para las muestras de tamaño 1000, 2000 y 3000, el valor de curtosis obtenido mediante la ecuación (2.19), usando los parámetros estimados del modelo, es más cercano al valor teórico de 288,95. Las diferencias relativas en estos casos fueron de aproximadamente 17,6%, 11,6% y 10,6%, respectivamente, en comparación con las diferencias del estimador muestral de Koziol, que fueron de 56,6%, 31,8% y 16,5%. En cambio, para la muestra de 4000, el estimador muestral fue el más preciso, con una diferencia de apenas 1,9%, frente al 14,7% de la estimación basada en el modelo.

En la cuarta y quinta fila de la Tabla 7 se observa el error relativo de la matriz de covarianza estimada con la ecuación (2.21) y con la función cov de R respectivamente. Se puede observar que, para cada una de las muestras, el error relativo es menor para la matriz de covarianza calculada con los parámetros estimados que la que fue calculada con las observaciones.

Podemos observar que, en general, la metodología propuesta en las Secciones 3-4 permite estimar adecuadamente los valores verdaderos de los parámetros cuando se trabaja con muestras entre 1000 y 4000 observaciones. Esta capacidad de estimación se refleja también en los valores de asimetría de Mardia y la curtosis de Koziol calculados mediante las ecuaciones (2.18) y (2.19). No obstante, es importante señalar que los resultados presentados en esta sección se basan en un único conjunto de datos aleatorios para cada tamaño de muestra. Por tanto, si bien ofrecen evidencia útil sobre el comportamiento de la metodología propuesta, no son suficientes por sí solos para establecer conclusiones definitivas. Sería necesario realizar estudios adicionales con múltiples réplicas para obtener resultados más robustos y generalizables.

Tabla 3: Media posterior e intervalos de credibilidad del 95 % de los elementos de Φ y Σ .

	Parámetro	Media	Intervalos de Credibilidad del 95 %
ϕ_1	0.960	0.965	[0.952,0.976]
ϕ_2	0.960	0.956	[0.939,0.970]
ϕ_3	0.960	0.962	[0.947,0.974]
ϕ_4	0.960	0.944	[0.925,0.960]
$\sigma_{11,\varepsilon\varepsilon}$	1.440	1.425	[1.189,1.701]
$\sigma_{12,\varepsilon\varepsilon}$	0.864	0.903	[0.763,1.062]
$\sigma_{13,\varepsilon\varepsilon}$	0.864	0.856	[0.717,1.016]
$\sigma_{14,\varepsilon\varepsilon}$	0.864	0.914	[0.778,1.069]
$\sigma_{22,\varepsilon\varepsilon}$	1.440	1.574	[1.339,1.846]
$\sigma_{23,\varepsilon\varepsilon}$	0.864	0.910	[0.768,1.073]
$\sigma_{24,\varepsilon\varepsilon}$	0.864	0.964	[0.829,1.116]
$\sigma_{33,\varepsilon\varepsilon}$	1.440	1.430	[1.196,1.706]
$\sigma_{34,\varepsilon\varepsilon}$	0.864	0.909	[0.774,1.062]
$\sigma_{44,\varepsilon\varepsilon}$	1.440	1.583	[1.376,1.819]
$\sigma_{11,\varepsilon\eta}$	-0.096	-0.112	[-0.150,-0.077]
$\sigma_{12,\varepsilon\eta}$	-0.072	-0.071	[-0.107,-0.036]
$\sigma_{13,\varepsilon\eta}$	-0.072	-0.079	[-0.115,-0.045]
$\sigma_{14,\varepsilon\eta}$	-0.072	-0.080	[-0.118,-0.044]
$\sigma_{21,\varepsilon\eta}$	-0.072	-0.090	[-0.129,-0.053]
$\sigma_{22,\varepsilon\eta}$	-0.096	-0.098	[-0.137,-0.061]
$\sigma_{23,\varepsilon\eta}$	-0.072	-0.077	[-0.116,-0.041]
$\sigma_{24,\varepsilon\eta}$	-0.072	-0.100	[-0.142,-0.061]
$\sigma_{31,\varepsilon\eta}$	-0.072	-0.090	[-0.129,-0.054]
$\sigma_{32,\varepsilon\eta}$	-0.072	-0.055	[-0.093,-0.020]
$\sigma_{33,\varepsilon\eta}$	-0.096	-0.089	[-0.128,-0.052]
$\sigma_{34,\varepsilon\eta}$	-0.072	-0.083	[-0.121,-0.047]
$\sigma_{41,\varepsilon\eta}$	-0.072	-0.073	[-0.110,-0.038]
$\sigma_{42,\varepsilon\eta}$	-0.072	-0.067	[-0.106,-0.031]
$\sigma_{43,\varepsilon\eta}$	-0.072	-0.065	[-0.102,-0.030]
$\sigma_{44,\varepsilon\eta}$	-0.096	-0.113	[-0.154,-0.074]
$\sigma_{11,\eta\eta}$	0.040	0.037	[0.025,0.051]
$\sigma_{12,\eta\eta}$	0.028	0.024	[0.015,0.034]
$\sigma_{13,\eta\eta}$	0.028	0.027	[0.017,0.038]
$\sigma_{14,\eta\eta}$	0.028	0.026	[0.017,0.037]
$\sigma_{22,\eta\eta}$	0.040	0.034	[0.023,0.050]
$\sigma_{23,\eta\eta}$	0.028	0.024	[0.015,0.036]
$\sigma_{24,\eta\eta}$	0.028	0.023	[0.015,0.034]
$\sigma_{33,\eta\eta}$	0.040	0.036	[0.024,0.052]
$\sigma_{34,\eta\eta}$	0.028	0.026	[0.017,0.037]
$\sigma_{44,\eta\eta}$	0.040	0.037	[0.026,0.052]

Tabla 4: Media posterior e intervalos de credibilidad del 95 % de los elementos de las primeras dos filas de los parámetros A_1 , A_2 , A_3 , A_4 y A_5 .

	Parámetro	Media	Intervalos de Credibilidad del 95 %
A_{11}^1	0.13	0.082	[0.028,0.135]
A_{12}^1	0.09	0.149	[0.102,0.196]
A_{13}^1	0.09	0.101	[0.051,0.150]
A_{14}^1	0.09	0.075	[0.028,0.122]
A_{11}^2	0.11	0.114	[0.065,0.164]
A_{12}^2	0.07	0.089	[0.048,0.130]
A_{13}^2	0.07	0.085	[0.042,0.128]
A_{14}^2	0.07	0.033	[-0.008,0.073]
A_{11}^3	0.09	0.107	[0.062,0.152]
A_{12}^3	0	-0.007	[-0.042,0.028]
A_{13}^3	0	-0.018	[-0.055,0.019]
A_{14}^3	0	-0.006	[-0.041,0.029]
A_{11}^4	0.07	0.112	[0.070,0.154]
A_{12}^4	0	-0.024	[-0.054,0.007]
A_{13}^4	0	-0.013	[-0.046,0.019]
A_{14}^4	0	0.009	[-0.021,0.040]
A_{11}^5	0.05	0.039	[0.000,0.078]
A_{12}^5	0	-0.007	[-0.034,0.020]
A_{13}^5	0	-0.003	[-0.031,0.026]
A_{14}^5	0	0.002	[-0.025,0.029]
A_{21}^1	0.09	0.046	[-0.007,0.099]
A_{22}^1	0.13	0.159	[0.103,0.215]
A_{23}^1	0.09	0.106	[0.053,0.158]
A_{24}^1	0.09	0.097	[0.047,0.147]
A_{21}^2	0.07	0.088	[0.041,0.134]
A_{22}^2	0.11	0.075	[0.023,0.126]
A_{23}^2	0.07	0.116	[0.070,0.162]
A_{24}^2	0.07	0.044	[0.000,0.088]
A_{21}^3	0	0.015	[-0.025,0.054]
A_{22}^3	0.09	0.081	[0.034,0.128]
A_{23}^3	0	-0.014	[-0.054,0.025]
A_{24}^3	0	-0.001	[-0.038,0.037]
A_{21}^4	0	-0.005	[-0.038,0.029]
A_{22}^4	0.07	0.07	[0.026,0.114]
A_{23}^4	0	0.007	[-0.027,0.041]
A_{24}^4	0	-0.007	[-0.039,0.025]
A_{21}^5	0	0.008	[-0.022,0.038]
A_{22}^5	0.05	0.04	[0.000,0.081]
A_{23}^5	0	-0.005	[-0.035,0.025]
A_{24}^5	0	0.012	[-0.017,0.040]

Tabla 5: Media posterior e intervalos de credibilidad del 95 % de los elementos de las últimas dos filas de los parámetros A_1 , A_2 , A_3 , A_4 y A_5 .

	Parámetro	Media	Intervalos de credibilidad del 95 %
A_{31}^1	0.09	0.09	[0.039,0.140]
A_{32}^1	0.09	0.083	[0.036,0.130]
A_{33}^1	0.13	0.17	[0.115,0.223]
A_{34}^1	0.09	0.059	[0.011,0.107]
A_{31}^2	0.07	0.068	[0.024,0.111]
A_{32}^2	0.07	0.065	[0.024,0.106]
A_{33}^2	0.11	0.164	[0.115,0.214]
A_{34}^2	0.07	0.055	[0.015,0.096]
A_{31}^3	0	-0.01	[-0.048,0.027]
A_{32}^3	0	-0.024	[-0.060,0.012]
A_{33}^3	0.09	0.092	[0.046,0.138]
A_{34}^3	0	0.01	[-0.025,0.045]
A_{31}^4	0	-0.003	[-0.035,0.029]
A_{32}^4	0	-0.009	[-0.040,0.022]
A_{33}^4	0.07	0.057	[0.015,0.100]
A_{34}^4	0	0.001	[-0.030,0.031]
A_{31}^5	0	-0.004	[-0.033,0.025]
A_{32}^5	0	0.009	[-0.018,0.036]
A_{33}^5	0.05	0.067	[0.027,0.107]
A_{34}^5	0	-0.009	[-0.036,0.018]
A_{41}^1	0.09	0.083	[0.030,0.136]
A_{42}^1	0.09	0.087	[0.035,0.138]
A_{43}^1	0.09	0.093	[0.040,0.146]
A_{44}^1	0.13	0.141	[0.086,0.196]
A_{41}^2	0.07	0.039	[-0.007,0.085]
A_{42}^2	0.07	0.081	[0.036,0.126]
A_{43}^2	0.07	0.06	[0.014,0.106]
A_{44}^2	0.11	0.094	[0.044,0.145]
A_{41}^3	0	0.024	[-0.015,0.064]
A_{42}^3	0	-0.003	[-0.041,0.035]
A_{43}^3	0	0.002	[-0.037,0.042]
A_{44}^3	0.09	0.115	[0.069,0.161]
A_{41}^4	0	0.001	[-0.033,0.034]
A_{42}^4	0	0.001	[-0.032,0.033]
A_{43}^4	0	0.022	[-0.013,0.056]
A_{44}^4	0.07	0.041	[-0.002,0.085]
A_{41}^5	0	0.005	[-0.025,0.034]
A_{42}^5	0	-0.015	[-0.044,0.014]
A_{43}^5	0	0.01	[-0.020,0.041]
A_{44}^5	0.05	0.036	[-0.004,0.077]

Tabla 6: Media posterior e intervalos de credibilidad del 95 % de ν y los elementos del término intercepción ν .

	Parámetro	Media	Intervalos de credibilidad del 95 %
ν	12	12.060	[9.582,15.410]
ν_1	0	0.042	[-0.009,0.093]
ν_2	0	0.01	[-0.045,0.065]
ν_3	0	0.018	[-0.032,0.069]
ν_4	0	0.013	[-0.043,0.069]

Tabla 7: En la primera fila se observa el estimador de asimetría muestral de Mardia. En la segunda y tercera fila se observa la curtosis de Koziol estimada con la ecuación (2.19) (el valor teórico 288.95) y el estimador de curtosis muestral de Koziol respectivamente (con 10000 errores simulados de tamaño 4000, el estimador de curtosis multivariada obtuvo una media de 302.37 y desviación estándar de 150.48). En la cuarta y quinta fila se observa el error relativo de la ecuación (2.21) y de la función cov de R respectivamente.

n	1000	2000	3000	4000
Asimetría muestral de Mardia	2.264	1.172	0.615	0.365
Curtosis de Koziol	238.15	255.86	258.19	246.39
Curtosis muestral de Koziol	452.59	380.72	336.65	294.42
$\ \hat{\Sigma}_{y_t} - \Sigma_{y_t}\ _F / \ \Sigma_{y_t}\ _F$	0.277	0.035	0.008	0.024
$\ \text{cov}(y_t) - \Sigma_{y_t}\ _F / \ \Sigma_{y_t}\ _F$	0.363	0.067	0.0271	0.051

5.2. Muestras pequeñas.

En esta parte de la sección se usa la misma configuración de parámetros que en la sección de muestras grandes para simular los conjuntos de datos. Como se mencionó al principio de la sección, se simulan conjuntos de datos de tamaño $n = 500, 625, 750, 850$, cada uno con diferentes órdenes $k = 1, 2, 3, 4, 5$, siendo un total de 20 conjuntos de datos simulados. Se usó una semilla igual a 1989 para que el análisis sea replicable.

La Tabla 8 presenta los errores de estimación relativos $\|\hat{\theta} - \theta\|_F / \|\theta\|_F$ y el error estándar de $\|\hat{\theta}\|_F$, entre paréntesis, de los parámetros estimados. Se observa que el error relativo para el parámetro estimado Σ en todos los tamaños de muestra y en todos los órdenes es menor que el error relativo de la muestra de tamaño 1000 con orden 5. Esto puede ser debido a que la muestra generada con la semilla 1989 es bien representativa del proceso de generación de datos. Cabe resaltar que la muestra de tamaño 850 (que es la más grande de esta parte de la sección) obtuvo los errores relativos más grandes. Por el contrario, los errores relativos más pequeños se observan mayoritariamente en la muestra de tamaño $n = 750$, aunque en ciertos órdenes específicos (como $k = 3, 4$) la muestra de tamaño $n = 625$ alcanza valores ligeramente menores. Sin embargo, los errores estándar en todos

los tamaños de muestra y en todos los órdenes son más grandes que los errores estándar en la muestra de tamaño 1000 con orden 5. En el caso del parámetro Φ , a medida que el tamaño de la muestra crece, el error relativo y el error estándar disminuyen en todos los tamaños de muestra y en todos los órdenes (excepto algunos casos). Se puede observar que, en general, las muestras de tamaño 750 y 850 tienen menor error relativo que la muestra de tamaño 1000; además, los errores estándar de la muestra de 850 son aproximados a los errores estándar de la muestra de tamaño 1000.

En el caso del parámetro β , la tendencia que se observó en la sección de muestras grandes se mantiene. Todos los errores relativos y errores estándar aumentan en todos los tamaños de muestra y en todos los órdenes; además, son mayores que los estimados en la muestra de tamaño 1000. En el caso del parámetro ν , el error relativo y el error estándar crecen a medida que el tamaño de la muestra decrece. Este resultado es similar a los expuestos en la sección de muestras grandes. En el orden $k = 5$, los errores relativos de las muestras de tamaño 500 y 850 son aproximadamente 27 y 16 veces más grandes que el de la muestra de tamaño 1000, respectivamente.

En la Tabla 9 se presenta el estimador de asimetría muestral de Mardia, la curtosis de Koziol calculada con los parámetros verdaderos, el estimador de la curtosis muestral de Koziol, el error relativo de la ecuación (2.21) y el error relativo de la función cov de R, para los tamaños de muestra $n = 500, 625, 750, 850$, cada uno con diferentes órdenes $k = 1, 2, 3, 4, 5$. Se observa que, en todos los tamaños de muestra y en todos los órdenes, el estimador de asimetría muestral de Mardia es menor que el calculado con la muestra de 1000 observaciones. Las muestras de tamaño 500 y 850 son las que tienen menor estimador de asimetría muestral de Mardia. Por otro lado, se observa que, a medida que el tamaño de la muestra crece, el estimador de la curtosis de Koziol se acerca al valor teórico (288.9), lo cual concuerda con los resultados expuestos en la sección de muestras grandes. En el caso del estimador de la curtosis muestral de Koziol, se observa que los valores más bajos corresponden a la muestra de tamaño 500, mientras que la muestra de tamaño 750 reporta los valores más altos. Por último, se observa que la ecuación (2.21) presenta errores relativos menores que la función cov de R en la mayoría de los casos, especialmente para las muestras de tamaño 625 y 500. Sin embargo, para $n = 850$, los errores de (2.21) son mayores en casi todos los órdenes.

Las Tablas 10–13 presentan la media posterior, la desviación estándar y los intervalos de credibilidad del 95 % correspondientes a ν , los elementos de Φ , la matriz de covarianza Σ , el término de intercepción ν y las matrices de coeficientes A_1, A_2, A_3, A_4 y A_5 . Estos parámetros se estimaron con la muestra de tamaño 850. En el caso de los parámetros de persistencia estimados $\phi_i, i = 1, \dots, 5$, se observa que son muy cercanos a los valores verdaderos. Esto también se refleja en los sesgos, los cuales son pequeños, variando entre $-0,015$ y $0,005$. Los intervalos de credibilidad del 95 % de los elementos $\sigma_{13,\varepsilon\eta}, \sigma_{14,\varepsilon\eta}, \sigma_{21,\varepsilon\eta}$ y $\sigma_{34,\varepsilon\eta}$ contienen el cero, a pesar de que los valores verdaderos son distintos de cero. Estos últimos son

Tabla 8: Errores de estimación relativos $\|\hat{\theta} - \theta\|_F / \|\theta\|_F$ y errores estándar $\|\hat{\theta}\|_F$ (entre paréntesis) para las muestras $n = 500, 625, 750, 850$, cada tamaño con diferentes órdenes $k = 1, 2, 3, 4, 5$.

	n	k				
		1	2	3	4	5
Σ	850	0.1916 (0.6220)	0.1950 (0.6623)	0.1925 (0.6174)	0.1790 (0.6087)	0.1796 (0.6187)
	750	0.0990 (0.6415)	0.1116 (0.6287)	0.1216 (0.6081)	0.1180 (0.6220)	0.0995 (0.5846)
	625	0.1162 (0.6778)	0.1344 (0.7045)	0.1195 (0.6745)	0.1084 (0.6420)	0.1096 (0.6523)
	500	0.1447 (0.5868)	0.1572 (0.5749)	0.1663 (0.6007)	0.1556 (0.5864)	0.1524 (0.6035)
	850	0.0110 (0.0158)	0.0104 (0.0156)	0.0106 (0.0167)	0.0113 (0.0160)	0.0103 (0.0161)
Φ	750	0.0133 (0.0187)	0.0104 (0.0177)	0.0108 (0.0186)	0.0101 (0.0179)	0.0109 (0.0182)
	625	0.0161 (0.0225)	0.0133 (0.0205)	0.0139 (0.0210)	0.0150 (0.0226)	0.0152 (0.0224)
	500	0.0445 (0.0392)	0.0388 (0.0368)	0.0385 (0.0370)	0.0394 (0.0405)	0.0388 (0.0380)
	850	0.4101 (0.0469)	0.4390 (0.0491)	0.4260 (0.0546)	0.4323 (0.0583)	0.4382 (0.0621)
	750	0.4349 (0.0501)	0.4186 (0.0528)	0.4117 (0.0591)	0.4209 (0.0640)	0.4306 (0.0677)
β	625	0.4494 (0.0563)	0.4490 (0.0593)	0.4499 (0.0659)	0.5989 (0.0759)	0.4682 (0.0753)
	500	0.4522 (0.0649)	0.4781 (0.0676)	0.4627 (0.0760)	0.4633 (0.0815)	0.4918 (0.0859)
	850	0.5286 (8.0678)	0.4876 (7.8517)	0.4613 (7.4465)	0.4836 (7.7421)	0.4485 (7.3591)
	750	0.4254 (7.0892)	0.4011 (6.7799)	0.4691 (7.8057)	0.4723 (7.8495)	0.4507 (7.5315)
	625	0.8449 (12.9035)	0.7926 (12.1304)	0.8284 (12.6374)	0.7441 (11.5609)	0.7272 (11.3698)
ν	500	0.9228 (14.5248)	0.9378 (14.3195)	0.8982 (13.9146)	0.8678 (13.6731)	0.7404 (11.7561)

pequeños, iguales a $-0,072$. Sin embargo, los parámetros estimados son cercanos al valor verdadero, variando entre $-0,027$ y $-0,05$.

En el caso de las matrices de coeficientes, todos los intervalos de credibilidad con un 95 % de confianza contienen los parámetros verdaderos, excepto para A_{11}^2 ,

Tabla 9: En la primera fila se observa el estimador de asimetría muestral de Mardia. En la segunda fila se observa la curtosis de Koziol (con los parámetros verdaderos es 288.9). En la tercera fila se presenta el estimador muestral de Koziol. En la cuarta y quinta fila se observa el error relativo de la ecuación (2.21) y de la función cov de R respectivamente.

	n	k				
		1	2	3	4	5
Asimetría muestral de Mardia	850	1.074	1.049	1.043	1.071	1.099
	750	1.583	1.603	1.572	1.573	1.569
	625	1.585	1.635	1.686	1.761	1.732
	500	0.934	0.892	0.925	0.962	0.986
Curtosis de Koziol	850	242.95	245.79	246.15	247.10	246.13
	750	222.95	225.19	222.14	222.06	220.66
	625	199.84	199.33	198.33	197.00	196.60
	500	158.65	158.82	159.74	157.13	164.13
Curtosis muestral de Koziol	850	198.57	198.61	198.51	197.85	200.65
	750	228.27	230.85	231.53	228.01	231.19
	625	198.74	203.89	203.52	193.02	206.04
	500	181.88	180.99	180.84	182.98	185.99
$\frac{\ \hat{\Sigma}_{y_t} - \Sigma_{y_t}\ _F}{\ \Sigma_{y_t}\ _F}$	850	0.127	0.064	0.017	0.072	0.145
	750	0.022	0.030	0.088	0.186	0.302
	625	0.051	0.089	0.161	0.458	0.358
	500	0.046	0.043	0.062	0.134	0.213
$\frac{\ \text{cov}(y_t) - \Sigma_{y_t}\ _F}{\ \Sigma_{y_t}\ _F}$	850	0.024	0.049	0.063	0.057	0.032
	750	0.079	0.031	0.020	0.037	0.061
	625	0.202	0.162	0.150	0.169	0.120
	500	0.109	0.080	0.114	0.148	0.179

$A_{12}^2, A_{34}^2, A_{33}^4$ y v_4 . Por otro lado, los intervalos de credibilidad del 95 % de los elementos estimados $A_{12}^2, A_{13}^2, A_{14}^2, A_{11}^5, A_{21}^1, A_{24}^1, A_{24}^2, A_{22}^4, A_{22}^5, A_{32}^2, A_{33}^4, A_{41}^1, A_{41}^2, A_{42}^2, A_{43}^2$ y A_{44}^5 contienen el cero cuando los parámetros verdaderos son distintos de cero. Estos parámetros verdaderos son pequeños, variando entre 0.05 y 0.09, lo cual es consistente con los resultados obtenidos en la muestra de tamaño 1000. No obstante, los valores estimados son cercanos a los parámetros verdaderos, lo cual se observa con mayor claridad al analizar el sesgo. En el caso de los elementos cuyo valor verdadero es cero, los estimadores presentan valores absolutos que oscilan entre 0.0002 y 0.086; de estos, 34 de 36 son menores o iguales a 0.0139. Por último, en la Tabla 13 se observa que el parámetro estimado ν es igual a 17.382, un valor considerablemente mayor que el parámetro verdadero, que es 12. Sin embargo, el intervalo de credibilidad del 95 % incluye el valor verdadero, lo que indica una estimación razonable pese al sesgo.

Podemos observar por los resultados presentados en esta parte de la sección que, con muestras pequeñas, la metodología propuesta en las Secciones 3–4 estima

Tabla 10: Media posterior e intervalos de credibilidad del 95 % de los elementos de Φ y Σ .

	Parámetro	Media	Sesgo	Intervalos de credibilidad del 95 %
ϕ_1	0.96	0.945	-0.015	[0.9165 , 0.967]
ϕ_2	0.96	0.966	0.006	[0.9467 , 0.982]
ϕ_3	0.96	0.950	-0.010	[0.9266 , 0.969]
ϕ_4	0.96	0.965	0.005	[0.9478 , 0.980]
$\sigma_{11,\varepsilon\varepsilon}$	1.44	1.698	0.258	[1.356 , 2.124]
$\sigma_{12,\varepsilon\varepsilon}$	0.864	1.021	0.157	[0.791 , 1.319]
$\sigma_{13,\varepsilon\varepsilon}$	0.864	1.042	0.178	[0.822 , 1.314]
$\sigma_{14,\varepsilon\varepsilon}$	0.864	1.036	0.172	[0.810 , 1.324]
$\sigma_{22,\varepsilon\varepsilon}$	1.44	1.712	0.272	[1.276 , 2.322]
$\sigma_{23,\varepsilon\varepsilon}$	0.864	1.028	0.164	[0.794 , 1.338]
$\sigma_{24,\varepsilon\varepsilon}$	0.864	1.032	0.168	[0.791 , 1.355]
$\sigma_{33,\varepsilon\varepsilon}$	1.44	1.653	0.213	[1.286 , 2.119]
$\sigma_{34,\varepsilon\varepsilon}$	0.864	1.016	0.152	[0.788 , 1.314]
$\sigma_{44,\varepsilon\varepsilon}$	1.44	1.620	0.180	[1.246 , 2.130]
$\sigma_{11,\varepsilon\eta}$	-0.096	-0.057	0.039	[-0.116 , 0.000]
$\sigma_{12,\varepsilon\eta}$	-0.072	-0.059	0.013	[-0.113 , -0.007]
$\sigma_{13,\varepsilon\eta}$	-0.072	-0.047	0.025	[-0.109 , 0.013]
$\sigma_{14,\varepsilon\eta}$	-0.072	-0.027	0.045	[-0.081 , 0.025]
$\sigma_{21,\varepsilon\eta}$	-0.072	-0.046	0.026	[-0.105 , 0.009]
$\sigma_{22,\varepsilon\eta}$	-0.096	-0.073	0.023	[-0.129 , -0.022]
$\sigma_{23,\varepsilon\eta}$	-0.072	-0.078	-0.006	[-0.144 , -0.016]
$\sigma_{24,\varepsilon\eta}$	-0.072	-0.056	0.016	[-0.110 , -0.006]
$\sigma_{31,\varepsilon\eta}$	-0.072	-0.058	0.014	[-0.115 , -0.005]
$\sigma_{32,\varepsilon\eta}$	-0.072	-0.088	-0.016	[-0.142 , -0.038]
$\sigma_{33,\varepsilon\eta}$	-0.096	-0.120	-0.024	[-0.189 , -0.060]
$\sigma_{34,\varepsilon\eta}$	-0.072	-0.050	0.022	[-0.104 , 0.000]
$\sigma_{41,\varepsilon\eta}$	-0.072	-0.090	-0.018	[-0.150 , -0.037]
$\sigma_{42,\varepsilon\eta}$	-0.072	-0.094	-0.022	[-0.150 , -0.044]
$\sigma_{43,\varepsilon\eta}$	-0.072	-0.089	-0.017	[-0.155 , -0.029]
$\sigma_{44,\varepsilon\eta}$	-0.096	-0.096	0.000	[-0.156 , -0.043]
$\sigma_{11,\eta\eta}$	0.04	0.044	0.004	[0.027 , 0.068]
$\sigma_{12,\eta\eta}$	0.028	0.027	-0.001	[0.016 , 0.043]
$\sigma_{13,\eta\eta}$	0.028	0.035	0.007	[0.021 , 0.054]
$\sigma_{14,\eta\eta}$	0.028	0.031	0.003	[0.019 , 0.047]
$\sigma_{22,\eta\eta}$	0.04	0.035	-0.005	[0.022 , 0.053]
$\sigma_{23,\eta\eta}$	0.028	0.029	0.001	[0.017 , 0.044]
$\sigma_{24,\eta\eta}$	0.028	0.023	-0.005	[0.013 , 0.036]
$\sigma_{33,\eta\eta}$	0.04	0.054	0.014	[0.034 , 0.082]
$\sigma_{34,\eta\eta}$	0.028	0.031	0.003	[0.019 , 0.047]
$\sigma_{44,\eta\eta}$	0.04	0.036	-0.004	[0.024 , 0.053]

relativamente bien los valores de los parámetros verdaderos. Sin embargo, es necesario realizar un análisis más detallado y profundo con muestras pequeñas. Por ejemplo, se usaron hiperparámetros para las distribuciones a priori que aportan mucha información sobre los parámetros verdaderos. En este sentido, es conveniente realizar un análisis más exhaustivo con diferentes distribuciones a priori y métodos alternativos para recolectar información sobre los parámetros verdaderos. Por otro lado, en este análisis solo se simuló un único conjunto de datos por tamaño y orden. Para evaluar con mayor precisión la robustez de la metodología propuesta, es fundamental simular varios conjuntos de datos del mismo tamaño y orden (al menos 1000) y, de esta manera, presentar diagnósticos estadísticos que resuman el desempeño de la metodología propuesta en muestras pequeñas.

6. EJEMPLO ILUSTRATIVO USANDO DATOS REALES

En esta sección presentamos una aplicación del modelo VAR-MSV-t de las ecuaciones (2.1)-(2.9) a datos de series de tiempo macroeconómicas clave para los Estados Unidos. El objetivo es estimar los parámetros del modelo y encontrar el orden que mejor se ajuste a los datos. Además, proporcionar una breve interpretación de los parámetros estimados.

6.1. Datos

La primera variable es la producción industrial: índice total (ajustado estacionalmente). Este índice mide la producción real de sectores como la minería, la manufactura, servicios públicos de electricidad y gas. Es un indicador macroeconómico importante para economistas e inversionistas debido a que se cree que las fluctuaciones dentro del sector industrial explican la mayor parte de la variación en el crecimiento económico general. Se escoge esta variable porque se cree que cambios en la demanda agregada conducen a que la industria cambie su producción, implicando cambios en la demanda de mano de obra que alteran las tasas de desempleo [47].

La segunda variable es la tasa de desempleo (ajustada estacionalmente). Mide el número de personas que buscan activamente un trabajo sin encontrarlo, como porcentaje o fracción de la fuerza laboral total. Se escoge esta variable debido a que, según la teoría económica, la mayoría de las personas obtienen la mayor parte de los ingresos trabajando en una profesión. Con estos ingresos las personas compran productos a las empresas. A su vez, una parte de los ingresos recibidos por las empresas se distribuye en el pago de los salarios de los trabajadores. Además, existe evidencia de que el desempleo responde de forma asimétrica a los cambios de producción [55].

La tercera variable es el índice del precio al consumidor para todos los consumidores urbanos. Es un índice de precios de una canasta de bienes y servicios pagados por los consumidores urbanos y constituye aproximadamente el 88 % de la población total. Los cambios porcentuales en el índice miden la tasa de infla-

Tabla 11: Media posterior, sesgo e intervalos de credibilidad del 95 % de los elementos de las últimas dos filas de los parámetros A_1 , A_2 , A_3 , A_4 y A_5 .

	Parámetro	Media	Sesgo	Intervalos de credibilidad del 95 %
A_{11}^1	0.13	0.101	-0.029	[0.025 , 0.177]
A_{12}^1	0.09	0.093	0.003	[0.025 , 0.161]
A_{13}^1	0.09	0.136	0.046	[0.065 , 0.208]
A_{14}^1	0.09	0.100	0.010	[0.027 , 0.173]
A_{11}^2	0.11	0.178	0.068	[0.112 , 0.244]
A_{12}^2	0.07	0.005	-0.065	[-0.049 , 0.060]
A_{13}^2	0.07	0.031	-0.039	[-0.025 , 0.087]
A_{14}^2	0.07	0.026	-0.044	[-0.032 , 0.083]
A_{11}^3	0.09	0.105	0.015	[0.044 , 0.166]
A_{12}^3	0	0.007	0.007	[-0.036 , 0.050]
A_{13}^3	0	0.021	0.021	[-0.024 , 0.065]
A_{14}^3	0	-0.008	-0.008	[-0.055 , 0.037]
A_{11}^4	0.07	0.100	0.030	[0.044 , 0.156]
A_{12}^4	0	0.001	0.001	[-0.035 , 0.037]
A_{13}^4	0	-0.005	-0.005	[-0.042 , 0.032]
A_{14}^4	0	0.002	0.002	[-0.035 , 0.040]
A_{11}^5	0.05	0.035	-0.015	[-0.016 , 0.086]
A_{12}^5	0	-0.010	-0.010	[-0.041 , 0.020]
A_{13}^5	0	-0.007	-0.007	[-0.038 , 0.025]
A_{14}^5	0	0.010	0.010	[-0.022 , 0.042]
A_{21}^1	0.09	0.055	-0.035	[-0.011 , 0.121]
A_{22}^1	0.13	0.165	0.035	[0.088 , 0.242]
A_{23}^1	0.09	0.149	0.059	[0.080 , 0.218]
A_{24}^1	0.09	0.064	-0.026	[-0.006 , 0.136]
A_{21}^2	0.07	0.068	-0.002	[0.015 , 0.121]
A_{22}^2	0.11	0.091	-0.019	[0.023 , 0.159]
A_{23}^2	0.07	0.091	0.021	[0.037 , 0.146]
A_{24}^2	0.07	0.031	-0.039	[-0.024 , 0.087]
A_{21}^3	0	-0.026	-0.026	[-0.068 , 0.017]
A_{22}^3	0.09	0.119	0.029	[0.058 , 0.181]
A_{23}^3	0	0.000	0.000	[-0.043 , 0.044]
A_{24}^3	0	0.001	0.001	[-0.044 , 0.045]
A_{21}^4	0	-0.003	-0.003	[-0.039 , 0.033]
A_{22}^4	0.07	0.046	-0.024	[-0.011 , 0.102]
A_{23}^4	0	0.002	0.002	[-0.035 , 0.038]
A_{24}^4	0	-0.004	-0.004	[-0.041 , 0.033]
A_{21}^5	0	0.012	0.012	[-0.018 , 0.043]
A_{22}^5	0.05	0.039	-0.011	[-0.014 , 0.092]
A_{23}^5	0	0.014	0.014	[-0.017 , 0.045]
A_{24}^5	0	0.001	0.001	[-0.031 , 0.032]

Tabla 12: Media posterior, sesgo e intervalos de credibilidad del 95 % de los elementos de las últimas dos filas de los parámetros A_1 , A_2 , A_3 , A_4 y A_5 .

	Parámetro	Media	Sesgo	Intervalos de credibilidad del 95 %
A_{31}^1	0.09	0.077	-0.013	[0.011 , 0.142]
A_{32}^1	0.09	0.105	0.015	[0.038 , 0.171]
A_{33}^1	0.13	0.150	0.020	[0.073 , 0.229]
A_{34}^1	0.09	0.114	0.024	[0.043 , 0.185]
A_{31}^2	0.07	0.061	-0.009	[0.009 , 0.113]
A_{32}^2	0.07	0.045	-0.025	[-0.007 , 0.097]
A_{33}^2	0.11	0.145	0.035	[0.076 , 0.214]
A_{34}^2	0.07	0.008	-0.062	[-0.046 , 0.062]
A_{31}^3	0	-0.008	-0.008	[-0.050 , 0.034]
A_{32}^3	0	0.002	0.002	[-0.040 , 0.043]
A_{33}^3	0.09	0.071	-0.019	[0.010 , 0.131]
A_{34}^3	0	0.012	0.012	[-0.031 , 0.055]
A_{31}^4	0	-0.003	-0.003	[-0.038 , 0.032]
A_{32}^4	0	-0.003	-0.003	[-0.037 , 0.032]
A_{33}^4	0.07	0.012	-0.058	[-0.045 , 0.069]
A_{34}^4	0	0.006	0.006	[-0.030 , 0.042]
A_{31}^5	0	0.006	0.006	[-0.024 , 0.036]
A_{32}^5	0	0.022	0.022	[-0.008 , 0.052]
A_{33}^5	0.05	0.056	0.006	[0.004 , 0.108]
A_{34}^5	0	0.003	0.003	[-0.027 , 0.034]
A_{41}^1	0.09	0.045	-0.045	[-0.021 , 0.110]
A_{42}^1	0.09	0.084	-0.006	[0.018 , 0.150]
A_{43}^1	0.09	0.127	0.037	[0.059 , 0.195]
A_{44}^1	0.13	0.197	0.067	[0.120 , 0.273]
A_{41}^2	0.07	0.049	-0.021	[-0.004 , 0.101]
A_{42}^2	0.07	0.043	-0.027	[-0.010 , 0.096]
A_{43}^2	0.07	0.046	-0.024	[-0.007 , 0.100]
A_{44}^2	0.11	0.086	-0.024	[0.018 , 0.155]
A_{41}^3	0	0.037	0.037	[-0.005 , 0.079]
A_{42}^3	0	0.009	0.009	[-0.033 , 0.051]
A_{43}^3	0	0.000	0.000	[-0.043 , 0.044]
A_{44}^3	0.09	0.098	0.008	[0.037 , 0.159]
A_{41}^4	0	-0.014	-0.014	[-0.049 , 0.021]
A_{42}^4	0	-0.006	-0.006	[-0.041 , 0.029]
A_{43}^4	0	0.007	0.007	[-0.029 , 0.043]
A_{44}^4	0.07	0.057	-0.013	[0.001 , 0.112]
A_{41}^5	0	-0.005	-0.005	[-0.035 , 0.025]
A_{42}^5	0	-0.010	-0.010	[-0.040 , 0.020]
A_{43}^5	0	0.006	0.006	[-0.024 , 0.037]
A_{44}^5	0.05	0.043	-0.007	[-0.008 , 0.094]

Tabla 13: Media posterior e intervalos de credibilidad del 95 % de ν y los elementos del término intercepción ν .

	Parámetro	Media	Sesgo	Intervalos de credibilidad del 95 %
ν	12	17.382	5.382	[11.104 , 28.111]
ν_1	0	0.000	0.000	[-0.077 , 0.077]
ν_2	0	-0.059	-0.059	[-0.134 , 0.015]
ν_3	0	0.011	0.011	[-0.060 , 0.083]
ν_4	0	-0.086	-0.086	[-0.156 , -0.016]

ción entre dos períodos de tiempo cualesquiera, por lo que se puede utilizar para reconocer períodos de inflación y deflación. Se escoge esta variable porque se cree que incrementos en la volatilidad de la inflación son asociados con un incremento en el desempleo y una reducción en el nivel de la producción industrial [41]. Además, existe evidencia de un efecto negativo de los choques de inflación en los rendimientos de las acciones [19].

La cuarta variable es el índice del mercado de valores S&P 500. Se considera un indicador del mercado de acciones de gran capitalización estadounidense e incluye 500 empresas líderes en la industria de la economía estadounidense. Se escogió esta variable por la creciente evidencia sobre la influencia de las variables financieras en las macroeconómicas [7, 63]. Además, se cree que cuando los choques al sistema financiero aumentan la asimetría de la información (una situación en la que una parte de un contrato financiero tiene información mucho menos precisa que la otra parte [39]), de modo que las fricciones financieras (imperfecciones en el mercado financiero [21]) aumentan drásticamente, se produce inestabilidad financiera y el sistema financiero ya no puede canalizar fondos hacia aquellos con oportunidades de producción productiva, con el resultado de que la economía puede experimentar una grave recesión económica [40].

Los datos son de frecuencia mensual, abarcando un período desde enero de 1947 hasta junio de 2019, consistiendo en 869 observaciones. Se usa la ecuación $r_t = 100 \log\left(\frac{y_t}{y_{t-1}}\right)$ para convertir las observaciones a rendimientos consistentes en: $r_{1t} = r_{PIt}$ = rendimiento de la producción industrial: índice total (ajustado estacionalmente), $r_{2t} = r_{TDt}$ = rendimiento de la tasa de desempleo (ajustada estacionalmente), $r_{3t} = r_{IPCt}$ = rendimiento del índice del precio al consumidor para todos los consumidores urbanos y $r_{4t} = r_{ISPt}$ = rendimiento del índice del mercado de valores S&P 500. De ahora en adelante se usará esta notación para referirse a los rendimientos de las series cuando se requiera.

En la Figura 2 se presentan los gráficos temporales mensuales de los rendimientos de las cuatro variables. Se observan algunos rendimientos con valores muy grandes en relación con los demás; al menos 10 valores por serie superan 3 veces la desviación estándar respecto a la media, por lo que, en un modelo con distribución normal, se les asignaría probabilidad casi cero [40]. Además, se observan períodos

de tiempo (por ejemplo, 1950-1960, 1970-1980 y 2000-2010) donde la variabilidad de los rendimientos es mayor que en otros, lo cual puede considerarse evidencia de que en esos períodos hay mayor volatilidad.

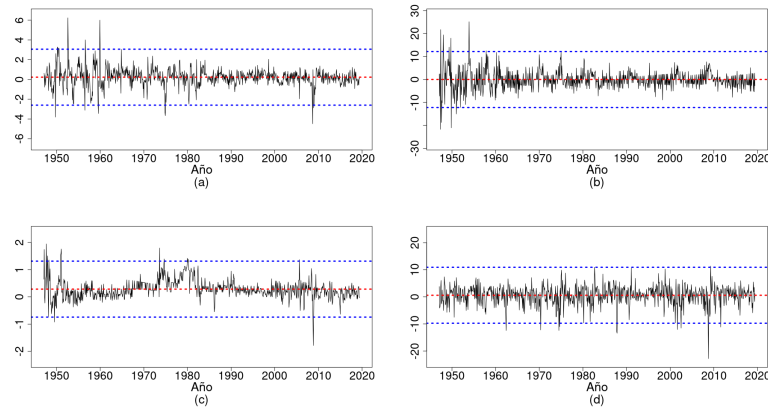


Figura 2: Gráficos temporales mensuales de los rendimientos desde febrero de 1947 hasta junio del 2019 de: (a) producción industrial, (b) tasa de desempleo, (c) índice del precio al consumidor, (d) índice del mercado de valores S&P 500. La línea roja es la media de la serie, mientras que la línea azul es tres veces la desviación estándar respecto a la media.

Para aumentar la evidencia de que hay períodos más variables, la Tabla 14 presenta la desviación estándar de los rendimientos de las cuatro variables por década (a excepción del primer y último intervalo). La desviación estándar de cada década se presenta en relación con la desviación estándar de la muestra completa, por lo que un valor inferior a uno indica un período con volatilidad relativamente baja [64]. Se observa que hay períodos donde la desviación estándar en relación con la desviación estándar de la muestra completa es mayor que uno, especialmente el período 1947-1960. Además, se observa que los aumentos en la variabilidad en los rendimientos del índice de precios al consumidor van acompañados con aumentos en la variabilidad en los rendimientos del índice S&P 500.

Estadísticos descriptivos de los rendimientos se presentan en la Tabla 15. Estos estadísticos incluyen resultados sobre la media, asimetría y curtosis. Se observa que hay exceso de curtosis en los rendimientos de las cuatro variables, variando desde 3.89 hasta 5.77, lo que aporta evidencia de que su distribución tiene colas más pesadas que la distribución normal. Para aportar mayor evidencia a la hipótesis de no normalidad, se realizó la prueba de Jarque–Bera (JB) de normalidad. El valor crítico de la prueba es 5.9914 y es menor que todos los valores críticos de la tabla, por lo que se rechaza la hipótesis de normalidad.

En las Tablas 16-17 se presentan las autocorrelaciones y autocorrelaciones cruzadas de las series r_{PI_t} , r_{TD_t} , r_{IPC_t} y r_{ISPI_t} . Se observa que las autocorrelaciones de la serie r_{IPC_t} en todos los rezagos son superiores a 0.3 por lo que está muy influen-

Tabla 14: Desviaciones estándar por década en relación con la desviación estándar de la muestra de los rendimientos de producción industrial, tasa de desempleo, índice del precio al consumidor e índice del mercado de valores S&P 500.

	1900				2000		
	47-60	61-70	71-80	81-90	91-00	01-10	11-19
r_{PIt}	1.631	0.825	1.021	0.719	0.511	0.857	0.502
r_{TDt}	1.759	0.829	0.789	0.654	0.609	0.693	0.588
r_{IPCt}	1.266	0.604	1.010	0.775	0.387	1.045	0.569
r_{ISPt}	0.874	0.946	1.102	1.084	0.832	1.288	0.805

Tabla 15: Estadísticos descriptivos de los rendimientos de producción industrial, tasa de desempleo, índice del precio al consumidor e índice del mercado de valores S&P 500.

	Media	Asimetría	Curtosis	JB
r_{PIt}	0.0023	0.1606	5.7760	1186.74
r_{TDt}	0.0000	0.4611	4.4331	727.50
r_{IPCt}	0.0028	0.4021	3.8903	560.19
r_{ISPt}	0.0063	-0.9980	3.9303	689.20

ciada por sus propios valores pasados. Las autocorrelaciones de la serie r_{ISPt} son las más pequeñas, variando entre -0.003 y 0.098 en los rezagos del 2 al 5 (excepto el primer rezago, que tiene un valor relativamente grande de 0.233). Esto concuerda con la teoría de que las series de tiempo financieras se comportan como ruido blanco. Algunas observaciones sobre las autocorrelaciones cruzadas son que varían entre -0.44 y 0.215. De las 66 autocorrelaciones cruzadas, 15 son menores que -0.1 mientras que solo 6 son mayores que 0.1, por lo que aproximadamente el 68 % de las autocorrelaciones cruzadas están entre -0.1 y 0.1. Este breve análisis de las autocorrelaciones y autocorrelaciones cruzadas sugiere que muchos de los elementos de las matrices de coeficientes serán aproximadamente cero y sus intervalos de credibilidad probablemente contengan el cero (la mayoría fuera de las diagonales). Se analizó cada serie de retornos aplicando un modelo autorregresivo univariado con volatilidad estocástica, efecto de apalancamiento cruzado y distribución t de Student. Se estimaron los parámetros por medio de la función `svtlsample` de la librería `stochvol` y `factorstochvol` de R [23, 27]. Se encontró que los parámetros estimados ν de cada serie son 16.8 para r_{PIt} , 14.8 para r_{TDt} , 19.4 para r_{IPCt} y 16.32 para r_{ISPt} . Esto sugiere la posibilidad de colas pesadas en los retornos.

6.2. Configuración de hiperparámetros

Los modelos VAR no son modelos parsimoniosos. Tienen una gran cantidad de coeficientes. En datos macroeconómicos mensuales, el número de observaciones de cada variable podría ser menor que mil. Sin información a priori, es difícil obtener estimaciones precisas de tantos coeficientes [29]. Para datos financieros,

Tabla 16: Correlaciones de las series temporales de los rendimientos r_{PIt} , r_{TDt} , r_{IPCt} y r_{ISPt} .

h	r_{PIt}	r_{TDt}	r_{IPCt}	r_{ISPt}
± 1	0.385	0.105	0.574	0.233
± 2	0.257	0.216	0.413	-0.003
± 3	0.212	0.146	0.382	0.027
± 4	0.155	0.093	0.364	0.059
± 5	0.047	0.145	0.355	0.098

Tabla 17: Correlaciones cruzadas de las series temporales de los rendimientos r_{PIt} , r_{TDt} , r_{IPCt} y r_{ISPt} .

h	r_{PIt}	r_{PIt}	r_{PIt}	r_{TDt}	r_{TDt}	r_{IPCt}
	r_{TDt}	r_{IPCt}	r_{ISPt}	r_{IPCt}	r_{ISPt}	r_{ISPt}
-5	-0.157	0.006	-0.025	-0.041	0.061	-0.049
-4	-0.194	0.031	-0.018	-0.080	0.052	-0.028
-3	-0.225	0.046	-0.017	-0.037	0.017	-0.048
-2	-0.265	0.048	0.007	-0.027	-0.031	-0.062
-1	-0.313	-0.016	0.027	-0.023	0.022	-0.119
0	-0.440	0.005	0.029	-0.035	-0.085	-0.043
1	-0.253	0.004	0.117	0.018	-0.095	0.049
2	-0.182	-0.040	0.215	-0.021	-0.103	-0.007
3	-0.108	-0.068	0.189	0.047	-0.149	-0.050
4	-0.067	-0.097	0.172	0.107	-0.134	-0.054
5	-0.057	-0.110	0.152	0.024	-0.111	-0.051

existe evidencia empírica de que cuando el número de observaciones es menor que mil, la elección de los hiperparámetros influye bastante en las distribuciones posteriores.

Nuestro modelo es un modelo VAR integrado con un modelo MSV y errores de cola pesada, por lo que la cantidad de parámetros es todavía mayor. Además, en nuestro conjunto de datos para este ejemplo solo contamos con 869 observaciones. Para obtener información a priori adecuada que permita una correcta elección de hiperparámetros de las distribuciones a priori ν , ϕ y Σ , ajustamos a los datos modelos autorregresivos univariados con volatilidad estocástica (ar-sv), volatilidad estocástica con efecto de apalancamiento (ar-svl) y volatilidad estocástica con efecto de apalancamiento y errores t (ar-svlt), y almacenamos sus parámetros. Luego, ajustamos un modelo VAR a los datos y estimamos su matriz de covarianza. Con esta información y la propiedad 9 del Apéndice A obtuvimos los hiperparámetros. Debido a que la estimación de la media posterior del parámetro ν es sensible a la elección de su distribución a priori [44, 45], ajustamos cinco configuraciones de hiperparámetros de la siguiente manera:

Configuración 1. $\Sigma \sim \text{IW}(2p, (2p\mathbf{R}_0)^{-1})$, $\nu \sim \text{Gamma}(1, 0,05)$, $(a_i, b_i) = (20, 1,5)$, $i = 1, \dots, p$, $\mathbf{R}_0^{21} = \mathbf{R}_0^{12} = \text{diag}(\hat{\sigma}_{11,\varepsilon\eta}, 0, 0, \hat{\sigma}_{44,\varepsilon\eta})$, $\mathbf{R}_0^{22} = \text{diag}(\hat{\sigma}_{11,\eta\eta}, \hat{\sigma}_{22,\eta\eta}, \hat{\sigma}_{33,\eta\eta}, \hat{\sigma}_{44,\eta\eta})$, $\mathbf{R}_0^{11} = \hat{\Sigma}_{\varepsilon\varepsilon}$, donde $\hat{\Sigma}_{\varepsilon\varepsilon}$, $\hat{\sigma}_{11,\varepsilon\eta}$, $\hat{\sigma}_{44,\varepsilon\eta}$, $\hat{\sigma}_{ii,\eta\eta}$, $i = 1, \dots, p$, son encontrados promediando la información obtenida del análisis de los tres modelos (ar-sv, ar-svl y ar-svlt) y la matriz de covarianza del modelo VAR ajustado a los datos. Los hiperparámetros de la priori de Litterman son los usuales de la literatura $\boldsymbol{\mu}_\beta = \mathbf{0}$, $\bar{\theta} = 0,2$, $\bar{\lambda} = 0,5$, mientras que los elementos $\bar{\sigma}_{ii}$, $i = 1, \dots, p$, son tomados de la diagonal de la matriz de covarianza de $\Sigma_{\varepsilon\varepsilon}^{(j)}$ en la iteración j .

Configuración 2. $\mathbf{R}_0 = \text{diag}(\hat{\sigma}_{11,\varepsilon\varepsilon}, \hat{\sigma}_{22,\varepsilon\varepsilon}, \hat{\sigma}_{33,\varepsilon\varepsilon}, \hat{\sigma}_{44,\varepsilon\varepsilon}, \hat{\sigma}_{11,\eta\eta}, \hat{\sigma}_{22,\eta\eta}, \hat{\sigma}_{33,\eta\eta}, \hat{\sigma}_{44,\eta\eta})$, donde $\hat{\Sigma}_{\varepsilon\varepsilon}$, $\hat{\sigma}_{ii,\eta\eta}$, $i = 1, \dots, p$, son encontrados con el análisis del modelo ar-svlt y la matriz de covarianza del modelo VAR ajustado a los datos, $\nu \sim \text{Gamma}(1,2, 0,08)$. Los demás parámetros son iguales a la configuración 1.

Configuración 3. $\nu \sim \text{Gamma}(1,2, 0,08)$. \mathbf{R}_0 es como la configuración 1 pero con el análisis del modelo ar-svlt y la matriz de covarianza del modelo VAR ajustado a los datos. Los demás parámetros son iguales a la configuración 1.

Configuración 4. $\nu \sim \text{Gamma}(0,01, 0,01)$, $\mathbf{a} = (22, 22, 22, 15)$, $\mathbf{b} = (1,1, 1,1, 1,1, 1,5)$, $\bar{\lambda} = 0,25080$, ajustado con la librería BVAR de R. Los demás parámetros son iguales a la configuración 3.

Configuración 5. $\mathbf{R}_0 = \text{diag}(0, 0, 0, 0, 0,8, 0,8, 0,8, 0,8)$. Los demás parámetros son iguales a la configuración 1.

El parámetro de ajuste se configuró $K = 0.014n$. Para evitar que el método se atasque firmemente debido al rechazo excesivo, se sigue la sugerencia de [53] incrementando K unas pocas iteraciones en intervalos regulares. Nosotros usamos incrementos regulares de $0.016n$ y $0.018n$.

6.3. Resultados

Como en el ejemplo presentado en la Sección 5, ajustamos las 869 observaciones al modelo de las ecuaciones (2.1)–(2.9). Primero, para establecer la información a priori más adecuada, supusimos que el orden VAR-MSV-t no puede exceder orden 5. Luego, estimamos el modelo VAR-MSV-t de orden 5 con las cinco diferentes configuraciones de hiperparámetros. Para cada modelo generamos 130000 muestras y descartamos las primeras 40000. Con las 90000 muestras restantes de cada modelo, estimamos los parámetros y calculamos la medida DIC. En la Tabla 18 se presentan las medidas DIC del modelo VAR-MSV-t de orden 5 con las 5 configuraciones de hiperparámetros. Se observa que la menor medida DIC fue para la configuración 3, por lo que la configuración 3 de hiperparámetros nos provee la información a priori más adecuada.

Como suponemos que el orden VAR-MSV-t no puede exceder el orden 5, estimamos los modelos VAR-MSV-t del orden 1 al 5 (cada modelo con la configuración 3 de hiperparámetros). Para cada modelo generamos 130000 muestras y

Tabla 18: Estimadores del DIC, desviaciones estándar y valores más grandes y más pequeños del DIC para el modelo VAR(5)-MSV-t con las 5 configuraciones de hiperparámetros.

Configuración	5	4	3	2	1
DIC	8500.85	8547.04	8441.59	8474.74	8716.74
DIC _{mín}	8500.52	8546.89	8441.08	8474.13	8716.42
DIC _{máx}	8501.08	8547.22	8442.15	8475.22	8717.05
Desviación	0.1764	0.1306	0.3315	0.3171	0.2097

descartamos las primeras 40000. Con las 90000 muestras restantes de cada modelo, estimamos los parámetros y calculamos la medida DIC. La Tabla 19 muestra las medidas DIC, el estimador de asimetría muestral de Mardia, la curtosis de Koziol y el estimador de curtosis muestral de Koziol de los residuales de todos los modelos estimados con la configuración 3 de los hiperparámetros. El modelo que tiene la menor medida DIC es el modelo VAR(5)-MSV-t. Además, se observa que el estimador de curtosis de Koziol es 258.54, valor que es 2.35 veces menor que el estimador de curtosis muestral de Koziol, el cual es igual a 606.70. Asimismo, el valor del estimador de asimetría muestral de Mardia obtenido a partir de los residuos es 2.59, lo que sugiere que el modelo captura adecuadamente la simetría teórica del proceso subyacente.

Tabla 19: Estimadores del DIC, desviaciones estándar y valores más grandes y más pequeños del DIC de los modelos VAR(i)-MSV-t, $i = 1, \dots, 5$, con la configuración de hiperparámetros 3. Además, se incluyen el estimador de asimetría muestral Mardia, curtosis de Koziol y estimador de curtosis muestral de Koziol de cada modelo.

VAR-MSV-t	5	4	3	2	1
DIC	8441.59	8565.68	8643.67	8719.72	8854.26
DIC _{mín}	8441.08	8565.43	8643.21	8719.20	8853.82
DIC _{máx}	8442.15	8566.01	8644.33	8720.31	8854.63
Desviación estándar	0.3315	0.1895	0.3580	0.3505	0.2651
Asimetría muestral de Mardia	2.5947	2.6223	2.2738	2.2396	2.2100
Curtosis de Koziol	258.54	268.86	253.75	245.77	227.98
Curtosis muestral de Koziol	606.70	575.65	561.77	550.39	522.96

En las Tablas 20-21 se presentan la media posterior, la desviación estándar y los intervalos de credibilidad del 95% de ν y de los elementos de Φ , Σ , ν , A_1 , A_2 , A_3 , A_4 y A_5 . Además, en la Tabla 22 se presentan los elementos de la matriz de correlación ρ (calculada a partir de la matriz de covarianza Σ). Cabe resaltar que solo presentamos los parámetros estimados cuyos intervalos de credibilidad no contienen el cero. Por último, en las Figuras 3-4 se presentan los gráficos de los pesos raros o extremos y la evolución de la volatilidad estimada, respectivamente.

Lo primero que observamos es que el parámetro estimado ν es igual a 23.19, un valor moderado que sugiere la presencia de colas más pesadas que las de la distribución normal. Aunque, para valores de ν cercanos a 23, la distribución t de Student y la normal son muy similares, las diferencias aún pueden ser relevantes en las colas, especialmente al modelar eventos extremos. Por otro lado, cuando ν es muy grande (empíricamente, mayor a 30), la evidencia sugiere que no hay movimientos de alta frecuencia y que los pesos generados por la distribución de $\lambda_t^{-1/2}$ tienden a concentrarse cerca de 1 [34]. Sin embargo, en la Figura 3 se observan catorce pesos que superan tres veces la desviación estándar respecto a la media, con magnitudes entre 1.2 y 1.31. El intervalo de credibilidad del 95 % para el valor de 1.31 es [1,013, 1,936], el cual no incluye el valor 1, mientras que el intervalo correspondiente para el valor 1.30 es [0,972, 2,050], conteniendo al 1 apenas en el extremo inferior. Esto sugiere que algunos de estos pesos inusuales podrían efectivamente diferenciarse de 1. Además, estos valores atípicos coinciden con eventos inusuales ocurridos en esas fechas, como se describe a continuación.

Los pesos de mayor magnitud corresponden a las fechas de febrero de 1986 y octubre de 2008. El peso de febrero de 1986 corresponde al período conocido como la gran moderación (1984–2006), un periodo donde no existía mucha variación en las variables macroeconómicas y que se puede apreciar en la Figura 4. En la Figura 5 se observa que, durante los años 1985–1986, la volatilidad estimada de los rendimientos de producción industrial y la tasa de desempleo se mantuvo claramente por debajo de 1. Para el índice de precios al consumidor, la volatilidad superó ligeramente el valor de 1 en el último trimestre de 1985. En cuanto al índice S&P 500, la volatilidad estimada también superó el valor de 1 en el último trimestre de 1985.

En febrero de 1986, los rendimientos de la tasa de desempleo alcanzaron un brusco 7,19 %, muy por encima de los rendimientos entre 1983 y enero de 1986, que variaban entre -7,18 % y 4,08 %, con una media de -1,29 %. La volatilidad estimada de los rendimientos de la tasa de desempleo tuvo un promedio de 0.82, con una desviación estándar de 0.02 en el periodo de 1983–1986, por lo que no refleja el rendimiento tan alto de febrero de 1986, pero la distribución de $\lambda_t^{-\frac{1}{2}}$ estimó un peso de magnitud extrema de 1.31 en ese mismo mes. En febrero, marzo y abril de 1986, los rendimientos del índice de precios al consumidor tuvieron una disminución brusca de -0,18 %, -0,54 % y -0,36 %, respectivamente, muy por debajo de los rendimientos entre 1983 y enero de 1986, que variaban entre 0,1 % y 0,71 %, con una media de 0,32 %. Por otro lado, la volatilidad estimada comenzó a aumentar en el cuarto trimestre de 1984 y superó el 1 en el cuarto trimestre de 1985, alcanzando su máximo de 1.35 en marzo de 1986. Este aumento gradual en la volatilidad estimada en los rendimientos del índice de precios al consumidor indica que en ese periodo hubo movimientos en los precios de la canasta de bienes y servicios, lo que puede reflejar inestabilidad económica o factores externos que afectan el comportamiento de los precios. Comenzando el tercer trimestre de 1984, la volatilidad estimada de los rendimientos de la producción industrial comenzó a

Tabla 20: Media posterior, desviación estándar e intervalos de credibilidad del 95 % de ν y de los elementos de las matrices Φ y Σ .

	Media	Desviación estándar	Intervalos de credibilidad del 95 %
ν	23.195	8.534	[12.570,44.176]
ϕ_1	0.930	0.023	[0.878,0.967]
ϕ_2	0.997	0.002	[0.990,0.999]
ϕ_3	0.940	0.017	[0.901,0.970]
ϕ_4	0.821	0.052	[0.701,0.902]
σ_{11}	0.341	0.060	[0.238,0.479]
σ_{12}	-0.493	0.091	[-0.683,-0.324]
σ_{15}	-0.083	0.024	[-0.133,-0.041]
σ_{17}	-0.051	0.024	[-0.100,-0.008]
σ_{22}	7.826	1.661	[4.598,11.077]
σ_{24}	-0.629	0.284	[-1.208,-0.096]
σ_{25}	0.262	0.111	[0.066,0.499]
σ_{26}	-0.084	0.033	[-0.150,-0.020]
σ_{33}	0.038	0.007	[0.026,0.052]
σ_{38}	0.023	0.009	[0.006,0.041]
σ_{44}	7.356	0.692	[6.072,8.791]
σ_{48}	-0.441	0.133	[-0.723,-0.197]
σ_{55}	0.133	0.041	[0.068,0.230]
σ_{66}	0.004	0.002	[0.002,0.008]
σ_{77}	0.099	0.028	[0.053,0.166]
σ_{88}	0.147	0.044	[0.077,0.252]

crecer lentamente, indicando un aumento leve en la variabilidad de la producción (aunque siempre por debajo de 1), alcanzando en marzo de 1986 el máximo de 1.

Nuestro modelo interpreta que en febrero de 1986 hubo un evento raro que proporcionó un incremento del 31 % en la magnitud de los choques de los rendimientos de las cuatro variables. Este evento raro puede ser debido a que en 1986 la economía de los Estados Unidos comenzaba su cuarto año de expansión después de la profunda recesión de 1981–1982. La inflación y el desempleo estaban disminuyendo lentamente. Se estaban logrando nuevos avances hacia el objetivo de la estabilidad de los precios. El comportamiento de los salarios y los precios siguió estando influenciado por el impulso antiinflacionario de las políticas implementadas [4]. Sin embargo, los efectos adversos de la caída de los precios del petróleo contribuyeron de manera importante a la marcada desaceleración de la inflación en 1986 y a un aumento repentino de la tasa de desempleo de casi el doble para los trabajadores mineros. Otro sector en el que la tasa de desempleo aumentó fue el manufacturero [51, 52].

El peso de octubre del 2008 corresponde al período conocido como la gran recesión (2007–2009). Se observa en la Figura 4, y con más detalle en la Figura 6, que

Tabla 21: Media posterior, desviación estándar e intervalos de credibilidad del 95 % de los elementos del vector término intercepción ν y de las matrices de coeficientes A_1, \dots, A_5 cuyos intervalos de credibilidad no contienen el cero.

	Media	Desviación estándar	Intervalos de credibilidad del 95 %
A_{11}^1	0.206	0.038	[0.132,0.280]
A_{12}^1	-0.019	0.006	[-0.031,-0.006]
A_{14}^1	0.013	0.006	[0.002,0.024]
A_{11}^2	0.157	0.034	[0.090,0.224]
A_{14}^2	0.015	0.005	[0.004,0.025]
A_{11}^3	0.100	0.032	[0.037,0.162]
A_{14}^3	0.014	0.005	[0.005,0.023]
A_{11}^4	0.082	0.031	[0.021,0.142]
A_{21}^1	-0.996	0.135	[-1.260,-0.731]
A_{22}^1	-0.138	0.035	[-0.208,-0.069]
A_{24}^1	-0.077	0.027	[-0.129,-0.025]
A_{21}^2	-0.333	0.124	[-0.579,-0.092]
A_{22}^2	0.069	0.032	[0.007,0.131]
A_{22}^5	0.105	0.030	[0.046,0.165]
ν_3	0.054	0.014	[0.027,0.080]
A_{33}^1	0.422	0.037	[0.351,0.494]
A_{33}^2	0.090	0.039	[0.014,0.166]
A_{33}^4	0.101	0.034	[0.034,0.169]
A_{33}^5	0.126	0.031	[0.065,0.188]
ν_4	0.832	0.170	[0.501,1.165]
A_{43}^1	-1.002	0.374	[-1.738,-0.272]
A_{44}^1	0.219	0.036	[0.149,0.289]

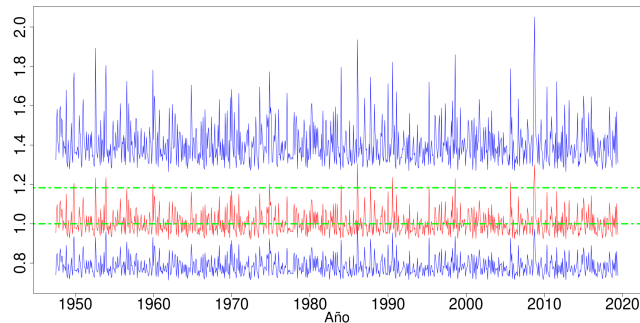


Figura 3: Gráfico temporal de $\lambda_t^{-1/2}$. La línea roja son los valores estimados, las líneas azules son sus respectivos intervalos de credibilidad del 95 %, la línea inferior verde corresponden a uno, mientras que la superior corresponde a tres veces la desviación estándar respecto a la media de los valores estimados.

durante el período 2006–2007 la tasa de desempleo se mantuvo estable, variando entre 4,4 % y 5 %, con una media de 4,61 %. Luego, en enero del 2008, comenzó a aumentar gradualmente, alcanzando un máximo de 10 % en octubre del 2009. En ese período, la volatilidad estimada en los rendimientos de la tasa de desempleo se mantuvo por debajo de 1, alcanzando su máximo de 0.94 en noviembre del 2006 y un mínimo de 0.72 en diciembre del 2009. En el caso de los rendimientos de la producción industrial, durante el período 2006–2007 los rendimientos eran estables, con una media de 1,18 % y una desviación estándar de 0,35 %. Sin embargo, en enero del 2008 se presentó una sucesión de rendimientos negativos que comenzó con -0,11 % y alcanzó un mínimo de -4,47 % en septiembre de ese mismo año. La volatilidad estimada en 2006–2007 fue inferior a 1, con un promedio de aproximadamente 0.76 y una desviación estándar de 0.06. Sin embargo, en enero del 2008 comenzó a aumentar gradualmente, alcanzando un máximo de 2.26 en octubre de ese mismo año, lo que refleja inestabilidad económica o factores externos que afectan el comportamiento de la producción. En 2006, la volatilidad estimada de los rendimientos del índice S&P 500 fue inferior a 1. Luego, en diciembre, comenzó a aumentar gradualmente, alcanzando 1 en abril del 2007 y un máximo de 2.32 en octubre del 2008. Este incremento sostenido en la volatilidad sugiere que los participantes del mercado financiero estaban reaccionando a algún evento que estaba ocurriendo.

Entre enero y octubre de 2006, la volatilidad estimada de los rendimientos del índice de precios al consumidor tuvo un promedio de 1.69 y una desviación estándar de 0.039. Luego, comenzó a decrecer, alcanzando un mínimo de 1.17 en julio de 2007. En junio del 2007 volvió a crecer gradualmente y, en febrero del 2008, aceleró el crecimiento, alcanzando un máximo de 3.2 en octubre de 2008. Este aumento acelerado en la volatilidad sugiere que, en ese período, hubo movimientos en los precios de la canasta de bienes y servicios, lo que refleja inestabilidad económica o factores externos que afectan el comportamiento de los precios.

Nuestro modelo interpreta que, en el período 2006–2010, hubo movimientos de baja frecuencia en las variables producción industrial, índice de precios al consumidor y el índice S&P 500. Además, sugiere que, en octubre de 2008, parte de la fuerza en los choques de los rendimientos de las variables se debió a un evento raro que les proporcionó un incremento del 30 % en su magnitud. Este evento raro se relaciona con el aumento de la volatilidad en octubre de 2008 en las tres variables, aunque en los rendimientos se refleja principalmente en el índice S&P 500. Este evento raro puede ser debido a que, en Estados Unidos, después de un período de varios años de aumento de las ventas y construcción de las viviendas y un rápido aumento en los precios de las viviendas, en 2006 los precios se desaceleraron drásticamente. Este evento provocó tensiones sustanciales importantes y volatilidad en los mercados financieros. Sumado a esto, aumentos en los precios del petróleo crudo y de algunas otras materias primas erosionaron los ingresos reales de los hogares y aumentaron los costos empresariales.

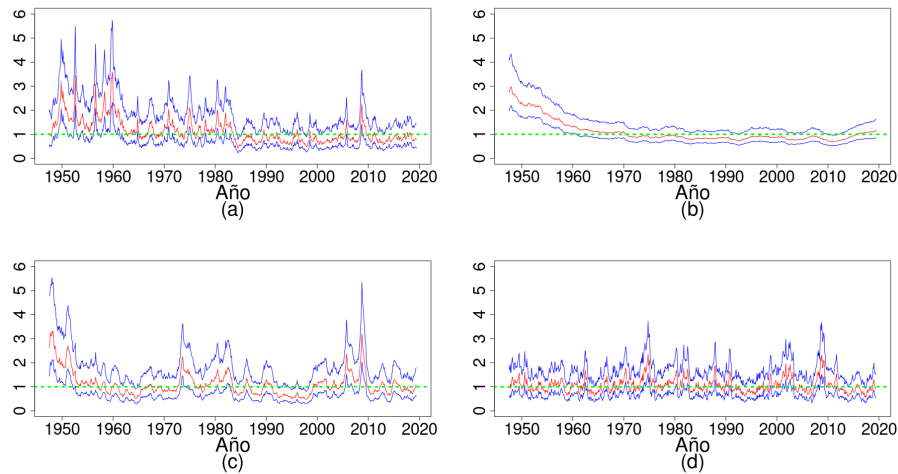


Figura 4: Gráficos temporales de la evolución de la volatilidad estimada para los rendimientos desde febrero de 1947 hasta junio del 2019 de: (a) Producción industrial: índice total (ajustado estacionalmente), (b) tasa de desempleo (ajustado estacionalmente), (c) índice del precio al consumidor para todos los consumidores urbanos e (d) índice del mercado de valores S&P 500. La línea roja es la volatilidad estimada, las líneas azules son sus respectivos intervalos de credibilidad del 95 % mientras que la línea verde corresponde a uno.

En el mercado laboral, la demanda de mano de obra desaceleró a principios de 2006 y en 2007 desaceleró aún más. Por otro lado, el comportamiento de la inflación básica en 2007 estuvo determinado por muchas de las mismas fuerzas que actuaron en 2006. Inicialmente, los efectos de los problemas en los mercados inmobiliario y financiero a otros sectores de la economía fueron limitados y la actividad económica general se mantuvo razonablemente bien, pero se desaceleró marcadamente en el cuarto trimestre de 2007 [5]. En 2008, la economía entró con poco impulso y, a mediados de ese mismo año, las tensiones en los mercados financieros se intensificaron. La economía, que en 2006–2007 se mantuvo razonablemente bien, experimentaba fuertes tensiones. Las condiciones en el mercado laboral empeoraron significativamente. Por otro lado, los precios de la energía y productos básicos cayeron bruscamente, provocando que las presiones inflacionarias disminuyeran apreciablemente. La inflación general se volvió negativa en el cuarto trimestre de 2008. En septiembre, varias empresas financieras se declararon en quiebra, entre ellas Lehman Brothers Holding. La quiebra de la empresa financiera Lehman Brothers Holding provocó que la liquidez en los mercados financieros a corto plazo desapareciera por un tiempo y los precios se desplomaran en todas las clases de activos [6].

Los parámetros de persistencia ϕ_i , $i = 1, \dots, p$, varían entre 0.8210 y 0.9966. En el caso de los rendimientos de la tasa de desempleo, el parámetro de persistencia

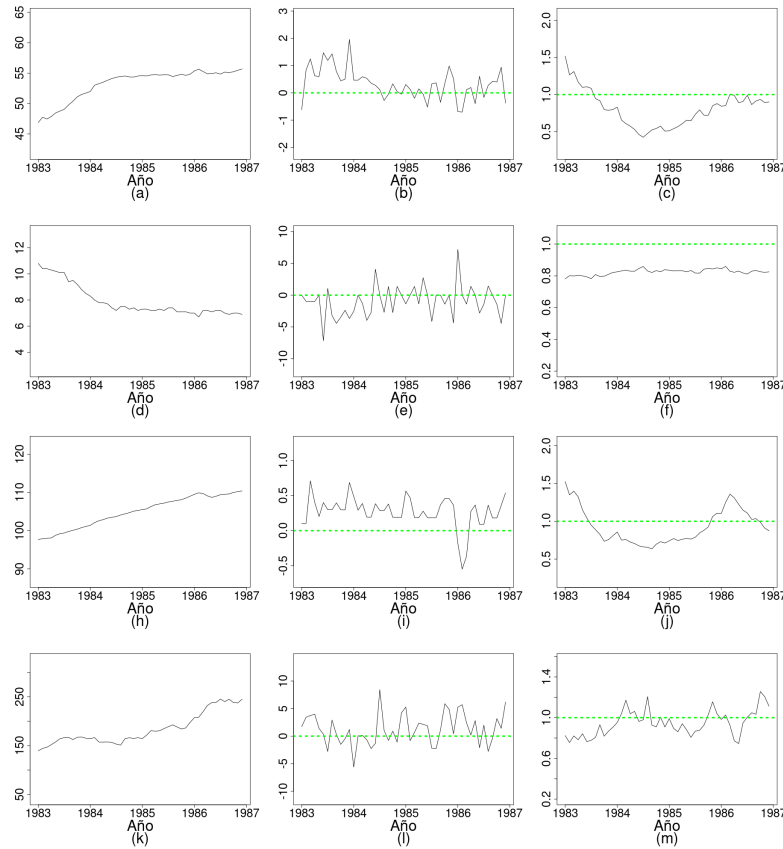


Figura 5: Gráficos temporales correspondientes al período 1983-1987. Las filas de los gráficos corresponden a las gráficas de las variables: producción industrial, tasa de desempleo, índice de precios al consumidor e índice S&P 500 respectivamente. Las columnas de los gráficos corresponden a las gráficas de: Las observaciones, rendimientos y volatilidad estimada respectivamente. Las líneas verdes horizontales corresponden al cero y al uno.

estimado es $\phi_2 = 0,9966$, un valor muy cercano a uno. Además, la varianza de la volatilidad estimada es $\sigma_{22,\eta} = 0,0039$, un valor cercano a cero, lo que indica que la evolución de la volatilidad decrece lentamente, sin movimientos bruscos, desde principios de 1947 hasta llegar a principios de 1970. Luego se mantiene en un promedio de aproximadamente 0.85, con una desviación estándar de 0.06, hasta finales de 2011. A partir de ese punto, crece gradualmente hasta finales de 2019. En el caso de los rendimientos del índice S&P 500, se observa que presenta la varianza de la volatilidad estimada más grande y el parámetro de persistencia estimado más pequeño, con valores $\sigma_{44,\eta} = 0,1467$ y $\phi_4 = 0,8210$, respectivamente. Esto indica que existe alta probabilidad de un choque relativamente grande en la volatilidad,

pero que perderá su fuerza rápidamente. Esto se observa en la Tabla 23, ya que un choque de la volatilidad pierde la mitad de su fuerza en aproximadamente 3.51 meses. Esto también se puede observar en la Figura 4-(d), donde los aumentos de la volatilidad decrecen rápidamente una vez que alcanzan su máximo.

Tabla 22: Media posterior, desviación estándar e intervalos de credibilidad del 95 % de los elementos de la matriz de correlación ρ cuyos intervalos de credibilidad no contienen el cero.

	Media	Desviación estándar	Intervalos de credibilidad del 95 %
$\rho_{12,\varepsilon\varepsilon}$	-0.304	0.034	[-0.369,-0.237]
$\rho_{11,\varepsilon\eta}$	-0.398	0.095	[-0.573,-0.208]
$\rho_{13,\varepsilon\eta}$	-0.281	0.116	[-0.499,-0.046]
$\rho_{24,\varepsilon\varepsilon}$	-0.083	0.036	[-0.153,-0.013]
$\rho_{21,\varepsilon\eta}$	0.262	0.099	[0.068,0.455]
$\rho_{22,\varepsilon\eta}$	-0.488	0.159	[-0.739,-0.127]
$\rho_{34,\varepsilon\eta}$	0.311	0.109	[0.086,0.510]
$\rho_{44,\varepsilon\eta}$	-0.429	0.105	[-0.626,-0.211]

La varianza y la persistencia de la volatilidad estimada del índice de precios al consumidor es $\sigma_{33,\eta\eta} = 0,0989$ y $\phi_3 = 0,9397$, respectivamente. En la Tabla 23, se observa que un choque de la volatilidad pierde la mitad de su fuerza en aproximadamente 11.16 meses. Esto se observa en la Figura 4-(c), donde aumentos en la volatilidad decrecen relativamente despacio una vez que alcanzan su máximo. Por último, la varianza y la persistencia de la volatilidad estimada de la producción industrial es $\sigma_{11,\eta\eta} = 0,1327$ y $\phi_1 = 0,9298$, respectivamente. En la Tabla 23, se observa que un choque de la volatilidad pierde la mitad de su fuerza en aproximadamente 9.53 meses. Esto se observa en la Figura 4-(a), donde aumentos grandes en la volatilidad decrecen relativamente despacio una vez que alcanzan su máximo.

El flujo de la volatilidad de los rendimientos de las cuatro variables varía entre 1.2523 y 1.6327, mientras que el porcentaje de la varianza general atribuible a la heterocedasticidad de los choques varía entre 27,03 % y 44,03 %. Los rendimientos del índice S&P 500 y la tasa de desempleo presentan los valores más bajos, 1.2523 y 1.3458, respectivamente, con porcentajes de varianza atribuida a la heterocedasticidad de 27,03 % y 33,10 %. En contraste, los rendimientos de la producción industrial y el índice de precios al consumidor presentan los valores más altos, 1.6327 y 1.5269, respectivamente, con porcentajes de varianza heterocedástica de 44,03 % y 40,15 %.

En la Tabla 21 se observa que los rendimientos de la producción industrial son explicados por los primeros cuatro rezagos pasados de sus propios rendimientos, por el primer rezago pasado de los rendimientos de la tasa de desempleo y por los primeros tres rezagos pasados de los rendimientos del índice S&P 500. A medida que el orden de sus propios rezagos aumenta, la magnitud de los coeficientes decrece gradualmente, desde $A_{11}^1 = 0,206$ hasta $A_{11}^4 = 0,082$. Estos coeficientes, además de

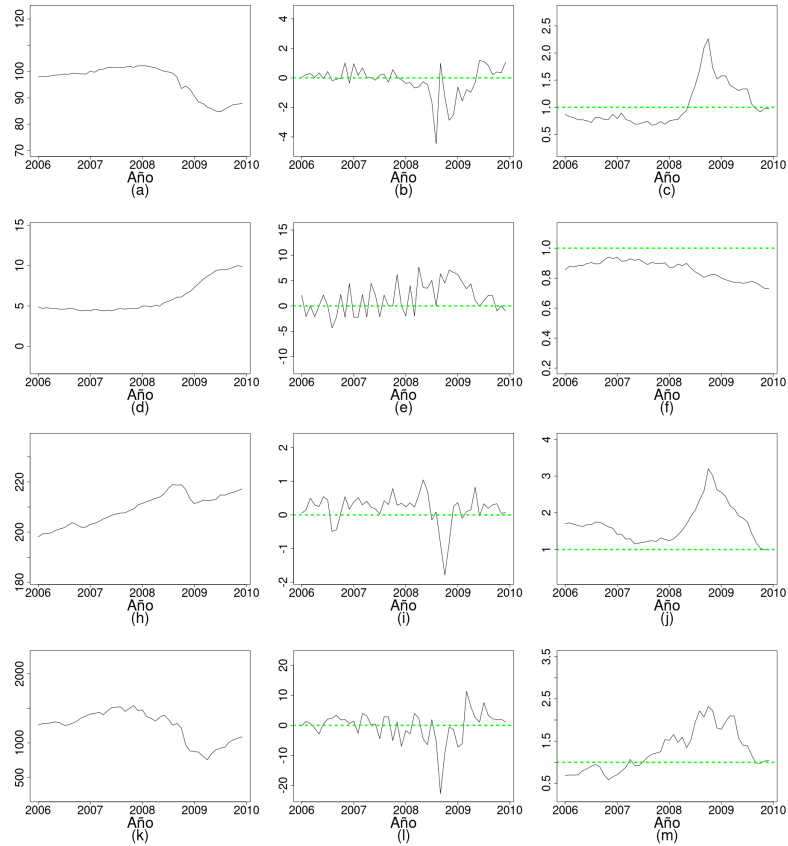


Figura 6: Gráficos temporales correspondientes al período 2006-2010. Las filas del gráfico corresponden a las gráficas de las variables: producción industrial, Tasa de desempleo, índice de precios al consumidor e índice S&P 500 respectivamente. Las columnas del gráfico corresponden a las gráficas de: Las observaciones, rendimientos y volatilidad estimada respectivamente. Las líneas verdes horizontales corresponden al cero y al uno.

ser positivos, son mayores que los coeficientes en valor absoluto de las otras series, lo que refleja una fuerte persistencia de la producción industrial en su dinámica temporal. A corto plazo, se observa un efecto débil, pero negativo del desempleo con un coeficiente $A_{12}^1 = -0,019$, coherente con el ajuste asimétrico entre producción y empleo. Por último, los coeficientes asociados a los rezagos del índice S&P 500, $A_{14}^1 = 0,013$, $A_{14}^2 = 0,015$ y $A_{14}^3 = 0,014$, aportan evidencia empírica de que las variables financieras influyen en las variables económicas.

Los rendimientos de la tasa de desempleo son explicados por el primer y segundo rezago de los rendimientos de la producción industrial, el primer, cuarto y quinto rezago de sus propios rendimientos pasados y el primer rezago de los rendimientos

Tabla 23: Flujo de la volatilidad, vida media de un choque en la volatilidad y porcentaje de la varianza que pertenece a la heterocedasticidad de los choques de los rendimientos de las variables.

	$\exp\{0,5\sigma_{\hat{u},\eta\eta}/(1-\phi_i^2)\}$	$-\log(2)/\log \phi_i $	$100(1-\sigma_{\hat{u},\varepsilon\varepsilon}/\sigma_{\hat{u},ww})$
Serie 1	1.6327	9.5310	44.0329
Serie 2	1.3458	209.4752	33.1061
Serie 3	1.5269	11.1630	40.1563
Serie 4	1.2523	3.5145	27.0377

del índice S&P 500. Los coeficientes asociados a los rezagos de la producción industrial son negativos, aproximadamente $A_{21}^1 = -0,996$ y $A_{21}^2 = -0,333$, y en valor absoluto, mayores que los rezagos de las otras series, lo que indica que la tasa de desempleo es explicada principalmente por la producción de los sectores industriales. En otras palabras, rendimientos negativos (positivos) en el primer y segundo rezago de la producción industrial implican un aumento (disminución) en los rendimientos actuales de la tasa de desempleo. Esto respalda la idea de que los cambios en la demanda agregada inducen ajustes en la producción, lo cual modifica la demanda de mano de obra y, por tanto, las tasas de desempleo. A corto y mediano plazo, los rendimientos de la tasa de desempleo también son explicados, aunque con menor fuerza, por sus propios rezagos. Por último, el primer rezago de los rendimientos del índice S&P 500 tiene un coeficiente estimado de $A_{24}^1 = -0,077$, lo que indica que rendimientos negativos (positivos) en el mes anterior tienden a aumentar (disminuir) los rendimientos de la tasa de desempleo, mostrando influencia de las variables financieras sobre el desempleo.

Los rendimientos del índice de precios al consumidor son explicados por un término de intercepción y sus propios rendimientos pasados (excepto el tercer rezago). Por otro lado, los rendimientos del índice S&P 500 son explicados por un término de intercepción, el primer rezago de los rendimientos del índice de precios al consumidor, $A_{43}^1 = -1,002$, y el primer rezago de sus propios rendimientos, $A_{44}^1 = 0,219$. En valor absoluto, el coeficiente del primer rezago de los rendimientos del índice de precios al consumidor es mayor que el de su propio rezago, por lo que los rendimientos del índice S&P 500 son explicados principalmente por los rendimientos del índice de precios al consumidor. Esto significa que un rendimiento positivo (negativo) del índice de precios al consumidor en el mes anterior implicaría una disminución (aumento) en los rendimientos del índice S&P 500. El coeficiente estimado $A_{43}^1 = -1,002$ indica un efecto negativo casi unitario del rendimiento pasado del índice de precios al consumidor sobre el rendimiento actual del S&P 500. Esto implica que un aumento (o caída) del 1% en la inflación mensual se asocia con una caída (o aumento) aproximada del 1% en los rendimientos del S&P 500. Este resultado evidencia una fuerte relación inversa entre la inflación pasada y el desempeño financiero actual.

En la Tabla 22 se observa que la correlación entre los choques de los rendimientos de la producción industrial y la tasa de desempleo es $\rho_{12,\varepsilon\varepsilon} = -0,304$, indicando una relación inversa. Esto sugiere que aumentos (disminuciones) en los choques de los rendimientos de la tasa de desempleo son influenciados por disminuciones (aumentos) en los choques de los rendimientos de la producción industrial, respectivamente. Por otro lado, los choques de los rendimientos de la producción industrial están correlacionados negativamente con los choques de la volatilidad del siguiente mes de sus propios rendimientos y los rendimientos del índice de precios al consumidor, siendo de $\rho_{11,\varepsilon\eta} = -0,398$ y $\rho_{13,\varepsilon\eta} = -0,281$, respectivamente. Esto quiere decir que la volatilidad del siguiente mes de los rendimientos de la producción industrial y el índice de precios al consumidor son influenciadas por disminuciones en los choques de los rendimientos de la producción industrial en el mes actual. Esto aporta evidencia de que los incrementos en la volatilidad de la inflación se asocian con una reducción en el nivel de producción industrial, y que los cambios en la demanda agregada inducen ajustes en la actividad productiva.

Los choques de los rendimientos de la tasa de desempleo están correlacionados con los choques de la volatilidad del siguiente mes de los rendimientos de la producción industrial y sus propios rendimientos, con correlaciones $\rho_{21,\varepsilon\eta} = 0,262$ y $\rho_{22,\varepsilon\eta} = -0,488$, respectivamente. Esto indica que aumentos actuales en los choques de la tasa de desempleo incrementan la incertidumbre sobre la actividad industrial futura, mientras que disminuciones actuales en los choques propios elevan la incertidumbre futura sobre el desempleo mismo. Además, los choques de la tasa de desempleo están débilmente correlacionados negativamente con los choques del índice S&P 500, $\rho_{24,\varepsilon\varepsilon} = -0,083$, lo que sugiere que disminuciones (aumentos) en los rendimientos bursátiles están asociadas, aunque levemente, con aumentos (disminuciones) en los choques del desempleo.

Los choques de los rendimientos del índice de precios al consumidor en el mes actual están correlacionados positivamente con los choques de la volatilidad del siguiente mes de los rendimientos del índice S&P 500, con un valor de $\rho_{34,\varepsilon\eta} = 0,311$. Esto quiere decir que la volatilidad del siguiente mes de los rendimientos del índice S&P 500 es influenciada por aumentos en los choques de los rendimientos del índice de precios al consumidor. Esto sugiere que los participantes del mercado reaccionarán en el mes siguiente a aumentos en los choques de los rendimientos del índice de precios al consumidor en el mes actual. Por otro lado, los choques de los rendimientos del índice S&P 500 del mes actual están correlacionados con los choques de la volatilidad del siguiente mes de sus propios rendimientos, siendo de $\rho_{44,\varepsilon\eta} = -0,429$. Esto quiere decir que la volatilidad del siguiente mes de los rendimientos del índice S&P 500 es influenciada por disminuciones en los choques de sus propios rendimientos en el mes actual. Esto sugiere que los participantes del mercado reaccionarán en el siguiente mes a disminuciones en los choques de los rendimientos en el mes actual.

7. CONCLUSIONES

En este artículo se propuso un modelo VAR-MSV- t para modelar series de tiempo multivariadas. El modelo VAR tiene el propósito de medir el efecto de los valores pasados de las variables involucradas en los valores actuales de cada variable. Por otro lado, la integración del modelo MSV tiene el propósito de capturar los movimientos de baja frecuencia, medir el efecto de los choques en el tiempo t de las variables endógenas sobre los choques de la volatilidad en el tiempo $t + 1$ y medir el efecto de un choque de la volatilidad actual sobre el pronóstico de la volatilidad. La integración de la distribución t de Student tiene el propósito de capturar los movimientos de alta frecuencia y, de esta forma, medir el efecto de los eventos raros en los choques de las variables endógenas.

Con el propósito de ajustar el modelo VAR-MSV- t a series de tiempo multivariadas, se propuso una metodología bayesiana para estimar los parámetros del modelo. Se aplicó la metodología bayesiana a tamaños de muestra grandes y pequeñas, simuladas con el modelo propuesto, para poder medir la precisión de los parámetros estimados. Se encontró que, con un tamaño de muestra igual a 2000, los parámetros se estiman con mucha precisión. Para poder encontrar el mejor orden VAR-MSV- t a partir de los datos, se propuso una metodología que consiste en calcular el Criterio de Información de Desviación (DIC) a cada modelo, comenzando desde el orden 1 hasta el orden máximo que se cree que no puede ser excedido. Se encontró que la metodología propuesta es eficiente para capturar el orden VAR-MSV- t en todas las muestras simuladas. En el caso de muestras pequeñas, se encontró que la metodología propuesta estima relativamente bien los parámetros del modelo.

El modelo se aplicó a series de tiempo macroeconómicas clave de los Estados Unidos, consistentes en: índice de producción industrial, tasa de desempleo e índice de precios al consumidor. Con el propósito de medir la influencia del mercado financiero sobre las variables macroeconómicas, se agregó el índice del mercado de valores S&P 500. Algunos de los hallazgos encontrados son: (a) La producción de los sectores industriales es influenciada por los rendimientos pasados del índice S&P 500 y de la tasa de desempleo. Se encontró que la mayor parte de su influencia proviene de sus propios rendimientos pasados, (b) La tasa de desempleo es influenciada por los rendimientos pasados del índice S&P 500 y sus propios rendimientos pasados. Se encontró que la mayor parte de su influencia proviene de los rendimientos pasados de la producción industrial, (c) Los precios de la canasta de bienes y servicios son influenciados por sus propios rendimientos pasados, (d) Los precios de las acciones de los mercados financieros son influenciados por los rendimientos pasados de los precios de la canasta de bienes y servicios y de sus propios rendimientos pasados.

También se encontró que los choques de los rendimientos de la producción industrial en el mes actual están correlacionados negativamente con la volatilidad de los rendimientos del siguiente mes de la producción industrial y del índice de

precios al consumidor. Además, se encontró que los choques de los rendimientos de la tasa de desempleo en el mes actual están correlacionados con la volatilidad de los rendimientos del siguiente mes de la producción industrial y de la tasa de desempleo. Por otro lado, la distribución t de Student capturó pesos de eventos raros, independientes de los choques de la volatilidad y de las variables endógenas, y coinciden con eventos que sucedieron en la economía de los Estados Unidos. Por último, se presentaron fórmulas de la asimetría de Mardia y la curtosis de Koziol del modelo. Se utilizaron datos simulados generados a partir del modelo propuesto y se encontró evidencia empírica de que, a medida que el tamaño de la muestra aumenta, tanto la asimetría de Mardia como la curtosis de Koziol convergen hacia sus valores teóricos.

Para trabajos futuros, se recomienda realizar un análisis más detallado y profundo sobre el comportamiento de la metodología en muestras pequeñas, siguiendo las sugerencias discutidas en la Sección 5. Además, sería valioso desarrollar un estudio empírico enfocado en la predicción tanto en muestras pequeñas como grandes, con el objetivo de proporcionar evidencia sobre si la incorporación de volatilidad estocástica y errores t mejora el desempeño predictivo del modelo.

8. AGRADECIMIENTOS Y FINANCIAMIENTO

Los autores desean agradecer al editor en jefe y a los dos revisores anónimos por sus valiosos comentarios y sugerencias, que ayudaron a mejorar sustancialmente la calidad de este artículo. También deseamos agradecer el apoyo institucional con el proyecto PI-699-DICIHT suscrita a la Dirección Científica, Humanística y Tecnológica, Universidad Nacional Autónoma de Honduras (UNAH).

A. APÉNDICE

Este apéndice tiene como propósito recordar algunas propiedades útiles que serán necesarias para las demostraciones de las ecuaciones (2.17)–(2.19) de la sección 2.

Observación 1. Si $\mathbf{x} = (x_1, \dots, x_p)^\top \sim N_p(\mathbf{0}, \mathbf{\Sigma})$, entonces

- i. $E(x_i x_j) = \sigma_{ij}$,
- ii. $E(x_i x_j x_k) = 0$,
- iii. $E(x_i x_j x_k x_l) = \sigma_{kl} \sigma_{ij} + \sigma_{kj} \sigma_{il} + \sigma_{jl} \sigma_{ik}$,
- iv. $E(x_i^2 x_j x_k) = \sigma_{ii} \sigma_{jk} + 2\sigma_{ij} \sigma_{ik}$,
- v. $E(x_i^2 x_j^2) = \sigma_{ii} \sigma_{jj} + 2(\sigma_{ij})^2$,
- vi. $E(x_i^3 x_j) = 3\sigma_{ii} \sigma_{ij}$,
- vii. $E(x_i^4) = 3\sigma_{ii}^2$,
- viii. $x_i + x_j + x_k + x_l \sim N(0, \sigma_{ii} + \sigma_{jj} + \sigma_{kk} + \sigma_{ll} + 2(\sigma_{ij} + \sigma_{ik} + \sigma_{il} + \sigma_{jk} + \sigma_{jl} + \sigma_{kl}))$.
Las distribuciones de $x_i + x_j$ y $x_i + x_j + x_k$ son análogas.

Observación 2. Aplicando recursividad en (2.2) y usando (2.7) se obtiene $\alpha_{t+1} = \Phi^t \alpha_1 + \sum_{i=1}^t \Phi^{i-1} \eta_{t+1-i} \sim N(0, \Phi^t \Sigma_0 \Phi^t + \sum_{i=1}^t \Phi^{i-1} \Sigma_{\eta} \Phi^{i-1}) = N(0, \Sigma_0)$.

Observación 3. $\text{cov}(\epsilon_t, \alpha_t) = \text{cov}(\epsilon_t, \Phi^{t-1} \alpha_1 + \sum_{i=1}^{t-1} \Phi^{i-1} \eta_{t-i}) = 0, t \geq 1$.

Observación 4. Si $x \sim N(\mu, \sigma^2)$, entonces $E[(x - \mu)^{n_1}] = \frac{n_1! \sigma^{n_1}}{2^{\frac{n_1}{2}} (\frac{n_1}{2})!}$ si n_1 es par y 0 si n_1 es impar. Note que, si n_1 es par, existe un entero n_1^* tal que $n_1 = 2n_1^*$ por lo que $\frac{n_1! \sigma^{n_1}}{2^{\frac{n_1}{2}} (\frac{n_1}{2})!} = \frac{(2n_1^*)! \sigma^{2n_1^*}}{2^{n_1^*} (n_1^*)!}$.

Observación 5. De [34] se tiene que si $\lambda \sim \text{Gamma}(\mu, \frac{\mu}{\Omega})$ entonces $z = \lambda^{-\frac{1}{2}}$ se distribuye Nakagami inversa con función de densidad de probabilidad

$$f(z|\mu, \Omega) = \frac{2}{\Gamma(\mu)} \left(\frac{\mu}{\Omega}\right)^{\mu} z^{-2\mu-1} \exp\left(-\frac{\mu}{\Omega z^2}\right), \quad z > 0, \mu > 0 \text{ y } \Omega > 0,$$

donde el r -ésimo momento es dado por $E(z^r) = \frac{1}{\Gamma(\mu)} \left(\frac{\mu}{\Omega}\right)^{\frac{r}{2}} \Gamma\left(\mu - \frac{r}{2}\right), \mu > \frac{r}{2}$. El caso particular donde $\mu = \frac{\nu}{2}$ y $\Omega = 1$ se obtiene $E(z^r) = \frac{1}{\Gamma(\frac{\nu}{2})} \left(\frac{\nu}{2}\right)^{\frac{r}{2}} \Gamma\left(\frac{\nu-r}{2}\right), \nu > r$. Por lo cual, el segundo y cuarto momento de $\lambda_t^{-\frac{1}{2}}$ en la ecuación (2.1) son dados por

$$E(z^2) = \frac{\left(\frac{\nu}{2}\right) \Gamma\left(\frac{\nu}{2} - 1\right)}{\Gamma\left(\frac{\nu}{2}\right)} = \frac{\nu}{\nu - 2}, \quad \nu > 2.$$

$$E(z^4) = \frac{\left(\frac{\nu}{2}\right)^2 \Gamma\left(\frac{\nu}{2} - 2\right)}{\Gamma\left(\frac{\nu}{2}\right)} = \frac{\nu^2}{(\nu - 2)(\nu - 4)}, \quad \nu > 4.$$

Observación 6. De [35] se tiene que si D y E son matrices tales que se puede formar el producto DE , entonces $(D \otimes D)(E \otimes E) = (DE \otimes DE)$.

Usando las Observaciones 1–6 y el hecho de que, si $x \in \mathbb{R}$, entonces $\exp(x) = e^x = \sum_{i=0}^{\infty} \frac{x^i}{i!}$, se desprenden las siguientes propiedades del modelo VAR-MSV-t de las ecuaciones (2.1)–(2.9). Como las demostraciones solo requieren cálculos elementales, se demuestra la primera propiedad; el resto de las propiedades se demostrarán solo cuando sea necesario.

Propiedad 1.

$$\begin{aligned} E(w_{it}w_{jt}) &= E\left(z^2 \exp\left(\frac{\alpha_{it}}{2}\right) \varepsilon_{it} \exp\left(\frac{\alpha_{jt}}{2}\right) \varepsilon_{jt}\right) = E\left(z^2 \exp\left(\frac{\alpha_{it} + \alpha_{jt}}{2}\right) \varepsilon_{it} \varepsilon_{jt}\right) \\ &= E\left(z^2 \sum_{n_1=0}^{\infty} \frac{(\alpha_{it} + \alpha_{jt})^{n_1} \varepsilon_{it} \varepsilon_{jt}}{2^{n_1} (n_1!)}\right) = E(z^2) \sum_{n_1=0}^{\infty} E\left(\frac{(\alpha_{it} + \alpha_{jt})^{n_1} \varepsilon_{it} \varepsilon_{jt}}{2^{n_1} (n_1!)}\right) \\ &= E(z^2) \sum_{n_1^*=0}^{\infty} \left(\frac{(2n_1^*)! ((\sigma_{0ii} + \sigma_{0jj} + 2\sigma_{0ij})^{\frac{1}{2}})^{2n_1^*}}{(2^{2n_1^*} (2n_1^*)!) (2^{n_1^*} (n_1^*)!)}\right) \sigma_{ij,\varepsilon\varepsilon} \\ &= \frac{\nu \sigma_{ij,\varepsilon\varepsilon}}{\nu - 2} \sum_{n_1^*=0}^{\infty} \frac{\left(\frac{\sigma_{0ii} + \sigma_{0jj} + 2\sigma_{0ij}}{8}\right)^{n_1^*}}{(n_1^*)!} = \frac{\nu \sigma_{ij,\varepsilon\varepsilon}}{\nu - 2} \exp\left(\frac{\sigma_{0ii} + \sigma_{0jj} + 2\sigma_{0ij}}{8}\right). \end{aligned}$$

Propiedad 2.

$$E(w_{it}w_{jt}w_{kt}) = E(z^3) \sum_{n_1=0}^{\infty} \left(\frac{1}{2^{n_1} (n_1!)}\right) E((\alpha_{it} + \alpha_{jt} + \alpha_{kt})^{n_1}) E(\varepsilon_{it} \varepsilon_{jt} \varepsilon_{kt}) = 0.$$

Propiedad 3.

$$E(w_{it}^2 w_{jt}) = E(z^3) \sum_{n_1=0}^{\infty} \left(\frac{1}{2^{n_1} (n_1!)}\right) E((2\alpha_{it} + \alpha_{jt})^{n_1}) E(\varepsilon_{it}^2 \varepsilon_{jt}) = 0.$$

Propiedad 4. La demostración de esta propiedad es similar a la Propiedad 1.

Sea $\sigma_1 = \sigma_{0ii} + \sigma_{0jj} + \sigma_{0kk} + \sigma_{0ll}$, $\sigma_2 = \sigma_{0ij} + \sigma_{0ik} + \sigma_{0il} + \sigma_{0jk} + \sigma_{0jl} + \sigma_{0kl}$ y $\sigma_3 = \sigma_{kl,\varepsilon\varepsilon} \sigma_{ij,\varepsilon\varepsilon} + \sigma_{k,j,\varepsilon\varepsilon} \sigma_{il,\varepsilon\varepsilon} + \sigma_{j,l,\varepsilon\varepsilon} \sigma_{ik,\varepsilon\varepsilon}$, entonces

$$E(w_{it}w_{jt}w_{kt}w_{lt}) = \frac{\nu^2 \sigma_3}{(\nu - 2)(\nu - 4)} \exp\left(\frac{\sigma_1 + 2\sigma_2}{8}\right).$$

Propiedad 5. La demostración de esta propiedad es similar a la Propiedad 1.

$$E(w_{it}^3 w_{jt}) = \frac{3\nu^2 \sigma_{ii,\varepsilon\varepsilon} \sigma_{ij,\varepsilon\varepsilon}}{(\nu - 2)(\nu - 4)} \exp\left(\frac{9\sigma_{0ii} + \sigma_{0jj} + 6\sigma_{0ij}}{8}\right).$$

Propiedad 6. La demostración de esta propiedad es similar a la Propiedad 1. Sea $\sigma_1 = \sigma_{0ii} + \sigma_{0jj} + 2\sigma_{0ij}$ y $\sigma_2 = \sigma_{ii,\varepsilon\varepsilon} \sigma_{jj,\varepsilon\varepsilon} + 2(\sigma_{ij,\varepsilon\varepsilon})^2$, entonces

$$E(w_{it}^2 w_{jt}^2) = \frac{\nu^2 \sigma_2}{(\nu - 2)(\nu - 4)} \exp\left(\frac{\sigma_1}{2}\right),$$

Propiedad 7. La demostración de esta propiedad es similar a la Propiedad 1. Sea $\sigma_1 = 4\sigma_{0ii} + \sigma_{0jj} + \sigma_{0kk}$, $\sigma_2 = 2\sigma_{0ij} + 2\sigma_{0ik} + \sigma_{0jk}$ y $\sigma_3 = \sigma_{ii,\varepsilon\varepsilon} \sigma_{jk,\varepsilon\varepsilon} + 2\sigma_{ij,\varepsilon\varepsilon} \sigma_{ik,\varepsilon\varepsilon}$, entonces

$$E(w_{it}^2 w_{jt} w_{kt}) = \frac{\nu^2 \sigma_3}{(\nu - 2)(\nu - 4)} \exp\left(\frac{\sigma_1 + 2\sigma_2}{8}\right).$$

Propiedad 8. La demostración de esta propiedad es similar a la Propiedad 1.

$$\begin{aligned} E(w_{it}^{k_1}) &= E\left(\left(z \exp\left(\frac{\alpha_{it}}{2}\right) \varepsilon_{it}\right)^{k_1}\right) = E\left[z^{k_1} \exp\left(\frac{k_1 \alpha_{it}}{2}\right) \varepsilon_{it}^{k_1}\right] \\ &= \begin{cases} \frac{1}{\Gamma(\frac{\nu}{2})} \left(\frac{\nu}{2}\right)^{\frac{k_1}{2}} \Gamma\left(\frac{\nu-k_1}{2}\right) \exp\left(\frac{k_1^2}{8} \frac{\sigma_{ii,\eta\eta}}{(1-\phi_i^2)}\right) \frac{k_1! (\sqrt{\sigma_{ii,\varepsilon\varepsilon}})^{k_1}}{2^{\frac{k_1}{2}} \left(\frac{k_1}{2}\right)!} & \text{si } k_1 \text{ es par.} \\ 0 & \text{si } k_1 \text{ es impar.} \end{cases} \end{aligned}$$

Nota: Se debe destacar que esta propiedad ya está demostrada en la literatura de modelos de volatilidad estocástica univariadas.

Propiedad 9. Por la propiedad 1 y 8 se obtiene la matriz de covarianza Σ_w de w_t .

Propiedad 10. La matriz de correlación de w_t es de la siguiente manera:

$$\rho_{ij,w_i w_j} = \exp\left(\frac{1}{8} \left(-\frac{\sigma_{ii,\eta\eta}}{(1-\phi_i^2)} - \frac{\sigma_{jj,\eta\eta}}{(1-\phi_j^2)} + \frac{2\sigma_{ij,\eta\eta}}{(1-\phi_i\phi_j)}\right)\right) \frac{\sigma_{ij,\varepsilon\varepsilon}}{\sqrt{\sigma_{ii,\varepsilon\varepsilon}\sigma_{jj,\varepsilon\varepsilon}}}.$$

Ahora se debe de demostrar que $-1 \leq \rho_{ij,w_i w_j} \leq 1$. Para poder lograr esto note que: 1) $-1 < \phi_i < 1 \implies 0 \leq \phi_i^2 < 1 \implies 0 < 1 - \phi_i^2 \leq 1$ y $0 < 1 - \phi_i\phi_j \leq 1$ y 2). El elemento ij de Σ_0 es $\sigma_{0ij} = \frac{2\sigma_{ij,\eta\eta}}{(1-\phi_i\phi_j)}$ entonces $-1 \leq \frac{\sigma_{0ij}}{\sqrt{\sigma_{0ii}\sigma_{0jj}}} \leq 1$. Entonces, realizando simples cálculos se tiene que

$$\begin{aligned} (\sqrt{\sigma_{0ii}} - \sqrt{\sigma_{0jj}})^2 \geq 0 &\implies -\sigma_{0ii} - \sigma_{0jj} + 2\sigma_{0ij} \leq 0 \\ \implies 0 &\leq \exp\left\{\frac{1}{8} \left(-\frac{\sigma_{ii,\eta\eta}}{(1-\phi_i^2)} - \frac{\sigma_{jj,\eta\eta}}{(1-\phi_j^2)} + \frac{2\sigma_{ij,\eta\eta}}{(1-\phi_i\phi_j)}\right)\right\} \leq 1 \\ \implies -1 &\leq \exp\left\{\frac{1}{8} \left(-\frac{\sigma_{ii,\eta\eta}}{(1-\phi_i^2)} - \frac{\sigma_{jj,\eta\eta}}{(1-\phi_j^2)} + \frac{2\sigma_{ij,\eta\eta}}{(1-\phi_i\phi_j)}\right)\right\} \frac{\sigma_{ij,\varepsilon\varepsilon}}{\sqrt{\sigma_{ii,\varepsilon\varepsilon}\sigma_{jj,\varepsilon\varepsilon}}} \leq 1. \end{aligned}$$

Propiedad 11. Sea $\Sigma_w = L_w L_w^T$, donde la matriz L_w es la descomposición de Cholesky. Considere el vector estandarizado $Y_w = L_w^{-1}(w_t - \mu_w) = L_w^{-1}w_t$. De [25, 28], y

por las propiedades 2-8 anteriores, se tiene que

$$\|E(\mathbf{w}_t \otimes^3)\|^2 = \sum_{i=1}^p \sum_{j=1}^p \sum_{k=1}^p E^2(w_{it}w_{jt}w_{kt}) = 0, \tag{A.1}$$

$$\|E(\mathbf{w}_t \otimes^4)\|^2 = \sum_{i=1}^p \sum_{j=1}^p \sum_{k=1}^p \sum_{l=1}^p E^2(w_{it}w_{jt}w_{kt}w_{lt}), \tag{A.2}$$

$$\begin{aligned} \|E(\mathbf{Y}_w \otimes^3)\|^2 &= \|E(\mathbf{L}_w^{-1} \mathbf{w}_t \otimes^3)\|^2 = \|(\mathbf{L}_w^{-1} \otimes^3) E(\mathbf{w}_t \otimes^3)\|^2 \\ &= \sum_{i=1}^{p^3} (\mathbf{L}_w^{1i} E(\mathbf{w}_t \otimes^3))^2 = 0, \\ \|E(\mathbf{Y}_w \otimes^4)\|^2 &= \|E(\mathbf{L}_w^{-1} \mathbf{w}_t \otimes^4)\|^2 = \|E(\mathbf{L}_w^{-1} \otimes^4)(\mathbf{w}_t \otimes^4)\|^2 \\ &= \|(\mathbf{L}_w^{-1} \otimes^4) E(\mathbf{w}_t \otimes^4)\|^2 = \sum_{i=1}^{p^4} (\mathbf{L}_w^{2i} E(\mathbf{w}_t \otimes^4))^2, \end{aligned}$$

donde \mathbf{L}_w^{1i} y \mathbf{L}_w^{2i} son la i -ésima fila de $\mathbf{L}_w^{-1} \otimes^3$ y $\mathbf{L}_w^{-1} \otimes^4$ respectivamente.

Nota: $\mathbf{A} \otimes^n = \underbrace{\mathbf{A} \otimes \mathbf{A} \otimes \dots \otimes \mathbf{A}}_{n\text{-productos}}$.

Por último, definimos \mathbf{J} , \mathbf{v}_c , \mathbf{A} y Σ_w de las ecuaciones (2.20)-(2.21). $\mathbf{J} = [\mathbf{I}_p : \mathbf{0}, \dots, \mathbf{0}]_{p \times pk}$, $\mathbf{v}_c = [\mathbf{v}^\top, \mathbf{0}^\top, \dots, \mathbf{0}^\top]_{pk \times 1}^\top$ y

$$\mathbf{A} = \begin{bmatrix} \mathbf{A}_1 & \mathbf{A}_2 & \dots & \mathbf{A}_{k-1} & \mathbf{A}_k \\ \mathbf{I}_p & \mathbf{0} & \dots & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_p & \dots & \mathbf{0} & \mathbf{0} \\ \vdots & & \ddots & \vdots & \vdots \\ \mathbf{0} & \mathbf{0} & \dots & \mathbf{I}_p & \mathbf{0} \end{bmatrix}_{pk \times pk}, \quad \Sigma_w = \begin{bmatrix} \Sigma_w & \mathbf{0}_{p \times (pk-p)} \\ \mathbf{0}_{(pk-p) \times p} & \mathbf{0}_{pk-p} \end{bmatrix}_{pk \times pk}.$$

REFERENCIAS

- [1] M. Asai y M. McAleer, *Asymetric multivariate stochastic volatility*. *Econometric Reviews* **25**(2005), no. 2-3, 453-473. DOI: [10.1080/07474930600712913](https://doi.org/10.1080/07474930600712913)
- [2] L. Benati, *The "Great Moderation" in the United Kingdom*. *Journal of Money, Credit and Banking* **40**(2008), no. 1, 121-147. DOI: [10.1111/j.1538-4616.2008.00106.x](https://doi.org/10.1111/j.1538-4616.2008.00106.x)
- [3] O. Blanchard y J. Simon, *The long and large decline in U.S. output volatility*. *Brookings Papers on Economic Activity* **40**(2001), no. 1, 135-164. DOI: [10.1353/eca.2001.0013](https://doi.org/10.1353/eca.2001.0013)
- [4] Board of Governors of the Federal Reserve System (U.S.), 1935- y Federal Reserve Board, 1914-1935, *Annual report of the Board of Governors of the Federal Reserve System (1986)*. 1986.

- [5] Board of Governors of the Federal Reserve System (U.S.), 1935- y Federal Reserve Board, 1914-1935, *Annual report of the Board of Governors of the Federal Reserve System (2007)*. 2007.
- [6] Board of Governors of the Federal Reserve System (U.S.), 1935- y Federal Reserve Board, 1914-1935, *Annual report of the Board of Governors of the Federal Reserve System (2008)*. 2008.
- [7] C.-W. Chiu, H. Mumtaz y G. Pintér, *Forecasting with VAR models: fat tails and stochastic volatility*. International Journal of Forecasting **33**(2017), no. 4, 1124-1143. DOI: [10.1016/j.ijforecast.2017.03.001](https://doi.org/10.1016/j.ijforecast.2017.03.001)
- [8] L. J. Christiano, *Comment*. Journal of Business & Economic Statistics **25** (2007), no. 2. Comentario de “On the Fit of New Keynesian Models”, 143-151. DOI: [10.1198/073500107000000061](https://doi.org/10.1198/073500107000000061)
- [9] T. E. Clark, *Real-time density forecasts from bayesian vector autoregressions with stochastic volatility*. Journal of Business & Economic Statistics **29**(2011), no. 3, 327-341. DOI: [10.1198/jbes.2010.09248](https://doi.org/10.1198/jbes.2010.09248)
- [10] T. E. Clark y F. Ravazzolo, *Macroeconomic forecasting performance under alternative specifications of time-varying volatility*. Journal of Applied Econometrics **30**(2014), no. 4, 551-575. DOI: [10.1002/jae.2379](https://doi.org/10.1002/jae.2379)
- [11] T. Cogley, *How fast can the new economy grow? A bayesian analysis of the evolution of trend growth*. Journal of Macroeconomics **27**(2005), no. 2, 179-207. DOI: [10.1016/j.jmacro.2003.11.018](https://doi.org/10.1016/j.jmacro.2003.11.018)
- [12] T. Cogley y T. J. Sargent, *Drifts and volatilities: monetary policies and outcomes in the post WWII US*. Review of Economic Dynamics **8**(2005), no. 2, 262-302. DOI: [10.1016/j.red.2004.10.009](https://doi.org/10.1016/j.red.2004.10.009)
- [13] V. Cúrdia, M. Del Negro y D. L. Greenwald, *Rare shocks, great recessions*. Journal of Applied Econometrics **29**(2014), no. 7, 1031-1052. DOI: [10.1002/jae.2395](https://doi.org/10.1002/jae.2395)
- [14] A. D’Agostino, L. Gambetti y D. Giannone, *Macroeconomic forecasting and structural change*. Journal of Applied Econometrics **28**(2013), no. 1, 82-101. DOI: [10.1002/jae.1257](https://doi.org/10.1002/jae.1257)
- [15] Q. Ding, J. Huang y H. Zhang, *The time-varying effects of financial and geopolitical uncertainties on commodity market dynamics: A TVP-SVAR-SV analysis*. Resources Policy **72**(2021), 102079. DOI: [10.1016/j.resourpol.2021.102079](https://doi.org/10.1016/j.resourpol.2021.102079)
- [16] T. Doan, R. Litterman y C. Sims, *Forecasting and conditional projection using realistic prior distributions*. Econometric Reviews **3**(1984), no. 1, 1-100. DOI: [10.1080/07474938408800053](https://doi.org/10.1080/07474938408800053)
- [17] J. Durbing y S. Koopman, *A simple and efficient simulation smoother for state space time series analysis*. Biometrika. **89**(2002), no. 3, 603-616. DOI: [10.1093/biomet/89.3.603](https://doi.org/10.1093/biomet/89.3.603)

- [18] G. Elliott, C. Granger y A. Timmermann, eds., *Handbook of economic forecasting*. Vol. 2A. Elsevier, Amsterdam, 2013.
- [19] E. F. Fama y G. Schwert, *Asset returns and inflation*. Journal of Financial Economics **5**(1977), no. 2, 115-146. DOI: [10.1016/0304-405X\(77\)90014-9](https://doi.org/10.1016/0304-405X(77)90014-9)
- [20] J. Gali y L. Gambetti, *On the sources of the Great Moderation*. American Economic Journal: Macroeconomics **1**(2009), no. 1, 26-57. DOI: [10.1257/mac.1.1.26](https://doi.org/10.1257/mac.1.1.26)
- [21] B. C. Greenwald y J. E. Stiglitz, *Financial market imperfections and business cycles*. The Quarterly Journal of Economics **108**(1993), no. 1, 77-114. DOI: [10.2307/2118496](https://doi.org/10.2307/2118496)
- [22] A. Harvey, E. Ruiz y N. Sheppard, *Multivariate stochastic variance models*. The Review of Economic Studies **61**(1994), 247-264. DOI: [10.2307/2297980](https://doi.org/10.2307/2297980)
- [23] D. Hosszejni y G. Kastner, *Modeling univariate and multivariate stochastic volatility in R with stochvol and factorstochvol*. Journal of Statistical Software **100**(2021), no. 12, 1-34. DOI: [10.18637/jss.v100.i12](https://doi.org/10.18637/jss.v100.i12)
- [24] T. Ishihara e Y. Omori, *Efficient bayesian estimation of a multivariate stochastic volatility model with cross leverage and heavy-tailed errors*. Computational Statistics & Data Analysis **56**(2012), no. 11, 3674-3689. DOI: [10.1016/j.csda.2010.07.015](https://doi.org/10.1016/j.csda.2010.07.015)
- [25] S. R. Jammalamadaka, E. Taufer y G. H. Terdik, *On multivariate skewness and kurtosis*. Sankhya A: The Indian Journal of Statistics **83**(2021), no. 2, 607-644. DOI: [10.1007/s13171-020-00211-6](https://doi.org/10.1007/s13171-020-00211-6)
- [26] P. Jong y N. Shephard, *The simulation smoother for time series models*. Biometrika. **82**(1995), no. 2, 339-350. DOI: [10.1093/biomet/82.2.339](https://doi.org/10.1093/biomet/82.2.339)
- [27] G. Kastner, *Dealing with stochastic volatility in time series using the R package stochvol*. Journal of Statistical Software **69**(2016), no. 5, 1-30. DOI: [10.18637/jss.v069.i05](https://doi.org/10.18637/jss.v069.i05)
- [28] T. Kollo, *Multivariate skewness and kurtosis measures with an application in ICA*. Journal of Multivariate Analysis **99**(2008), no. 10, 2328-2338. DOI: [10.1016/j.jmva.2008.02.033](https://doi.org/10.1016/j.jmva.2008.02.033)
- [29] G. Koop y D. Korobilis, *Bayesian multivariate time series methods for empirical macroeconomics*. Foundations and Trends[®] in Econometrics **3**(2010), no. 4, 267-358. DOI: [10.1561/0800000013](https://doi.org/10.1561/0800000013)
- [30] S. Koopman, *Disturbance smoother for state space models*. Biometrika. **80**(1993), no. 1, 117-126. DOI: [10.1093/biomet/80.1.117](https://doi.org/10.1093/biomet/80.1.117)
- [31] J. A. Koziol, *A note on measures of multivariate kurtosis*. Biometrical Journal **31**(1989), no. 5, 619-624. DOI: [10.1002/bimj.4710310517](https://doi.org/10.1002/bimj.4710310517)
- [32] R. Litterman, *Forecasting with Bayesian vector autoregressions: five years of experience*. Journal of Business & Economic Statistics. **4**(1986), no. 1, 25-38. DOI: [10.1080/07350015.1986.10509491](https://doi.org/10.1080/07350015.1986.10509491)

- [33] N. Loperfido, *Some remarks on Koziol's kurtosis*. Journal of Multivariate Analysis **175**(2020), 104565. DOI: [10.1016/j.jmva.2019.104565](https://doi.org/10.1016/j.jmva.2019.104565)
- [34] F. Louzada, P. Ramos y D. Nascimento, *The inverse Nakagami-m distribution: a novel approach in reliability*. IEEE Transactions on Reliability **67**(2018), no. 3, 1030-1042. DOI: [10.1109/TR.2018.2829721](https://doi.org/10.1109/TR.2018.2829721)
- [35] H. Lütkepohl, *Handbook of Matrices*. John Wiley, Baftins Lane, Chichester, 1996.
- [36] H. Lütkepohl, *New introduction to multiple time series analysis*. Springer Science & Business Media, Berlin, 2005. DOI: [10.1007/978-3-540-27752-1](https://doi.org/10.1007/978-3-540-27752-1)
- [37] K. V. Mardia, *Measures of multivariate skewness and kurtosis with applications*. Biometrika **57**(1970), no. 3, 519-530. DOI: [10.1093/biomet/57.3.519](https://doi.org/10.1093/biomet/57.3.519)
- [38] H. S. Migon, D. Gamerman y F. Louzada, *Statistical Inference, An Integrated Approach, Second Edition*. CRC Press, 2014. DOI: [10.1201/b17229](https://doi.org/10.1201/b17229)
- [39] F. S. Mishkin, *Understanding financial crises: a developing country perspective*. Working Paper 5600, National Bureau of Economic Research. 1996. DOI: [10.3386/w5600](https://doi.org/10.3386/w5600)
- [40] F. S. Mishkin, *Monetary policy strategy: lessons from the crisis*. Working Paper 16755, National Bureau of Economic Research. 2011. DOI: [10.3386/w16755](https://doi.org/10.3386/w16755)
- [41] D. J. Mullineaux, *Unemployment, industrial production, and inflation uncertainty in the United States*. The Review of Economics and Statistics **62**(1980), no. 2, 163-169. DOI: [10.2307/1924741](https://doi.org/10.2307/1924741)
- [42] H. Mumtaz, *A generalized stochastic volatility in mean VAR*. Economics Letters **173**(2018), 10-14. DOI: [10.1016/j.econlet.2018.08.044](https://doi.org/10.1016/j.econlet.2018.08.044)
- [43] H. Mumtaz, *A generalised stochastic volatility in mean VAR. An updated algorithm*. (2020), no. 908.
- [44] J. Nakajima e Y. Omori, *Leverage, heavy-tails and correlated jumps in stochastic volatility models*. Computational Statistics & Data Analysis **53**(2009), no. 6, 2335-2353. DOI: [10.1016/j.csda.2008.03.015](https://doi.org/10.1016/j.csda.2008.03.015)
- [45] J. Nakajima e Y. Omori, *Stochastic volatility model with leverage and asymmetrically heavy-tailed error using GH skew Student's t-distribution*. Computational Statistics & Data Analysis **56**(2012), no. 11, 3690-3704. DOI: [10.1016/j.csda.2010.07.012](https://doi.org/10.1016/j.csda.2010.07.012)
- [46] S. Ng y J. H. Wright, *Facts and challenges from the Great Recession for forecasting and macroeconomic modeling*. Journal of Economic Literature **51**(2013), no. 4, 1120-54. DOI: [10.1257/jel.51.4.1120](https://doi.org/10.1257/jel.51.4.1120)
- [47] A. M. Okun, *Potential GNP: its neasurement and significance*. American Statistical Association **10**(1963), 98-103.
- [48] E. R. Ortega y M. H. Veiga, *Modelos de volatilidad estocástica: una alternativa atractiva y factible para modelizar la evolución de la volatilidad*. Anales de estudios económicos y empresariales **18**(2008), 9-68.

- [49] G. C. Primiceri, *Time varying structural vector autoregressions and monetary policy*. The Review of Economic Studies **72**(2005), no. 3, 821-852. DOI: [10.1111/j.1467-937X.2005.00353.x](https://doi.org/10.1111/j.1467-937X.2005.00353.x)
- [50] M. L. V. Rivera, *Modelos autorregresivos vectoriales integrados con volatilidad estocástica multivariada*. Tesis de maestría. Universidad Nacional Autónoma de Honduras, (2022).
- [51] S. E. Shank, *Employment up, unemployment stable in the first half of 1986*. Monthly Labor Review **109**(1986), no. 8, 3-8.
- [52] S. E. Shank y S. E. Haugen, *The employment situation during 1986: job gains continue, unemployment dips*. Monthly Labor Review **110**(1987), no. 2, 3-10.
- [53] N. Shephard y M. Pitt, *Likelihood analysis of non-Gaussian measurement time series*. Biometrika **84**(1997), no. 3, 653-667. DOI: [10.1093/biomet/84.3.653](https://doi.org/10.1093/biomet/84.3.653)
- [54] N. Shephard y M. K. Pitt, *Time varying covariances: a factor stochastic volatility approach*. Bayesian Statistics 6, Proceedings of the Sixth Valencia International Meeting. (edited by J.M. Bernardo, J.O. Berger, A.P. Dawid and A.F.M Smith). Oxford University Press, Oxford, 1999, 547-570. DOI: [10.1093/oso/9780198504856.001.0001](https://doi.org/10.1093/oso/9780198504856.001.0001)
- [55] P. Silvapulle, I. A. Moosa y M. J. Silvapulle, *Asymmetry in Okun's law*. Canadian Journal of Economics / Revue canadienne d'Economie **37**(2004), no. 2, 353-374. DOI: [10.1111/j.0008-4085.2004.00006.x](https://doi.org/10.1111/j.0008-4085.2004.00006.x)
- [56] C. A. Sims, *Money, income, and causality*. American Economic Review **62**(1972), no. 4, 540-552.
- [57] C. A. Sims, *The role of approximate prior restrictions in distributed lag estimation*. Journal of the American Statistical Association **67**(1972), no. 337, 169-175. DOI: [10.1080/01621459.1972.10481221](https://doi.org/10.1080/01621459.1972.10481221)
- [58] C. A. Sims, *Comparison of interwar and postwar business cycles: monetarism reconsidered*. American Economic Review **70**(1980), no. 2, 250-257. DOI: [10.3386/w0430](https://doi.org/10.3386/w0430)
- [59] C. A. Sims, *Macroeconomics and reality*. Econometrica **48**(1980), no. 1, 1-48. DOI: [10.2307/1912017](https://doi.org/10.2307/1912017)
- [60] M. K. P. So, W. K. Li y K. Lam, *Multivariate modelling of the autoregressive random variance process*. Journal of Time Series Analysis **18**(1997), no. 4, 429-446. DOI: [10.1111/1467-9892.00060](https://doi.org/10.1111/1467-9892.00060)
- [61] M. K. So, K. Lam y W. Li, *An empirical study of volatility in seven Southeast Asian stock markets using ARV models*. Journal of Business Finance & Accounting **24**(1997), no. 2, 261-276. DOI: [10.1111/1468-5957.00104](https://doi.org/10.1111/1468-5957.00104)
- [62] D. J. Spiegelhalter, N. G. Best, B. P. Carlin y A. Van Der Linde, *Bayesian measures of model complexity and fit*. Journal of the Royal Statistical Society: Series B (Statistical Methodology) **64**(2002), no. 4, 583-639. DOI: [10.1111/1467-9868.00353](https://doi.org/10.1111/1467-9868.00353)

- [63] J. H. Stock y M. W. Watson, *Forecasting output and inflation: the role of asset prices*. Journal of Economic Literature **41**(2003), no. 3, 788-829. DOI: [10.1257/002205103322436197](https://doi.org/10.1257/002205103322436197)
- [64] J. H. Stock y M. W. Watson, *Has the business cycle changed and why?* NBER Macroeconomics Annual **17**(2002), 159-230. DOI: [10.3386/w9127](https://doi.org/10.3386/w9127)
- [65] C. A. C. Torres y M. L. V. Rivera, *Modelos autorregresivos vectoriales integrados con volatilidad estocástica multivariada aplicado a la economía de Estados Unidos en el periodo de 1948–2019*. Aglala **15**(2024), no. 2, 116-142.
- [66] C. A. C. Torres y M. L. V. Rivera, *curtosis-asimetria-var-msv-t*. Consultado el 23 de junio de 2025. 2025. URL: <https://github.com/MarvinVillafranca/curtosis-asimetria-var-msv-t>.
- [67] K. Triantafyllopoulos, *Time-varying vector autoregressive models with volatility*. Journal of Applied Statistics **38**(2011), no. 2, 369-382. DOI: [10.1080/02664760903406512](https://doi.org/10.1080/02664760903406512)
- [68] R. S. Tsay, *Analysis of Financial Time Series: Third Edition*. John Wiley & Sons, Hoboken, New Jersey, 2010. DOI: [10.1002/9780470644560](https://doi.org/10.1002/9780470644560)
- [69] H. Uhlig, *Bayesian vector autoregressions with stochastic volatility*. Econometrica **65**(1997), no. 1, 59-73. DOI: [10.2307/2171813](https://doi.org/10.2307/2171813)
- [70] J. Yu, *On leverage in a stochastic volatility model*. Journal of econometrics **127**(2005), no. 2, 165-178. DOI: [10.1016/j.jeconom.2004.08.002](https://doi.org/10.1016/j.jeconom.2004.08.002)